

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA QUÍMICA

ÁREA DE CONCENTRAÇÃO

DESENVOLVIMENTO DE PROCESSOS QUÍMICOS

**MODELAGEM HÍBRIDO NEURONAL APLICADA A PROCESSOS
FERMENTATIVOS**

Autora: Layse Harumi Pereira Harada

Orientador: Rubens Maciel Filho

Co-orientadora: Aline Carvalho da Costa

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-São Paulo

Fevereiro, 2001

UNICAMP
BIBLIOTECA CENTRAL
SECÃO CIRCULANTE

UNICAMP

20672702

UNIDADE 3c
N.º CHAMADA:
T/ UNICAMP
H212m
V. Er.
TOMBO BC/ 46765
PROC. 16-392/01
C D
PREC. R\$ 11,00
DATA 27/10/01
N.º CPD

CM00161034-1

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

H212m Harada, Layse Harumi Pereira
Modelagem híbrido neuronal aplicada a processos fermentativos
/ Layse Harumi Pereira Harada. --Campinas, SP: [s.n.], 2001.

Orientadores: Rubens Maciel Filho, Aline Carvalho da Costa.
Dissertação (mestrado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Química.

1. Simulação (Computadores). 2. Modelos matemáticos. 3.
Redes neurais (Computação). 4. Fermentação. 5. Otimização
matemática. I. Maciel Filho, Rubens. II. Costa, Aline Carvalho da.
III. Universidade Estadual de Campinas. Faculdade de Engenharia
Química. IV. Título.

Dissertação de Mestrado defendida por Layse Harumi Pereira Harada e aprovada em 22 de Fevereiro de 2001 pela banca examinadora constituída pelos doutores:



Prof. Dr. – Rubens Maciel Filho (Orientador)



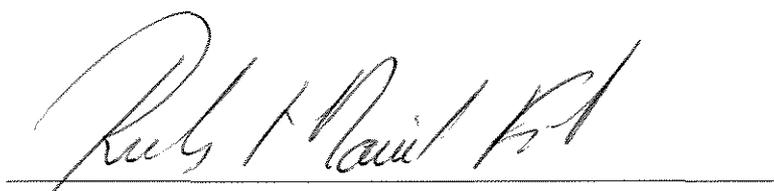
Prof(a). Dr(a). Liliane Maria Ferrareso Lona – FEQ/UNICAMP (Titular)



Dr(a). Dalva Janine Rita – RHODIA/ CPP (Titular)

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Este exemplar corresponde à versão final da Dissertação de Mestrado em Engenharia Química defendida por Layse Harumi Pereira Harada em Fevereiro de 2001.

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

Rubens Maciel Filho – Orientador

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

DEDICATÓRIA

Aos meus pais, Vivaldo e Neves.

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

AGRADECIMENTOS

Agradeço ao professor Dr. Rubens pela oportunidade, confiança e apoio durante a realização deste trabalho.

À Aline, pela dedicação, orientação e paciência para a concretização desse trabalho.

À Mami, pela força, orações, conselhos e amor desde que cheguei em Campinas.

Ao Meu irmão, Kenji, pelo amor, apoio e conselhos valiosos de sempre mesmo à distância.

Ao meu irmão Henry, ao Rodolfo e a Marci pelo amor, apoio e torcida desde sempre.

Ao Tio Geraldo pela confiança e “trotocínio” para que o Mestrado se tornasse realidade.

A tia Margarida pela força, amizade durante todo esse período.

Ao Luiz Meleiro, pela ajuda, amizade e paciência no desenvolvimento deste trabalho.

A Isabelita dos patins, Dri, Edler (cachorrão), Éderson, Marco, Rubão (meu amigão), Kelerson, Pops, Camile pela grande amizade e apoio de todas as horas.

Ao Edwin (minduim), Favinha, Urso, Luís, Pastor, Celeste, Richard, Ramón, Kathia, Lucilene, Evelyn, Kiki pela amizade, toques, e alegria de todos os dias.

A Célia, Rosa, Marta, Rita e Rafael pela ajuda nas documentações e impressões desse trabalho.

À FAPESP, pelo apoio financeiro na realização deste projeto.

E a todos dos laboratórios LOPCA/ LDPS, pelo coleguismo e descontração das horas de trabalho e àqueles que contribuíram direta ou indiretamente para a realização deste projeto.

EPÍGRAFE

“Sorte é o que acontece quando a preparação encontra a oportunidade.”

(Elmer Letterman)

RESUMO

Um dos principais problemas no controle e otimização de processos biotecnológicos é a construção de modelos confiáveis para o sistema. Para estes processos, o desenvolvimento de modelos detalhados baseado em princípios fundamentais e intensos estudos cinéticos são freqüentemente caros e consomem tempo. Assim, seria de grande vantagem encontrar algum modo simples e rápido que os descrevessem, para posterior uso em otimização e controle. Foram propostos muitos métodos recentemente para alcançar esta meta. Um deles é o uso de redes neuronais. O treinamento de uma rede neuronal, porém, requer um número grande de dados experimentais. Além disso, a interpretação de tais modelos é difícil.

Outra alternativa é o uso de um modelo híbrido neuronal no qual os aspectos do problema cujo comportamento quantitativo é bem compreendido são descritos através de equações matemáticas determinísticas, enquanto as redes neuronais descrevem a cinética. É esperado que estes modelos tenham um melhor desempenho que as redes neuronais do tipo caixa preta, já que a generalização e extrapolação estão limitadas às partes incertas do processo e o modelo básico sempre é consistente com os princípios fundamentais. Além disso, menos dados são requeridos para o treinamento.

Há muitas estruturas de redes neuronais citadas na literatura e não há nenhum método proposto que defina a melhor estrutura a ser usada para um determinado caso. Na maioria dos artigos presentes na literatura que aplicam redes neuronais para bioprocessos, a rede neuronal do tipo feedforward (FNN) é usada. Uma estrutura que não foi muito explorada é a rede *funcional link* (FLN). Esta rede demonstrou ter boa capacidade de aproximação não-linear apesar da estimação de seus pesos ser linear. Devido à estimação linear, seu treinamento é rápido, requer baixo esforço computacional e a convergência é garantida.

O objetivo do trabalho presente é determinar um modelo para um processo extrativo de produção de álcool de uma maneira simples e rápida. Este modelo deve ser suficientemente confiável para fins de otimização e estudos de controle. Um modelo híbrido neuronal é construído e redes neuronais são utilizadas para descrever as cinéticas desconhecidas. Uma comparação é feita entre o uso das FLNs e das FNNs na descrição da cinética de processo.

É demonstrado que as redes neuronais do tipo *funcional link* permitem o desenvolvimento de um modelo híbrido neuronal com uma vantagem significativa quando

comparados com redes neurais artificiais convencionais. O modelo híbrido apresenta bom desempenho e é mais simples que o obtido fazendo uso de redes neurais do tipo *feedforward*. Como a estimação dos pesos é linear para as redes FLNs, o seu uso no modelo híbrido pode possibilitar uma implementação simples de um esquema adaptativo, no qual os pesos são reestimados on-line baseados nos dados do processo.

ABSTRACT

One of the most severe problems in the control and optimization of biotechnological process is the construction of reliable models of the system. For these processes, the development of detailed models based on fundamental principles and intense kinetic studies is frequently expensive and time consuming. Thus, it would be of great advantage to find some simple and rapid way of describing them, accurately enough for optimization and control. Many methods have been proposed in recent years to achieve this goal. One of them is the use of neural networks. The training of a neural network, however, requires a large number of experimental data. Furthermore, the interpretation of such models is difficult.

Another alternative is the use of hybrid neural models, in which the aspects of the problem whose quantitative behavior is well understood are described by deterministic mathematical equations, while neural networks describe the kinetics. These models are expected to perform better than "black-box" neural network models, since generalization and extrapolation are confined only to the uncertain parts of the process and the basic model is always consistent with first principles. Besides, significantly fewer data are required for their training.

There are many neural network structures cited in the literature and there are no methods proposed to define the best structure to be used for a given case. In the majority of the papers applying neural networks to bioprocesses, the feedforward neural network (FNN) is used. A structure that has not been much explored is the functional link network (FLN). This network has been shown to have a good non-linear approximation capability although the estimation of its weights is linear. Due to the linear estimation, its training is rapid, requires low computational effort and the convergence is guaranteed.

The objective of the present work is to determine a model for an extractive ethanol production process in a simple and rapid way. This model must be sufficiently reliable to be used for posterior optimization and control studies. A hybrid neural model is constructed, using neural networks to describe the unknown kinetics. A comparison is made between the use of the FLNs and the FNNs to describe the process kinetics.

It is shown that the functional link networks enables the development of hybrid neural models with significant advantage when compared to conventional artificial neural networks. The hybrid model presents a good performance and is much simpler than the one

using the FNNs. As the estimation of the network weights in the FLNs is linear, their use in the hybrid model enables simple implementation of an adaptive scheme in which weights are reestimated on-line based on process data.

SUMÁRIO

DEDICATÓRIA V

AGRADECIMENTOS VI

EPÍGRAFE VII

RESUMO VIII

ABSTRACT X

SUMÁRIO XII

LISTA DE FIGURAS XIV

LISTA DE TABELAS XVI

NOMENCLATURA XVII

CAPÍTULO I XX

1. INTRODUÇÃO 1

CAPÍTULO II 3

2. REVISÃO BIBLIOGRÁFICA 4

2.1 FERMENTAÇÃO ALCOÓLICA EXTRATIVA 4

REDES NEURONAIIS 7

 Características Gerais 10

 Arquiteturas de Rede 12

Redes Funcional Link (FLN)	17
MODELO HÍBRIDO	20
CAPÍTULO III	25
3. SIMULAÇÃO DO PROCESSO FERMENTATIVO	26
3.1 ANÁLISE DO COMPORTAMENTO DINÂMICO DO PROCESSO	31
Estado estacionário de referência do processo convencional	31
Estudo dinâmico do processo	33
CAPÍTULO IV	40
4. DESENVOLVIMENTO, TREINAMENTO E ANÁLISE DO DESEMPENHO DAS REDES NEURONAIS	41
4.1. REDE NEURONAL <i>FEEDFORWARD</i>	41
TREINAMENTO:	41
VALIDAÇÃO	45
4.2 REDE <i>FUNCIONAL LINK</i>	51
TREINAMENTO:	52
VALIDAÇÃO	56
CAPÍTULO V	60
5. DESENVOLVIMENTO E ANÁLISE DO MODELO HÍBRIDO	61
CAPÍTULO VI	68
6. CONCLUSÕES	69
SUGESTÕES PARA TRABALHOS FUTUROS	71
REFERÊNCIAS	72

LISTA DE FIGURAS

FIGURA 1 ESTRUTURA DE UMA REDE NEURONAL ARTIFICIAL	11
FIGURA 2: ESQUEMA DE PROCESSAMENTO EM UM ÚNICO NEURÔNIO	12
FIGURA 3: REDE NEURONAL ARTIFICIAL <i>FEEDFORWARD</i> .	13
FIGURA 4: TREINAMENTO DE UMA REDE.	14
FIGURA 5: CORREÇÃO DOS PESOS DAS CONEXÕES NO TREINAMENTO DA REDE.	15
FIGURA 6: FUNÇÕES DE ATIVAÇÃO.	15
FIGURA 7: ESTRUTURA DE UMA FLN (<i>FUNCTIONAL-LINK NETWORK</i>)	19
FIGURA 8: FLUXOGRAMA GERAL DO PROCESSO DE FERMENTAÇÃO ALCOÓLICA ACOPLADA A UM <i>FLASH</i> A VÁCUO.	26
FIGURA 9: PERFIL DA CONCENTRAÇÃO E TEMPERATURA PARA OS PROCESSOS CONVENCIONAL E EXTRATIVO	32
FIGURA 10: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE: (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM F_0	35
FIGURA 11: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM R.	37
FIGURA 12: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM r .	38
FIGURA 13: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM S_0 .	39
FIGURA 14- PERTURBAÇÕES ALEATÓRIAS EM S_0	43
FIGURA 15- PERTURBAÇÕES ALEATÓRIAS EM T_0	43
FIGURA 16: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 341	43
FIGURA 17: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 351	44
FIGURA 18: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 361	44
FIGURA 19: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 371	44
FIGURA 20: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 381	45
FIGURA 21: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 391	45
FIGURA 22: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 3101	45
FIGURA 23: PERTURBAÇÃO ALEATÓRIA EM S_0	46
FIGURA 24: PERTURBAÇÃO ALEATÓRIA EM T_0	46
FIGURA 25: ERRO NO TREINAMENTO	47
FIGURA 26: ERRO NA VALIDAÇÃO	48
FIGURA 27: VALIDAÇÃO PARA A ESTRUTURA 341	48
FIGURA 28: VALIDAÇÃO PARA A ESTRUTURA 351	48
FIGURA 29: VALIDAÇÃO PARA A ESTRUTURA 361	49
FIGURA 30: VALIDAÇÃO PARA A ESTRUTURA 371	49
FIGURA 31: VALIDAÇÃO PARA A ESTRUTURA 381	49
FIGURA 32: VALIDAÇÃO PARA A ESTRUTURA 391	50
FIGURA 33: VALIDAÇÃO PARA A ESTRUTURA 3101	50
FIGURA 34: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 1	54
FIGURA 35: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 2	54
FIGURA 36: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 3	55
FIGURA 37: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 4	55
FIGURA 38: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 5	55
FIGURA 39: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 6	56
FIGURA 40. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 1	56
FIGURA 41. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 2	57
FIGURA 42. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 3	57
FIGURA 43. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 4	57
FIGURA 44. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 5	58
FIGURA 45. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 6	58
FIGURAS 46: PERTURBAÇÃO ALEATÓRIA EM S_0	61
FIGURAS 47: PERTURBAÇÃO ALEATÓRIA EM T_0	61
FIGURA 48: PERFIL DE CONCENTRAÇÃO DE BIOMASSA, X.	62
FIGURA 49: PERFIL DE CONCENTRAÇÃO DE SUBSTRATO, S.	62

FIGURA 50: PERFIL DE CONCENTRAÇÃO DE PRODUTO, P.	62
FIGURA 51: PERFIL DE TEMPERATURA, T.	63
FIGURA 52: PERFIL DA TAXA ESPECÍFICA DE CRESCIMENTO, μ .	63
FIGURA 53: PERFIL DA CONCENTRAÇÃO DE CÉLULAS, X	64
FIGURA 54: PERFIL DA CONCENTRAÇÃO DE ART, S	64
FIGURA 55: PERFIL DA CONCENTRAÇÃO DE ETOH, P	64
FIGURA 56: PERFIL DA TEMPERATURA, T	65
FIGURA 57: PERFIL DA TAXA DE CRESCIMENTO ESPECÍFICA, μ	65

LISTA DE TABELAS

TABELA 1: EXPANSÃO POLINOMIAL DE GRAU SEIS.	19
TABELA 2: PARÂMETROS USADOS NAS EQUAÇÕES	30
TABELA 3: ESTADO ESTACIONÁRIO DE REFERÊNCIA DO PROCESSO CONVENCIONAL	31
TABELA 4: ESTADO ESTACIONÁRIO DE REFERÊNCIA DO PROCESSO EXTRATIVO	33
TABELA 5: VARIÁVEIS SELECIONADAS PARA O ESTUDO DINÂMICO	34
TABELA 6. ERROS DO TREINAMENTO E VALIDAÇÃO PARA OS DIFERENTES NÚMEROS DE NEURÔNIOS NA CAMADA INTERNA.	47
TABELA 7 – DESVIO PADRÃO RESIDUAL ESCRITO EM FORMA DE PORCENTAGEM DA MÉDIA DOS VALORES CALCULADOS PELO MODELO DETERMÍNISTICO	66

NOMENCLATURA

ART – açúcares redutores totais

CC – *cascade correlation*

cor – desempenho da rede FLN

Cp – capacidade calorífica

D – taxa de diluição (h^{-1})

E – Energia de ativação

e – vetor de erros da rede FFN

F – vazão volumétrica de substrato(L/h)

f(.) – função de ativação

FFN – *feedforward neural network*

FLN – *funcional link neural network*

G - gradiente

H – matriz hessiana

h – série de monômios

J – matriz jacobiana

Ks – constante de saturação de substrato (kg/m^3)

M – espaço de dimensão maior da entrada da rede FLN

MIMO – *multiple input multiple output*

MLP – *multilayer perceptron*

n – expoente do termo de inibição pelo produto

N – número de dados experimentais

n_g – grau de monômios

n_x – tamanho do vetor de entrada da rede FLN

P - concentração de produto (g/L)

Pmáx – concentração do produto quando o crescimento celular cessa

R – taxa de reciclo de células

r – taxa de reciclo de etanol

RBF – *radial basis function*

RNN – *recursive neural network*

S – concentração de substrato (g/L)

Stt - o quadrado da diferença entre os valores experimentais e a média desses valores.

SEE – o quadrado da diferença entre os valores experimentais e os valores calculados pela rede FLN

SISO – *single input single output*

T – temperatura (°C)

t – tempo

w – peso de conexão entre os nós da rede FFN

X – concentração de biomassa(g/L)

x_e – vetor de entradas da rede FLN

y – saída calculada pela rede

y – saída da rede

y_e – saída experimental da rede FLN

$\overline{y_e}$ - média dos valores experimentais para a rede FLN

$Y_{p/s}$ – rendimento do produto baseado no consumo de substrato(kg/kg)

$Y_{x/s}$ – rendimento do crescimento de células baseado no consumo do substrato (kg/kg)

Letras Gregas

ρ - density (kg/m³)

σ - taxa cinética de consumo de substrato (kg/[m³.h])

π - taxa cinética de formação de etanol (kg/[m³.h])

μ - taxa específica de crescimento (kg/[m³.h])

ΔH – variação de entalpia

$\mu_{\text{máx}}$ – taxa máxima de crescimento (h⁻¹)

Δt – variação do tempo

Subscritos

A – referente à água

c – referente à saída da centrífuga (fase pesada)

c1 – referente à saída da centrífuga após sangria(fase pesada)

E – referente à saída da centrífuga (fase leve)

F –referente à alimentação do fermentador

i – referente à rede FLN

j – referente à rede FFN

L – referente à saída do tanque *flash*

LR – referente ao reciclo de etanol para o fermentador

LS – referente à saída do tanque *flash* enviado à destilaria

o – referente ao estado inicial

P – referente à sangria realizada após a centrifuga

r – referente ao reciclo de células

V – referente à saída do tanque *flash* enviado à retificação

Sobrescritos

(k) –referente às redes neuronais FLN e FFN

T – transposta da matriz

CAPÍTULO I

INTRODUÇÃO

Um dos maiores desafios no controle e otimização de bioprocessos é a determinação de modelos confiáveis do sistema que apresentem soluções compatíveis com as suas aplicações. Este é um ponto crítico nas aplicações em tempo real e esta dificuldade está associada à natureza complexa do metabolismo microbiano e à natureza altamente não linear da sua cinética. A dificuldade é maior quando se trata de processos em batelada ou batelada alimentada, já que, sendo estes processos transientes, suas variáveis apresentam, normalmente, grandes mudanças com o tempo. Assim, um modelo adotado para otimização e/ou controle deve descrever o comportamento do processo na faixa de condições operacionais desejada e ainda apresentar uma solução rápida e que não exija um esforço computacional exagerado.

Sendo a modelagem de bioprocessos difícil, ela é normalmente demorada e cara, e sendo assim, nem sempre modelos detalhados baseados em princípios fundamentais e estudos cinéticos intensos estão prontamente disponíveis. Assim, surge a necessidade de se encontrar maneiras simples e rápidas de desenvolver modelos confiáveis que possam ser usados para a otimização e controle destes processos.

Muitas opções têm sido propostas nos últimos anos para atingir esta meta, entre elas o uso de redes neuronais (DI MASSIMO *et al.*, 1992; SIMUTIS *et al.*, 1993; GLASSEY *et al.*, 1994) que oferecem uma ferramenta para o uso direto dos dados do processo para gerar relações de entrada-saída. Uma outra opção é o uso de modelos híbrido neuronais. Como foi demonstrado por PSICHOGIOS e UNGAR (1992), o fator crítico na determinação do comportamento dinâmico de bioprocessos é a cinética desconhecida. Nos modelos híbrido neuronais, os aspectos do problema cujo comportamento quantitativo é bem entendido são descritos por equações matemáticas determinísticas (as equações de balanço), enquanto redes neuronais descrevem a cinética. A vantagem destes modelos quando comparados a modelos somente baseados em redes neuronais é que eles são mais fáceis de analisar e interpretar, necessitam de menos dados para o treinamento das redes, são desenvolvidos mais rapidamente, interpolam e extrapolam melhor, já que a incerteza é embutida somente em uma parte do modelo (na cinética, por exemplo) e podem ser mais facilmente usados de forma adaptativa, já que o esforço para reestimação dos pesos da rede é menor (COSTA *et al.*, 1998).

O objetivo deste trabalho é estudar metodologias para a obtenção rápida e simples de modelos para bioprocessos, que devem ser confiáveis o suficiente para serem utilizados para otimização e controle. Pretende-se, ainda, estudar a aplicação das redes neuronais do tipo *functional link* (FLN) (COSTA, 1998). O caso de estudo será um processo de fermentação alcoólica extrativo. Para isso, as seguintes etapas serão cumpridas: simulação do processo de fermentação alcoólica extrativo; estudo das redes neuronais dos tipos *feedforward* e *functional link*, que serão usadas para descrever os parâmetros cinéticos do bioprocessos estudado; determinação de um modelo híbrido neuronal para a fermentação alcoólica e estudo do seu desempenho na descrição do comportamento dinâmico do processo.

CAPÍTULO II

2. REVISÃO BIBLIOGRÁFICA

2.1 Fermentação Alcoólica Extrativa

Em função da necessidade de aumentar a produção de álcool sem ter de aumentar a capacidade volumétrica instalada, é preciso aumentar a produtividade, ou seja, o teor de álcool no vinho, necessitando-se para isto concentrações maiores de açúcares na alimentação. Como a quantidade de etanol formada é uma consequência da concentração de açúcares na corrente que alimenta o fermentador, num processo em que se usam elevadas concentrações de açúcares permite-se a obtenção de um alto teor alcoólico. Isto ocasiona condições inadequadas à fisiologia da célula de levedura, pois, sendo o álcool um inibidor de fermentações, ocasiona fermentações incompletas, reduzindo o rendimento e a produtividade da fermentação alcoólica (SILVA, 1998).

Desta forma, num processo usando altas concentrações de açúcares, a eliminação do produto inibidor é essencial para se conseguir alta produtividade e alto rendimento, já que a grande restrição do processo convencional é o álcool produzido inibir as células de leveduras. Quando o substrato açucarado é fermentado e ocorre o aumento da concentração de etanol de 45 g/L para 120 g/L, a taxa específica de produção de etanol e a taxa de crescimento específico de células de leveduras são severamente diminuídas, de acordo com AIBA *et al.* (1968).

Nos últimos 20 anos, vários sistemas combinando fermentação com sistema de separação foram propostos para minimizar o problema da inibição pelo produto no processo de fermentação alcoólica. Entre eles podem ser citados:

a) Fermentação sob vácuo (RAMALINGHAM & FINN, 1977): o sistema de usar vácuo para remover álcool foi transformado em patente e publicado por Boeckeler em 1948. O sistema é operado a vácuo para manter a temperatura no fermentador em níveis desejados. O etanol tem uma volatilidade bem acima da água, facilitando a sua separação em pressão e temperatura reduzidas. MAIORELLA *et al.* (1983) alertaram que na fermentação a vácuo, a concentração de alguns produtos não voláteis, se não forem eliminados do processo por sangria, também inibe o processo, limitando o rendimento e a produtividade final da fermentação.

b) Fermentação com sistemas de duas fases líquidas (KÜHN, 1980): são obtidos misturando dois solventes imiscíveis. Os solventes são misturados em proporções adequadas, duas fases são formadas, a fase superior rica em álcool, enquanto a fase inferior conterá o concentrado de células e uma pequena parte do álcool produzido. A fase superior pode ser liberada do álcool através da destilação e posteriormente retornada para o concentrado celular (fase inferior).

c) Fermentação acoplada ao sistema de destilação *flash* à vácuo: este processo é um melhoramento da fermentação sob vácuo, pois segundo LEE *et al.* (1981) três dos grandes problemas do processo sob vácuo são a dificuldade de adicionar o oxigênio no fermentador, manter condições estéreis e comprimir uma grande quantidade de CO₂. Com estas restrições foi proposto um processo de fermentação na pressão atmosférica conectada com o *flash* a vácuo. SILVA (1998) modelou e simulou este sistema de fermentação alcoólica contínua extrativa.

d) Perextração (MATSUMURA & MARKI, 1986): técnica que combina um solvente extrator e uma membrana. O termo perextração é a junção das palavras permeável (membrana) e extração (solvente).

e) Fermentação acoplada a ultrafiltração (FERRAS *et al.*, 1986): esta técnica foi usada na fermentação de produção do solvente acetobutílico para eliminação dos produtos inibidores. Este sistema pode futuramente ser usado na retirada do etanol do meio de fermentação.

f) Fermentação acoplada a um sistema de extração líquido-líquido (SLAPACK *et al.*, 1987): o produto inibidor é seletivamente removido por solventes. O grande inconveniente da extração líquido-líquido está relacionado aos solventes. As propriedades para um solvente ideal são: alto coeficiente de distribuição em relação ao produto inibidor, baixa volatilidade em relação ao produto inibidor; ser não tóxico; relativamente barato. Conseguir agrupar todas essas propriedades em um só solvente é o grande problema a ser resolvido.

g) Fermentação acoplada ao sistema de adsorção e dessorção (SLAPACK *et al.*, 1987): esta técnica está sendo estudada para a retirada do produto inibidor da fermentação, pois o

sistema de adsorção utiliza-se de adsorventes seletivos, usando dois estágios. O etanol é primeiramente separado por adsorção em materiais sólidos (zeólitas com propriedades hidrofóbicas), e depois é removido dos poros do adsorvente por calor ou absorção gasosa.

h) Pervaporação (VICENTE *et al.*, 1993): tem como princípio a difusão das substâncias voláteis através de uma membrana hidrofóbica. A técnica consiste em usar uma membrana para remover seletivamente o produto inibidor formado na fermentação. No caso da fermentação alcoólica, o mosto líquido se difunde na superfície da membrana, e o álcool é então evaporado e removido por um fluxo fluido gasoso ou líquido e, posteriormente, o removido é condensado. Os principais tipos de membranas são de politetrafluoroetileno (PTFE) e POLY gel membranes.

i) Fermentação acoplada ao sistema de absorção (stripping) em coluna de recheio (TAYLOR *et al.*, 1996): o mosto em fermentação é continuamente enviado para a coluna empacotada, onde o etanol é removido. A técnica de absorção estudada utilizou o CO₂ produzido na fermentação como gás absorvedor.

j) Fermentação acoplada ao sistema de destilação *flash* na pressão atmosférica: ISHIDA & SHIMIZU (1996) estudaram a eliminação da inibição pelo produto através da fermentação em batelada repetida acoplada ao sistema de destilação *flash* a pressão atmosférica. Utilizaram modelagem matemática desenvolvida com base nos dados experimentais, otimizando o processo e verificando que a produtividade deste processo é superior ao processo batelada repetida convencional em quatro vezes.

MINIER e GOMA (1982) realizaram um estudo do processo de fermentação alcoólica extrativo acoplado a um sistema de extração líquido-líquido. Os autores propõem que o dodecanol é um solvente em potencial durante a fermentação com células imobilizadas e concluem que a técnica de separação por membrana para separar e recuperar o etanol do dodecanol é possível diminuindo o custo energético da fermentação.

CYSEWSKI e WILKE (1977) descrevem a fermentação alcoólica sob vácuo e com reciclo de células para *Saccharomyces cerevisiae* à temperatura de 35°C. Os autores afirmaram que esse tipo de fermentação tem a vantagem de eliminar a inibição causada pelo produto formado, o que permite soluções concentradas de açúcar serem fermentadas à

elevadas taxas. O reciclo de células em conjunto faz com que a produtividade seja doze vezes maior do que a fermentação convencional. À pressão atmosférica, a fermentação com reciclo de células aumenta a produtividade do etanol. A desvantagem é o acúmulo de componentes não voláteis no fermentador.

WOOLEY *et al.* (1999) descrevem os novos rumos que a pesquisa e o desenvolvimento de processos de produção de etanol a partir de biomateriais. Utilizam o software comercial Aspen Plus para simulação do processo e produção de etanol através da conversão enzimática. Os autores concluem que o bio-álcool tem grande potencial na concorrência atual e apresentam uma otimização do processo de forma que o custo de produção fosse viável à concorrência do mercado atual.

Redes Neurais

Redes Neurais Artificiais são técnicas computacionais relacionadas à inteligência artificial que apresentam um modelo matemático inspirado na estrutura neuronal de organismos inteligentes. Possuem aplicações nas mais diversas áreas, como engenharia e medicina, substituindo modelos determinísticos que não são bem conhecidos ou cujo tempo de processamento computacional seja relativamente grande. Além da estrutura similar às redes neurais de animais, as artificiais também adquirem conhecimento através da experiência.

Backpropagation.

É interessante observar que, apesar dos grandes desenvolvimentos com redes neurais terem acontecido na década de 80, a introdução das técnicas neurais em aplicações de engenharia química se concentraram no final da década de 80 segundo HAYKIN (1999).

RUMELHART *et al.* (1986) demonstraram o método *backpropagation* de treinamento de redes neurais, o qual ajusta os pesos das conexões dentro da rede para minimizar a medida da diferença entre o vetor saída real e o vetor predito pela rede.

BHAT e McAVOY (1990) aplicaram redes neurais como ferramenta para modelagem dinâmica e controle de processos químicos. O algoritmo *backpropagation* foi utilizado na modelagem da resposta dinâmica do pH dentro de um reator de mistura perfeita (CSTR). Esses autores chamaram atenção para o fato de que uma rede com grande número

de pesos tem forte tendência em “sobre ajustar” os pesos e, portanto, o desempenho da rede com dados não usados para o treinamento é normalmente ruim.

Em 1990, THIBAUT *et al.*, investigaram o uso de redes neurais de estrutura *feedforward* na modelagem dinâmica e predição das principais variáveis de um bioprocessamento.

LINKO e ZHU (1991) realizaram o estudo da dinâmica de um bioprocessamento em batelada e batelada alimentada através da programação de uma rede neuronal de estrutura *feedforward* para predição das variáveis do processo. O sistema proposto estimou precisamente e rapidamente as variáveis com ou sem ruído mesmo sob a dinâmica variante do processo.

CLERAN *et al.* (1991) compararam o desempenho de modelos obtidos pelo Group Method of Data Handling (GMDH) e redes neurais *feedforward* para a identificação e predição da taxa de produção de CO₂ durante a fermentação alcoólica para a produção de vinho e concluíram que modelos preditivos obtidos por redes neurais são geralmente mais precisos.

KARIM e RIVERA (1992) realizaram um estudo comparativo entre dois tipos diferentes de redes neurais, *feedforward* e recorrente, para estimar as variáveis significativas de um processo de fermentação alcoólica por *Zymomonas mobilis*. Os dois tipos de rede descreveram bem os dados experimentais, porém as redes recorrentes apresentaram melhor capacidade de generalização.

CHEN e BILLINGS (1992) estudaram a identificação de sistemas não lineares usando redes neurais artificiais. Nesta referência, três diferentes arquiteturas de redes foram utilizadas, *MultiLayer Perceptron* (MLP), a do tipo *Radial Basis Function* (RBF) e a do tipo *Funcional Link* (FLN) e as vantagens e desvantagens dessas estruturas são discutidas usando dados reais e simulados.

YOU e NIKOLAAU (1993) aplicaram uma rede neuronal recursiva (RNN) para a modelagem estática e dinâmica não linear de um processo. Resultados da simulação mostraram que esta rede pode ser treinada tanto a partir de relações do estado estacionário quanto dinâmicas do processo batelada ou contínuo para sistemas SISO (*Single input and single output*) e MIMO (*Multiple input and multiple output*). Uma comparação do desempenho desta rede neuronal com a rede *feedforward* foi apresentada e os autores mostram que, embora as RNNs apresentem bons resultados, o tempo requerido para o seu treinamento é bem maior.

IORDACHE *et al.* (1993) apresentaram uma nova estrutura de rede neuronal organizada como um modelo poliestocástico para a caracterização da cinética de um processo de fermentação, comparando seu desempenho com uma rede neuronal de estrutura *feedforward* (*Hamming NN*). Os autores concluíram que a vantagem da nova estrutura de rede neuronal proposta é mais ou menos relacionada à hierarquia durante a classificação de dados, o que permite uma drástica redução do esforço experimental, além de apresentar flexibilidade sob condições reais.

SAVKOVIC-STEVANOVIC (1994) publicou um estudo de redes neuronais aplicado a dados de produção industrial. Os métodos de treinamento Poweel e a regra delta generalizada foram utilizados para diferentes topologias de redes *feedforward*. Os dois métodos de treinamento apresentaram pequena diferença no desempenho e foi verificado que com a adição da camada interna na estrutura *feedforward* o desempenho melhorava.

GLASSEY *et al.* (1994) utilizaram rede neuronal artificial *feedforward* na predição da concentração de biomassa para diferentes taxas de alimentação em um processo de fermentação industrial para produção de proteína recombinante obtendo resultados com níveis de precisão aceitáveis. Esses autores chamam atenção para a necessidade de se escolher um conjunto de treinamento de forma a abranger toda a região na qual o modelo é válido de forma a evitar problemas de interpolação e extrapolação.

SAXÉN e SAXÉN (1996) estudaram um modelo neuronal da cinética de um processo de fermentação em batelada por *Sacharomyces cerevisiae* comparando o desempenho das redes *feedforward* e recorrente e as mesmas foram capazes de descrever o crescimento bifásico. As redes do tipo recorrente descreveram melhor o processo estudado, uma vez que capacidade dinâmica é requerida para modelar as fases “lag” que ocorrem quando as células mudam seu metabolismo.

LEAL *et al.* (1997) realizaram a comparação entre o uso de rede neuronal artificial e de fusão de dados para estimar a biomassa na produção de antibiótico. O algoritmo de treinamento da rede utilizado foi o Levenberg-Marquardt. O objetivo do trabalho foi testar a capacidade de generalização e o desempenho na presença de ruído e falha de sensor. Os parâmetros estimados foram de exatidão comparável a erros de instrumentação para os dois métodos estudados.

KARIM *et al.* (1997) realizaram um estudo de diferentes tipos de redes neuronais para controle de bioprocessos relacionados com a produção de álcool combustível. Vários tipos de redes foram utilizados, sendo que as redes do tipo *Radial Basis Function* (RBF)

não se mostraram adequadas quando extrapolação é desejada. Entretanto, a *MultiLayer Perceptron* (MLP), Recorrente e *Cascade Correlation* (CC) demonstraram melhor capacidade de predição.

ACUÑA *et al.* (1998) utilizaram redes neuronais na elaboração de modelos estáticos e dinâmicos para estimar a concentração de biomassa *online* durante a cultura em batelada conduzida a temperatura e pH controlados. Os modelos estáticos poderiam fornecer estimações parciais diante dos valores iniciais de concentração de biomassa errados. Os modelos dinâmicos, entretanto, apesar de terem o desempenho parecido com os modelos estáticos, oferecem respostas mais estáveis devido a uma ação contínua vinda da metodologia de treinamento e método associado para estimação da biomassa.

HUSSAIN (1999) realizou uma revisão das aplicações de redes neuronais no controle de processos químicos, tais como: sensor de análise de dados, detecção de falhas e identificação de processos não lineares.

HENRIQUES *et al.* (1999) propuseram um modelo híbrido usando dados experimentais de fermentação em batelada e batelada alimentada. A rede neuronal funcional link modificada foi utilizada para estimar a cinética do processo.

Características Gerais

Uma rede neuronal artificial é composta por várias unidades de processamento, cujo funcionamento é bastante simples. Essas unidades, também chamadas de neurônios ou nós, geralmente são conectadas por canais de comunicação que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma Rede Neuronal Artificial vem das interações entre as unidades de processamento da rede.

A maioria dos modelos de redes neuronais possui alguma regra de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados. Em outras palavras, elas aprendem através de exemplos.

Arquiteturas neuronais são tipicamente organizadas em camadas, como pode ser visto na Figura 1.

Usualmente as camadas são classificadas em três grupos:

- camada de Entrada: onde os padrões são apresentados à rede;
- camadas Intermediárias ou Internas: onde é feita a maior parte do processamento, através das conexões
- camada de Saída: onde o resultado final é concluído e apresentado.

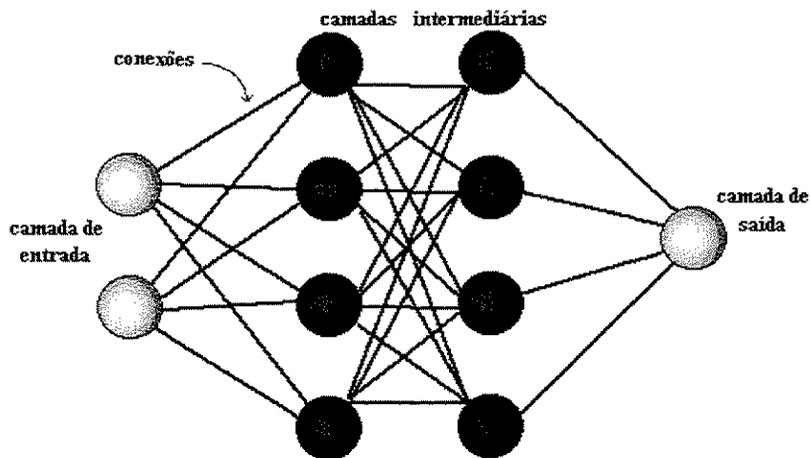


FIGURA 1 ESTRUTURA DE UMA REDE NEURONAL ARTIFICIAL

A operação de uma unidade de processamento, localizada nas camadas intermediárias ou de saída é mostrada na Figura 2 e pode ser resumida da seguinte maneira:

- sinais são apresentados à entrada;
- cada sinal é multiplicado por um número, ou peso, que indica a sua influência na saída da unidade;
- é feita a soma ponderada dos sinais que produz um nível de atividade;
- Esta soma ponderada é passada através de uma função de ativação;
- se este nível de atividade exceder um certo limite (threshold), a unidade produz uma determinada resposta de saída.

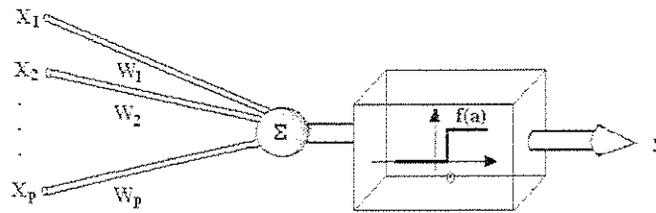


FIGURA 2: ESQUEMA DE PROCESSAMENTO EM UM ÚNICO NEURÔNIO

A camada de entrada somente distribui o sinal para a camada interna adjacente.

Uma rede neuronal é especificada principalmente pela sua topologia, pelas características dos nós e pelas regras de treinamento.

Arquiteturas de Rede

Redes Feedforward (FNN)

A estrutura de uma rede neuronal *feedforward* é mostrada na Figura 3. Neste tipo de rede o sinal de entrada é propagado através das camadas de processamento antes que a saída da rede seja calculada. Cada camada é composta de um número de nós ou neurônios, e cada um desses nós é composto por um simples combinador linear, com uma função de transferência adequada que calcula a saída do neurônio a partir do sinal de entrada. Cada neurônio tem conexões de entrada com os nós da camada anterior somente, e a saída do nó é transmitida para os nós da próxima camada. Cada nó tem um vetor peso associado que transforma linearmente seu vetor de entrada segundo Brown e Harris (1994).

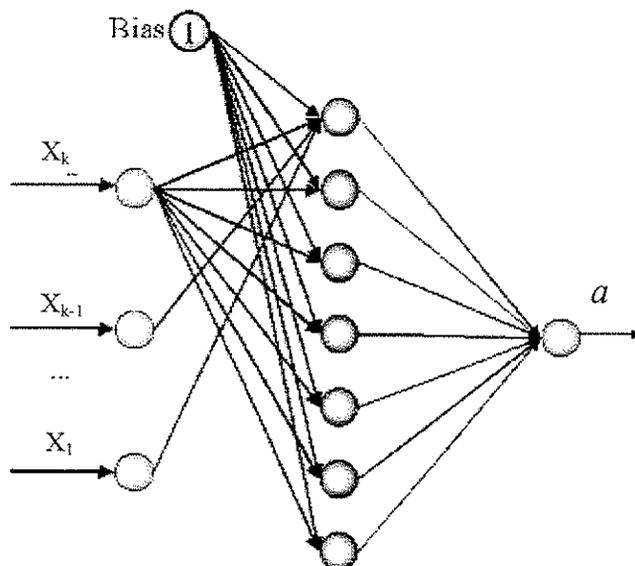


FIGURA 3: REDE NEURONAL ARTIFICIAL *FEEDFORWARD*.

Treinamento de uma Rede Neuronal *Feedforward*

A propriedade mais importante das redes neuronais é a habilidade de aprender de seu ambiente e com isso melhorar seu desempenho. Isso é feito através de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. O treinamento age modificando os pesos das conexões de forma a minimizar o erro entre a saída calculada pela rede e o valor real, que pode também ser um valor calculado por um modelo determinístico. O aprendizado ocorre quando a rede neuronal atinge uma solução generalizada para uma classe de problemas.

Denomina-se algoritmo de aprendizado a um conjunto de regras bem definidas para a solução de um problema de aprendizado. Existem muitos tipos de algoritmos de aprendizado específicos para determinados modelos de redes neuronais, estes algoritmos diferem entre si principalmente pelo modo como os pesos são modificados.

Backpropagation

É o método mais utilizado em engenharia química. Durante o treinamento com o algoritmo *backpropagation*, a rede opera em uma sequência de dois passos, como pode ser observado na Figura 4. Primeiro, um padrão é apresentado à camada de entrada da rede. A

atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se esta não estiver correta, o erro é calculado. O erro é propagado a partir da camada de saída até a camada de entrada, e os pesos das conexões das unidades das camadas internas vão sendo modificados conforme o erro é retropropagado, conforme mostra a Figura 5.

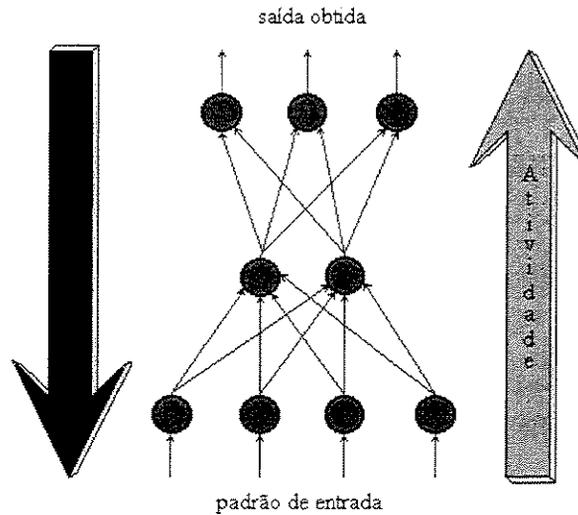


FIGURA 4: TREINAMENTO DE UMA REDE.

A relação entre as entradas (saídas de elementos da camada anterior) e a saída de um nó genérico, o nó i da camada k ($y_i^{(k)}$), é dada por

$$z_i^{(k)} = \sum_{j=1}^{n_{k-1}} w_{ij}^{(k)} y_j^{(k-1)} + \theta_i^{(k)} \quad (1)$$

$$y_i^{(k)} = f(z_i^{(k)}) \quad (2)$$

onde $w_{ij}^{(k)}$ é o peso de conexão entre os nós, $y_j^{(k-1)}$ é a saída do nó j da camada anterior,

$\theta_i^{(k)}$ é o *bias* e $f(\cdot)$ é chamada de função ativação do nó, sendo tipicamente escolhida como

$$f(z) = \frac{1}{1 + \exp(-z)} \quad (3)$$

ou

$$f(z) = \tanh(z) = \frac{1 - \exp(-2z)}{1 + \exp(-2z)} \quad (4)$$

Outras funções de ativação podem ser escolhidas de acordo com o problema, como mostrado na Figura 6.

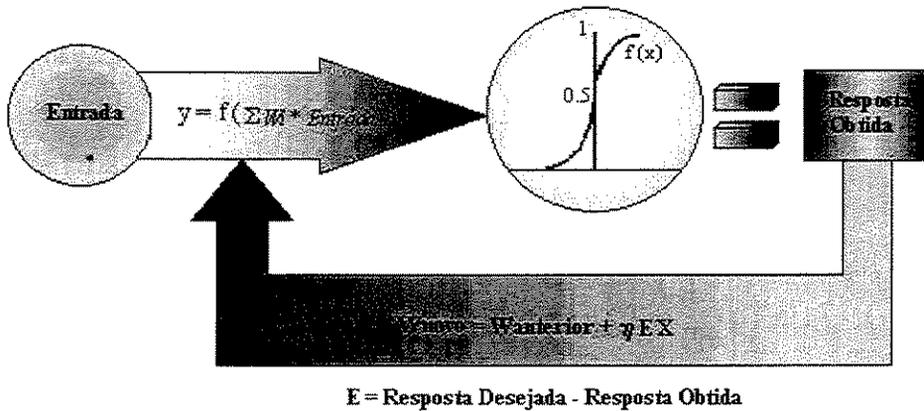


FIGURA 5: CORREÇÃO DOS PESOS DAS CONEXÕES NO TREINAMENTO DA REDE.

As relações acima são válidas para nós nas camadas internas ou de saída. A camada de entrada somente distribui o sinal para a camada interna adjacente.

Uma vez escolhida a função de ativação dos nós, o treinamento da rede consiste na estimação dos valores dos pesos de conexão entre as camadas adjacentes, além da determinação do número de camadas e de neurônios em cada uma delas. O número de nós na camada de entrada é dado pelo número de entradas no processo que se quer modelar e o número de nós na saída pelo número de saídas do processo. O número de camadas internas e de neurônios em cada uma delas, no entanto, devem ser calculados por tentativa e erro, sendo o número ótimo aquele que resulta em um menor erro na validação, quando dados de entrada não usados no treinamento são alimentados à rede e o vetor de saídas calculado pela rede é comparado ao vetor de saída real.

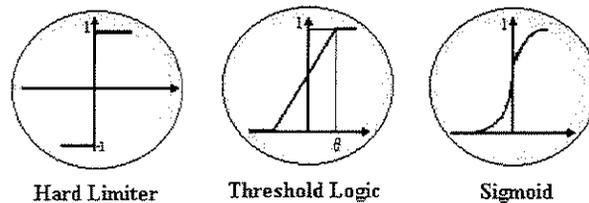


FIGURA 6: FUNÇÕES DE ATIVAÇÃO.

Existem vários métodos de otimização para a minimização do erro (DEMUTH & BEALE, 1994):

- método gradiente
- método gradiente conjugado
- método quasi-Newton
- método Levenberg-Marquardt

Neste trabalho o algoritmo de treinamento a ser utilizado será o Levenberg-Marquardt, pois o mesmo garante uma velocidade de treinamento maior em relação aos outros métodos acima citados.

O algoritmo Levenberg-Marquardt, como os métodos quasi-Newton, foi desenvolvido para aproximar uma velocidade de treinamento de segunda ordem sem ter que computar a matriz hessiana. Quando o desempenho da função tem a forma de uma soma de quadrados (típico de um treinamento de redes *feedforward*), a matriz hessiana pode ser aproximada como:

$$H=J^T J \quad (5)$$

E o gradiente pode ser computado como:

$$G= J^T e \quad (6)$$

Onde J é a matriz jacobiana, que contém a primeira derivada dos erros da rede com relação aos pesos e aos “biases”, e e é um vetor dos erros da rede.

Este método foi utilizado no treinamento da rede neuronal para estimar a cinética do processo de fermentação alcoólica.

Utilização

Depois que a rede estiver treinada e o erro estiver em um nível satisfatório, ela poderá ser utilizada como um modelo preditivo. Para isto, a rede deverá ser utilizada apenas no modo progressivo (*feed-forward*). Ou seja, novas entradas são apresentadas à camada de entrada, são processadas nas camadas intermediárias e os resultados são apresentados na camada de saída, como no treinamento, mas sem a retropropagação do erro. A saída apresentada é o modelo dos dados, na interpretação da rede.

Limitações

As redes neuronais que utilizam *backpropagation*, assim como muitos outros tipos de redes neuronais artificiais, podem ser vistas como "caixas pretas", na qual quase não se sabe porque a rede chega a um determinado resultado, uma vez que os modelos não apresentam justificativas para suas respostas. Neste sentido, muitas pesquisas vêm sendo realizadas visando a extração de conhecimento de redes neuronais artificiais, e a criação de procedimentos explicativos, onde se tenta justificar o comportamento da rede em determinadas situações.

É muito difícil definir a arquitetura ideal da rede de forma que ela seja tão grande quanto o necessário para conseguir obter as representações necessárias e ao mesmo tempo pequena o suficiente para se ter um treinamento mais rápido. Não existem regras claras para se definir quantas unidades devem existir nas camadas intermediárias, quantas camadas, ou como devem ser as conexões entre essas unidades.

Redes Funcional Link (FLN)

Uma maneira de evitar o treinamento não linear, inerente às redes *feedforward*, é fazer inicialmente alguma transformação ou expansão não linear das entradas da rede e então combinar os termos resultantes linearmente. A estrutura obtida tem uma boa capacidade de aproximação não linear e ainda assim a estimação dos pesos da rede continua sendo um problema linear. PAO (1989) se referiu a esta estrutura como *functional link network* (FLN).

Uma rede neuronal do tipo *funcional link* (PAO, 1989) tem estrutura similar a uma rede neuronal *feedforward* de três camadas, exceto pelo fato de empregar termos polinomiais ou trigonométricos na camada interna. Nós lineares são usados dentro das camadas de entrada e saída. O uso desse tipo de função dentro da camada interna tem uma longa história dentro da comunidade de modelagem não linear onde um número pequeno de termos polinomiais de baixa ordem ou os termos dominantes dentro da série de Fourier têm sido usados para introduzir não linearidades dentro do algoritmo linear convencional, o que evita o treinamento não linear.

Essas redes neuronais são algoritmos de aproximação universal; elas podem aproximar um função não linear contínua dentro de uma precisão arbitrária, dado um

número suficiente de nós na camada interna. Os termos polinomiais têm uma saída diferente de zero para o espaço de entrada total e são chamados funções base de generalização global. Uma aplicação com sucesso dessas redes requer um conjunto de funções base que podem representar a função desejada adequadamente dentro do domínio das entradas, mas sem aumentar o número de parâmetros da rede. Modificações no peso ou introdução (ou remoção) de um novo termo dentro da camada interna afetam a saída da rede globalmente e, então, não fica claro como a estrutura deve ser escolhida ou que tipo de relação é armazenado dentro da rede. O sucesso da rede depende criticamente das representações armazenadas dentro dos nós da camada interna (BROWN & HARRIS, 1994).

A estrutura geral de uma FLN é mostrada na Fig. 7, onde \mathbf{x}_e é o vetor de entradas da rede e $y_i(\mathbf{x}_e)$ é uma das saídas. A camada interna faz uma expansão funcional das entradas da rede, transformando o espaço das entradas, de dimensão n_1 , num espaço de dimensão maior, M ($M > n_1$). A camada de saída consiste em m nós, sendo cada nó, de fato, um combinador linear. A relação entrada-saída da FLN é definida como

$$y_i(\mathbf{x}_e) = \sum_{j=1}^M w_{ij} h_j(\mathbf{x}_e), \quad 1 \leq i \leq m \quad (7)$$

Há várias maneiras de se fazer a expansão funcional. A mais fácil e mais usada é a expansão em polinômios. Neste caso, o resultado da expansão ($h_j(\mathbf{x}_e)$) é uma série de monômios de \mathbf{x}_e . Os monômios gerados para uma expansão polinomial de grau máximo igual a seis são mostrados na Tabela 1.

O número de monômios gerados pode ser calculado por

$$M = \frac{(n_x + n_g)!}{n_x! n_g!} \quad (8)$$

onde: n_x = tamanho do vetor das entradas
 n_g = grau dos monômios

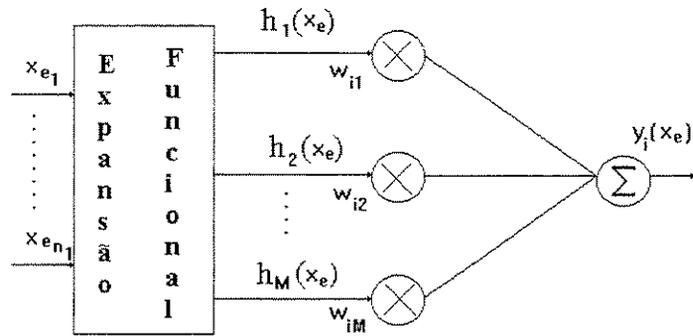


FIGURA 7: ESTRUTURA DE UMA FLN (*FUNCTIONAL-LINK NETWORK*)

Tabela 1: Expansão polinomial de grau seis.

Grau	Monômios
0	1
1	x_i $i=1, n_x$
2	$x_i x_j$ $i=1, n_x; j=i, n_x$
3	$x_i x_j x_k$ $i=1, n_x; j=i, n_x; k=j, n_x$
4	$x_i x_j x_k x_l$ $i=1, n_x; j=i, n_x; k=j, n_x;$ $l=k, n_x$
5	$x_i x_j x_k x_l x_m$ $i=1, n_x; j=i, n_x; k=j, n_x;$ $l=k, n_x; m=l, n_x$
6	$x_i x_j x_k x_l x_m x_n$ $i=1, n_x; j=i, n_x; k=j, n_x;$ $l=k, n_x; m=l, n_x; n=m, n_x$

Entre as vantagens das FLNs sobre as FNNs está o fato de que a rede resultante é linear nos pesos, os quais podem ser estimados por métodos lineares e, logo, de convergência garantida.

HENRIQUE e LIMA (1996) propuseram uma modificação na estrutura das FLNs na qual a saída descrita pela eq. (7) é transformada por uma função de ativação não linear inversível. A nova saída é

$$y_i(\mathbf{x}_e) = f_i\left(\sum_{j=1}^M w_{ij}h_j(\mathbf{x}_e)\right) \quad (9)$$

onde $f(\bullet)$ é uma função não linear inversível como, por exemplo, as dadas nas eqs. 3 e 4.

Esta modificação é feita com o objetivo de aumentar a capacidade de aproximação não linear das FLNs. Neste caso, o treinamento é feito usando-se uma saída transformada, que é igual à saída original da FLN transformada pela inversa da função de ativação f . Por exemplo, se f é uma função logaritmo neperiano, a saída transformada que é usada para o treinamento é $\exp(y)$. Desta forma, a estimação dos pesos continua a ser linear (COSTA, 1998).

O desempenho da rede é medido por (MILTON e ARNOLD, 1990)

$$\text{cor} = \left(1 - \frac{\text{SEE}}{S\tau\tau}\right)100\% \quad (10)$$

onde $\text{SEE} = \sum_{k=1}^N (y_e(k) - y(k))^2$, $S\tau\tau = \sum_{k=1}^N (y_e(k) - \overline{y_e})^2$, $y_e(k)$ é a saída experimental k , $y(k)$ é a saída da rede correspondente, $\overline{y_e}$ é o valor médio das saídas experimentais e N é o número de dados experimentais.

Modelo Híbrido

O Modelo Híbrido é outra forma de se obter modelos confiáveis, que representem as características mais representativas do processo de maneira que possam ser usados para otimização e controle de processos. Acredita-se que esses modelos são mais confiáveis justamente por estarem baseados em princípios fundamentais do processo (balanços de massa e/ ou energia) além das redes neuronais.

PSICHOGIOS e UNGAR (1992) desenvolveram um modelo híbrido neuronal para simular um bioreator operando em batelada alimentada. Eles demonstraram uma nova alternativa para obtenção de modelos confiáveis e úteis para modelagem de processos a partir de considerações físicas (balanços de massa e energia) mas também incluindo termos difíceis de modelar a partir de princípios básicos. Neste estudo, os resultados do modelo

híbrido neuronal foram superiores àqueles obtidos para a rede neuronal usada como “caixa preta”.

FU e BARFORD (1994) desenvolveram um modelo estruturado híbrido para estudos de simulação e controle, no qual o uso de redes neuronais foi incorporado para aumentar a precisão de uma simulação complexa. O treinamento da rede adotado foi o reforçado (*reinforcement learning*), uma vez que o vetor das saídas desejadas não é conhecido.

SCHUBERT *et al.* (1994) desenvolveram uma técnica aperfeiçoada de modelagem híbrido neuronal para processos bioquímicos composta de um conjunto de equações diferenciais dinâmicas, uma rede neuronal artificial e um sistema especialista nebuloso para estimação, predição, otimização da taxa de alimentação e controle do processo de produção de fermento biológico.

AZEVEDO *et al.* (1997) apresentaram um estudo comparativo da modelagem híbrido neuronal com os métodos convencionais, a aproximação determinística e formulações baseadas em diferentes topologias de redes neuronais artificiais para a produção de fermento biológico em um fermentador operando em batelada alimentada. Neste trabalho, o modelo híbrido foi considerado superior aos outros métodos.

COSTA *et al.* (1998) e ALVES *et al.* (1998) apresentaram um modelo híbrido neuronal para fermentação em batelada alimentada para a produção de etanol com o objetivo de implementar um controle ótimo adaptativo. A rede neuronal utilizada foi do tipo funcional link (FLN). A simulação do processo mostrou que o esquema adaptativo proposto obteve resultados próximos do ótimo mesmo na presença de erros de modelagem.

SHENE *et al.* (1999) compararam o desempenho de uma rede neuronal do tipo “caixa preta” com o modelo híbrido na predição das variáveis mais significativas da fermentação em batelada por *Zymomonas mobilis* CP4. Neste trabalho, a rede neuronal sozinha obteve um resultado superior ao encontrado para o modelo híbrido.

COSTA *et al.* (1999) desenvolveram um modelo híbrido na fermentação alcoólica para ser utilizado como estratégia de otimização para bioreatores que operam em batelada alimentada. Neste estudo foi utilizada a rede neuronal funcional link (FLN) para a predição das taxas cinéticas.

QI *et al.* (1999) utilizou um modelo híbrido neuronal para modelagem de reator de leito fixo com parede refrigerada, no qual uma rede neuronal de estrutura *feedforward* com três camadas foi adotada para estimar os parâmetros do processo de oxidação do benzeno à

anidrido maleico. Os resultados do modelo híbrido foram comparados com modelos pseudo-homogêneos bidimensionais e sua capacidade de predição foi idêntica, além de sua estrutura mais simplificada facilitar sua implementação para otimização on-line e controle em tempo real.

FU e BARFORD (1994) propuseram um modelo híbrido para a descrição do cultivo de células de hibridoma para a produção de anticorpos monoclonais. A rede neuronal utilizada na estimação dos parâmetros desconhecidos do processo fez uso do treinamento reforçado para uma estrutura de uma rede feedforward de três camadas positiva mais uma feedback negativa. A melhor topologia foi determinada de acordo com o erro obtido entre as predições do modelo com os valores experimentais para uma estrutura de sete variáveis como entrada e quatorze como saída. Eles concluíram que o modelo híbrido apresentou boa capacidade de predição para o sistema em estudo.

ZORZETTO *et al.* (2000) discorreram uma análise comparativa entre diferentes estruturas de modelo para produção em batelada de cerveja. Quatro modelos, dois de estrutura caixa preta e dois com estrutura caixa cinza, ou seja, modelos híbridos. Dados experimentais foram utilizados para avaliação do desempenho dos modelos caixa preta. Aos modelos híbridos, equações de balanço foram acrescentados e em seguida, foi avaliado seu desempenho na predição das variáveis do processo. Eles concluíram que os modelos híbridos são superiores na sua capacidade de predição uma vez que somente o comportamento pouco conhecido do processo, a cinética, é descrito por redes neuronais. Este trabalho apresentou que um modelo caixa preta proposto apresentou resultados melhores para testes de extrapolação que para os de interpolação, o que não é esperado para esse tipo de estrutura.

BENGLILOU *et al.* (2000) apresentaram uma comparação entre os modelos híbridos, os modelos de redes neuronais e os modelos baseados em princípios fundamentais na predição de variáveis e suas capacidades de extrapolação e interpolação. Eles concluíram que os modelos híbridos apresentam um melhor desempenho de predição e capacidade de interpolação e extrapolação pois constituem-se das características tanto de redes neuronais quanto de princípios fundamentais. A melhor estrutura obtida foi com dez neurônios na camada interna para cinco variáveis de entrada na estimação do pH dentro de um CSTR. Para aumentar o domínio de aplicabilidade, os autores adicionaram às variáveis

um desvio de 5% usando o sistema pseudo randômico binário (PRBS) para as entradas principais do processo.

CHEN *et al.* (2000) propuseram um modelo híbrido utilizando uma metodologia que permite identificar separadamente os modelos cinéticos das taxas de reação mesmo que elas ocorram simultaneamente dentro do reator. Essa metodologia permite a identificação de pseudo-coeficientes estequiométricos do modelo a partir da identificação das taxas de reação. Esse procedimento é baseado na transformação de estado que permite que as funções cinéticas sejam completamente separadas uma das outras. Isso significa que cada reação biológica que ocorre dentro do reator pode ser representada separadamente pelo modelo de rede neuronal como se fosse uma só, apesar de todas as reações envolvidas acontecerem ao mesmo tempo. Esse modelo foi testado num bioprocessamento em modo batelada alimentada considerando o substrato como agente limitante. Os dados experimentais foram ajustados por splines cúbicas para depois serem submetidos ao treinamento. Funções do tipo radial foram usadas e a melhor topologia de rede encontrada foi a com quatro neurônios na camada interna.

WILLIS *et al.* (1990) apresentaram uma introdução de redes neuronais como identificadores de dados e também como controladores de processos. Os autores simularam dados de fermentação em batelada e também de destilação e concluíram que redes neuronais têm grande potencial em reduzir os problemas de processo.

LE LANN *et al.* (1999) aplicaram uma metodologia própria baseada em modelos de tendência para a determinação da estequiometria e da cinética, a otimização das condições operacionais e o controle de um reator batelada através de um modelo de controle preditivo linear e não linear.

FARZA *et al.* (1997) propuseram observadores não lineares para estimação de parâmetros em bioprocessos. Esses observadores estimam on-line as taxas cinéticas dentro de um bioreator. Eles não precisam de modelos de taxas de reação e conseguem estimar com precisão essas variáveis. Esse observador não foi testado e o projeto apresentado assume medidas das variáveis “full time”, o que nem sempre é observado na prática.

FEYO DE AZEVEDO *et al.* (1997) realizaram a modelagem e a comparação entre uma aproximação convencional, diferentes estruturas de redes neuronais e uma estrutura híbrida na produção em batelada alimentada de fermento para panificação em escala

laboratorial. Eles concluíram que o modelo híbrido tem mais vantagens sobre o modelo convencional e sobre as redes neurais, pois estas requerem um tempo de treinamento muito grande além não apresentar bom desempenho quando os dados do processo estão fora da faixa de treinamento.

CAPÍTULO III

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

3. SIMULAÇÃO DO PROCESSO FERMENTATIVO

Dentre todas as alternativas propostas para extração do etanol durante a fermentação, o sistema que acopla o fermentador a um tanque *flash* a vácuo apresenta, dentro da realidade brasileira, a maior quantidade de aspectos positivos. Além deste sistema ser o que mais se ajusta às unidades brasileiras, pode-se operar o *flash* a uma temperatura que elimine a necessidade de trocadores de calor no estágio fermentativo, o que reduz drasticamente os custos fixos e de manutenção (SILVA, 1998). Um esquema geral do processo está mostrado na Figura 8.

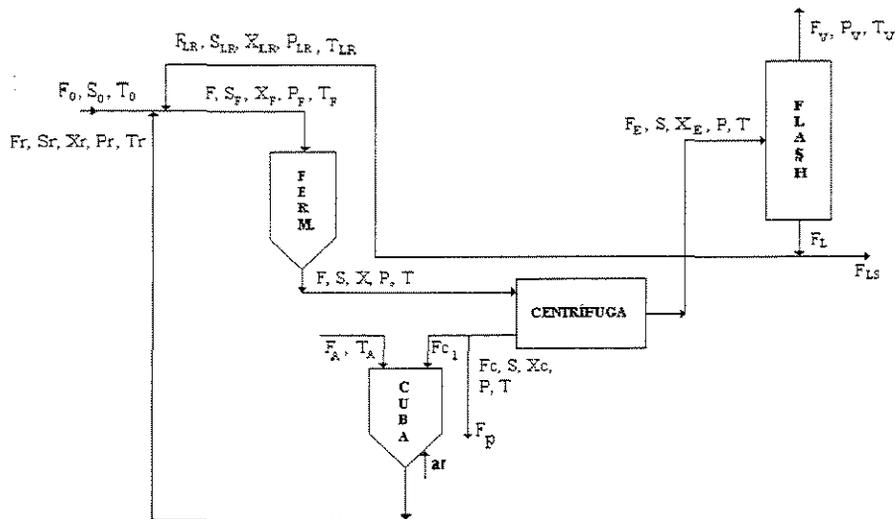


FIGURA 8: FLUXOGRAMA GERAL DO PROCESSO DE FERMENTAÇÃO ALCOÓLICA ACOPLADA A UM *FLASH* A VÁCUO.

Este consiste em quatro unidades interligadas:

- fermentador (unidade de produção de etanol)
- centrífuga (unidade de separação de células)
- cubagem (unidade de tratamento do fermento)
- tanque *flash* a vácuo (unidade de separação da mistura etanol-água)

O uso da centrífuga e da cubagem para tratamento do fermento procura simular condições idênticas às industriais para a separação de células.

O processo inicia-se com a alimentação do substrato no fermentador, que já contém uma concentração inicial de células de levedura. Dentro do fermentador, o substrato é consumido e etanol é produzido. A corrente de saída do fermentador, contendo substrato,

células e etanol, é enviada para a centrífuga, onde é separada em duas fases, uma pesada contendo a maior parte das células, e outra leve, praticamente isenta de células. A fase que contém as células é enviada para a unidade de tratamento do fermento (cuba), após se fazer uma sangria para remover uma pequena parte do fermento. Na cuba, o creme de leveduras é diluído com água, na proporção de um volume de água para um volume de creme, e adiciona-se ácido sulfúrico sob forte agitação, até que se atinja pH entre 2 e 2,5. O fermento tratado é, então, enviado para o fermentador para um novo ciclo do processo. O tempo ideal de repouso do fermento na cuba está entre uma e duas horas.

Antes de se atingir o estado estacionário, a operação do processo extrativo é semelhante ao processo convencional, ou seja, o mosto delevedurado (fase leve de saída da centrífuga) é enviado para a destilaria (para a separação da mistura etanol-água). Quando se atinge o estado estacionário do processo convencional, o sistema de separação a vácuo é acionado. O mosto delevedurado é conduzido para o tanque *flash*, onde ocorre parte da separação da mistura etanol-água. Com o objetivo de eliminar o trocador de calor, o tanque *flash* é operado numa temperatura entre 28 e 30°C e numa pressão associada de 30 a 40 mmHg. A parte vaporizada é enviada para uma coluna de retificação juntamente com uma parte do líquido e a outra parte do líquido retorna ao fermentador. Este retorno de uma parte do líquido tem como objetivo manter a concentração de etanol no fermentador em um valor no qual ele possa atuar como anti-séptico.

Para se determinar as concentrações e vazões de entrada no fermentador, faz-se um balanço de massa global no processo. Para isso algumas aproximações (ANDRIETTA, 1994) são feitas:

- as concentrações de substrato e produto na saída da centrífuga, tanto na fase leve quanto na pesada, são iguais às concentrações na saída do fermentador.
- fixa-se a concentração de células de levedura no reciclo. Para se obter este valor fixo, varia-se a vazão de água (F_A) que dilui o creme de levedura. Para manter a vazão (F_r) constante, utiliza-se uma purga (F_p). Esta purga permite uma constante renovação celular e serve para retirar produtos secundários acumulados no fermentador.

As seguintes equações foram obtidas

$$F_r = FR \quad (11)$$

onde R é a taxa de reciclo de células.

$$F = F_0 + F_{LR} + F_r \quad (12)$$

Substituindo-se a eq. (11) na eq. (12) obtém-se a equação que descreve a vazão de alimentação do fermentador

$$F_r = FR \quad (13)$$

O tempo de residência pode ser descrito por

$$t_{re} = \frac{V_r}{F} \quad (14)$$

onde V_r é o volume do fermentador

Um balanço de biomassa na centrífuga leva a

$$FX = F_E X_E + F_c X_c \quad (15)$$

e o balanço global leva a

$$F_c = F - F_E \quad (16)$$

Substituindo-se a eq. (15) na eq. (16) obtém-se a vazão de alimentação do tanque *flash*

$$F_E = \frac{F(X - X_c)}{X_E - X_c} \quad (17)$$

Do balanço de biomassa na cuba de tratamento do fermento pode-se obter a vazão de alimentação da cuba

$$F_{c1} = \frac{F_r X_r}{X_c} \quad (18)$$

e do balanço global na cuba pode-se obter a vazão de água alimentada

$$F_A = F_r - F_{c1} \quad (19)$$

Do balanço de massa para substrato na cuba obtém-se a concentração de substrato no reciclo da cuba

$$S_r = \frac{F_{c1} S}{F_r} \quad (20)$$

e do balanço para produto obtém-se a concentração de produto no reciclo da cuba

$$P_r = \frac{F_{c1} P}{F_r} \quad (21)$$

Do balanço de massa para substrato no tanque *flash* obtém-se a concentração de substrato no reciclo do *flash*

$$S_{LR} = \frac{F_E S}{F_L}$$

UNICAMP (22)
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

e do balanço para biomassa obtém-se a concentração de biomassa no reciclo do *flash*

$$X_{LR} = \frac{F_E X_E}{F_L} \quad (23)$$

A vazão de reciclo do *flash* é calculada por

$$F_{LR} = rF_L \quad (24)$$

onde r é a taxa de reciclo do tanque *flash*.

Um balanço para substrato na alimentação do fermentador determina a concentração de substrato na alimentação

$$S_F = \frac{F_0 S_0 + F_{LR} S_{LR} + F_r S_r}{F} \quad (25)$$

O balanço para produto determina a concentração de produto na alimentação

$$P_F = \frac{F_r P_r + F_{LR} P_{LR}}{F} \quad (26)$$

e o balanço para biomassa determina a concentração de biomassa na alimentação do fermentador

$$P_F = \frac{F_r X_r + F_{LR} X_{LR}}{F} \quad (27)$$

Através de um balanço de energia na cuba pode-se calcular a temperatura da corrente de reciclo da cuba

$$T_r = \frac{F_A T_A + F_{cl} T}{F_r} \quad (28)$$

e através de um balanço de energia na alimentação do fermentador determina-se a temperatura da corrente de alimentação

$$T_F = \frac{F_0 T_0 + F_r T_r + F_{LR} T_{LR}}{F} \quad (29)$$

As equações de balanço no fermentador são

$$\frac{dX}{dt} = \mu X - D(X - X_F) \quad (30)$$

$$\frac{dS}{dt} = D(S_F - S) - \sigma X \quad (31)$$

$$\frac{dP}{dt} = \pi X - D(P - P_F) \quad (32)$$

$$\frac{dT}{dt} = D(T_F - T) + \frac{\sigma X \Delta H}{\rho C_P} \quad (33)$$

onde X, S, P e T são as concentrações de biomassa, substrato e produto e temperatura no fermentador; μ , σ e π são as taxas cinéticas específicas de crescimento, consumo de substrato e formação de produto, dadas pelas equações de 34 a 36; X_F , S_F e P_F são as concentrações de biomassa, substrato e produto na alimentação e D é a taxa de diluição.

$$\mu = \mu_{\max} \left(\frac{S}{K_s + S} \right) \left(1 - \frac{P}{P_{\max}} \right)^n \quad (34)$$

$$\sigma = \frac{\mu}{Y_{X/S}} \quad (35)$$

$$\pi = \mu \frac{Y_{P/S}}{Y_{X/S}} \quad (36)$$

Os parâmetros envolvidos nas equações acima são dados por (SILVA, 1998)

Tabela 2: Parâmetros usados nas equações

μ_{\max}	$Z \exp(-E/(R_g T))$
P_{\max}	$638.1 \exp(-0.05741 T)$
n	3.0
$Y_{X/S}$	0.445
$Y_{P/S}$	0.033
K_s	1.6

As constantes para o cálculo de μ_{\max} são dadas por

$$E = 1.54 \cdot 10^4 \text{ cal/mol}$$

$$Z = 4.5 \cdot 10^{10} \text{ h}^{-1}$$

O modelo descrito acima foi usado no trabalho de SILVA (1998).

O modelo híbrido foi simulado em FORTRAN 90 usando um algoritmo de Runge-Kutta de 4ª ordem para resolver as equações (30)-(33):

Considera-se que a dinâmica do tanque "flash" é bastante rápida em comparação com a da fermentação. Assim, pode-se assumir uma condição de estado pseudo-estacionário neste tanque. Os balanços de massa no tanque "flash" são dados por:

$$F_E = F_V + F_L \quad (38)$$

$$F_E x_{Ei} = F_V y_i + F_L x_i \quad (39)$$

O equilíbrio líquido-vapor da mistura etanol-água é calculado por:

$$K_i = \frac{y_i}{x_i} = \gamma_i \frac{p_i^{\text{sat}}}{p} \quad (40)$$

3.1 ANÁLISE DO COMPORTAMENTO DINÂMICO DO PROCESSO

Na análise dinâmica do processo, mudanças nas variáveis de entrada podem alterar de forma significativa as condições das variáveis na saída do processo, deslocando os valores destas variáveis para outro patamar de estado estacionário. A determinação do impacto das mudanças das condições operacionais tanto do ponto de vista qualitativo quanto quantitativo é importante para a definição da política operacional, assim como para projetar um sistema de controle com viabilidade prática de implementação.

Estado estacionário de referência do processo convencional

O processo extrativo só tem início quando o processo convencional atinge o estado estacionário. Através de simulação em computador foram obtidos os valores do estado estacionário de referência do processo convencional mostrados na Tabela (3) a partir dos seguintes dados iniciais: $X=30\text{g/L}$, $S=0.1\text{ g/L}$, $P=0\text{ g/L}$ e $T=30\text{ °C}$. A Figura (9) apresenta o perfil da concentração e temperatura tanto para o processo convencional quanto para o processo extrativo.

A partir dos valores do estado estacionário do processo convencional realizaram-se as simulações do processo extrativo. Os valores das variáveis para obtenção do estado estacionário de referências estão de acordo com Silva (1998) para faixas ótimas de operação do processo extrativo, e são: $S_0=180\text{ g/L}$, $R=0.35$, $r=0.45$, $F_0=100\text{ L/h}$. Esses valores foram selecionados devido ao bom rendimento e produtividade alcançados, de 86.046% e 20.66 g/Lh, respectivamente.

Tabela 3: Estado estacionário de referência do processo convencional

Variáveis	Valores
X_S	34,536g/L
S_S	30,312g/L
P_S	49,619g/L
T_S	35,347°C

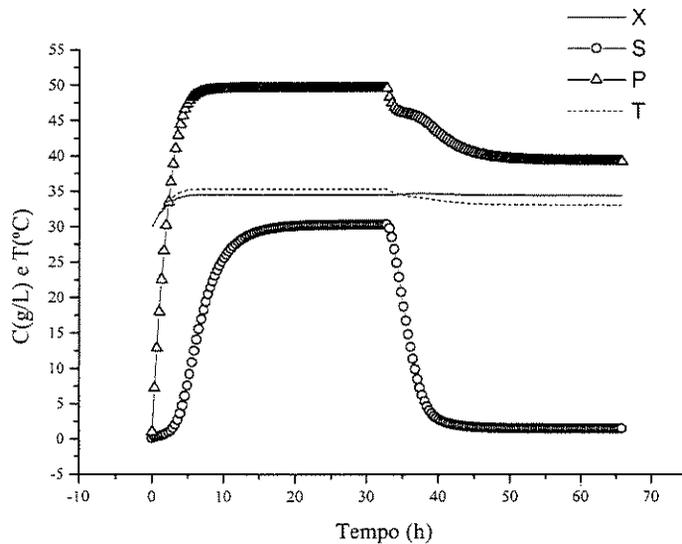


FIGURA 9: PERFIL DA CONCENTRAÇÃO E TEMPERATURA PARA OS PROCESSOS CONVENCIONAL E EXTRATIVO

A dinâmica do processo foi estudada sob diferentes condições de operação e a Figura (9) representa o perfil típico do processo em condições otimizadas conforme Silva (1998). Nesta figura, a primeira fase da fermentação representa uma fermentação contínua convencional até que o estado estacionário seja alcançado, quando então entra em operação o *flash*. A partir do instante de operação do *flash* indicado pelo tempo final do processo convencional e nova contagem de tempo para o processo extrativo, verificam-se mudanças significativas no perfil do processo: queda gradual da temperatura (35°C para 33°C), queda acentuada da concentração de substrato (ART) para níveis compatíveis com a atividade industrial. O valor de saída da concentração de ART no reator ($S_S=1,47$ g/L) permanece um pouco acima do operado nos processos convencionais, que é de no máximo 1 g/L. Há uma diferença marcante entre o processo em estudo e o processo convencional. Enquanto no processo convencional toda concentração de ART no efluente do reator é perdida, pois esta concentração é enviada para a coluna de destilação (mosto delevedurado), no esquema do processo extrativo se consegue recuperar continuamente praticamente 50% deste valor, que retorna ao reator através do fluxo de reciclo do tanque *flash* (F_{LR}). A concentração de células (X_S) mantém-se praticamente constante, enquanto que a concentração de etanol (P_S) diminui gradualmente até estabilizar-se em níveis aceitáveis, ficando na concentração desejada de 40 g/L segundo SILVA (1998). A temperatura dentro do reator é mantida

dentro de níveis desejados para o emprego da *Saccharomyces cerevisiae* como agente transformador do açúcares em álcool. Há um abaixamento significativo de 2 °C na temperatura do reator (T_S) em relação ao valor de início do processo extrativo. Este abaixamento de temperatura é devido à contribuição das vazões de reciclo da cuba de tratamento e principalmente do tanque *flash*, que apresenta temperaturas em torno de 30°C. Sabe-se que no decorrer da fermentação alcoólica, a temperatura do meio em fermentação apresenta tendência a elevar-se, porque ocorrem reações exotérmicas. No processo proposto por SILVA(1998), o único modo de controlar a temperatura do meio reagente é através da mistura dos fluxos provenientes do reciclo da cuba de tratamento e do reciclo do tanque *flash*, pois não se utiliza trocador de calor acoplado ao reator, como é utilizado no processo convencional. O resultado obtido indica o bom funcionamento do sistema de extração *flash* à vácuo. Como já comentado, o processo sempre se inicia com a operação convencional. Neste instante inicial, o reator possui um volume que contém uma concentração de células de levedura (X_S), uma quantidade desprezível de etanol (P_S) e uma concentração remanescente de açúcares fermentescíveis também desprezível (S_S).

A Tabela (4) lista os valores do estado estacionário de referência para o processo extrativo. Com estes valores, inicia-se o estudo dinâmico do processo extrativo frente às perturbações nas variáveis de entrada.

Tabela 4: Estado estacionário de referência do processo extrativo

Variáveis	Valores
X_S	34,446g/L
S_S	1,476g/L
P_S	39,246g/L
T_S	33,066°C
Rendimento	86,046 %
Produtividade	20,66g/Lh

Estudo dinâmico do processo

Este estudo dinâmico é similar ao realizado por SILVA(1998) e foi repetido neste trabalho para testar o funcionamento do modelo e para a familiarização com o comportamento do processo frente a perturbações em diferentes entradas.

Através das variações degrau nas variáveis de entrada selecionadas, vazão de alimentação (F_0), taxas de reciclo de células (R) e do tanque *flash* (r) e concentração de substrato na alimentação (S_0), pode-se verificar os efeitos destas perturbações nas variáveis de saída e observar os comportamentos que estas variáveis apresentam, tais como: comportamentos regulares e irregulares, não apresentando a característica de espelhamento esperada.

Neste estudo, as variáveis de entrada selecionadas foram sujeitas à perturbações do tipo degrau de $\pm 10\%$ e o efeito sobre as variáveis de saída são visualizados nas Figuras 10, 11, 12 e 13. A Tabela 5 mostra as variáveis de entrada e saída selecionadas:

Tabela 5: Variáveis selecionadas para o estudo dinâmico

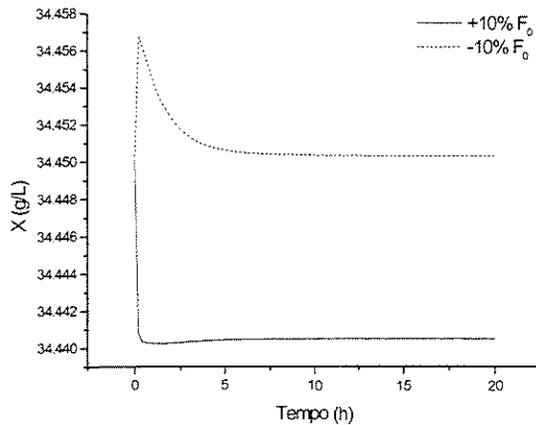
Variáveis de Entrada	Variáveis de Saída
F_0 (L/h)	X (g/L)
R	S (g/L)
r	P (g/L)
S_0 (g/L)	T (°C)

As Figuras 10, 11, 12 e 13 apresentam o perfil de concentração e temperatura no reator das variáveis de saída X , S , P e T frente às perturbações das variáveis de entrada F_0 , R , r e S_0 .

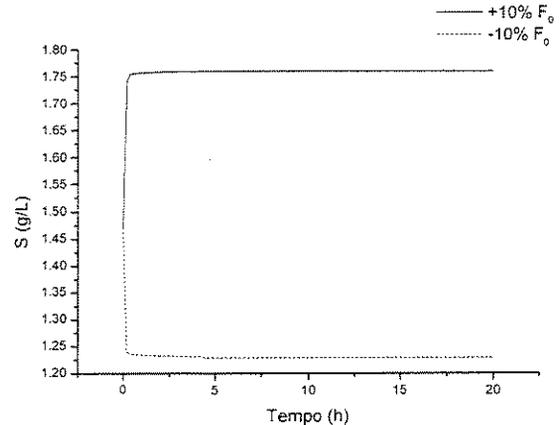
Na Figura 10, onde é mostrado o perfil das variáveis de saída para a perturbação degrau de $\pm 10\%$ em F_0 , observa-se que para a concentração de biomassa, a resposta é decrescente em relação ao estado estacionário de referência para perturbações de $+10\%$ em F_0 . Para perturbações de -10% , X inicialmente aumenta e depois diminui, estabilizando-se num novo estado estacionário um pouco acima do de referência. Para a concentração de substrato, as respostas são simétricas havendo aumento da concentração de S para a perturbação de $+10\%$ e diminuição da concentração de S para a perturbação de -10% . A concentração do produto diminui quando se perturba F_0 em $+10\%$ e aumenta quando a perturbação é -10% . Finalmente, a temperatura diminui para perturbação de $+10\%$ em F_0

e, para perturbação de -10% , inicialmente aumenta e depois diminui, se estabilizando num estado estacionário abaixo do de referência.

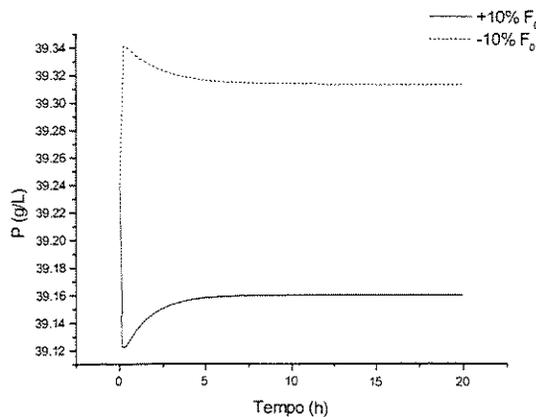
Verifica-se na Figura 10, dois comportamentos irregulares tanto para o perfil de concentração de biomassa quanto para a temperatura. No caso da temperatura, o aumento na vazão faz com que o aumento na quantidade de material reagente aumente a temperatura inicialmente devido às reações exotérmicas, alcançando um estado estacionário abaixo do estado de referência.



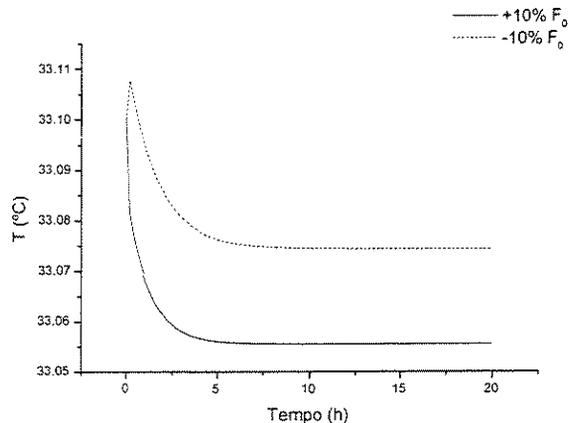
(a)



(b)



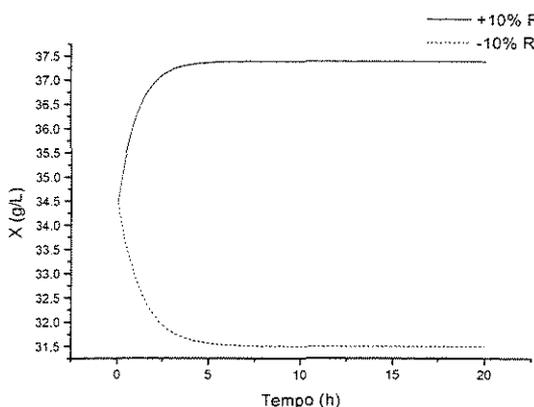
(c)



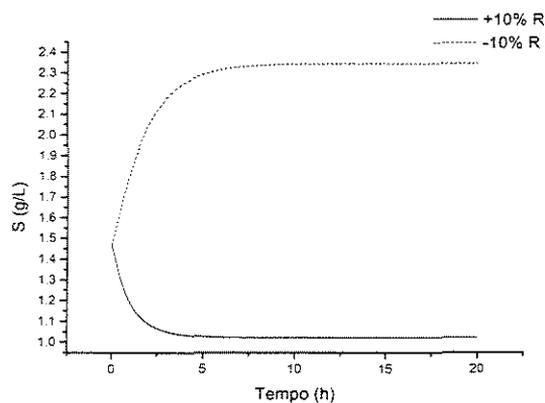
(d)

FIGURA 10: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE: (a) X , (b) S , (c) P E (d) T PARA PERTURBAÇÃO DE GRAU EM F_0

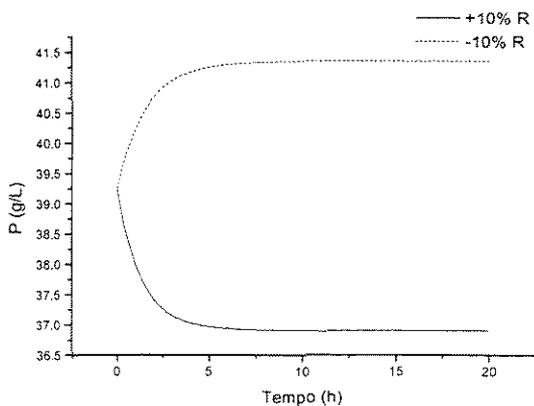
A Figura 11 apresenta o perfil das variáveis de saída frente à perturbação degrau em R. Nota-se que, para todas as quatro variáveis de saída estudadas, as respostas foram similares porém inversas. Para a variável X, a perturbação de +10% leva a um estado estacionário acima do de referência e a perturbação de -10% a um valor de estado estacionário abaixo do de referência. Verifica-se na figura 11(a), que para um aumento na taxa de reciclo de células, há um aumento de células para dentro do reator até que seja alcançado um estado estacionário acima do estado de referência. Para as outras variáveis (S, P e T) o comportamento é inverso. No caso da concentração de substrato, o aumento de células dentro do reator faz com que o substrato seja consumido até que alcance um estado estacionário abaixo do estado de referência porém quando diminui a concentração de células há um aumento na concentração de substrato até que alcance um estado de referência acima do estado de referência, visto na figura 11(b).



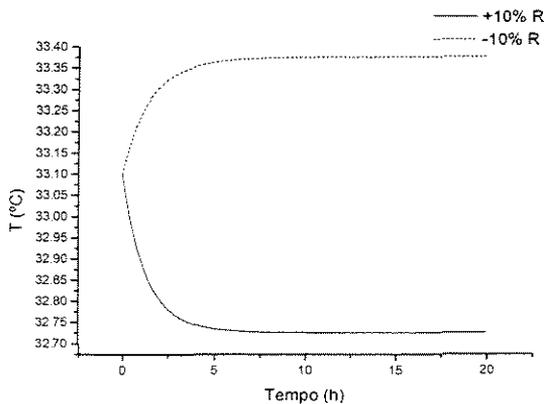
(a)



(b)



(c)

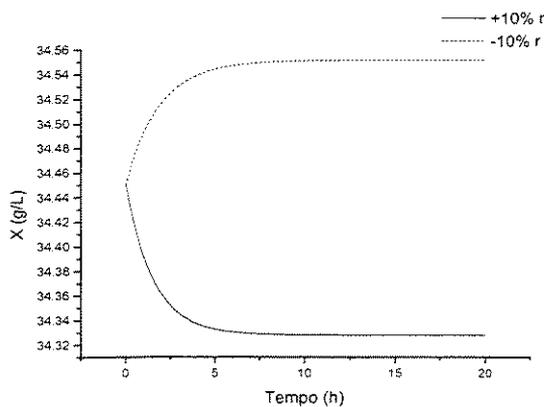


(d)

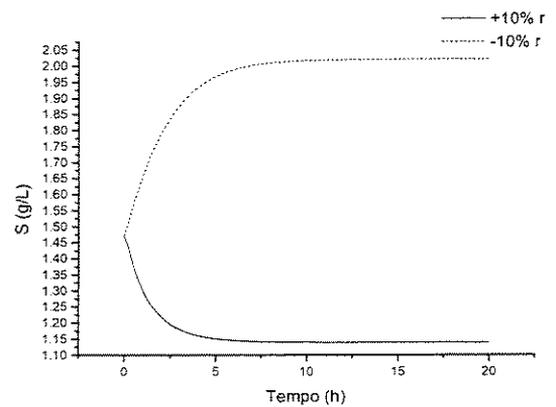
FIGURA 11: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM R.

Os perfis de concentração e de temperatura frente à perturbação de $\pm 10\%$ em r estão mostrados na Figura 12. Verifica-se que as respostas foram similares para as variáveis em estudo e as mesmas apresentaram o mesmo comportamento tanto para perturbação -10% da taxa de reciclo do tanque *flash*, r , afastando-se positivamente do ponto de referência, quanto para a perturbação $+10\%$, afastando do ponto de referência negativamente.

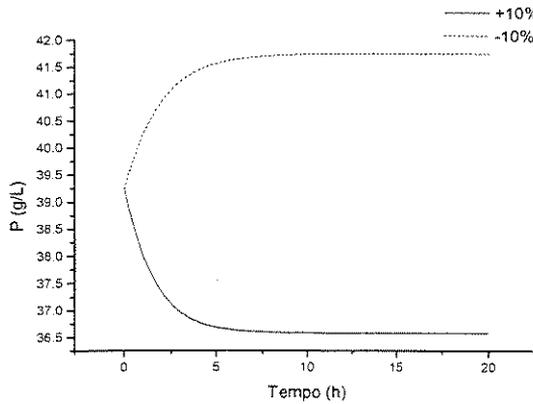
Na figura 12(c), verifica-se que para o aumento da taxa de reciclo de etanol, a concentração de etanol cai até que alcance um estado estacionário inferior ao de referência, que pode ser devido a inibição do produto na fermentação, com o aumento da concentração de etanol dentro do fermentador.



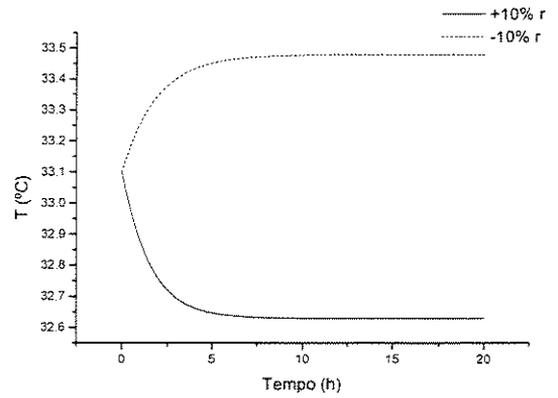
(a)



(b)



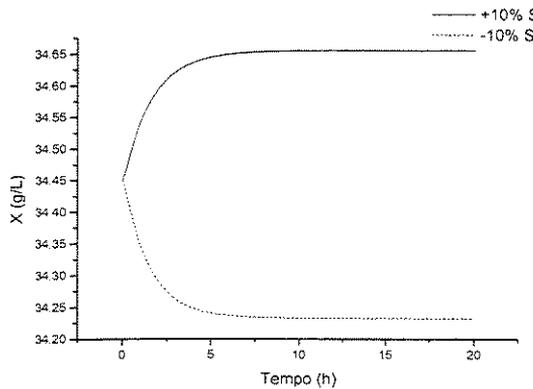
(c)



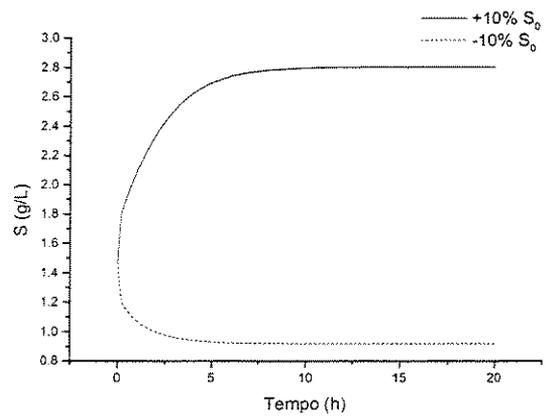
(d)

FIGURA 12: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM r .

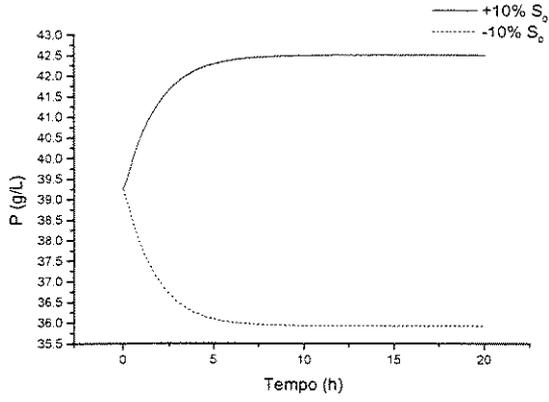
Na Figura 13 são mostrados os perfis das variáveis de saída frente à perturbação na variável S_0 . As variáveis X, S, P e T se comportaram com desvio positivo do ponto de referência para a perturbação de +10% enquanto se deslocaram em sentido oposto para a perturbação de -10% na concentração de substrato na alimentação.



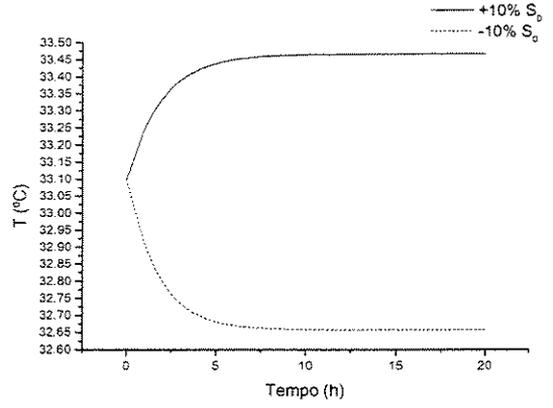
(a)



(b)



(c)



(d)

FIGURA 13: PERFIL DE CONCENTRAÇÃO E TEMPERATURA NO REATOR DE (a) X, (b) S, (c) P E (d) T PARA PERTURBAÇÃO DEGRAU EM S_0 .

CAPÍTULO IV

4. DESENVOLVIMENTO, TREINAMENTO E ANÁLISE DO DESEMPENHO DAS REDES

NEURONAIS

As redes neuronais foram utilizadas nesse caso-estudo com o objetivo de descrever a cinética do processo, e com a finalidade de generalizar mais o modelo híbrido que se pretende propor. Duas estruturas de redes neuronais, *feedforward* e *functional link*, serão utilizadas para descrever a cinética do processo num modelo híbrido e uma comparação entre o desempenho das duas redes será apresentado.

4.1. Rede Neuronal *Feedforward*

Um rede neuronal *feedforward* com três camadas foi desenvolvida, treinada e analisada para o caso em estudo. Optou-se pelo uso de uma única camada interna baseando-se em resultados apresentados na literatura. Segundo Hornik et al. (1989) alguns autores mostraram que qualquer função contínua pode ser descrita por uma rede neuronal com uma única camada interna. Outros autores dizem que duas camadas internas são necessárias (Wang, 1990). O número de camadas internas deve ser suficiente para atingir a precisão desejada. Porém, se muitos neurônios e camadas são usados, a rede se torna sobre-ajustada, ou seja, capaz de ajustar bem os dados de treinamento mas incapaz de generalização usando entradas não usadas no treinamento. Além disso um número de neurônios muito grande aumenta muito o tempo de treinamento.

Treinamento:

O treinamento foi realizado no software Matlab 5.3, que apresenta um *toolbox* específico para redes neuronais. O algoritmo de treinamento foi o Levenberg-Maquardt, que é um *backpropagation* generalizado. A função de ativação utilizada foi a sigmoidal, que é dada pela equação (3):

As entradas da rede são as variáveis de estado do processo que apresentam influência na cinética: as concentrações de substrato e produto e a temperatura. A saída da

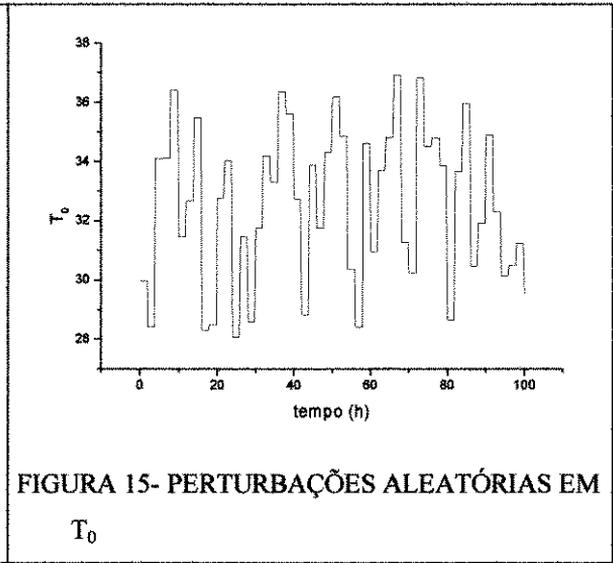
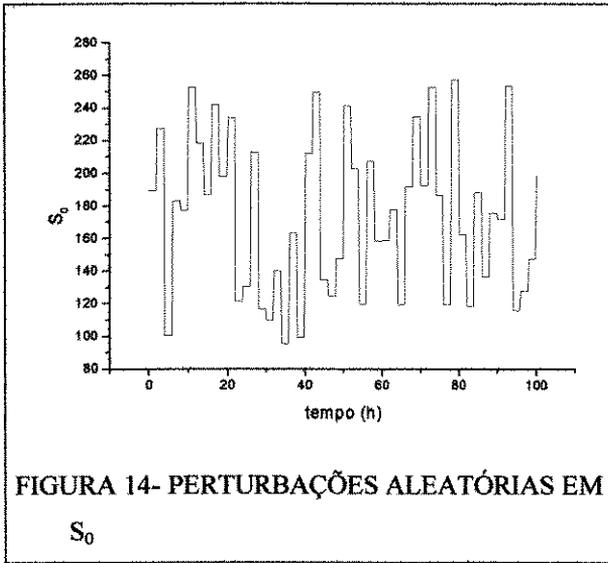
rede é a taxa específica de crescimento. Esta taxa cinética foi escolhida para ser descrita pela rede pois todas as outras taxas cinéticas dependem dela. Num caso de treinamento a partir de dados experimentais, as variáveis de estado estão disponíveis. A saída, no entanto, não é mensurável, mas pode ser estimada a partir da discretização da equação de balanço de biomassa, que é dada pela equação (30).

A discretização desta equação leva a:

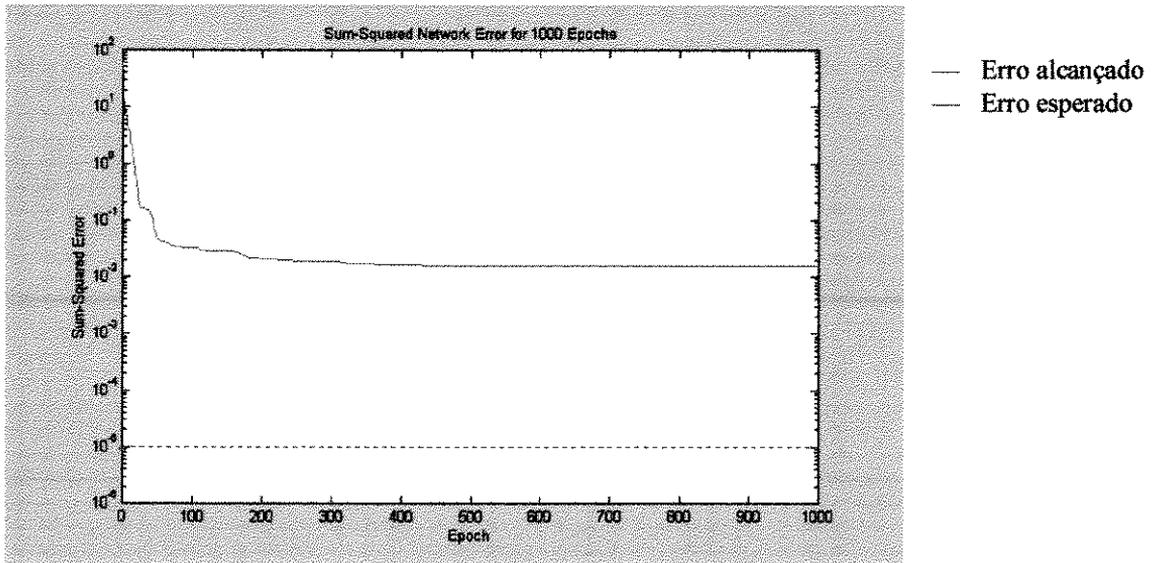
$$\mu = \frac{X(t + \Delta t) - X(t)}{\Delta t \cdot X(t)} + D(X(t) - X_F(t)) \quad (37)$$

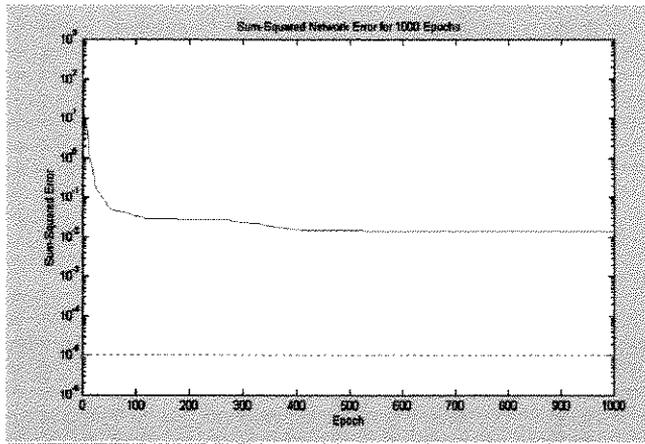
Normalmente os dados experimentais de concentração apresentam ruído e devem ser filtrados antes do cálculo das taxas, ou os resultados serão bastante ruins, como citado por Costa (1998). Neste caso, como os dados foram obtidos a partir de simulação, este problema não é encontrado.

Para gerar uma série temporal de dados de treinamento e validação, o modelo da fermentação alcoólica extrativa proposto por Silva (1998) foi usado. Para treinamento foram feitas perturbações aleatórias de duas em duas horas na concentração inicial do substrato e na temperatura inicial do processo. Estas perturbações foram feitas nos seguintes intervalos: $95 < S_0 < 260$ g/L e $28 < T_0 < 37$ °C. As figuras 14 e 15 mostram estas perturbações. Foi gerado um arquivo com 300 pares de treinamento. Este tipo de arquivo de treinamento foi usado com sucesso por Fonseca (1998). As variáveis utilizadas para perturbações aleatórias no processo foram selecionadas de acordo com a influência que exercem sobre as variáveis do processo de acordo com Costa *et al.* (2000), que realizou um planejamento fatorial para determinar quais as variáveis que exercem maior influência sobre o processo de fermentação alcoólica extrativa.



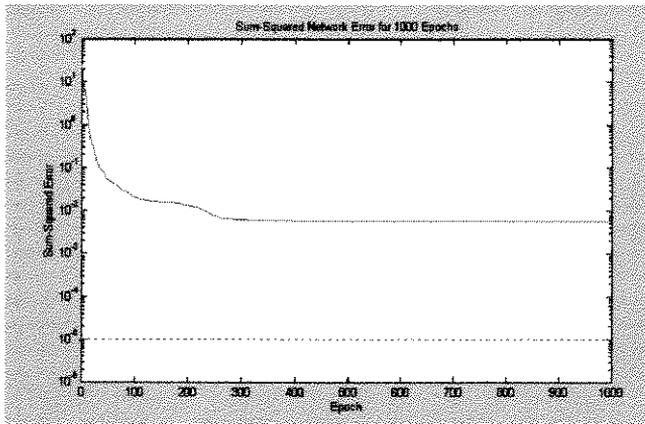
As figuras 16, 17, 18, 19, 20, 21 e 22 apresentam os gráficos do erro do treinamento *versus* o número de iterações para diferentes números de neurônios na camada interna durante o treinamento.





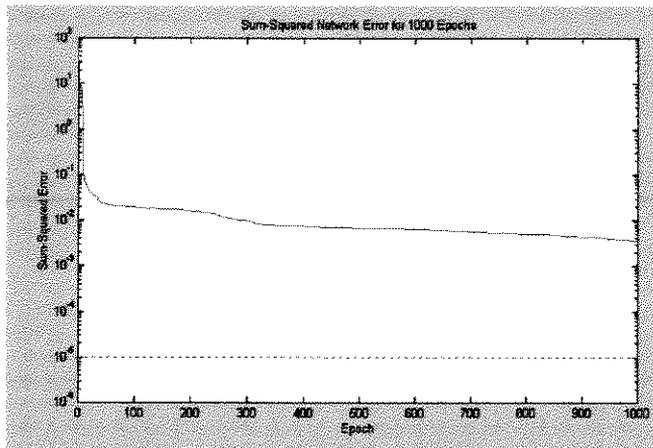
— Erro alcançado
 - - Erro esperado

FIGURA 17: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 351



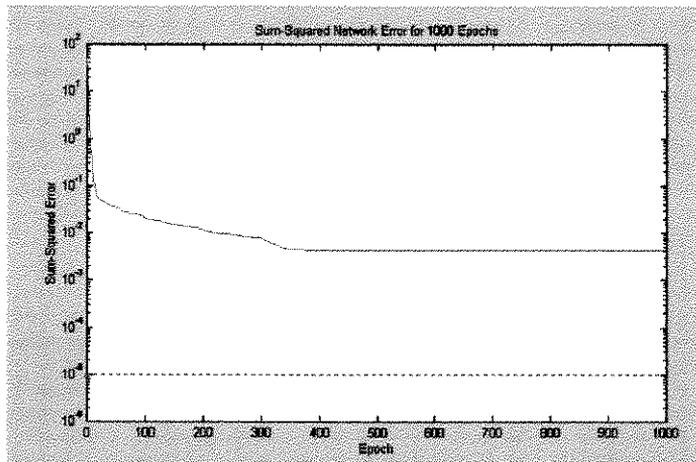
— Erro alcançado
 - - Erro esperado

FIGURA 18: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 361



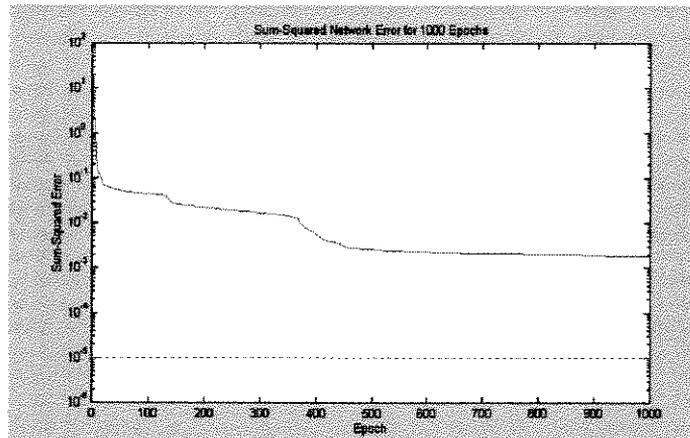
— Erro alcançado
 - - Erro esperado

FIGURA 19: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 371



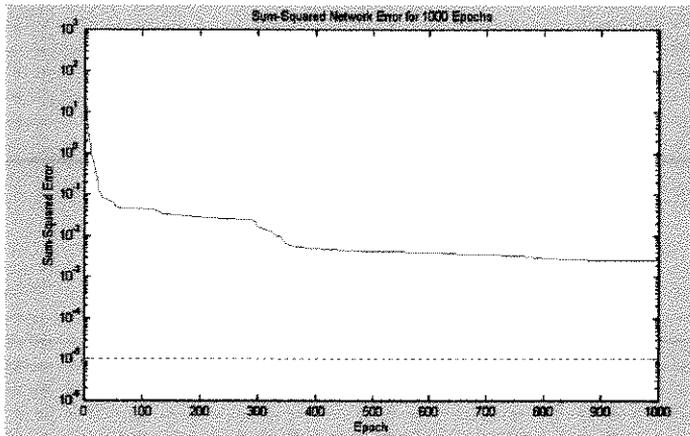
— Erro alcançado
 - - - Erro esperado

FIGURA 20: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 381



— Erro alcançado
 - - - Erro esperado

FIGURA 21: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 391



— Erro alcançado
 - - - Erro esperado

FIGURA 22: GRÁFICO DE TREINAMENTO PARA A TOPOLOGIA 3101

Validação

O arquivo de validação utilizou outra faixa para as perturbações aleatórias em S_0 e T_0 , sendo que essas variáveis estão dentro do seguinte intervalo: $100 < S_0 < 150$ g/L e $28 < T_0 < 31$ °C. Estas perturbações estão mostradas nas figuras 23 e 24.

Realizou-se o estudo do número de neurônios na camada interna. Para isso, iniciou-se o treinamento com uma rede pequena, e foi-se aumentando o número de neurônios. A estrutura escolhida deve ser aquela que resulta em um menor erro na validação, ou seja, para dados que não foram usados no treinamento, o que significa uma melhor capacidade de generalização. A tabela 6 apresenta os erros obtidos no treinamento e na validação para diferentes números de neurônios na camada interna.

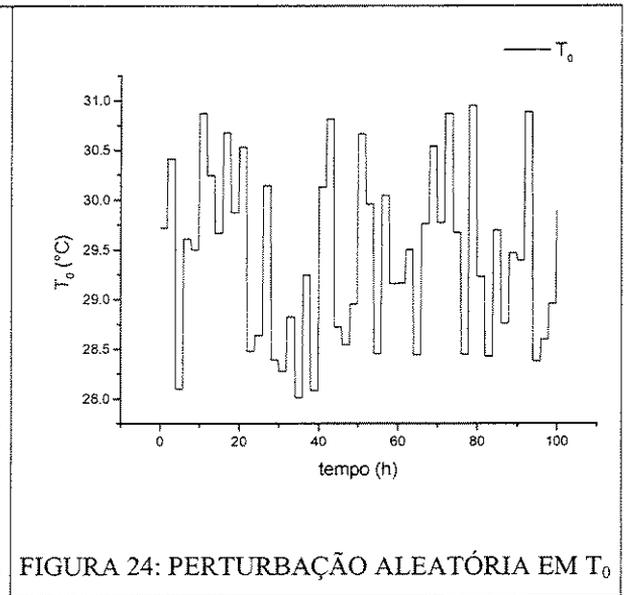
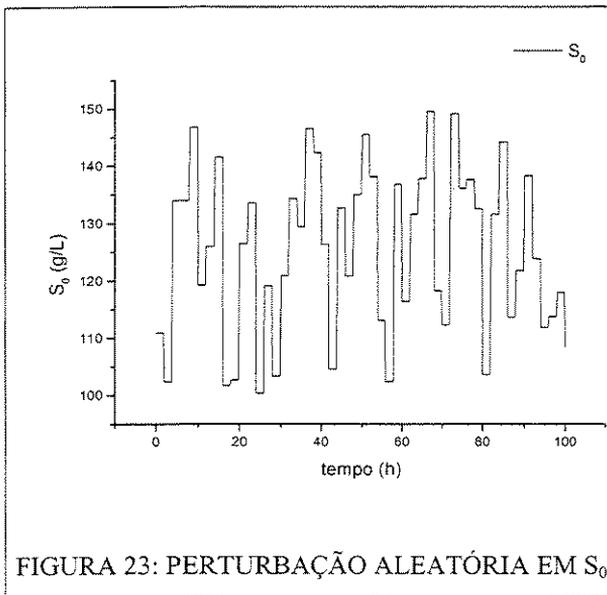


Tabela 6. Erros do Treinamento e Validação para os diferentes números de neurônios na camada interna.

No. Neurônios na camada interna	Treino	Validação
4	1.56E-02	1.020E-05
5	1.43E-02	5.015E-06
6	5.4E-03	3.190E-05
7	3.10E-03	4.165E-06
8	4.5E-03	1.480E-02
9	2.60E-03	1.622E-05
10	3.60E-03	5.570E-02

Nas figuras 25 e 26 estão mostrados os erros da tabela 6 em forma gráfica.

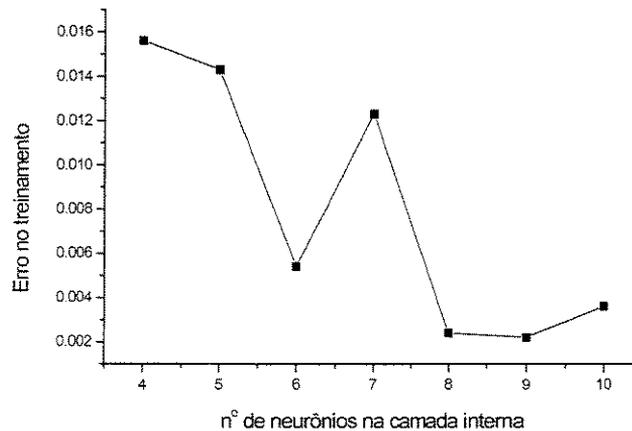


FIGURA 25: ERRO NO TREINAMENTO

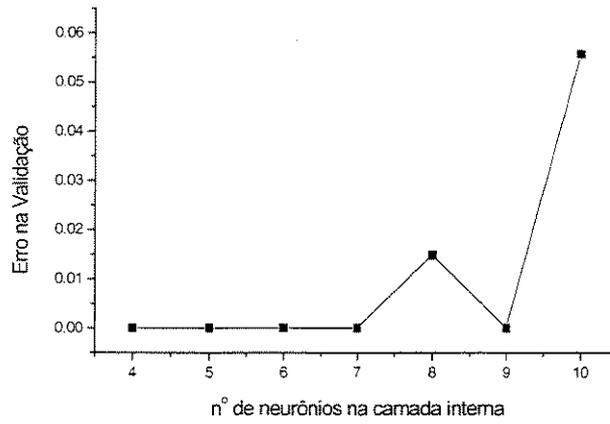


FIGURA 26: ERRO NA VALIDAÇÃO

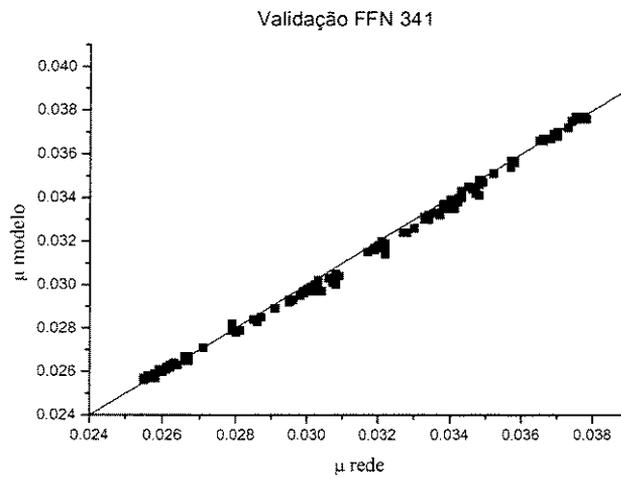


FIGURA 27: VALIDAÇÃO PARA A ESTRUTURA 341

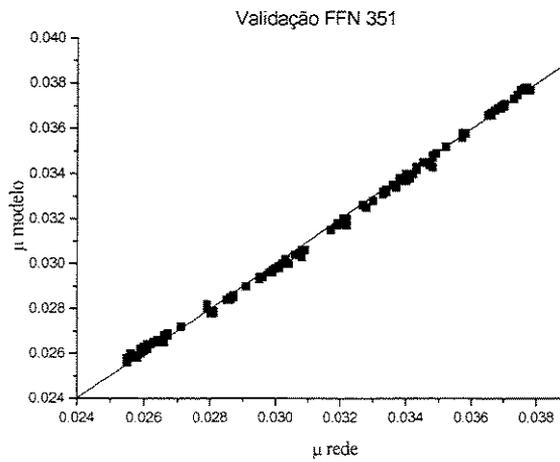


FIGURA 28: VALIDAÇÃO PARA A ESTRUTURA 351

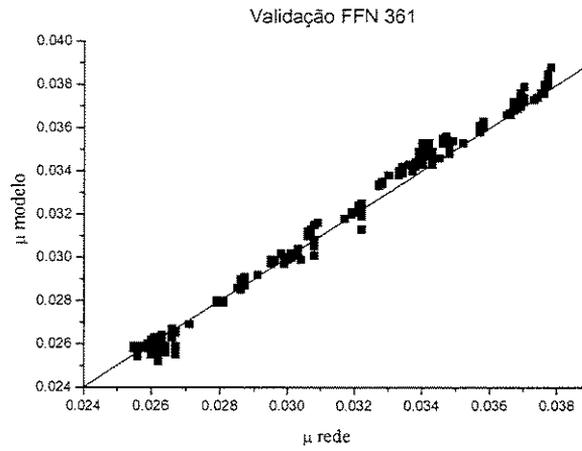


FIGURA 29: VALIDAÇÃO PARA A ESTRUTURA 361

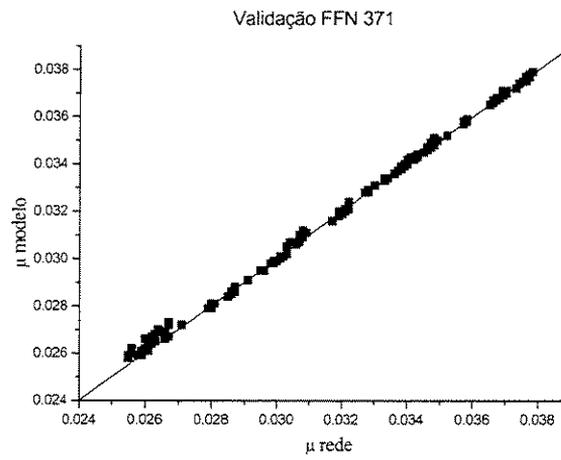


FIGURA 30: VALIDAÇÃO PARA A ESTRUTURA 371

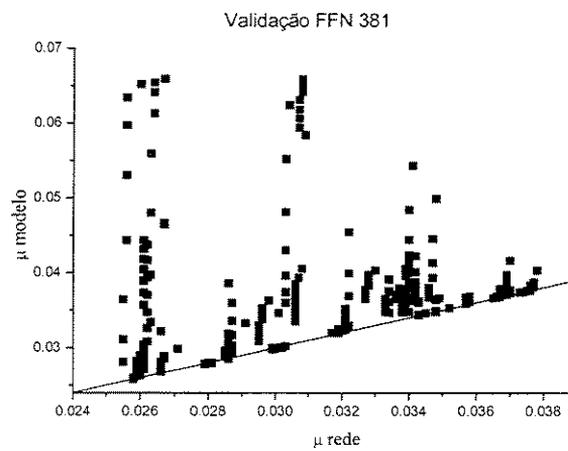


FIGURA 31: VALIDAÇÃO PARA A ESTRUTURA 381

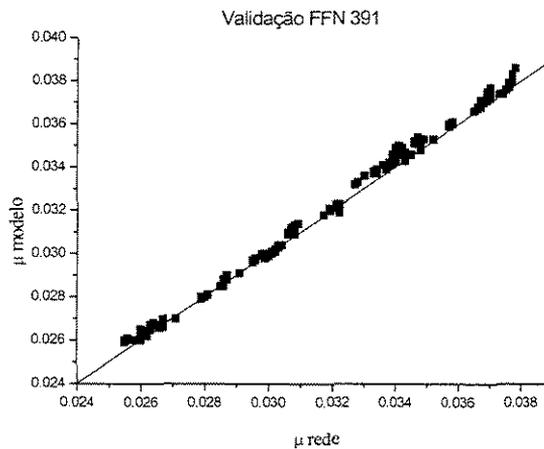


FIGURA 32: VALIDAÇÃO PARA A ESTRUTURA 391

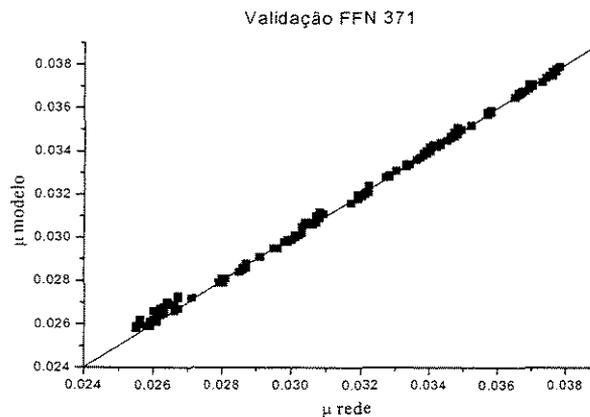


FIGURA 33: VALIDAÇÃO PARA A ESTRUTURA 3101

Após a análise da tabela 6 e das figuras 25 e 26, nota-se que, para a fase de treinamento, o erro diminui com o aumento do número de neurônios na camada interna até 6 neurônios, em seguida aumenta bruscamente, voltando a diminuir para a estrutura de 8 neurônios e tende a se estabilizar. Para a fase de validação, os erros até 7 neurônios para a camada interna são baixos, mas aumenta para a estrutura de 8 neurônios, voltando a diminuir para 9 neurônios. Para 10 neurônios, o erro volta a aumentar bastante. O menor erro alcançado na validação foi obtido para 7 neurônios na camada interna, mas, dado o baixo número de entradas e saídas, 7 neurônios na camada interna parece ser um número muito grande. Como a ordem de grandeza dos erros para 5 neurônios na camada interna é a mesma, acredita-se que a melhor topologia seja a de 5 neurônios, uma vez que o aumento do número de neurônios nessa camada pode levar à construção de redes que sobreajustam

os pesos não sendo, portanto, adequadas para serem usadas no modelo híbrido pois os valores preditos não seriam confiáveis. As redes para 1, 2 e 3 neurônios na camada interna não convergiram.

As figuras 27 a 33 mostram a validação da FNN para quatro até dez neurônios na camada interna. Foram plotados o valor de μ esperado (calculado pelo modelo) *versus* o valor de μ calculado pela rede. Observa-se que, de todas as estruturas, as redes com 5 e 7 neurônios apresentam melhor desempenho como já foi discutido anteriormente. As estruturas com 8 e 10 neurônios na camada interna, representadas pelas figuras 31 e 33, respectivamente, apresentam a característica de sobreajuste dos dados. O aumento dos neurônios na camada interna não é o único parâmetro para o comportamento de sobreajuste da rede, o valor inicial dos “biases” também influencia. No programa desenvolvido em Matlab 5.3., o chute inicial desses valores era determinado aleatoriamente e não se tinha controle sobre isso, portanto, dependendo do valor do “bias” inicial, o programa poderia convergir ou não, supõe-se que isto aconteceu com a topologia com 8 neurônios na camada interna, onde o valor inicial dos bias não gerou um bom resultado.

4.2 Rede *Funcional Link*

Nas redes do tipo *feedforward*, as entradas de um nó são primeiro combinadas linearmente através de certos pesos antes de passar por uma função de ativação não linear. É esta função de ativação dos nós que realmente confere à rede a sua capacidade de aproximação não linear. A não linearidade, entretanto, dificulta a estimação dos pesos da rede. Dessa forma, regras de treinamento não linear devem ser usadas, a velocidade de treinamento geralmente é muito baixa e mínimos locais podem causar problemas (CHEN e BILLINGS, 1992).

Uma maneira de evitar o treinamento não linear é fazer inicialmente alguma transformação ou expansão não linear das entradas da rede e então combinar os termos resultantes linearmente. A estrutura obtida tem uma boa capacidade de aproximação não linear e ainda assim a estimação dos pesos da rede continua sendo um problema linear. PAO (1989) se referiu a esta estrutura como *funcional link network* (FLN).

Neste trabalho, as FLNs descreverão a taxa específica de crescimento para o

processo em estudo no modelo híbrido neuronal e o seu desempenho será comparado com os da rede FNNs.

Treinamento:

Algumas modificações foram aplicadas à esta rede com a finalidade de melhorar seu desempenho.

A primeira foi proposta por HENRIQUE e LIMA (1996). A saída dada pela equação (7) é transformada por uma função de ativação linear inversível. A nova equação é dada pela equação (9).

Outra modificação é que as entradas reais, x_e , são transformadas em um número maior (n_z) de entradas auxiliares, z , $\left(x_{e_1}, x_{e_2}, \dots, \frac{1}{x_{e_1}}, \frac{1}{x_{e_2}}, x_{e_1}^{0.5}, x_{e_2}^{0.5} \text{ etc} \right)$ antes que a expansão funcional seja realizada. Essas entradas auxiliares são escolhidas de acordo com o problema. Então, uma expansão polinomial, de ordem máxima igual a seis, é feita nestas novas entradas. O número de monômios gerados pode ser calculado pela equação (8).

Essas modificações aumentam a capacidade de aproximação não linear da rede e a estimação dos parâmetros continua sendo um problema linear (COSTA, 1998).

Conforme proposto por Henrique e Lima (1996), um método baseado no estimador ortogonal é usado para eliminar os nós não significativos durante o treinamento da rede (BILLINGS *et al.*, 1989). Esse método reduz significativamente o tamanho e a complexidade da rede, evitando o sobreajuste dos dados. COSTA (1998) apresenta a descrição detalhada desse estimador.

O desempenho da rede é medido pela equação (10) (MILTON e ARNOLD, 1990) dado pela equação (10).

As entradas da rede, (x_e) são as variáveis de estado do processo: concentrações de substrato e produto e a temperatura. A saída ($y(x_e)$) é a taxa cinética a ser calculada. As variáveis de estado normalmente estão disponíveis, seja por medição direta ou pelo uso de estimadores, como, por exemplo, o filtro de Kalman estendido. As saídas não são mensuráveis, mas podem ser estimadas a partir dos dados experimentais pela discretização das equações do balanço de massa. Como este trabalho é de simulação, as entradas da rede não apresentam ruído. SIMUTIS *et al.* (1993) ressaltou, entretanto, que em situações reais

isto não é verdade e algum tipo de algoritmo tem que ser aplicado para gerar valores de taxas cinéticas confiáveis a partir de dados de concentração com ruído. Outra abordagem é o uso de um integrador para calcular as concentrações de substrato e produto a partir de valores de taxas cinéticas produzidas pela rede. Portanto, durante o treinamento, o erro a ser minimizado é dado pela diferença entre as concentrações e temperatura calculadas e as suas medidas reais. Como a integração é numericamente mais estável que a diferenciação de conjuntos de dados com ruído, uma melhor estabilidade do processo de treinamento foi alcançada.

A função de ativação inversa foi a que apresentou os melhores resultados que aqui serão apresentados.

A taxa específica de crescimento é dada pela equação (34)

A entrada da rede é dada por

$$\mathbf{x}_e = [S \quad P \quad T]$$

A saída da rede é

$$y(\mathbf{x}_e) = [\mu]$$

O vetor z escolhido é definido como

$$\mathbf{z} = \left[S \quad \frac{1}{S} \quad P^3 \quad \frac{1}{P} \quad \ln(T) \quad \frac{1}{\ln(T)} \right] \quad (38)$$

A função de ativação usada foi

$$f\left(\sum_{j=1}^M \mathbf{w}_{ij} h_j(\mathbf{z})\right) = \frac{1}{\sum_{j=1}^M \mathbf{w}_{ij} h_j(\mathbf{z})} \quad (39)$$

A função de ativação e o vetor z foram determinados após vários testes. A determinação do vetor z talvez seja a maior dificuldade encontrada para o treinamento de uma rede do tipo *funcional link*. Normalmente, é escolhido de forma a aumentar a aproximação não linear da rede neuronal.

O número de monômios gerado é função do tamanho do vetor z e do grau da expansão polinomial. Neste caso, o melhor resultado foi quatro. O número de monômios gerado foi 210. O estimador ortogonal foi usado para eliminar os monômios pouco significantes e estimar os pesos.

As figuras 34, 35, 36, 37, 38 e 39 mostram os resultados obtidos para o mesmo arquivo de dados usado durante o treinamento da rede, usando como função de ativação, a inversa, e variando o grau de monômios de 1 até 6, para uma tolerância da ordem de

0.99999.

Em azul está representado a taxa de crescimento específica calculada pelo modelo determinístico e em vermelho a estimativa da rede neuronal do tipo *funcional link*.

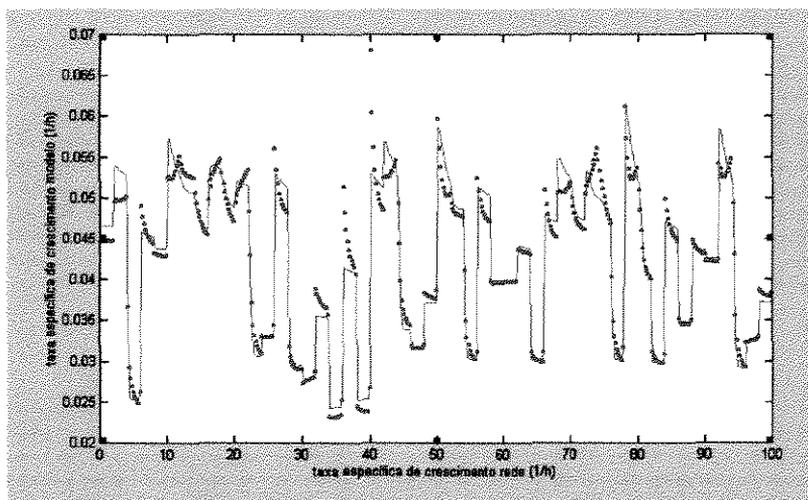


FIGURA 34: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 1

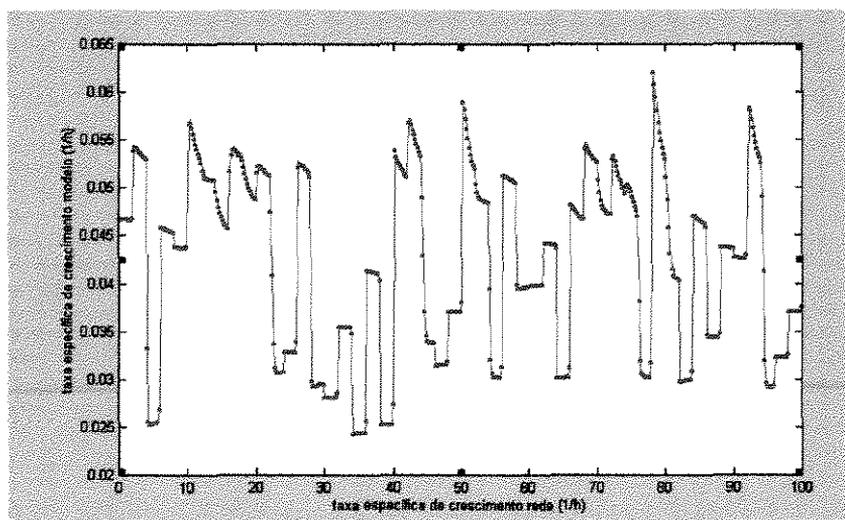


FIGURA 35: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 2

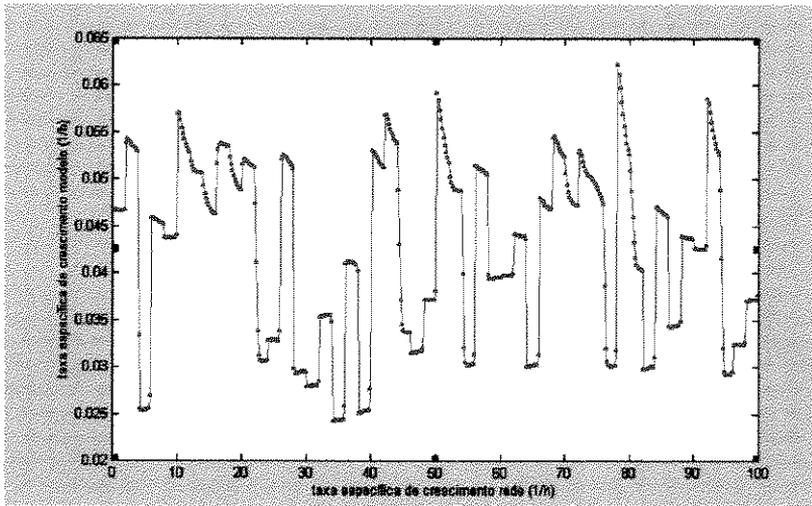


FIGURA 36: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 3

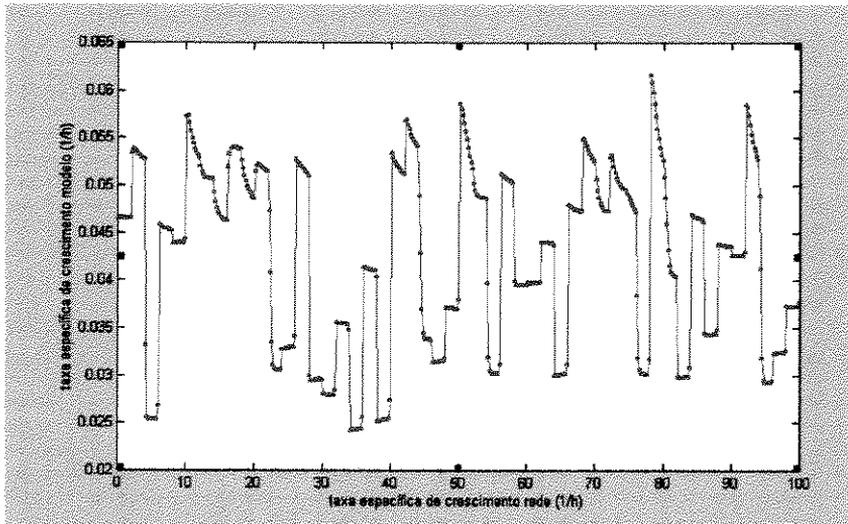


FIGURA 37: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 4

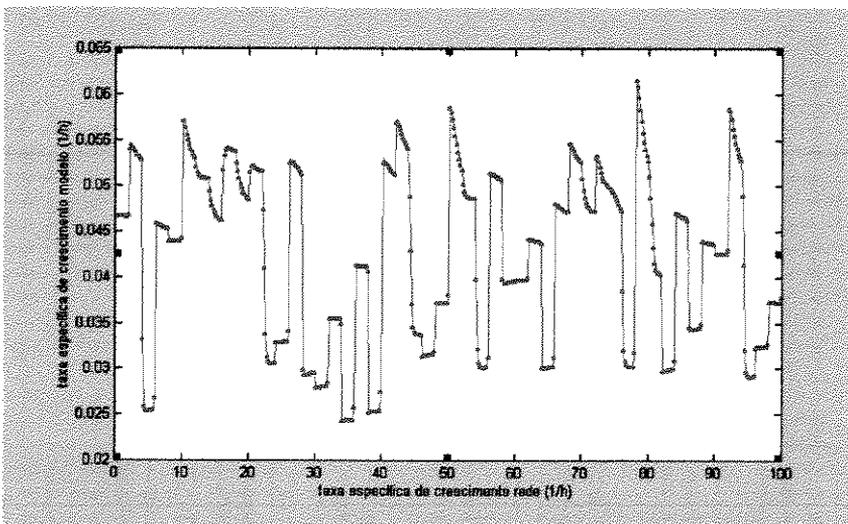


FIGURA 38: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 5

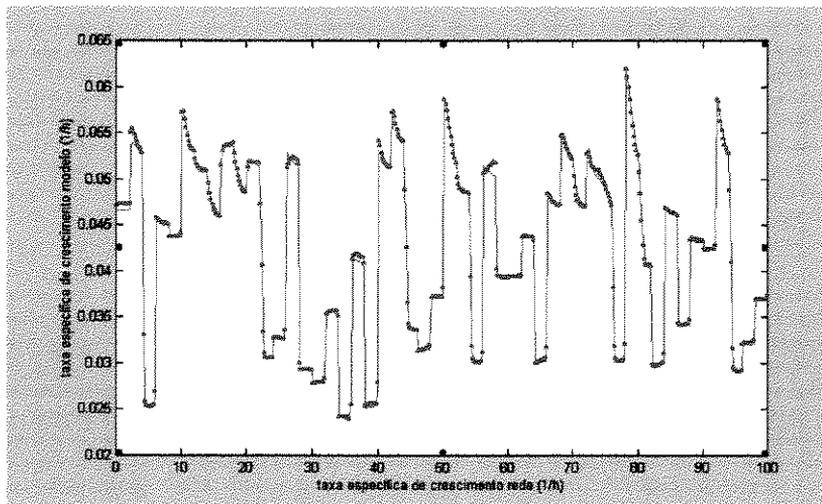


FIGURA 39: TREINAMENTO PARA O GRAU DE MONÔMIOS IGUAL A 6

Verifica-se que com grau de monômios igual a 1, a rede não conseguiu descrever bem a cinética. Para as redes com grau de monômios igual a 2 e 3, os resultados foram equivalentes. A menor rede que descreveu bem o comportamento da cinética foi a com grau de monômios igual a 4.

Validação

Para a validação desta rede o mesmo conjunto de dados diferente usado para a validação das FNNs foi utilizado. Os resultados obtidos são demonstrados pelas figuras 40, 41, 42, 43, 44 e 45.

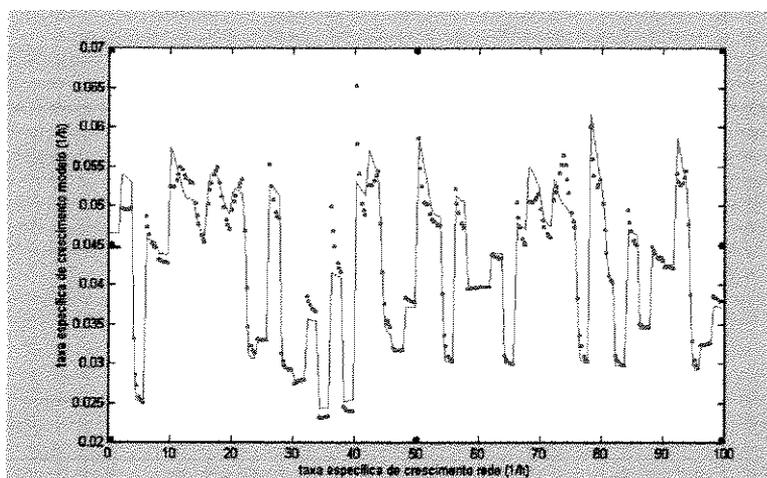


FIGURA 40. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 1

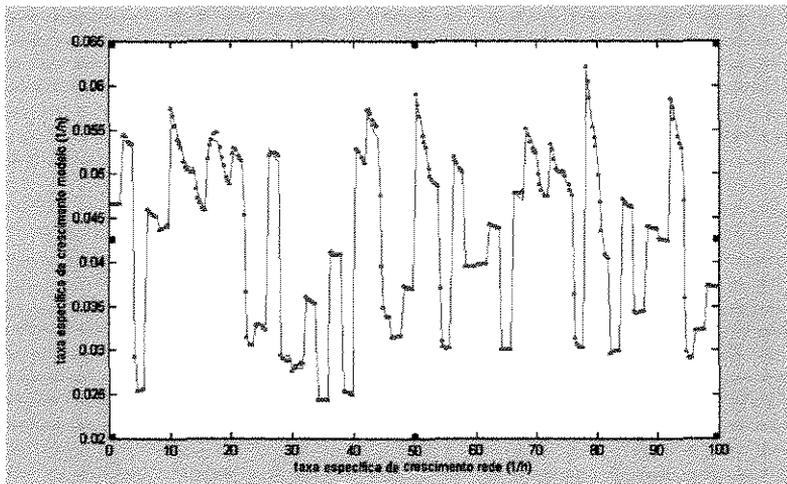


FIGURA 41. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 2

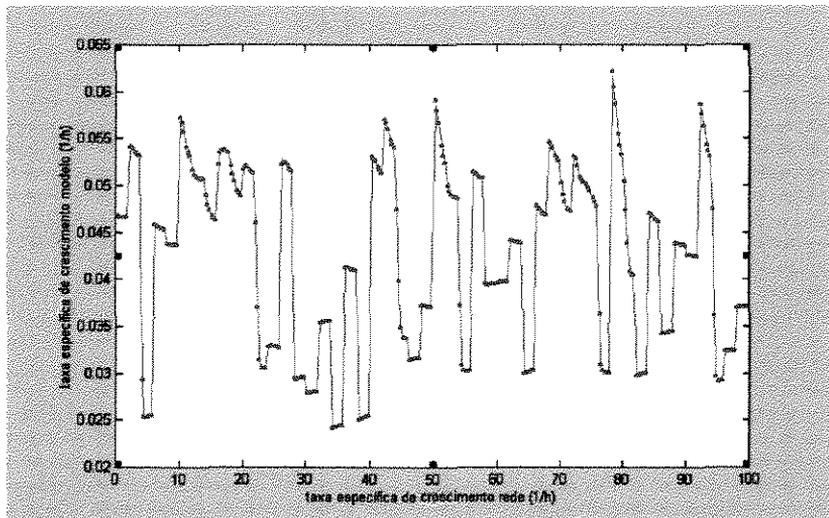


FIGURA 42. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 3

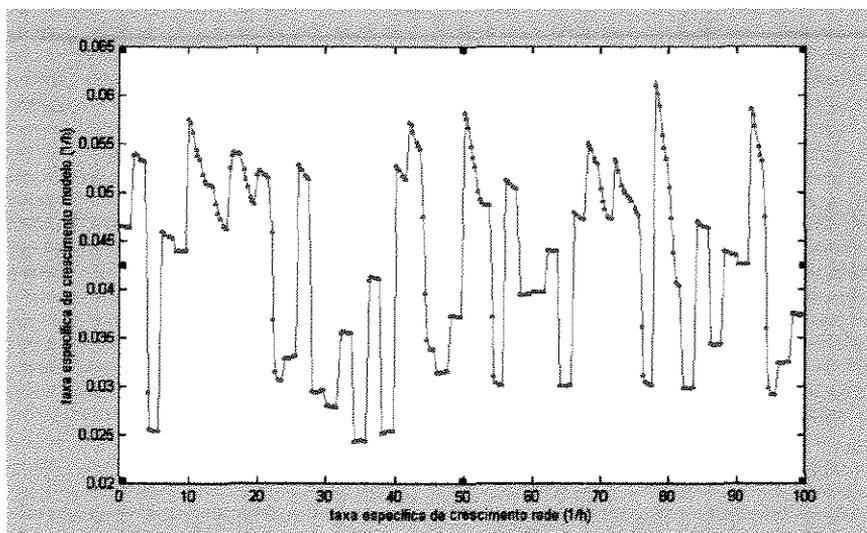


FIGURA 43. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 4

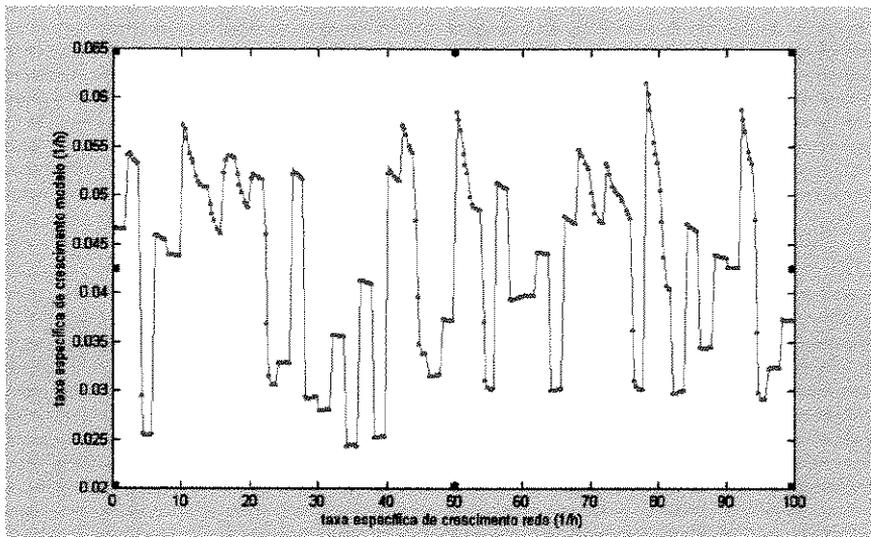


FIGURA 44. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 5

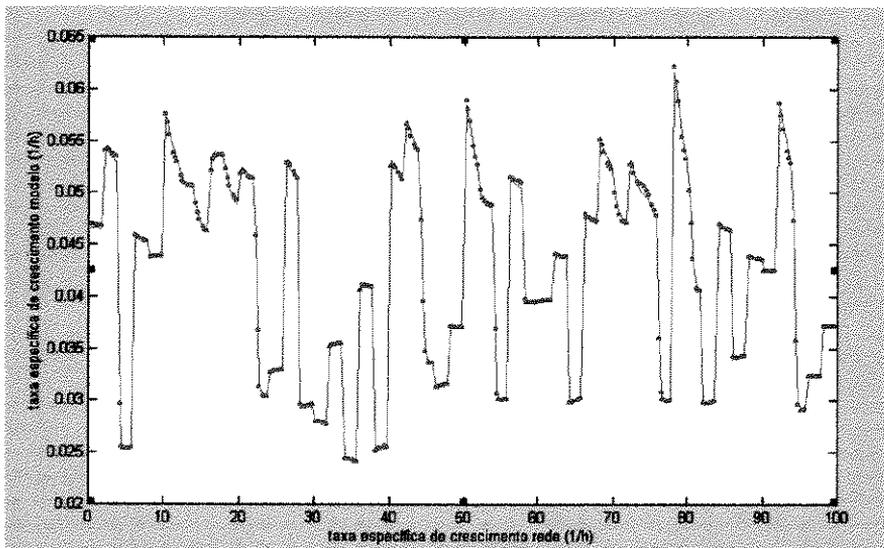


FIGURA 45. VALIDAÇÃO PARA A REDE COM GRAU DE MONÔMIO IGUAL A 6

Após vários testes, a rede com grau de monômios igual a 4 foi escolhida como padrão para o modelo híbrido uma vez que teve bom desempenho na descrição da cinética gerando a seguinte equação.

$$\mu_{FLN} = \frac{1}{9.2463 \times 10^3 \frac{1}{P \ln(T)} - 4.7981 \times 10^{-10} \frac{P^6}{S^2} + 2.3596 \times 10^{-8} P^6 \ln^2(T) + 1.2682 \times 10^{-2} \frac{P^2 \ln(T)}{S}}$$

$$+ \frac{1}{-1.0818 \times 10^{-7} \frac{\ln^3(T)}{S} - 1.4394 \times 10^{-20} \frac{1}{S^2 \ln^2(T)} + 1.1211 \times 10^5 \frac{1}{\ln^4(T)} - 4.9361 \times 10^{-1} \frac{P^3}{S \ln^2(T)}}$$

$$\begin{aligned}
& + \frac{1}{1.9430 \times 10^1 P^6 \ln(T) + 1.6948 \times 10^2 P^{12} - 7.4106 \times 10^{-3} \frac{1}{SP^3}} \\
& + \frac{1}{-4.3306 \times 10^{-2} \frac{1}{S^4} - 4.1145 \times 10^{-10} \frac{P^6 \ln(T)}{S} + 1.1118 \times 10^{-6} \frac{P^6}{\ln^2(T)}}
\end{aligned} \tag{40}$$

O número de parâmetros gerados para estimar a cinética através da rede FNN é igual a 22, considerando os bias e pesos gerados. A equação obtida para a cinética através da rede FLN gerou 14 termos, demonstrando um desempenho comparável e um pouco superior ao da FNN, como será apresentado e discutido no próximo capítulo.

Realizando uma primeira comparação a rede FLN apresenta algumas vantagens sobre a rede FNN porque ela apresenta um treinamento rápido, e devido a maior dificuldade encontrada no desenvolvimento de modelos que utilizam a rede FNN acoplada. A rede FLN gera uma equação, que pode possibilitar uma implementação simples de um esquema adaptativo onde os parâmetros são reestimados on-line baseados nos dados do processo enquanto a FNN se utiliza um modelo interno sobre o qual não se tem muito conhecimento.

CAPÍTULO V

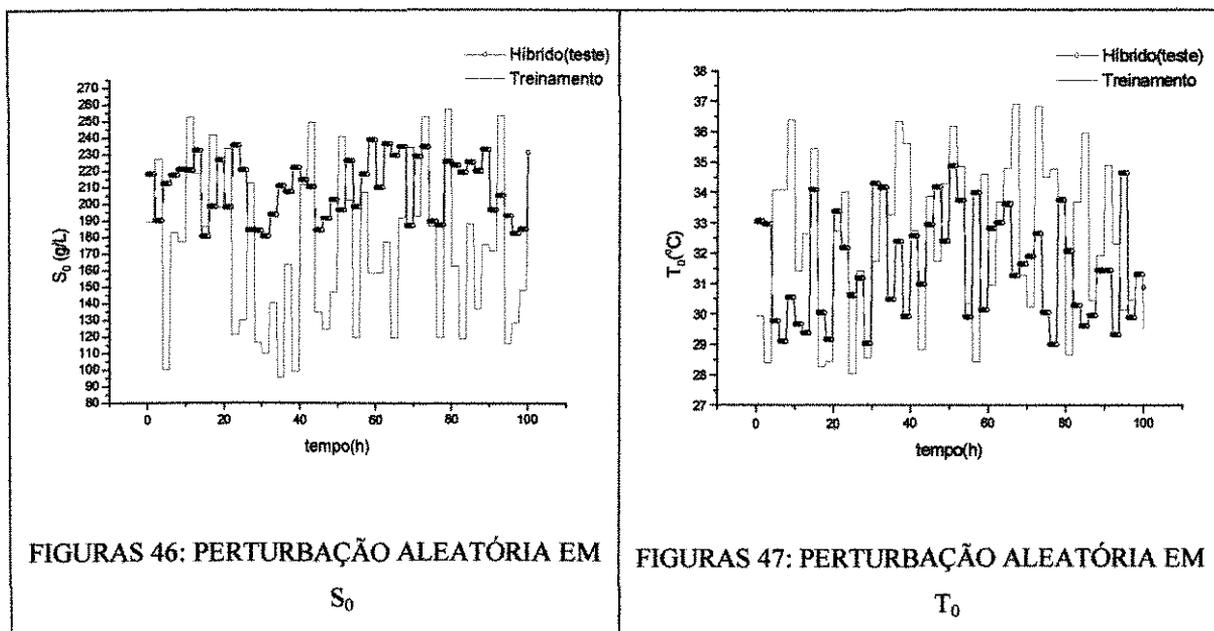
5. DESENVOLVIMENTO E ANÁLISE DO MODELO HÍBRIDO

O processo a ser descrito é a fermentação alcoólica extrativa (Silva, 1998). Um esquema do processo é mostrado na figura 8.

O modelo híbrido usa as equações de balanço de massa e energia dadas pelas equações (30)-(33):

Para o teste do modelo híbrido, perturbações aleatórias em S_0 e T_0 diferentes daquelas realizadas para a validação foram feitas e o intervalo para essas variáveis está dentro da seguinte faixa: $180 < S_0 < 240$ g/L e $29 < T_0 < 35^\circ\text{C}$, dados estes que saem um pouco da faixa para qual a rede foi treinada.

As figuras 46 e 47 mostram as perturbações aleatórias em S_0 e T_0 realizadas de 2 em 2 horas para o teste do modelo híbrido junto com o arquivo de treinamento original, para se visualizar as diferenças entre os dois.



As figuras 48, 49, 50, 51 e 52 mostram a comparação entre o desempenho do modelo determinístico e do modelo híbrido para o processo, quando se fazem as perturbações aleatórias mostradas nas figuras 46 e 47.

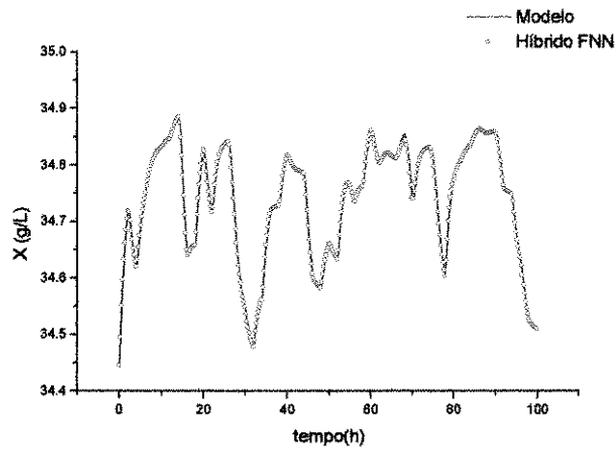


FIGURA 48: PERFIL DE CONCENTRAÇÃO DE BIOMASSA, X.

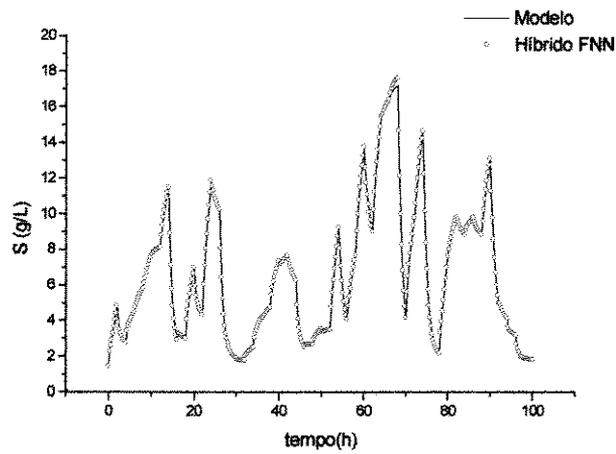


FIGURA 49: PERFIL DE CONCENTRAÇÃO DE SUBSTRATO, S.

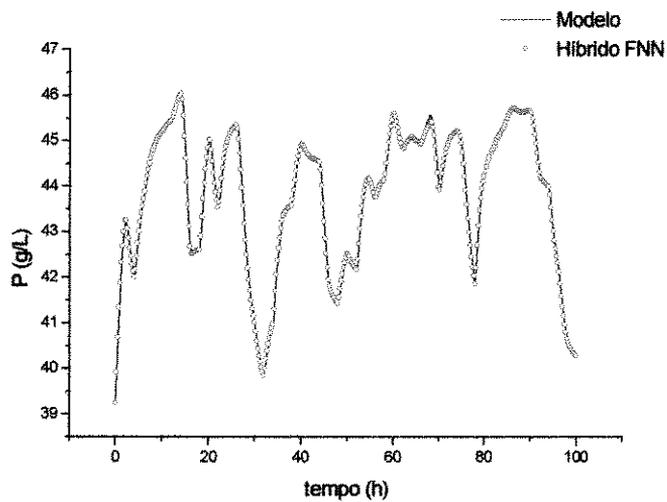


FIGURA 50: PERFIL DE CONCENTRAÇÃO DE PRODUTO, P.

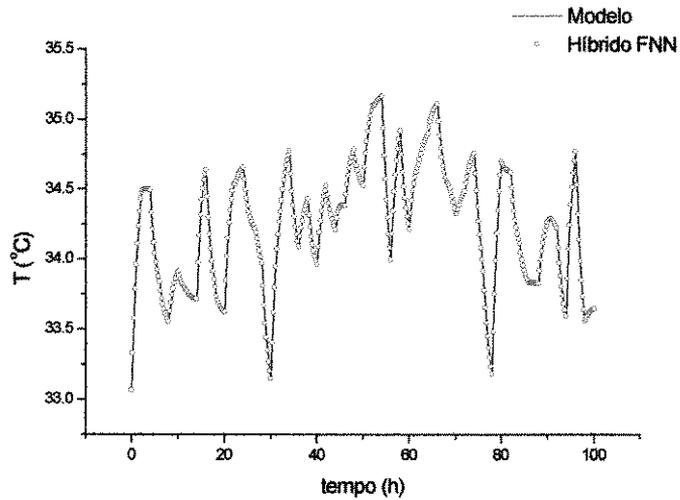


FIGURA 51: PERFIL DE TEMPERATURA, T .

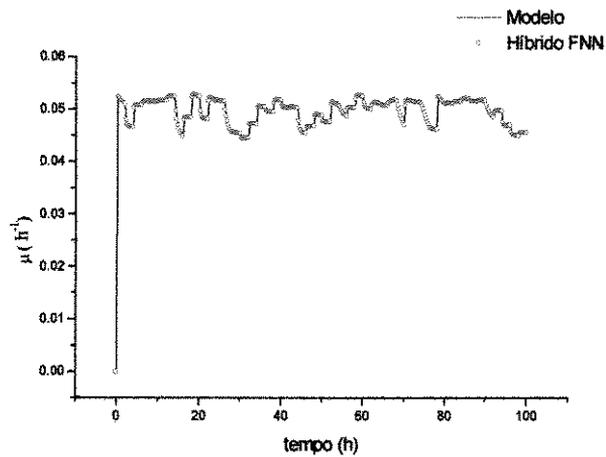


FIGURA 52: PERFIL DA TAXA ESPECÍFICA DE CRESCIMENTO, μ .

O desempenho do modelo híbrido composto pelas equações (30) a (33) e pela rede do tipo *funcional link* é mostrado nas Figuras 53 a 57.

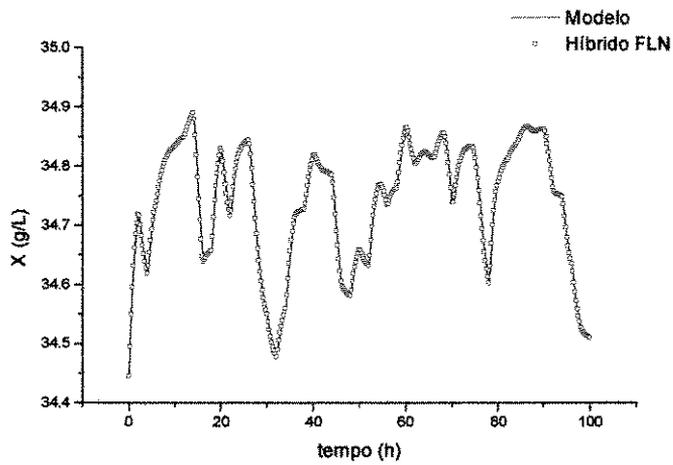


FIGURA 53: PERFIL DA CONCENTRAÇÃO DE CÉLULAS, X

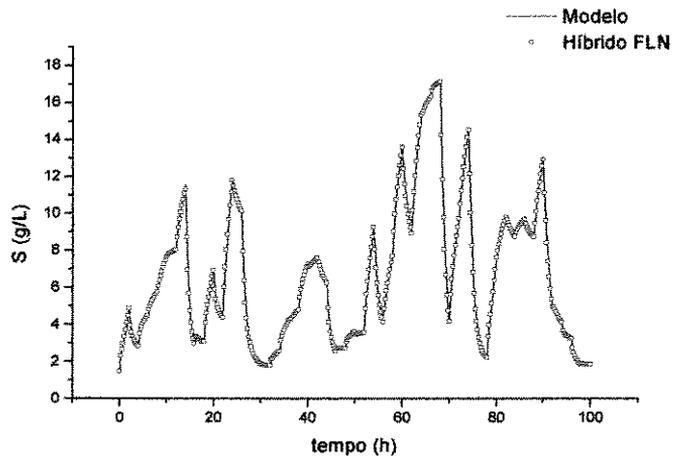


FIGURA 54: PERFIL DA CONCENTRAÇÃO DE ART, S

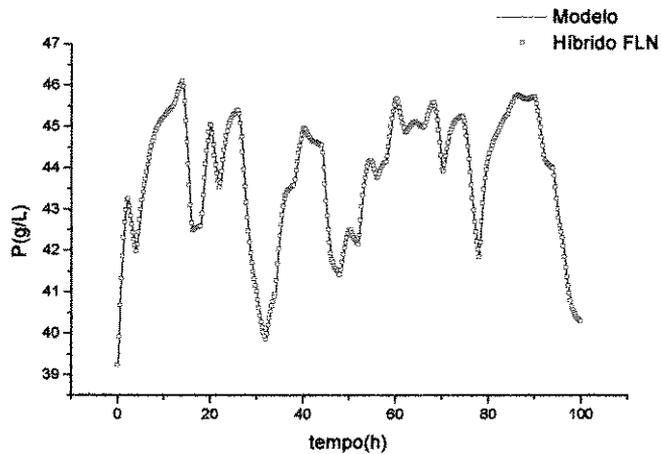


FIGURA 55: PERFIL DA CONCENTRAÇÃO DE ETOH, P

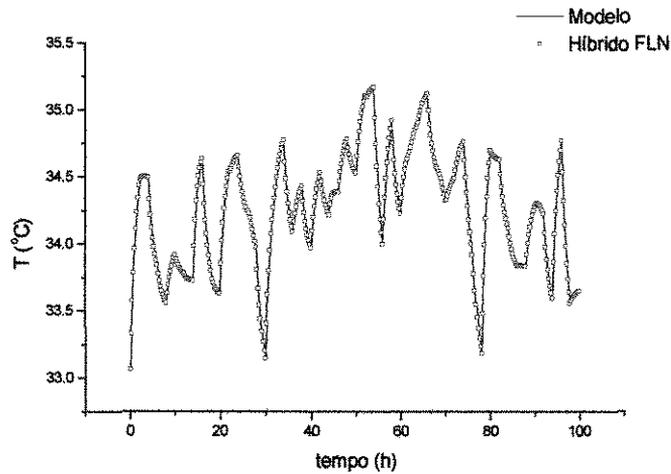


FIGURA 56: PERFIL DA TEMPERATURA, T

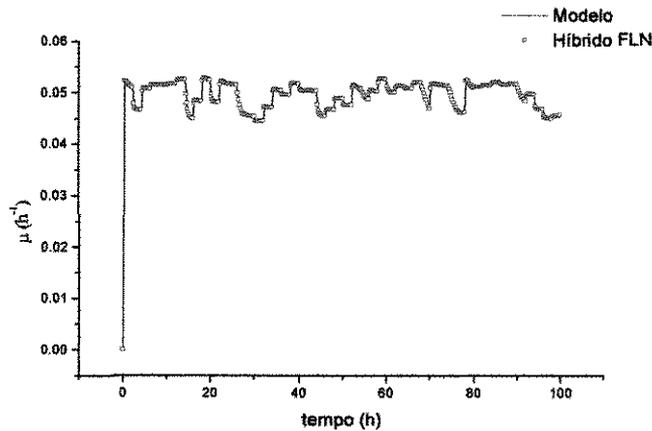


FIGURA 57: PERFIL DA TAXA DE CRESCIMENTO ESPECÍFICA, μ

A qualidade de predição de um modelo pode ser caracterizado usando o desvio padrão residual (DPR) representado pela equação (41). Este desvio fornece uma indicação da precisão da predição, como foi sugerido por CLERAN *et al.* (1991):

$$\text{DPR} = \sqrt{\frac{\sum_{i=1}^n (y_i - y_{pi})^2}{n}} \quad (41)$$

onde y_i é o valor calculado pelo modelo determinístico, y_{pi} é o valor calculado pelo modelo híbrido e n é o número de dados utilizados.

Como a magnitude de DPR vai variar dependendo da magnitude da variável a ser predita, é mais fácil analisar o DPR em forma de porcentagem da média dos valores calculados pelo modelo determinístico (\bar{y}_i) (ATALA *et al.*, 2001).

$$\text{DPR}(\%) = \frac{\text{DPR}}{y_i} \cdot 100 \quad (42)$$

Os resultados são mostrados na Tabela 7. Observa-se que os desvios são bem pequenos variando de 5.88E-03 até 1.26 para a rede FNN, e de 7.47E-03 até 9.37E-01 para a rede FLN. Desvios abaixo de 10% são considerados aceitáveis dentro da engenharia de bioprocessos. Acredita-se que os valores pequenos obtidos foram em virtude do trabalho ser de simulação e os dados não apresentarem ruído.

Tabela 7 – Desvio Padrão Residual escrito em forma de porcentagem da média dos valores calculados pelo modelo determinístico

DPR(%)	DPR(%)	DPR(%)	DPR(%)	DPR(%)	DPR(%)	DPR(%)	DPR(%)	DPR(%)	DPR(%)
X (FNN)	S (FNN)	P (FNN)	T (FNN)	μ (FNN)	X (FLN)	S (FLN)	P (FLN)	T (FLN)	μ (FLN)
5.88E-03	1.26	6.92E-02	1.10E-02	1.40E-01	3.96E-03	9.37E-01	4.75E-02	7.47E-03	7.60E-02

Para a concentração de biomassa, X, os dois modelos apresentam desempenho similar. Já para a concentração de substrato, S, o modelo híbrido acoplado a rede FLN apresenta um desempenho superior ao apresentado pela rede FNN. Para a concentração de produto, P, os modelos apresentam desempenho similares. Para a descrição da temperatura, T, o modelo híbrido acoplado à FLN tem um desempenho melhor que aquele obtido para o modelo híbrido acoplado à FNN. E para a descrição da cinética do processo, o modelo híbrido que utiliza a FLN demonstrou melhor resultado do que o obtido para o modelo híbrido que utiliza a FNN.

Por outro lado, observando as Figuras 48 a 52, que apresenta a comparação entre o modelo determinístico e o modelo híbrido acoplado a rede *feedforward* e as Figuras 53 a 57, onde é demonstrado o desempenho do modelo híbrido acoplado à rede *funcional link* com o modelo híbrido, verifica-se que tanto o modelo híbrido que utiliza a rede neuronal *feedforward* quanto o modelo híbrido que utiliza a rede do tipo *funcional link* podem ser

utilizados para prever o desempenho do processo e para o projeto de controladores de processo, uma vez que os desvios apresentados foram bem pequenos e abaixo de 10% para ambos os modelos. Porém, a rede *funcional link* apresentou um desempenho superior ao da rede *feedforward* na descrição do comportamento da concentração de substrato, da temperatura e da taxa específica de crescimento.

CAPÍTULO VI

6. CONCLUSÕES

A revisão bibliográfica realizada procurou descrever os itens de grande importância no entedimento dos estudos subsequentes, principalmente, no estudo da otimização do processo proposto, através do uso de redes neuronais para descrição da cinética do processo.

O modelo determinístico que representa o processo foi baseado em cinética fermentativa e balanços de massa e energia. Esta modelagem do processo é representada pelo sistema de equações diferenciais ordinárias para reproduzir os perfis de células, substrato, produto e temperatura no reator, utilizando o método de Runge-Kutta de 4ª ordem através de um programa em linguagem Fortran 90. Um estudo foi realizado para entender o comportamento dinâmico do processo, o que foi útil na fase de desenvolvimento do modelo híbrido.

SILVA(1998) mostrou que as condições utilizadas para as variáveis de entrada ($S_0=180$ g/L, $R=0.35$, $r=0.45$ e $F_0=100$ L/h) levaram a uma alta performance do sistema de operação, para o qual o rendimento foi de 86,04% e produtividade de 21 g/L.

Um modelo híbrido foi proposto para o processo de fermentação alcoólica contínua e extrativa. Este modelo híbrido é composto das equações do balanço de massa e energia no fermentador, de uma rede neuronal que descreve a taxa específica de crescimento e dos balanços de massa no processo extrativo. As redes neuronais utilizadas foram as do tipo *feedforward* e *funcional link*.

Um estudo da influência do número de neurônios na camada interna da rede neuronal *feedforward* no desempenho do processo foi apresentado e, após vários testes, concluiu-se que a rede com 5 neurônios apresentou bons resultados para descrever a taxa específica de crescimento e o modelo híbrido construído com essa estrutura foi capaz de descrever bem a dinâmica do processo de fermentação alcoólica extrativa.

Uma rede neuronal pouco conhecida dentro da engenharia química, FLN, foi testada na descrição da cinética do processo e seu desempenho comparado ao da rede FNN, através de um desvio padrão residual(DPR).

A rede FLN apresentou desempenho um pouco superior àquele obtido para a rede *feedforward* porém com algumas vantagens: como treinamento desta estrutura de rede é linear, tem convergência garantida e o tempo gasto é bem pequeno, a FLN possibilita uma

simples implementação de um esquema adaptativo onde os parâmetros são reestimados on-line baseados nos dados do processo.

SUGESTÕES PARA TRABALHOS FUTUROS

- Um estudo do processo de fermentação alcoólica extrativo na prática comparando os resultados reais com os obtidos através desta simulação.
- Realizar o estudo da influência do valor inicial dos bias na convergência do treinamento da rede do tipo *feedforward*.
- Utilizar dados reais para o treinamento das topologias de redes aqui apresentadas.
- Fazer uso da colocação ortogonal para diminuir a propagação do erro.

REFERÊNCIAS

- ACUÑA, G.; LATRILLE, E.; BÉAL, C. e CORRIEU, G., “Static and dynamic neural network models for estimating biomass concentration during thermophilic lactic acid bacteria batch cultures”, *Journal of Fermentation and Bioengineering*, vol. 85, no. 6, 615-622, 1998.
- AIBA, S.; SHODA, M.; NAGATANI, M., "Kinetics of product inhibition in alcohol fermentation, *Biotechnology and Bioengineering*, vol. 10, 845-864, 1968.
- ALVES, T. L. M.; COSTA, A. C.; HENRIQUES, A. W. S. e LIMA, E. L., “ Adaptive Optimal Control of Fed-batch Alcoholic Fermentation”, *Applied biochemistry and biotechnology* , vol. 70-72, 463-477, 1998.
- ANDRIETTA, S. R. “Modelagem, simulação e controle de fermentação alcoólica contínua em escala industrial”. Tese de Doutorado, Faculdade de Engenharia de Alimentos – FEA, Universidade Estadual de Campinas, Campinas, 1994.
- AZEVEDO, S. F.; DAHM, B.; OLIVEIRA, F. R., “Hybrid modelling of biochemical process: a comparison with the conventional approach”, *Computers chem. Engng*, vol. 21, S751-S756, 1997.
- BENQLILOU, C., RUIZ, D., ESPUÑA, A. and PUIGJANER, L.; Enhancement of prediction and extrapolation properties by using hybrid neural network model. *Proceedings of CHISA 2000*.
- BILLINGS, S. A., CHEN, S. e KORENBERG, M. J.; Identification of MIMO non-linear systems using a forward-regression orthogonal estimator, *Int. J. Control*, vol. 49, no. 6, 2157-2189, 1989.
- BHAT, N. e McAVOY, T. J., “Use of neural nets for dynamic modeling and control of chemical process systems”, *Computers chem. Engng.*, vol. 14, no. 4/5, 573-583, 1990.
- BROWN e HARRIS, “Neurofuzzy Adaptive Modelling and Control”, **Editora**, 1994.
- CHEN, L., BERNARD, O., BASTIN, G., ANGELOV, P.; Hybrid modelling of biotechnological processes using neural networks. *Control Engineering Practice*, 8, 821-827, 2000.
- CHEN, S. e BILLINGS, S. A., “Neural networks for nonlinear dynamic system modelling and identification”, *Int. J. Control*, vol. 56, no. 2, 319-346, 1992.
- CLERAN, Y.; THIBAUT, J.; CHERUY, A.; CORRIEU, G., “Comparison of prediction performances between models obtained by the group method of data handling and

- neural networks for the alcoholic fermentation rate in enology”, *Journal of Fermentation and Bioengineering*, vol. 71, no. 5, 356-362, 1991.
- COSTA, A. C., “Controle Ótimo Adaptativo em Processos Fermentativos”, Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro-RJ, 1998.
- COSTA, A. C., DECHECHI, E. C., SILVA, F. L. H., MAUGERI FILHO, F. and MACIEL FILHO, R.; “Simulated Dynamics and Control of an Extractive Alcoholic Fermentation”. *Applied Biochemistry and Biotechnology*, vol 84-6, SPR 577-593, 2000.
- COSTA, A. C.; ALVES, T. L. M; HENRIQUES, A. W. S., MACIEL FILHO, R. e LIMA, E. L., “An adaptative optimal control scheme based on hybrid neural modelling”, *Computers chem. Engng.*, vol. 22, suppl., S859-S862, 1998.
- COSTA, A. C.; HENRIQUES, A. S. W.; ALVES, T. L. M.; MACIEL FILHO, R. e LIMA, E. L., “A hybrid neural model for optimization of fed-batch fermentations”, *Brazilian Journal of Chemical Engineering*, vol. 16, no. 1, 53-63, March, 1999.
- CYSEWSKI, G. R. and WILKE, C. R.; Rapid ethanol fermentations using vacuum and cell recycle. *Biotechnology and bioengineering*, vol. XIX, 1125-1143, 1977.
- DEMUTH, H. e BEALE, M. “NEURAL NETWORK TOOLBOX USER’S GUIDE *For Use With Matlab*”, The Math Works, Inc., January, 1994.
- DI MASSIMO, C., MONTAGUE, G.A., WILLIS, M.J. *et al.* “Towards Improved Penicillin Fermentation via Artificial Neural Networks”, *Comp. Chem. Engng.*, v. 16, pp. 283-291, 1992.
- FARZA, M., HAMMOURI, H., OTHMAN, S. and BUSAWON, K.; Nonlinear observers for parameter estimation in bioprocesses. *Chemical Engng. Science*, vol. 52, no. 23, pp. 4251-4267, 1997.
- FERRAS, E.; MINIER, M.; GOMA, G., “Acetonobutylic fermentation: improvement of performances by coupling continuous fermentation and ultrafiltration”. *Biotechnology and Bioengineering*, vol. 28, 523-533, 1986.
- FEYO DE AZEVEDO, S., DAHM, B. and OLIVEIRA, F. R.; Hybrid Modelling of Biochemical Processes: A comparison with the conventional approach. *Computers chem. Engng.*, vol. 21, Suppl.,pp. S751-S756, 1997.

- FONSECA, E.F., "Controle Preditivo Baseado em Redes Neurais de Fermentação Alcoólica Contínua", Tese de Mestrado, UFPE, Recife, PE, 1998.
- FOTOPOULOS, J.; GEORGAKIS, C. e STENGER Jr., H. G., "Uncertainty issues in the modeling and optimization of batch reactors with tendency models, Chemical Engineering Science, vol. 49,no. 24B, 5533-5547, 1994.
- FU, P. e BARFORD, J. P., "Hybrid modelling combining a detailed metabolic simulation and neural networks approaches to complex biochemical processes", Proceedings of PSE, 571-574, 1994.
- GLASSEY, J.; MONTAGUE, G. A.; WARD, A. C. e KARA, B. V., "Artificial neural network based experimental design procedures for enhancing fermentation development", Biotechnology and bioengineering, vol. 44, 397-405, 1994.
- HAYKIN, S. Neural Networks A Comprehensive Foundation Second Edition Prentice Hall Inc., 1999
- HENRIQUE, H. M. e LIMA, E. L., "Model Structure Determination in Neural Network Models", submetido, 1996.
- HENRIQUES, A. W. S.; COSTA, A. C.; ALVES, T. L. M. e LIMA, E. L., "Optimization of fed-batch processes: challenges and solutions, Brazilian Journal of Chemical Engineering, vol. 16, no. 2, 171-177, 1999.
- HORNIK, K., M. Stinchcombe, H. White, "Multilayer Feedforward Networks are Universal Approximators", *Neural Networks*, v 2, pp 359-366, 1989.
- HUSSAIN, M. A., "Review of the applications of neural networks in chemical process control – simulation and online implementation", *Artificial Intelligence in Engineering*, 13, 55-68, 1999.
- IORDACHE, O.; CORRIOU, J. P.; GARRIDO-SANCHEZ, L.; FONTEIX, C.; TONDEUR, D., "Neural network frames. Application to biochemical kinetic diagnosis", *Computers chem. Engng.*, vol. 17, no. 11, 1101-1113, 1993.
- ISHIDA, K. & SHIMIZU, K. , "Novel repeated batch operation for flash fermentation system: experimental data and mathematical modeling. *Journal of Chemical Technology & Biotechnology*, vol. 66, 340-346,1996.
- KARIM, M. N. e RIVERA, S. L., "Comparison of *feedforward* and recurrent neural networks for bioprocess state estimations, Proceedings of ESCAPE, S369-S377, 1992.

- KARIM, M. N.; YOSHIDA, T.; RIVERA, S. L.; SAUCEDO, V. M.; EIKENS, B. e GYO-SEOP OH., "Global and local neural network models in biotechnology: Application to different cultivation processes", *Journal of Fermentation and Bioengineering*, vol. 83, no. 1, 1-11, 1997.
- KÜHM, I., "Communications to the editor: Alcoholic fermentation in an aqueous two-phase system". *Biotechnology and Bioengineering*, vol.22, 2393-2398, 1980.
- LE LANN, M. V., CABASSUD, M., CASAMATTA, G.; Modeling, Optimization and Control of Batch Chemical Reactors in fine chemical production. *Annual Reviews in Control*, 23, 25-34, 1999.
- LEAL, R. R.; BUTLER, P.; LANE, P. e PAYNE, P. A., "Data fusion and artificial neural networks for biomass estimation", *IEE Proc.-Sci, Meas. Technol.* Vol. 144, no. 2, March, 1997.
- LEE, J. H.; WOODARD, J. C.; PAGAN, R. J.; ROGERS, P. L., "Vacuum fermentation for ethanol production using strains of *Zymomonas mobilis*". *Biotechnology Letters* , vol.3, no.4, 177-182, 1981.
- LINKO, P. e ZHU, Y.," Neural network programming in bioprocess variable estimation and state prediction", *J. of biotechnology*", 21, 253-270, 1991.
- MAIORELLA, B.; BLANCH, H. W.; WILKE, C. R., "By- production inhibition effects on ethanolic fermentation by *Saccharomyces cerevisiae*". *Biotechnology and Bioengineering*, vol.25, 534-541, 1983.
- MATSUMURA, M. & MÄRKL, H., "Elimination of ethanol inhibition by perstraction". *Biotechnology and Bioengineering*, vol.28, 534-541, 1986.
- MILTON, J. S. e ARNOLD, J. C., "Introduction to Probability and Statistics", McGraw Hill. New York, 1990.
- MINIER, M. and GOMA, G.; Ethanol Production by extractive fermentation. *Biotechnology and bioengineering*, vol. XXVI, pp. 1565-1579, 1982.
- PAO, Y.-H., "Adaptative Pattern Recognition and Neural Networks. Reading. Mass: Addison-Wesley, 1989.
- PENGCHENG FU and BARFORD, J. P.; Hybrid modelling combining a detailed metabolic simulation and neural network approaches to complex biochemical processes. *Proceedings of PSE '94*, 571-574.

- PSICHOGIOS, D. C. e UNGAR, L. H., "A hybrid neural network-first principles approach to process modeling, *AIChE Journal*, vol. 38, no. 10, 1499-1511, 1992.
- QI, H., ZHOU, X.G., LIU, L. H.; YUAN, W. K., "A hybrid neural network-first principles model for fixed-bed reactor", *Chem. Eng. Science*, 54, 2521-2526, 1999.
- RAMALINGHAM, A. E R.K. FINN, "The vacuform process, a new approach to fermentation of alcohol", *Biotechnol. Bioeng.*, 19 (1977), 583-589.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J., "Learning representations by back-propagating errors", *Nature*, v. 323, 533-536, 1986.
- SAVKOVIC-STEVANOVIC, J., "Neural networkks for process analysis optimization: modeling and applications", *Computers chem. Engng.*, vol. 18, no. 11/12, 1149-1155, 1994.
- SAXÉN, B. e SAXÉN, H., "A neural network based model of biorection kinetics", *The Canadian Journal of Chemical Engineering*, vol. 74, 124-131, 1996.
- SCHUBERT, J., SIMUTIS, R., DORS, M., HAVLIK, I. e LÜBBERT, A., "Bioprocess Optimization and control: application of hybrid modelling", *Journal of Biotechnology*, 35, 51-68, 1994.
- SHENE, C.; DIEZ, C. e BRAVO, S., "Neural networks for prdiction of state of *Zymomonas mobilis* CP4 batch fermentations", *Computers and Chem Engineering*, 23, 1097-1108, 1999.
- SILVA. F. L. H., "Modelagem, Simulação e Controle de Fermentação Alcoólica Contínua Extrativa", Tese de Doutorado, FEA/UNICAMP, Campinas-SP, 1998.
- SIMUTIS, R., HAVLIK, I., LÜBBERT, A. "Fuzzy-Aided Neural Network for Real-Time State Estimation and Process Prediction in the Alcohol Formation Step of Production-Scale Beer Brewing", *J. Biotechnol.*, v. 27, pp. 203-215, 1993.
- SLAPACK, G. E.; RUSSEL, I.; STEWART, G. G., "Thermophilic microbes in ethanol production". Florida: CRC Press, 1987.
- TAYLOR, F.; KURANTZ, M. J.; GOLDBERG, N.; CRAIG Jr., J. C., "Control of packed column fouling in the continuous fermentation and stripping of ethanol". *Biotechnology and Bioengineering*, vol. 51, 33-39, 1996.
- THIBAUT, J.; VAN BREUSEGEM, V.; CHÉRUUY, A., "On-Line prediction of fermentation variables using neural networks", *Biotechnology and Bioengineering*, vol. 36, 1041-1048, 1990.

- VICENTE, M. S.; GONZO, E. E.; GOTTIFREDI, J. C., "Separation of ethanol-water mixtures by pervaporation through poly gel membranes". *Latin American Applied Research*, Buenos Aires, vol. 23, no. 3, 173-178, 1993.
- WANG, J., "Approximated Canonical Decomposition of Non-Linearity", *Int. J. Cont.*, v 56, pp 655-672, 1990.
- WILLIS, M. J., DI MASSIMO, C., MONTAGUE, G. A., THAM, M. T., MORRIS, A. J.; Artificial neural networks in process engineering. *IEE Proceedings-D*, vol. 138, no. 3, May 1991.
- WOOLEY, R., RUTH, M., GLASSNER, D. and SHEEHAN, J.; Process design and costing of bioethanol technology: a tool for determining the status and direction of research and development. *Biotechnol. Prog.* 15, 794-803, 1999.
- YOU, Y. e NIKOLAAU, M., "Dynamic process modeling with recurrent neural networks", *AIChE Journal*, vol. 39, no. 10, 1654-1667, 1993.
- ZORZETTO, L. F. M., MACIEL FILHO, R. and WOLF-MACIEL, M. R.; Process modelling development through artificial neural networks and hybrid models. *Computers and Chemical Engineering*, 24, 1355-1360, 2000.