

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA QUÍMICA

ÁREA DE CONCENTRAÇÃO

DESENVOLVIMENTO DE PROCESSOS BIOTECNOLÓGICOS

**MODELAGEM DINÂMICA DE BIOPROCESSOS POR
MODELOS HÍBRIDOS**

Autor: **Vera Lúcia Reis de Gouvêa** (*Engenheira de Alimentos, UFV, 1997*)

Orientador: Prof. Dr. **Rubens Maciel Filho**

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

Campinas - São Paulo

Outubro - 2000

UNICAMP

BIBLIOTECA CENTRAL

SEÇÃO CIRCULANTE



540101045

UNIDADE BC
 N.º CHAMADA:
T/UNICAMP
G745m
 V. _____ Ex. _____
 TOMBO BC/ 43453
 PROC. 16-392101
 C D
 PREC.º R\$ 11,00
 DATA 25/01/01
 N.º CPD _____

CM-00153430-9

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

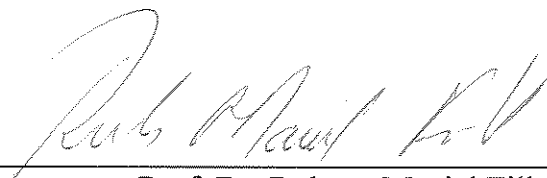
G745m

Gouvêia, Vera Lúcia Reis de
 Modelagem dinâmica de bioprocessor por modelos
 híbridos / Vera Lúcia Reis de Gouvêia.--Campinas,
 SP: [s.n.], 2000.

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

1. Redes neurais (Computação). 2. Álcool. 3.
 Modelos matemáticos. 4. Fermentação. I. Maciel
 Filho, Rubens. II. Universidade Estadual de Campinas.
 Faculdade de Engenharia Química. III. Título.

Dissertação de tese de Mestrado defendida por Vera Lúcia Reis de Gouvêia e aprovada em 30 de Outubro de 2000 pela banca examinadora constituída pelos doutores:



Prof. Dr. Rubens Maciel Filho
Orientador



Prof. Dr. Francisco Maugeri Filho
Titular



Prof. Dr. Silvio Roberto Andrietta
Titular

Prof. Dr. Ângela Maria Moraes
Suplente

Agradecimentos

- A Deus e aos meus pais por tudo.
- A todos os amigos e familiares pelo incentivo.
- Ao meu marido Wanderley pelo auxílio, paciência e amor.
- Ao professor Dr. Rubens Maciel Filho pelo apoio nas horas difíceis e orientação.
- Aos amigos do LOPCA pelo companheirismo e troca de informações.
- A FAPESP pelo auxílio financeiro, na forma de bolsa de estudo e reserva técnica, e pela sugestões ao trabalho.

Sumário

LISTA DE FIGURAS	X
LISTA DE TABELAS.....	xii
NOMENCLATURA	xiv
RESUMO	xviii
ABSTRACT	xix
1. INTRODUÇÃO.....	1
1.1 ORGANIZAÇÃO DO TRABALHO	2
2. REVISÃO DA LITERATURA	3
2.1. INTRODUÇÃO.....	3
2.2. PROCESSOS BIOTECNOLÓGICOS.....	3
2.3. PROCESSO FERMENTATIVO PARA PRODUÇÃO DE ETANOL	4
2.3.1. <i>Histórico</i>	4
2.3.2. <i>Produção de Etanol Via Fermentativa</i>	5
2.4. REDES NEURAIS ARTIFICIAIS	8
2.4.1. <i>Introdução</i>	8
2.4.2. <i>Histórico</i>	9

2.4.3. Neurônios Biológicos e Artificiais	10
2.4.4. Passos para o Desenvolvimento de uma Rede Neural Artificial.....	13
2.4.5. Considerações Finais.....	14
2.4.6. Aplicações	15
2.5. MODELAGEM HÍBRIDA	17
2.6. CONCLUSÃO	19
3. DESCRIÇÃO DO PROCESSO FERMENTATIVO PARA PRODUÇÃO DE ETANOL.....	20
3.1. INTRODUÇÃO.....	20
3.2. ETAPAS DO BIOPROCESSO	20
3.3. VARIÁVEIS	22
4. MODELAGEM MATEMÁTICA DO PROCESSO FERMENTATIVO PARA PRODUÇÃO DE ETANOL	25
4.1. INTRODUÇÃO	25
4.2. MODELO E PARÂMETROS CINÉTICOS.....	26
4.3. BALANÇO DE MASSA PARA O REATOR I	28
4.3.1. Balanço global	29
4.3.2. Balanço de Substrato	29
4.3.3. Balanço de Produto	29
4.3.4. Balanço de Células	29
4.4. BALANÇO DE ENERGIA	30
4.4.1. Fluido Reagente no reator i.....	30
4.4.2. Fluido Reagente no trocador de calor i.....	30
4.4.3. Fluido de Resfriamento no trocador de calor i.....	30
4.5. BALANÇOS DE MASSA APLICADOS AO SISTEMA	30
4.6. RENDIMENTO E PRODUTIVIDADE	32
4.7. CONCLUSÃO.....	32
5. ESTUDO DAS VARIÁVEIS DO BIOPROCESSO	33
5.1. INTRODUÇÃO.....	33

5.2. DINÂMICA DO BIOPROCESSO	33
5.3. ANÁLISE DE SENSIBILIDADE	36
5.4. ANÁLISE DA LINEARIDADE ENTRE AS VARIÁVEIS DE ENTRADA E SAÍDA	38
5.4.1. Princípio da Superposição	38
5.4.2. Fator de Escala	41
5.5. CONCLUSÃO	44
6. MODELAGEM DO PROCESSO FERMENTATIVO POR REDE NEURAL ARTIFICIAL	45
6.1. INTRODUÇÃO	45
6.2. CARACTERÍSTICAS DAS REDES NEURAS ARTIFICIAIS UTILIZADAS	46
6.3. TREINAMENTO DAS REDES NEURAS ARTIFICIAIS	47
6.3.1. Caso A	47
6.3.2. Caso B	53
6.4. CONCLUSÃO	59
7. MODELAGEM HÍBRIDA	60
7.1. INTRODUÇÃO	60
7.2. CÁLCULO DA VELOCIDADE ESPECÍFICA DE CRESCIMENTO DA <i>SACCHAROMYCES</i> <i>CEREVISAE</i> COM REDE NEURAL	60
7.3. MODELO HÍBRIDO	65
7.4. CONCLUSÃO	71
8. CONCLUSÕES GERAIS	72
9. REFERÊNCIAS BIBLIOGRÁFICAS	74
APÊNDICE	79

Lista de Figuras

Figura 1. Esquema geral do processos fermentativo para obtenção de Etanol	6
Figura 2. Constituintes da célula neural	11
Figura 3. Unidade de processamento de uma Rede Neural Artificial (Neurônio artificial).....	12
Figura 4. Tipo de Modelos:(a) Modelo determinístico (b); Modelo “Caixa Preta” ; (c) e (d) Modelos Híbridos “Caixa Cinza”	18
Figura 5. Esquema geral da construção de um modelo híbrido	18
Figura 6. Esquema do processo fermentativo para produção de etanol em escala industrial proposto por ANDRIETTA	22
Figura 7. Variáveis no processo	23
Figura 8. Reator de mistura perfeita i.....	28
Figura 9. Perfil de $S(4)$ para perturbação degrau em T_w	34
Figura 10. Perfil de $S(4)$ para perturbação degrau em RR, S_0, F_0 e TJE	34
Figura 11. Perfil de Produtividade e Rendimento para perturbação degrau em S_0	35
Figura 12. Perfil de $X(4)$ para perturbação degrau em F_0, S_0, TJE , e RR	35
Figura 13. Teste do princípio da superposição de $S(4)$ para as variáveis de entrada S_0 e F_0	39
Figura 14. Teste do princípio da superposição de $S(4)$, $REND$ e $PROD$ para as variáveis de entrada S_0 e RR	40
Figura 15. Teste do princípio do fator de escala de $S(4)$ para a variável de entrada S_0	41

Figura 16. Teste do princípio do fator de escala de REND, PROD e X(4), para a variável de entrada S0	42
Figura 17. Teste do princípio do fator de escala de S(4), REND e PROD, para a variável de entrada F0	43
Figura 18. Teste do princípio do fator de escala de X(4) para a variável de entrada F0.....	44
Figura 19. Fluxograma do programa de treinamento das Redes Neurais Artificiais	47
Figura 20. Configuração geral da Rede Neural Artificial (Caso A).....	48
Figura 21. Evolução do E_{QG} e E_{Rel} , para S(4), durante o treinamento.....	51
Figura 22. Gráficos de comparação das variáveis S(4), X(4), REND e PROD do modelo determinístico e da Rede Neural.....	52
Figura 23. Configuração geral da Rede Neural Artificial (Caso B).....	53
Figura 24. Fluxograma do programa de utilização dos pesos ajustados (Modelo “caixa preta”) ..	56
Figura 25. Perfil de S(4), X(4) e P(4) para perturbação degrau em F0 (Modelo determinístico e Redes Neurais Artificiais).....	57
Figura 26. E_{rel} para cada perturbação em F0	58
Figura 27. Configuração geral da Rede Neural Artificial para o cálculo de μ	61
Figura 28. Exemplo de uma superfície de erro	63
Figura 29. Valores de μ obtidos pela Rede Neural Artificial e pelo modelo	64
Figura 30. Evolução do E_{QG} durante o treinamento	65
Figura 31. E_{rel} para cada perturbação em F0	67
Figura 32. Perfis de REND e PROD para perturbação degrau de +20% em F0	68
Figura 33. Perfis S(1), S(2), S(3), S(4), X(1) e X(2) para perturbação degrau de +20% em F0.....	69
Figura 34. Perfis X(3), X(4), P(1), P(2), P(3) e P(4) para perturbação degrau de +20% em F0.....	70

Lista de Tabelas

Tabela 1. Produção de Etanol e açúcar de cana-de-açúcar 1996/97	5
Tabela 2. Comparação estrutural entre um neurônio biológico e um artificial	13
Tabela 3. Valores das variáveis operacionais de processo.....	22
Tabela 4. Variáveis independentes trabalhadas na simulação.....	23
Tabela 5. Variáveis dependentes trabalhadas na simulação.....	24
Tabela 6. Valores das variáveis independentes.....	24
Tabela 7. Valores dos Parâmetros Cinéticos.....	28
Tabela 8. Valores das constantes utilizadas	28
Tabela 9. Sensibilidade das variáveis do processo sobre as variáveis de saída	37
Tabela 10. Intervalo das variáveis de treinamento da rede (caso A).....	49
Tabela 11. Configurações de rede iniciais e respectivos erros (caso A).	50
Tabela 12. Erros para os pares de treinamento.....	51
Tabela 13. Erros para os pares de validação	51
Tabela 14. : Amostra do conjunto de dados para treinamento da Rede Neural Artificial Dinâmica.....	54
Tabela 15. Intervalo das variáveis de treinamento da rede (caso B).....	54
Tabela 16. Algumas configurações de rede iniciais e respectivos erros (caso B).....	55
Tabela 17. Intervalo das variáveis independentes e dependente.....	61

Tabela 18. Configurações de algumas redes e respectivos E_{rel} em % do valor de μ para dados de treinamento e validação.	62
Tabela 19. Valor de E_{rel} em % para S(4).	66
Tabela 20. Valor de E_{rel} em % para S, X, P, REND e PROD.	68

Nomenclatura

LETRAS LATINAS

a	<i>Constante de ajuste da dependência de P_{max} com a temperatura ($^{\circ}C$)</i>
A	<i>Constante de ajuste da dependência de μ_{max} com a temperatura ($^{\circ}C$)</i>
Cp_i	<i>Calor específico do fluido reagente ($cal.g^{-1}.^{\circ}C^{-1}$)</i>
Cp_j	<i>Calor específico do fluido de resfriamento ($cal.g^{-1}.^{\circ}C^{-1}$)</i>
E	<i>Energia de Ativação ($cal.mol^{-1}$)</i>
E_{GG}	<i>Erro quadrático global</i>
E_{Rel}	<i>Erro relativo médio (%)</i>
F	<i>Vazão do meio de alimentação ($m^3.h^{-1}$)</i>
F_a	<i>Vazão de água no tanque de tratamento ($m^3.h^{-1}$)</i>
F_{c_i}	<i>Vazão de fluido reagente no trocador de calor i ($m^3.h^{-1}$)</i>
F_{j_i}	<i>Vazão de fluido de resfriamento no trocador de calor i ($m^3.h^{-1}$)</i>
F_l	<i>Vazão de leite de levedura ($m^3.h^{-1}$)</i>
F_{l_1}	<i>Vazão de leite de levedura após sangria ($m^3.h^{-1}$)</i>
F_r	<i>Vazão de reciclo ($m^3.h^{-1}$)</i>
F_s	<i>Vazão de sangria de levedura ($m^3.h^{-1}$)</i>
F_v	<i>Vazão de vinho delevedurado ($m^3.h^{-1}$)</i>
$F_{W(i)}$	<i>Vazão de alimentação do reator i ($m^3.h^{-1}$)</i>
K_0	<i>Constante do modelo de dependência de P_{max} com a temperatura ($g.l^{-1}$)</i>
K_S	<i>Constante de afinidade ao substrato ($g.l^{-1}$)</i>

m	<i>Potência do termo de inibição pela concentração celular</i>
n	<i>Potência do termo de inibição pelo produto</i>
P	<i>Concentração de etanol ($\text{g}\cdot\text{l}^{-1}$)</i>
P_f	<i>Concentração de etanol na saída do último reator ($\text{g}\cdot\text{l}^{-1}$)</i>
P_i	<i>Concentração de etanol no reator i ($\text{g}\cdot\text{l}^{-1}$)</i>
P_{max}	<i>Concentração de etanol onde o crescimento celular cessa ($\text{g}\cdot\text{l}^{-1}$)</i>
P_r	<i>Concentração de etanol no reciclo ($\text{g}\cdot\text{l}^{-1}$)</i>
$Prod$	<i>Produtividade ($\text{g}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$)</i>
P_w	<i>Concentração de etanol na entrada do primeiro reator ($\text{g}\cdot\text{l}^{-1}$)</i>
$Rend$	<i>Rendimento do processo (%)</i>
R_G	<i>Constante dos gases ($\text{cal}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$)</i>
r_P	<i>Velocidade de formação de produto ($\text{g}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$)</i>
RR	<i>Taxa de reciclo</i>
r_S	<i>Velocidade de consumo de substrato ($\text{g}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$)</i>
r_X	<i>Velocidade de crescimento celular ($\text{g}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$)</i>
S	<i>Concentração de ART ($\text{g}\cdot\text{l}^{-1}$)</i>
S_0	<i>Concentração de ART no meio de alimentação ($\text{g}\cdot\text{l}^{-1}$)</i>
S_f	<i>Concentração de ART na saída do último reator ($\text{g}\cdot\text{l}^{-1}$)</i>
S_i	<i>Concentração de ART no reator i ($\text{g}\cdot\text{l}^{-1}$)</i>
S_r	<i>Concentração de ART no reciclo ($\text{g}\cdot\text{l}^{-1}$)</i>
S_w	<i>Concentração de ART na entrada do primeiro reator ($\text{g}\cdot\text{l}^{-1}$)</i>
t	<i>Tempo (h)</i>
T_a	<i>Temperatura de alimentação do reator 1 ($^{\circ}\text{C}$)</i>
T_{c_i}	<i>Temperatura do fluido reagente na saída do trocador i ($^{\circ}\text{C}$)</i>
T_i	<i>Temperatura do reator i ($^{\circ}\text{C}$)</i>
T_{j_e}	<i>Temperatura do fluido de resfriamento na entrada do trocador ($^{\circ}\text{C}$)</i>
T_{j_i}	<i>Temperatura do fluido de resfriamento na saída do trocador i ($^{\circ}\text{C}$)</i>

T_W	<i>Temperatura na entrada do primeiro reator ($^{\circ}\text{C}$)</i>
U	<i>Coefficiente global de troca térmica ($\text{Kcal.h}^{-1}.\text{ }^{\circ}\text{C}^{-1}.\text{m}^{-2}$)</i>
V	<i>Volume do reator (m^3)</i>
VE	<i>Valor da variável de entrada</i>
VS	<i>Valor da variável de saída</i>
X	<i>Concentração de células (g.l^{-1})</i>
X_i	<i>Concentração de células no reator i (g.l^{-1})</i>
X_l	<i>Concentração de células no leite de levedura (g.l^{-1})</i>
X_{max}	<i>Concentração de células onde o crescimento celular cessa (g.l^{-1})</i>
X_r	<i>Concentração de células no reciclo (g.l^{-1})</i>
X_V	<i>Concentração de células no vinho delevedurado (g.l^{-1})</i>
X_W	<i>Concentração de células na entrada do primeiro reator (g.l^{-1})</i>
$Y_{P/S}$	<i>Rendimento teórico em etanol</i>
$Y_{X/S}$	<i>Rendimento teórico em biomassa</i>

LETRAS GREGAS

β	<i>Taxa de Aprendizado</i>
α	<i>Termo de momento</i>
ΔH	<i>Calor de reação (Kcal.Kg^{-1})</i>
ρ_i	<i>Densidade do fluido reagente no reator i (Kg.m^{-3})</i>
ρ_j	<i>Densidade do fluido de resfriamento (Kg.m^{-3})</i>
μ	<i>Velocidade específica de crescimento do microrganismo (h^{-1})</i>
μ_{max}	<i>Velocidade específica de crescimento máxima do microrganismo (h^{-1})</i>

ABREVIATURAS

<i>ART</i>	<i>Açúcares Redutores Totais</i>
<i>RNA</i>	<i>Rede Neural Artificial</i>
<i>NEU</i>	<i>Número de neurônios na camada intermediária</i>
<i>LMDT</i>	<i>Média logarítmica das temperaturas</i>

Resumo

Os bioprocessos estão ganhando cada vez mais espaço nos processos produtivos por sua versatilidade, alta taxa de produção em produtos específicos e alto valor agregado dos produtos gerados. A característica comum destes processos é que apresentam um comportamento complexo onde diversos mecanismos estão envolvidos para a obtenção de produtos específicos.

Este projeto de pesquisa tem como principal objetivo o desenvolvimento de modelos híbridos para descrição do comportamento dinâmico de bioprocessos. Estes modelos híbridos são desenvolvidos acoplando-se Redes Neurais Artificiais (RNA) a modelos determinísticos sendo úteis, como representações matemáticas, para aplicações em controle e otimização em tempo real e principalmente como “*soft-sensors*”.

O bioprocesso tratado neste trabalho, como caso de estudo, é a produção de etanol via fermentativa utilizando-se o microrganismo *Saccharomyces cerevisiae*. O processo é representado por um modelo determinístico, constituído por um sistema de equações diferenciais ordinárias, resolvido pelo método Runge-Kutta de 4ª ordem. Este modelo foi devidamente validado e é utilizado como fonte de dados para estudos da dinâmica, e de desenvolvimento de modelos baseados em RNA e híbridos.

Para construção do modelo híbrido (modelo “Caixa Cinza”) foi treinada uma RNA que prediz a velocidade específica de crescimento quando fornecidos os valores das concentrações de produto, células e substrato. Após o ajuste esta RNA foi acoplada ao modelo determinístico fornecendo os perfis dinâmicos das variáveis: S, X e P (para os quatro reatores), rendimento e produtividade. Foi desenvolvido também um modelo tipo “Caixa Preta” do bioprocesso. Este modelo, baseado em RNA, fornece o perfil dinâmico de S(4), X(4) e P(4) durante 20h de processo. Todos os modelos foram devidamente testados sendo comprovadas suas eficiências.

Palavras-chave: Redes Neurais Artificiais, Etanol, Modelagem Híbrida.

Abstract

The biological processes are each time more important in the productive processes due to its versatility, high production rate in specific products and high value of the generated products. These processes present a complex behavior where several mechanisms are involved to obtain specific products.

The main goal of this research project is the development of hybrid models for description of the dynamic behavior of biological processes. These models are developed by coupling Artificial Neural Networks (ANN) and deterministic models. The hybrid models will be used as mathematical representations of the process, for applications in real time control and optimization and mainly as soft-sensors.

The biological process studied in this work is related with the fermentative ethanol production using the microorganism *Saccharomyces cerevisiae*. The process is represented by a deterministic model, constituted by a system of differential ordinary equations, that is solved by 4th order Runge-Kutta method. This model was validated properly in industrial scale and it is used as data source for studies of the dynamics of the process, development of the models with ANN as well as of the hybrid models.

For the construction of the hybrid model (“Gray Box” model) it was accomplished the training of a ANN that predicts the specific speed of growth of the microorganism when supplied the values of the product, cells and substrate concentrations.

After adjustment of the ANN, it is coupled to the deterministic model supplying the dynamic profiles of the variables: S, X and P (for the four reactors), efficiency and productivity. It was also developed a “Black Box” model of the biological process. This model, based in ANN, supplies the dynamic profile of S(4), X(4) and P(4) during 20h of process. All the models were properly tested and their efficiency were verified.

Key Words: Artificial Neural Networks, Ethanol, Hybrid Models.

Capítulo 1

Introdução

A utilização de bioprocessos vem crescendo de maneira significativa nos últimos anos requerendo tecnologias cada vez mais sofisticadas e grande número de pesquisas. Este trabalho vem contribuir no avanço do desenvolvimento de modelos para uso em sistemas de controle e otimização de bioprocessos. O assunto é considerado como uma das áreas de ponta de pesquisa em Engenharia Química. De fato, é uma nova etapa para o desenvolvimento e implementação de técnicas de otimização, controle avançado e operação assistida por computador.

O bioprocessos a ser tratado neste trabalho, como estudo de caso, é a produção de etanol via fermentativa utilizando-se o microorganismo *Saccharomyces cerevisiae*. O processo fermentativo para a obtenção do etanol será representado por um modelo determinístico, constituído por um sistema de equações diferenciais ordinárias, que pode ser resolvido numericamente por um algoritmo baseado no método Runge-Kutta de 4ª ordem. Este modelo já foi devidamente validado com dados experimentais da Usina Guarani, e será utilizado como fonte de dados para estudos de dinâmica, e de

desenvolvimento dos modelos de Redes Neurais Artificiais e híbridos.

O principal objetivo deste projeto de pesquisa é o estudo do desenvolvimento e implementação de modelos híbridos na descrição do comportamento dinâmico de bioprocessos. Um modelo híbrido será desenvolvido acoplando-se Redes Neurais Artificiais com modelo determinístico. Este modelo será útil, como representação matemática, para aplicações em controle e otimização em tempo real e principalmente como “*soft-sensors*” (sofisticado sistema de monitoramento que fornece o valor indireto de uma variável de difícil mensuração através de variáveis que podem ser medidas durante o processo).

1.1 - Organização do Trabalho

O capítulo 2 apresenta uma revisão da literatura consultada, dando destaque aos aspectos teóricos sobre Redes Neurais Artificiais . Em seguida, no capítulo 3, é feita uma breve descrição do bioprocessos.

A modelagem matemática do bioprocessos está contida no capítulo 4. No capítulo 5 é feito um estudo das variáveis do sistema. Dois modelos tipo “caixa preta” desenvolvidos utilizando-se Redes Neurais Artificiais são detalhados no capítulo 6. A modelagem híbrida é abordada no capítulo 7, que inclui o treinamento de uma Rede Neural Artificial para o cálculo da velocidade específica de crescimento da levedura. O trabalho possui ainda um Apêndice que contém fundamentos teóricos introdutórios sobre Redes Neurais Artificiais.

Capítulo 2

Revisão da Literatura

2.1 - Introdução

Neste capítulo será feita uma breve revisão sobre as características dos processos biotecnológicos em geral e o desenvolvimento do processo fermentativo para produção de etanol. Além disso, será abordado também a teoria sobre Redes Neurais Artificiais e Modelagem Híbrida.

2.2 - Processos Biotecnológicos

Os processos biotecnológicos estão ganhando cada vez mais espaço nos processos produtivos por sua versatilidade, alta taxa de produção em produtos específicos e alto valor agregado dos produtos por eles obtidos. Um processo biotecnológico pode utilizar microrganismos (células vivas), células mortas, organelas ou compostos isolados das células (enzimas), como agente de transformação bioquímica. Independentemente do tipo de agente, a característica comum destes processos é que apresentam um comportamento

complexo onde diversos mecanismos estão envolvidos para a obtenção de produtos específicos. Estes mecanismos basicamente envolvem o conhecimento da cinética e dos fenômenos de transporte e podem, dependendo do tipo de modelo a ser desenvolvido, considerar os aspectos intercelulares, intracelulares ou ambos. Estas características tornam os bioprocessos bastante complexos.

Os bioprocessos são, por natureza, variáveis com o tempo e não-lineares. Segundo KARIM et al (1997), a natureza tempo-variante se deve ao fato de que os microrganismos sofrem, ao longo do tempo, mudanças fisiológicas e morfológicas e mutações. Já a não-linearidade apresentada pelos bioprocessos é fruto de vários fatores, dentre eles: cinética celular de crescimento e formação de produtos, limitações termodinâmicas do processo, efeitos de transferência de massa e calor, agitação do fluido no reator, interação entre os fenômenos, dentre outros.

Os modelos biológicos são comumente obtidos dos princípios básicos aplicando-se teoria de conservação de massa e energia acrescidos da termodinâmica, cinética e correlações de transferência de massa e energia. Existe também uma dificuldade de medida “*on line*” dos produtos, uma vez que estes podem ser excretados pelos microrganismos (extracelular) ou permanecer na célula (intracelular).

2.3 - Processo Fermentativo para Produção de Etanol

2.3.1 - Histórico

A produção de etanol (álcool etílico) via fermentativa no Brasil é feita utilizando-se, principalmente, a cana-de-açúcar. A cana-de-açúcar foi introduzida no Brasil no século XVII e desde então passou a ser uma atividade importante chegando a ocupar grandes áreas de plantações de café.

O PRÓALCOOL (Programa de Incentivo à Produção de Álcool) foi criado pelo governo na 1ª crise do petróleo. A utilização do álcool como combustível agradou os usineiros pois o preço do açúcar sofria de constantes oscilações no mercado internacional naquela época. A perspectiva da utilização do álcool como combustível deu um enorme impulso na sua produção e no aprimoramento de técnicas. Na área agrícola houve melhoramento das variedades das canas cultivadas, das técnicas de plantio e colheita. Na

indústria desenvolveram-se novos processos, novos microrganismos foram testados, dentre outras melhorias, visando o aumento de rendimento e produtividade. A tabela 1 mostra alguns dados da produção de etanol no Brasil.

Tabela 1: Produção de Etanol e açúcar de cana-de-açúcar 1996/97.

	Cana-de-Açúcar	Açúcar	Etanol
Área cultivada (milhões hectare)	4,2	1,51	2,69
Produção total	273 (milhões de ton)	13,95 (milhões de ton)	13,90 (milhões de m ³)
Produtividade	65 t/ha	7,6 t/ha	5170 l/ha
Produtividade (por ton de cana)		0,138 t	79,5 litros ¹

Fonte: Datagro 1997,n° 2.Datagro 1997,n° 4.Datagro 1996,n° 10.Datagro 1996,n° 6.

Em 1996 existiam 352 unidades produzindo açúcar e álcool: 246 no centro-Sul e 106 no Nordeste. Do total das unidades, 196 produziam açúcar e álcool, 135 apenas álcool e 21 apenas açúcar (Datagro 1996, n° 15).

As indústrias devem, cada vez mais, se aperfeiçoar para não perder mercado. O apoio governamental ainda se faz necessário para manter o Brasil como detentor das tecnologias mais avançadas para produção fermentativa de etanol utilizando cana-de-açúcar. O etanol é considerado um combustível alternativo, renovável e com um futuro promissor nesta época de preocupação ecológica.

2.3.2 - Produção de Etanol Via Fermentativa

A fermentação alcoólica requer uma série de reações que podem ser resumidas em:



Os microrganismos mais utilizados na produção de álcool são as leveduras, principalmente a *Saccharomyces cerevisiae* (LIMA,1982). Elas funcionam como catalisadores biológicos.

A matéria-prima, fonte de glicose, pode variar e são enquadradas em 3 grandes grupos:

¹ Média incluindo a região nordeste. Na região Centro-Sul esta média sobe para cerca de 87 litros (citação ANDRIETTA,2000).

- Açucaradas (ex.: melão, cana-de-açúcar, beterraba açucareira)
- Amiláceas e feculentas (ex.: mandioca, milho, babaçu, batata-doce)
- Celulósicas (ex.: bagaço residual, madeira e seus resíduos)

As fontes denominadas açucaradas são mais vantajosas pois o microrganismo obtém a glicose mais facilmente. Existem microrganismos que produzem etanol diretamente das fontes amiláceas e feculentas não necessitando de prévio tratamento da matéria-prima. No Brasil, a grande extensão territorial e o cultivo histórico da cana-de-açúcar levou à sua utilização como principal fonte de matéria-prima. Além disso, estudos comprovam que a cana-de-açúcar possui maior produtividade em álcool por área colhida por ano se comparada a outras fontes de glicose.

O processo fermentativo para produção de etanol pode ser dividido em três fases:

- **Preparo da matéria-prima:** inclui moagem da cana, peneiragem, decantação (a quente) e resfriamento. Gera como resíduo o bagaço e a torta.
- **Transformação:** é o processo de fermentação no qual o açúcar é transformado em etanol pelos microrganismos.
- **Separação:** destilação com finalidade de separar o álcool do vinhoto.

A Fig. 1 apresenta um esquema geral do processo.

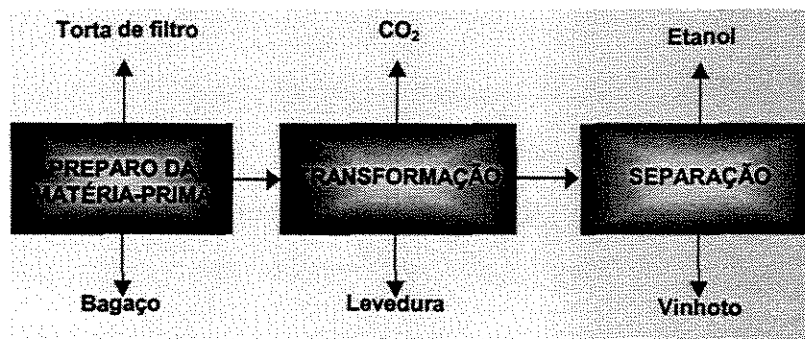


Figura 1: Esquema geral do processo fermentativo para obtenção de Etanol

Os processos fermentativo em escala industrial evoluíram, e estão evoluindo, desde a década de 70. Segundo ANDRIETTA (1998), a evolução dos processos de

fermentação alcoólica pode ser dividida em:

Processos de batelada alimentada: Utilizado na década de setenta com grande evolução nos anos 80, resultando em melhor desempenho e produtividade.

Processos de fermentação contínua: apresentam maior produtividade que os processos em batelada e batelada alimentada principalmente pelo fato de que não há paradas para descarga, lavagem e enchimento das dornas. O processo de fermentação contínua também se encontra em evolução:

1ª geração: evoluíram paralelamente com os processos de batelada alimentadas e não passaram de adaptações destes. Apesar do baixo custo possuíam sérios problemas de concepção. Foram responsáveis pelo atraso na implementação dos sistemas contínuos.

2ª geração: surgiram no final dos anos 80 introduzindo conceitos de engenharia aos projetos. Estes processos apresentam melhor desempenho e estabilidade oferecendo maior confiança que os anteriores. O ponto negativo é que não levam em conta fatores cinéticos importantes na otimização do processo.

3ª geração: apareceram no início dos anos 90. Levam em conta fatores cinéticos da levedura e não possuem problemas de fluxo. Levaram os processos contínuos a terem maior produtividade, menor consumo de insumos, melhor aproveitamento da massa celular excedente para secagem, maior facilidade para automação e possibilidade de operação em regime permanente.

ANDRIETTA (1998) expõe também o que poderia vir a ser uma evolução dos processos fermentativos contínuos que não necessitam da utilização de centrífugas para separação da levedura do vinho. As centrífugas são responsáveis por cerca de 25% do custo total de uma destilaria. No entanto estes processos ainda estão sendo estudados.

O processo convencional de fermentação possui características inibitórias devido ao fato do etanol (produto) inibir o crescimento celular. Esta inibição afeta de forma negativa o rendimento e a produtividade do processo. Para resolver este problema trabalhos recentes estão sendo desenvolvidos para extrair o etanol durante o processo diminuindo assim sua influência inibitória (SILVA, 1998).

No capítulo 4 será mostrada a descrição de um processo industrial para produção fermentativa de etanol.

Para uma revisão mais detalhada dos processos utilizados na produção de etanol, tipos de reatores comumente empregados e modelos cinéticos propostos, pode ser consultado o trabalho desenvolvido por KALIL (1997).

2.4 - Redes Neurais Artificiais

Nesta sessão serão abordados aspectos introdutórios das Redes Neurais Artificiais bem como sua utilização em várias áreas incluindo processos biotecnológicos. No Apêndice se encontram fundamentos teóricos introdutórios sobre Redes Neurais Artificiais. Nele estão descritos o funcionamento, a arquitetura, os tipos de redes, o treinamento e regras de atualização de pesos.

2.4.1 - Introdução

Segundo FEIGENBAUM (1981) “a inteligência artificial é a parte da ciência da computação ligada ao “*design*” de sistemas computacionais inteligentes, ou seja, sistemas que exibem características que podem ser associadas com o comportamento da inteligência humana”. Assim, as Redes Neurais Artificiais se enquadram nas técnicas de inteligência artificial.

Redes Neurais Artificiais (ANN - *Artificial Neural Networks*) são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes que adquirem conhecimento através da experiência (MENDES, 1997).

Por BHAGAT (1990), “redes neurais imitam o processo de aprendizagem humano. Elas podem identificar e aprender padrões que correlacionam um conjunto de dados de entrada a valores correspondentes na saída. Uma vez treinadas estas redes podem ser usadas para prever as saídas de um dado processo, sendo fornecidos apenas os dados de entrada”. Segundo CHITRA, BULSON e MORRELL (1995) as Redes Neurais Artificiais não são necessariamente seqüenciais nem determinísticas. São compostas de vários elementos de processamento simples que normalmente não fazem mais do que somar

entradas de forma ponderada. Ao invés de executar uma série de instruções, uma rede neural responde em relação às entradas que lhe são fornecidas. O resultado final consiste em um estado geral da rede obtido depois que ela atinge a condição de estado estacionário, o qual correlaciona padrões entre os dados de entrada e as saídas correspondentes ou os valores de saída desejados. A rede final pode ser usada para prever os resultados que seriam obtidos para um outro conjunto de dados de entrada.

De fato, as redes neurais são altamente versáteis, têm o potencial de resolver problemas não-lineares e complexos, mesmo que os dados existentes sobre o processo sejam imprecisos ou contenham ruídos. Dentro deste contexto se constituem em ferramentas de grande potencial para a representação do comportamento de bioprocessos, especialmente quando predições dinâmicas são requeridas.

A aplicação das Redes Neurais Artificiais está aumentando cada vez mais, sendo elas utilizadas também na modelagem de equipamentos. Sua utilização é adequada principalmente quando o sistema em questão é muito complexo (apresentando incertezas em relação a fluidodinâmica do processo, cinética ou parâmetros do modelo) e/ou quando o tempo de processamento gasto para a simulação através de modelos determinísticos é elevado (BATISTA, 1996).

2.4.2 - Histórico

As mais importantes publicações sobre Redes Neurais Artificiais foram desenvolvidas por McCULLOCK e PITTS (1943), HEBB (1949), e ROSENBLATT (1958). O modelo de MCCULLOCK e PITTS (ou modelo MCP) é baseado na plausibilidade neurofisiológica de um neurônio desempenhando uma função de soma e “*threshold*”, onde os pesos nas conexões entre os nós correspondem às sinapses inibidoras e excitadoras de um neurônio real. HEBB, em 1949, introduziu um modelo básico de rede de auto-organização no qual ele propôs uma regra (regra de Hebb) para o ajuste de pesos. A descrição do “*Perceptron*” de uma única camada por ROSENBLATT (1958) introduziu o conceito de aprendizado em Redes Neurais Artificiais. ROSENBLATT introduziu o modelo “*Perceptron*” de aprendizado supervisionado.

O “*Perceptron*” não durou muito, MINSKY publicou um livro no final da

década de 60 (MINSKY e PAPERT, 1969) demonstrando formalmente as limitações do “*Perceptron*” de uma única camada. Com isto as Redes Neurais Artificiais foram deixadas de lado. Os anos 60 e 70 não foram muito produtivos para redes neurais no entanto importantes trabalhos sobre modelos de redes neurais em visão, memória, controle e auto-organização foram realizados.

Um reinício da área aconteceu com a publicação dos trabalhos de HOPFIELD (1982) que utilizou redes simétricas para otimização e de RUMELHART, HINTON e WILLIAMS (1986b) que introduziram o método “*Backpropagation*”. Estes algoritmos mostraram que Redes Neurais Artificiais podem ser utilizadas sem as limitações defendidas por MARVIN MINSKY e PAPERT.

Um grande interesse pelas Redes Neurais Artificiais voltou a ocorrer após a publicação de trabalhos clássicos na área como: a descrição do modelo de Hopfield (HOPFIELD, 1982), Máquinas de Boltzmann (ACKLEY, HINTON e SEJNOWSKY, 1985) e o algoritmo “*Backpropagation*” (RUMELHART, HINTON e WILLIAMS, 1986a, 1986b). MENDES (1997) cita também outros trabalhos importantes como os desenvolvidos por: WILLSHAW, BUNEMAN e LONGUETT-HIGGINS (1969); ANDERSON (1972); KOHONEN (1972); FUKUSHIMA (1975); CARPENTER e GROSSBERG (1987a, 1987b) e KANERVA (1984).

Sistemas neurais sem peso (ALEKSANDER, 1966) também tiveram um papel importante na história das redes Neurais Artificiais, pois proporcionaram modelos de rápida aprendizagem e fácil implementação. Como referência de modelo sem peso pode-se citar o WISARD (ALEKSANDER, THOMAS e BOWDEN, 1984), o qual se tornou disponível como uma máquina comercial no início dos anos 80 e ainda está em uso atualmente em aplicações de reconhecimento de padrões.

2.4.3 - Neurônios Biológicos e Artificiais

2.4.3.1 - Neurônios Biológicos

O sistema nervoso é formado por células especiais denominadas neurônios. São os neurônios que controlam todo o corpo humano recebendo e enviando mensagens. A Fig. 2 mostra um neurônio típico com suas respectivas partes.

Os dendritos são um conjunto de terminais de entrada, o corpo central é composto pelo(s) núcleo(s) e organelas e os axônios são longos terminais de saída. A comunicação dos neurônios ocorre através de sinapses. Sinapse é definida como a região onde dois neurônios entram em contato e através da qual os impulsos nervosos são transmitidos entre eles (MENDES, 1997).

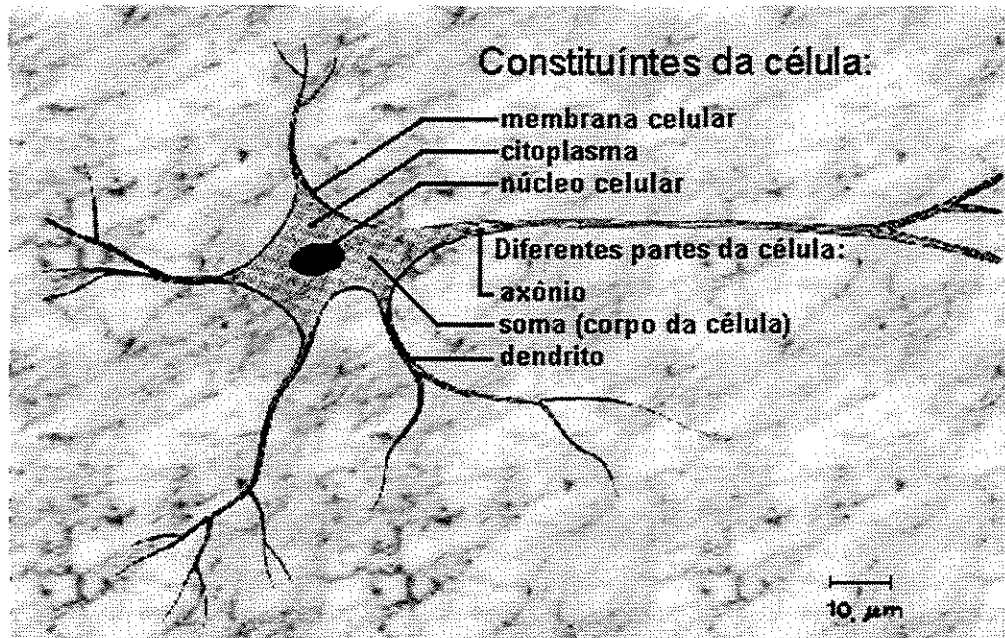


Figura 2: Constituintes da célula neural (MENDES, 1997)

Para exemplificar esta comunicação consideremos que um neurônio A recebe impulsos nervosos em um determinado momento. Estes impulsos são processados e, quando atingem um dado limiar de ação, o neurônio A dispara produzindo uma substância neurotransmissora que flui do corpo celular para o axônio que está conectado a um dendrito de um neurônio B. O neurotransmissor pode diminuir ou aumentar a polaridade da membrana pós-sináptica do neurônio A, inibindo ou excitando a geração dos pulsos no neurônio B. Este processo depende de vários fatores como a geometria da sinapse e o tipo de neurotransmissor (MENDES, 1997).

O cérebro humano possui entre 10^{10} e 10^{11} neurônios. Cada neurônio forma, em média, de 10^3 a 10^4 sinapses. Isto implica na formação de redes complexas com um número de sinapses na faixa de 10^{14} (ANDERSON, 1997).

O estudo cada vez maior do sistema nervoso humano possibilitou aos pesquisadores obter uma rede neural que imitasse o funcionamento do sistema biológico, o

que culminou na existência das Redes Neurais Artificiais .

2.4.3.2 - Neurônios Artificiais

Como já dito, Redes Neurais Artificiais são formadas por elementos de processamento que são os neurônios artificiais. A Fig. 3 mostra um esquema deste neurônio.

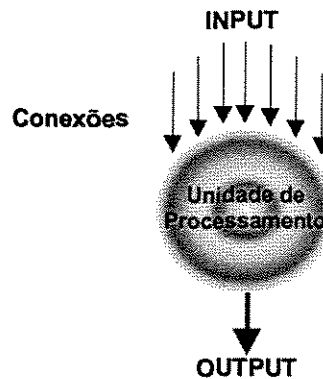


Figura 3: Unidade de processamento de uma Rede Neural Artificial (Neurônio artificial).

Os valores de entrada (*input*) são originados de conexões com outros neurônios ou de uma entrada externa à rede. São análogos aos dendritos dos neurônios biológicos. A unidade de processamento (corpo celular) é responsável pelas operações matemáticas necessárias (este processo está explicado no Apêndice). Os valores de saída (*output*) fazem conexão com outros “dendritos” e são análogos aos axiônios do neurônio biológico.

A tabela 2 mostra, de maneira mais detalhada, a comparação estrutural dos neurônios biológicos e artificiais.

Tabela 2: Comparação estrutural entre um neurônio biológico e um artificial

BIOLÓGICO		ARTIFICIAL	
Dendritos	São as ligações que conectam um neurônio a outro e fornecem a entrada (ativação ou inibição) para o neurônio.	Entradas e Saídas	As entradas são as saídas de outro neurônio ou podem ser uma entrada externa alimentada à rede. As entradas e saídas correspondem aos dendritos e aos axônios, respectivamente.
Axônio	É uma longa fibra neural que leva a informação de saída de um neurônio aos dendritos de um outro neurônio.		
Sinapse	É uma junção de um dendrito e um axônio. É a conexão através da qual passa uma carga eletro-química do axônio ao dendrito	Pesos	São as quantidades associadas às conexões entre as unidades de processamento.
Corpo Celular	É no corpo celular que a informação de entrada vinda dos dendritos é convertida para uma saída que é levada através do axônio até os dendritos de um outro neurônio.	Função Soma	Esta função combina as entradas ativadas (são as entradas e os pesos) em uma ativação única que será recebida pelo neurônio.
		Função de Transferência	Esta função converte a ativação única recebida pelo neurônio em uma saída. Esta informação de saída se transforma na informação de entrada de um outro neurônio

2.4.4 - Passos para o Desenvolvimento de uma Rede Neural Artificial

1º Passo: Coleta dos dados

Os dados colhidos devem ser significativos e cobrir o problema amplamente. Deve-se lembrar que a rede aprende com estes dados e se caso eles forem restritos às operações normais a rede não conseguirá trabalhar eficientemente em situações fora de controle.

2º Passo: Divisão dos dados

Os dados devem ser divididos em duas categorias: um conjunto será usado no treinamento e o outro para verificar se a rede está bem treinada. Este último conjunto é chamado de dados de teste ou validação. Um subconjunto de dados extraído dos dados de treinamento pode ser utilizado para verificar a capacidade de generalização durante o treinamento (MENDES, 1997).

Todos os dados devem ser colocados em ordem aleatória para evitar tendências. Podem também, dependendo do caso, sofrer normalizações, escalonamento e conversão de formato. Estes tratamentos podem tornar os dados mais apropriados para se trabalhar.

3° Passo: Definir a configuração da rede

Este passo inclui a seleção do paradigma neural, a determinação da topologia da rede e determinação dos parâmetros do algoritmo de treinamento e funções de ativação.

4° Passo: Treinamento

É a fase em que os pesos são ajustados. Deve-se definir os valores iniciais de cada peso (geralmente aleatório). O modo de treinamento utilizado normalmente é o padrão pois armazena quantidade de dados e é menos susceptível a problemas de mínimos locais. Porém, o modo “*batch*” torna o treinamento mais estável. O tempo de treinamento depende de vários fatores e um critério de parada pode ser utilizado. Um critério bastante utilizado é o número máximo de ciclos, no entanto, deve-se levar em consideração a taxa de erro média e a capacidade de generalização. O ponto ótimo de parada deve ser aquele em que a rede tem um erro mínimo e uma capacidade de generalização máxima.

5° Passo: Teste da rede

Nesta fase são utilizados dados que não foram usados no treinamento da rede. A análise destes dados permite a avaliação do desempenho da rede. Pode-se verificar se há neurônios com peso pequeno que podem ser descartados ou com pesos significativos que podem estar associados a um “*over-training*” da rede. O “*over-training*” faz com que a rede se especialize demais nos dados de treinamento perdendo a sua capacidade de generalização.

6° Passo: Integração da rede

Após o treino e validação a rede pode, então, ser acoplada ao ambiente operacional da aplicação para o qual ela foi criada. O sistema deve periodicamente monitorar o desempenho da rede fazendo sua manutenção e indicando se há necessidade de retreinamento.

2.4.5 - Considerações Finais

A taxa de aprendizado β da rede varia de 0 a 1. Quanto maior o valor de β mais rápido será o treinamento. No entanto, valores altos podem levar à oscilação do modelo. Sendo assim, é preciso ter uma taxa alta o suficiente para agilizar o treinamento mas que não leve à

oscilação.

Uma maneira de aumentar a taxa de aprendizado sem levar à oscilação é modificar a regra delta generalizada para incluir o termo momento (α). O termo momento leva em consideração o efeito de mudanças anteriores de pesos na direção do movimento atual no espaço de pesos. Adicionando o termo de momento à equação de correção dos pesos pode-se melhorar o tempo de treinamento do algoritmo de “*backpropagation*” e a estabilidade do processo. Torna-se útil em superfícies de erro que contenham longas gargantas, com curvas acentuadas ou vales com descidas suaves pois a rede consegue atingir o fundo dos vales na superfície de erro.

Um fenômeno chamado “paralisia da rede” ocorre quando o processo de treinamento fica extremamente lento. Este efeito é consequência do ajuste dos pesos em valores muito altos o que força os neurônios a trabalharem com altos valores de saída em uma região onde a derivada da função de ativação é pequena. Para contornar esta situação deve-se, geralmente, reduzir a taxa de aprendizado.

Para finalizar é interessante ressaltar um trecho extraído do livro de KOVÁCS (1996): *“Para o problema da aproximação de funções com redes neurais artificiais multicamadas, sabemos que sempre existe uma solução, apenas não sabemos como determiná-la. Esta talvez seja uma das chaves do sucesso da aplicação da teoria de redes neurais às mais diversas áreas de conhecimento: é um excelente substituto para ignorância.”*

2.4.6 - Aplicações

As Redes Neurais Artificiais são utilizadas em várias áreas. Sua primeira aplicação foi no processamento de sinal tentando diminuir o ruído de linhas telefônicas. Outras aplicações a serem citadas são: Controle de processos, Reconhecimento de padrões e voz; Produção de voz (leitura de um texto pelo computador); Medicina; Setor financeiro. NOVA (1992) aplicou Redes Neurais Artificiais em classificação de obras de arte (imagens de pinturas) quanto à sua autenticidade.

Na engenharia química seu uso vem crescendo muito nos últimos anos, tendo diversos trabalhos publicados. Apenas para exemplificar, pode-se citar o

desenvolvimento de software usando modelos determinísticos e redes neurais para o processo de craqueamento catalítico feito por BATISTA (1996), previsão de comportamento de reator de síntese de amônia (YAMAMOTO,1993), aplicação em controle (ENDER e MACIEL FILHO, 2000), desenvolvimento de “*soft sensor*” (MELEIRO e MACIEL FILHO, 2000) e para a construção de modelos com diferentes níveis de conhecimento (ZORZETTO, et al, 2000).

Segundo BATISTA (1996), Redes Neurais Artificiais são úteis na modelagem de sistemas complexos que apresentam incertezas em relação à fluidodinâmica do processo, cinética ou parâmetros do modelo, ou ainda, quando o tempo de processamento da simulação com modelos determinísticos é elevado.

KARIM et al (1997) mostram aplicações de Redes Neurais Artificiais em processos biotecnológicos. A primeira aplicação analisa a fermentação de açúcar por *Z.mobilis* produzindo álcool. Através da medição de variáveis tais como: temperatura, potencial de oxirredução, percentagem de CO₂ e densidade óptica obteve-se valores “*on-line*” de biomassa, etanol e concentração de glicose. Outra aplicação estudou a produção de anticorpos monoclonais (Mab) usando hibridoma. Diferentes redes foram utilizadas para prever as concentrações de Mab e células viáveis. Continuando a análise do trabalho de KARIM et al (1997), os autores ressaltam o uso de Redes Neurais Artificiais na modelagem de sistemas biológicos, mas, segundo eles, deve-se tomar cuidado com o tipo de rede a ser empregada (topologia, arquitetura, etc), seleção dos dados e números de parâmetros, levando-se em conta a acurácia desejada.

A aplicação de Redes Neurais Artificiais realmente é muito ampla e atinge vários setores. Uma rápida procura de trabalhos sobre o assunto comprova a afirmativa anterior. No entanto, isto não quer dizer que o campo de trabalho esteja saturado, pelo contrário, ainda há muito o que se descobrir. Na referência ARBIB (1995) se encontra uma coletânea dos artigos de grande importância na área de Redes Neurais Artificiais desde os primeiros trabalhos de MCCULLOCK e PITTS (1943).

Assim como em outras áreas, a aplicação de Redes Neurais Artificiais em processos biotecnológicos está crescendo bastante, abrindo um campo de trabalho muito grande e com grande potencial de aplicação.

2.5 - Modelagem Híbrida

O desenvolvimento de modelos puramente mecanísticos ou determinísticos (Fig. 4a) suficientemente detalhados para descrever com sucesso os principais fenômenos que ocorrem no processo é uma atividade cara e muitas vezes demorada. Nestes modelos todas as equações de balanço, taxas cinéticas e parâmetros necessários estão descritos por relações matemáticas bem definidas. A grande vantagem destes modelos é o conhecimento detalhado do processo que eles fornecem.

Por outro lado, o uso de modelos totalmente caixa preta, como é o caso de Redes Neurais Artificiais, pode levar a um uso inadequado das informações disponíveis (ZORZETTO et al, 2000). A utilização destes modelos (Fig. 4b) torna o processo desconhecido pois não se sabe as leis e relações que o regem. Sabe-se apenas a reação do processo após determinada perturbação. Não há justificativa, em princípio, do ponto de vista matemático para as respostas do modelo.

Os modelos híbridos, também chamados de modelos “caixa cinza” combinam equações determinísticas com técnicas de modelos “caixa preta” (que, em princípio, serão reservadas para as partes não muito esclarecidas do sistema), de forma a obter um uso eficiente das informações disponíveis. A razão para isto é que existem aspectos do processo que podem ser acuradamente descritos por equações relativamente simples, como os balanços de massa e energia. Pode-se gerar os modelos híbridos utilizando diferentes níveis de informação conforme mostrado em Fig. 4c e Fig. 4d. Normalmente, as análises têm mostrado que os modelos de Redes Neurais Artificiais puros apresentam um bom desempenho na faixa das condições de operação do processo, que foram utilizadas no seu treinamento. No entanto, a inclusão do conhecimento mecanístico, através das equações determinísticas, aumenta as capacidades de extrapolação e interpolação do modelo.

O modelo híbrido deve ser desenvolvido após definida a estrutura ótima da rede e seu treinamento. A rede treinada e validada é acoplada ao modelo determinístico.

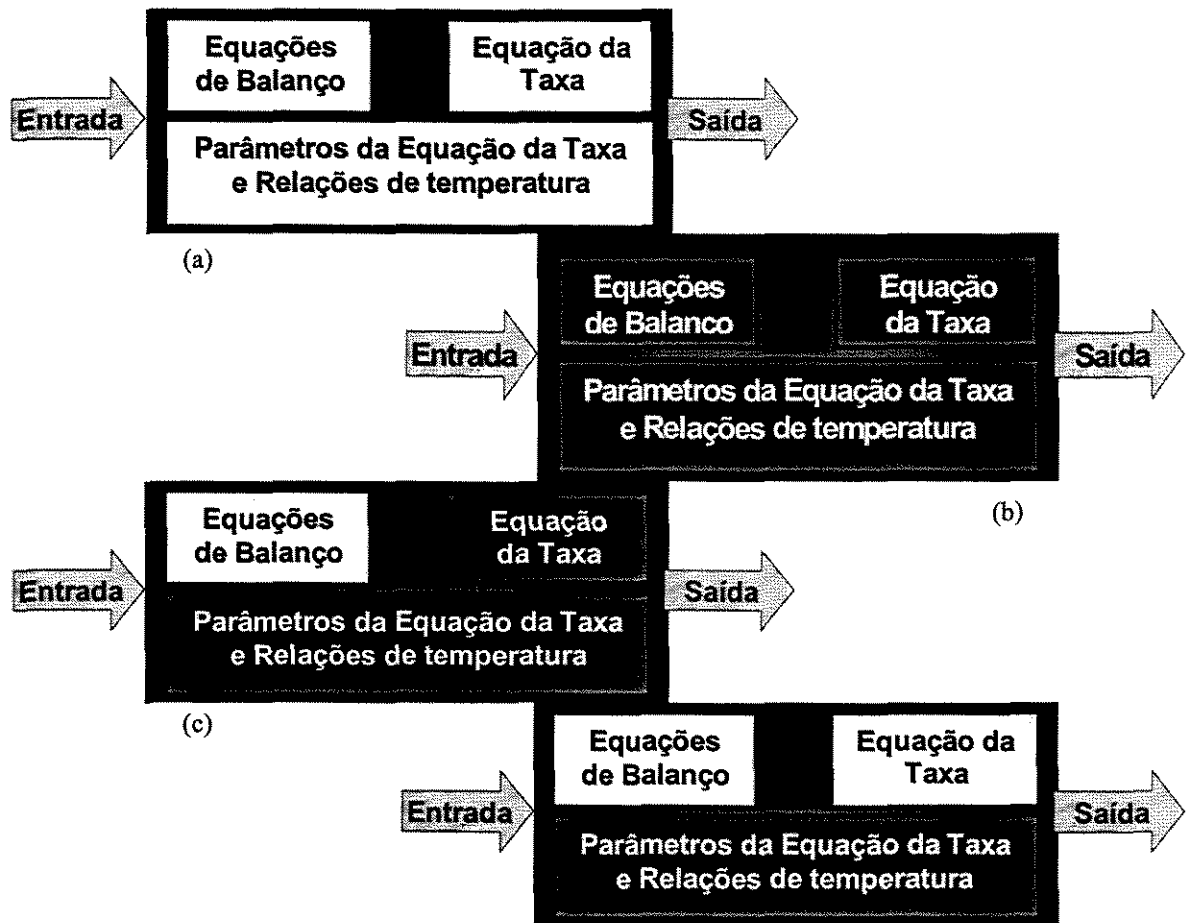


Figura 4: Tipo de Modelos: (a) Modelo determinístico. (b) Modelo “Caixa Preta” (ex.: Redes Neurais Artificiais) (c) e (d) Modelos Híbridos “Caixa Cinza”

Dentro deste contexto será desenvolvido um modelo híbrido (de acordo com a Fig.5) para representar o sistema fermentativo de produção de etanol e modelos tipo “caixa preta”, visando a obtenção de modelos úteis para aplicações em controle e otimização em tempo real e também para serem utilizados como “*soft-sensor*”.



Figura 5: Esquema geral da construção de um modelo híbrido.

2.6 - Conclusão

Pelo exposto neste capítulo pôde-se comprovar a viabilidade de se utilizar Redes neurais Artificiais na modelagem de bioprocessos. Sendo os bioprocessos variáveis com o tempo e não-lineares as Redes Neurais Artificiais com sua alta versatilidade, tem o potencial de resolver estes problemas complexos.

Com base nesta revisão será aplicado o conceito de Redes Neurais Artificiais na obtenção de um modelo híbrido e modelos tipo “caixa preta” para o processo fermentativo de produção de etanol. Este processo foi escolhido como estudo de caso por apresentar um modelo matemático validado industrialmente. O capítulo seguinte apresenta uma descrição deste bioprocessos.

Capítulo 3

Descrição do Processo Fermentativo para Produção de Etanol

3.1 - Introdução

Este capítulo tem como objetivo apresentar uma descrição resumida do bioprocesso utilizado para o desenvolvimento da modelagem híbrida e dos modelos tipo “caixa preta”.

3.2 - Etapas do Bioprocesso

O processo esquematizado na Fig. 6 foi apresentado por ANDRIETTA (1994) e representa um processo em escala industrial. Consiste de 4 reatores de mistura perfeita ligados em série com reciclo de células.

O primeiro reator recebe o mosto provindo da etapa de preparação da matéria-prima e as células recicladas saídas da unidade de tratamento ácido. Cada reator

possui um sistema de resfriamento pois a conversão da glicose em etanol, realizada pelo microrganismo, é um processo exotérmico. O controle da temperatura é importante pois ela possui grande influência no comportamento da levedura. O vinho fermentado é enviado para a unidade de separação. Este vinho é constituído principalmente de etanol, água, leveduras e açúcares não consumido. A unidade de separação é composta por centrífugas de pratos que operam com velocidade angular de 8000rpm. O vinho fermentado entra na centrífuga com cerca de 30g/l a 45g/l de células (base seca). A separação fornece dois produtos: o vinho delevedurado que é mandado para a unidade de destilação; e o creme de levedura que segue para a unidade de tratamento ácido.

O vinho delevedurado deve ter no máximo 3g/l de células e ao sair da centrífuga é enviado ao tanque pulmão da unidade de destilação (dorna volante). A concentração de etanol no vinho delevedurado é de 9 a 12%² e, dependendo do processo de destilação, fornecerá como produto final álcool hidratado ou anidro (ANDRIETTA, 1994).

O creme de levedura, destinado à unidade de tratamento ácido, contém de 160 a 200g/l de células (base seca). Esta unidade é constituída de 2 a 4 tanques agitados. O creme é primeiramente diluído com água de boa qualidade na proporção 1:1 sendo o pH ajustado com ácido sulfúrico concentrado (2,0 a 2,5). A duração desta etapa é de 1 a 2h. A adição da água tem como objetivos: a quebra do poder tampão do creme diminuindo a quantidade de ácido adicionado; diminuição da concentração de produtos inibitórios ao crescimento do microrganismo; adição de O₂ ao meio. A finalidade do tratamento ácido é fazer uma seleção de células eliminando contaminantes e células “velhas”. Como já dito, todo o processo é contínuo e as células retornam ao primeiro reator.

O mosto é uma mistura de caldo-de-cana tratado e melão (provindo da fabricação do açúcar). O caldo-de-cana recebe um aquecimento seguido de decantação na sua preparação. A finalidade deste tratamento é eliminar sólidos insolúveis (terra, bagacilho, etc.) que podem interferir no processo fermentativo bem como sobrecarregar as centrífugas.

² Atualmente este valor é de 6,5-9,5 % (citação ANDRIETTA, 2000).

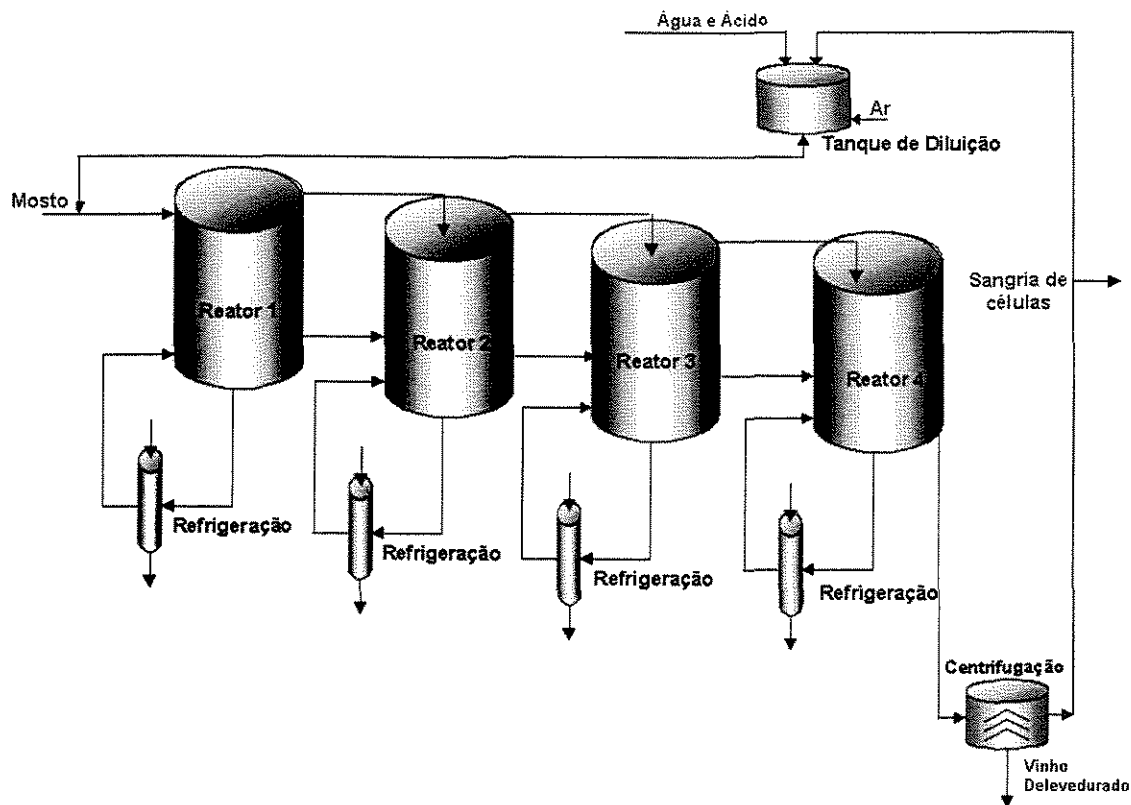


Figura 6: Esquema do processo fermentativo para produção de etanol em escala industrial proposto por ANDRIETTA (1994).

3.3 - Variáveis

A tabela 3 mostra os valores operacionais das principais variáveis do processo. Estes são valores em que o processo deve trabalhar para que se consiga manter sua lucratividade e eficiência. O valor de 180 g/l para concentração de ART no mosto depende da taxa de reciclo que neste caso é de 0,3. A Fig. 7 apresenta as variáveis no processo.

Tabela 3: Valores das variáveis operacionais de processo (ANDRIETTA, 1994)

Variáveis	Valores
Concentração de ART no mosto	180,0 g/l
Concentração de células no creme	180,0 g/l
Concentração de células no reciclo	90,0 g/l
Concentração de células no vinho	3,0 g/l
Temperatura do fluido refrigerante (água)	28,0 °C ³

³ Com a utilização de um sistema de resfriamento em circuito fechado este valor sobe para 30 °C (citação ANDRIETTA, 2000).

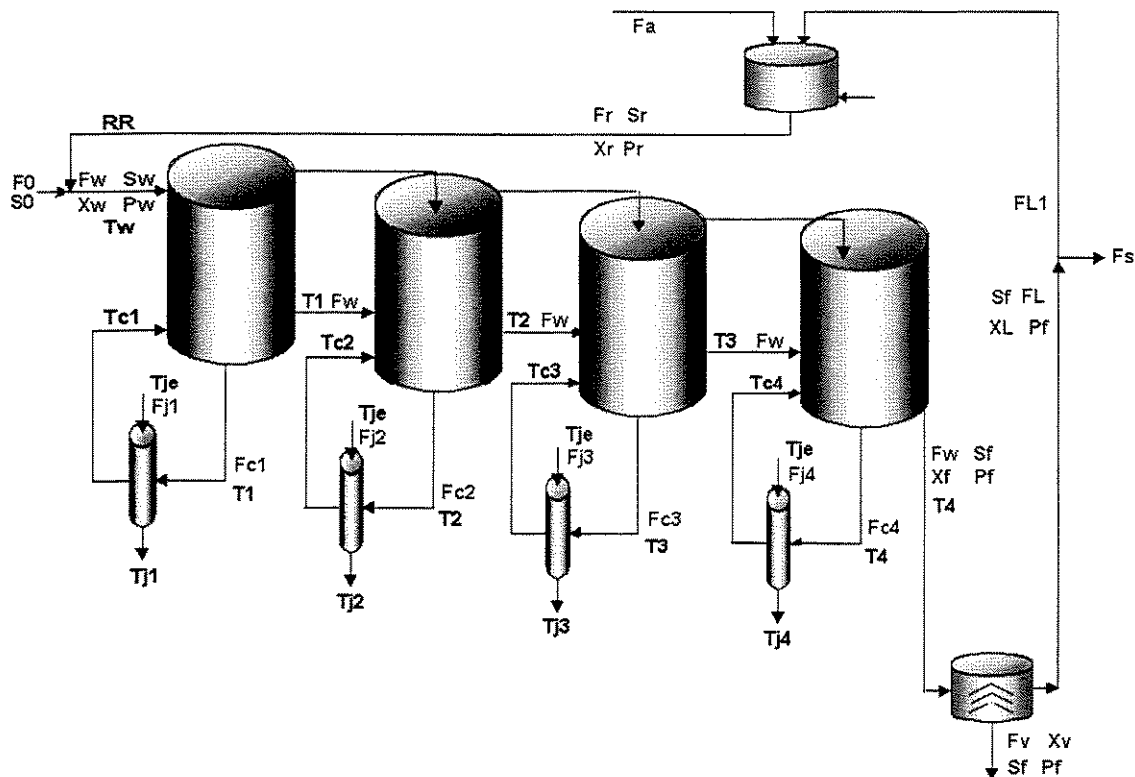


Figura 7: Variáveis no processo.

As tabelas 4 e 5 mostram as principais variáveis independentes e dependentes do processo as quais foram utilizadas na modelagem. A tabela 6 apresenta os valores em que as variáveis independentes devem trabalhar durante o processo.

Tabela 4: Variáveis independentes trabalhadas na simulação:

Variável	Descrição	Observações
F0	Vazão do meio de alimentação na entrada do sistema (m^3h^{-1})	
S0	Concentração de Açúcares Redutores Totais no meio de alimentação ($g l^{-1}$)	Deve estar em torno de 180g/l para de chegar à concentração de etanol desejada.
RR	Taxa de reciclo	Fração entre a vazão do creme de levedura tratado e a vazão total de alimentação do fermentador. Valor mais utilizado é de 0,3 ⁴ .
Tw	Temperatura de alimentação do primeiro reator (°C)	Engloba a temperatura do mosto vinda dos resfriadores e do creme de levedura do reciclo. Seu valor é fixado em 28°C
TJE	Temperatura do fluido de resfriamento na entrada do trocador i (°C)	Depende da temperatura da água do rio utilizado caso a indústria não possua um sistema de resfriamento em circuito fechado. O valor médio da temperatura de rios é em torno de 28°C no estado de São Paulo.

⁴ Atualmente utiliza-se valores entre 0,35 e 0,45 devido a maior sangria de levedura (citação ANDRIETTA, 2000).

Tabela 5: Variáveis dependentes trabalhadas na simulação

Variável	Descrição
S(i)	Concentração de ART no reator i (g l^{-1})
X(i)	Concentração de células no reator i (g l^{-1})
P(i)	Concentração de etanol no reator i (g l^{-1})
T(i)	Temperatura no reator i ($^{\circ}\text{C}$)
TC(i)	Temperatura do fluido reagente na saída do trocador i ($^{\circ}\text{C}$)
FJ(i)	Vazão do fluido de resfriamento no trocador de calor i, (m^3h^{-1})
FW	Vazão de alimentação do primeiro reator (m^3h^{-1})
FL1	Vazão de leite de levedura após a sangria (m^3h^{-1})
Rendimento	Quantidade de etanol produzido pela fermentação alcoólica em relação a quantidade de açúcares que pode ser convertida em álcool pela multiplicação do valor estequiométrico 0,511g/g. (%)
Produtividade	Quantidade de etanol formado em relação a uma unidade de volume do reator por unidade de tempo. ($\text{g. l}^{-1} .\text{h}^{-1}$)

Tabela 6: Valores das variáveis independentes (ANDRIETTA, 1994)

Variáveis independentes manipuladas	Valores
F0	100.0 ($\text{m}^3.\text{h}^{-1}$)
RR	0.30
0	180.0 (g/l)
TW	28.0 ($^{\circ}\text{C}$)
TJE	28.0 ($^{\circ}\text{C}$)

No próximo capítulo será mostrado a modelagem matemática deste bioprocessos. Esta modelagem foi utilizada para o desenvolvimento de um *software* feito em FORTRAN 90 que simula toda a planta mostrada na Fig.7. Os dados para treinamento das Rede Neurais Artificiais foram obtidos deste modelo como será demonstrado nos capítulos seguintes.

Capítulo 4

Modelagem Matemática do Processo Fermentativo para Produção de Etanol

4.1 - Introdução

O modelo matemático do processo permite fazer sua simulação reproduzindo o comportamento dinâmico do processo após sofrer perturbações. As equações do modelo determinístico que representam o processo descrito na Fig. 7 são obtidas por balanços de massa e energia. O modelo cinético empregado na modelagem foi descrito por LEE (1983) sendo seus parâmetros obtidos por ANDRIETTA (1994). O modelo foi considerado não segregado e não estruturado, ou seja, a célula é tratada como um soluto na modelagem. As equações diferenciais geradas foram resolvidas pelo método de Ruge-Kutta de 4ª ordem implementado em linguagem FORTRAN 90. A modelagem matemática de bioprocessos é considerada complexa requerendo modelos cinéticos adequados e parâmetros confiáveis.

4.2 - Modelo e Parâmetros Cinéticos

Um modelo cinético microbiano para aplicação industrial deve ser simples (rapidez), ter poucos parâmetros para ajuste e descrever com precisão o processo. A conversão de ART em etanol por um microrganismo envolve grande número de reações enzimáticas. Uma descrição completa deste modelo cinético levaria à necessidade de muitos parâmetros o que tornaria o modelo complexo e inviável para aplicação industrial. Para contornar este problema são feitas simplificações. Considerar o modelo como não segregado e não estruturado já reduz bastante o número de parâmetros no ajuste.

LEE (1983) propôs um modelo cinético derivado do modelo de “Monod” que se mostrou adequado em escala industrial quando a dinâmica é relativamente lenta (como os processos fermentativos).

$$\mu = \mu_{max} \left(1 - \frac{P}{P_{max}}\right)^n \left(1 - \frac{X}{X_{max}}\right)^m \left(\frac{S}{K_S + S}\right) \quad \text{Equação 1}$$

O cálculo de μ_{max} inclui a temperatura (T) e segue a equação de Arrhenius:

$$\mu_{max} = Ae^{-\frac{E}{RT}} \quad \text{Equação 2}$$

Segundo ANDRIETTA (1994), abaixo da temperatura crítica de 32°C P_{max} é constante, acima desta temperatura seu cálculo é feito pela Eq. 3.

$$P_{max} = K_0 e^{aT} \quad \text{Equação 3}$$

Por definição: $\mu = \frac{1}{X} \frac{dX}{dt}$

Ou seja:

$$\frac{dX}{dt} = \mu X \quad \text{Equação 4}$$

Substituindo a equação 1 em 4 tem-se:

$$\frac{dX}{dt} = \mu_{max} X \left(1 - \frac{P}{P_{max}}\right)^n \left(1 - \frac{X}{X_{max}}\right)^m \left(\frac{S}{K_S + S}\right) \quad \text{Equação 5}$$

A equação acima descreve a variação de células com o tempo. Sabe-se que:

$$\frac{dX}{dt} = Y_{X/S} \left(-\frac{dS}{dt} \right) \quad \text{Equação 6}$$

$Y_{X/S}$ é uma constante calculada experimentalmente. Relacionando as equações 5 e 6 e sabendo que o substrato é consumido:

$$Y_{X/S} = \frac{dX/dt}{-dS/dt}$$

$$-\frac{dS}{dt} = \frac{\mu_{\max} X}{Y_{X/S}} \left(1 - \frac{P}{P_{\max}} \right)^n \left(1 - \frac{X}{X_{\max}} \right)^m \left(\frac{S}{K_S + S} \right) \quad \text{Equação 7}$$

Analogamente, tem-se que: $Y_{P/S} = \frac{dP/dt}{-dS/dt}$

Resultando em:

$$\frac{dP}{dt} = \frac{\mu_{\max} X Y_{P/S}}{Y_{X/S}} \left(1 - \frac{P}{P_{\max}} \right)^n \left(1 - \frac{X}{X_{\max}} \right)^m \left(\frac{S}{K_S + S} \right) \quad \text{Equação 8}$$

Resumindo, para o processo analisado tem-se as seguintes relações:

$$r_X = \frac{dX}{dt} = \mu_{\max} X \left(1 - \frac{P}{P_{\max}} \right)^n \left(1 - \frac{X}{X_{\max}} \right)^m \left(\frac{S}{K_S + S} \right) \quad \text{Equação 9}$$

$$r_S = \frac{dS}{dt} = -\frac{\mu_{\max} X}{Y_{X/S}} \left(1 - \frac{P}{P_{\max}} \right)^n \left(1 - \frac{X}{X_{\max}} \right)^m \left(\frac{S}{K_S + S} \right) \quad \text{Equação 10}$$

$$r_P = \frac{dP}{dt} = \frac{\mu_{\max} X Y_{P/S}}{Y_{X/S}} \left(1 - \frac{P}{P_{\max}} \right)^n \left(1 - \frac{X}{X_{\max}} \right)^m \left(\frac{S}{K_S + S} \right) \quad \text{Equação 11}$$

Os parâmetros cinéticos e as constantes utilizadas nas equações diferenciais são mostrados nas tabelas 7 e 8.

Tabela 7: Valores dos Parâmetros Cinéticos (ANDRIETTA, 1994)

Parâmetros	Valores
μ_{\max}	0,41 litro/hora
P_{\max}	103 grama/litro ($T \leq 32^{\circ}\text{C}$)
X_{\max}	100 grama/litro
n	3,0
m	1,0
K_s	1,6

Tabela 8: Valores das constantes utilizadas

Constantes	Valores
E	$1,54 \cdot 10^4 \text{ cal.mol}^{-1}$
A	$4,50 \cdot 10^{10}$
K_0	895,6 g/l
a	$-0,0676 \text{ }^{\circ}\text{C}^{-1}$

Segundo DECHECHI (1998), a aplicação de um atraso linear na velocidade específica de crescimento do microrganismos melhora a aplicação de um modelo considerado não estruturado e não segregado para o estudo dinâmico de um processo. Assim, foi aplicada a equação 12 para a correção de μ . A correção é feita a cada passo de integração aplicado para a resolução das equações diferenciais.

$$\mu_{\text{corrigido}} = \mu \left(\frac{S_{\text{anterior}}}{S_{\text{atual}}} \right)$$

Equação 12

4.3 - Balanço de massa para o reator *i*

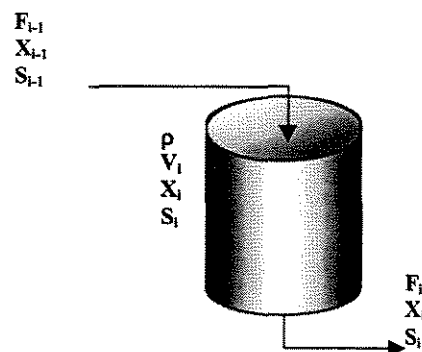


Figura 8: Reator de mistura perfeita *i*.

4.3.1 - Balanço global

$$\frac{d(V_i \rho_i)}{dt} = F_{i-1} \rho_{i-1} - F_i \rho_i$$

Considerando ρ_i constante em cada reator:

$$\frac{dV_i}{dt} = \frac{F_{i-1} \rho_{i-1}}{\rho_i} - F_i \quad \text{Equação 13}$$

4.3.2 - Balanço de Substrato

O balanço para o substrato é representado pela seguinte expressão:

$$\frac{d(V_i S_i)}{dt} = F_{i-1} S_{i-1} - F_i S_i + r_S V_i$$

Substituindo a Eq. 10 na equação acima:

$$\frac{d(V_i S_i)}{dt} = F_{i-1} S_{i-1} - F_i S_i - \frac{V_i \mu_{\max} X_i}{Y_{X/S}} \left(1 - \frac{P_i}{P_{\max}}\right)^n \left(1 - \frac{X_i}{X_{\max}}\right)^m \left(\frac{S_i}{K_S + S_i}\right) \quad \text{Equação 14}$$

4.3.3 - Balanço de Produto

Analogamente o balanço para o produto é representado pela seguinte expressão:

$$\frac{d(V_i P_i)}{dt} = F_{i-1} P_{i-1} - F_i P_i + r_P V_i$$

$$\frac{d(V_i P_i)}{dt} = F_{i-1} P_{i-1} - F_i P_i - \frac{V_i \mu_{\max} X_i Y_{P/S}}{Y_{X/S}} \left(1 - \frac{P_i}{P_{\max}}\right)^n \left(1 - \frac{X_i}{X_{\max}}\right)^m \left(\frac{S_i}{K_S + S_i}\right) \quad \text{Equação 15}$$

4.3.4 - Balanço de Células

O balanço de células proposto por ANDRIETTA (1994) engloba células viáveis e mortas não incluindo a taxa de morte celular. Ou seja, é considerado que a concentração de células incapazes de crescer e reproduzir desprezível.

$$\frac{d(V_i X_i)}{dt} = F_{i-1} X_{i-1} - F_i X_i + r_X V_i$$

$$\frac{d(V_i X_i)}{dt} = F_{i-1} X_{i-1} - F_i X_i - V_i \mu_{\max} X \left(1 - \frac{P_i}{P_{\max}}\right)^n \left(1 - \frac{X_i}{X_{\max}}\right)^m \left(\frac{S_i}{K_S + S_i}\right) \quad \text{Equação 16}$$

4.4 - Balanço de energia

4.4.1 - Fluido Reagente no reator *i*

$$\frac{\rho_i C_{p_i} d(V_i T_i)}{dt} = F_{i-1} T_{i-1} \rho_i C_{p_i} - F_i T_i \rho_i C_{p_i} + F c_i (T_{c_i} - T_i) \rho_i C_{p_i} + V_i \Delta H r_s \quad \text{Equação 17}$$

O termo $V_i \Delta H r_s$ provem da geração de calor do crescimento do microrganismo para o reator da Fig. 8. Segundo SHULER e KARGI (1992), entre 40% e 50% da energia obtida pelo microrganismo é transformada em energia biológica (ATP) durante o metabolismo aeróbico. O restante da energia é perdida para o meio na forma de calor. Para o reator estudado temos que:

$$\frac{d(V_i T_i)}{dt} = F_{i-1} T_{i-1} - F_i T_i + F c_i (T_{c_i} - T_i) + \frac{V_i \Delta H X_i}{\rho_i C_{p_i} Y_{x/s}} \mu_i \quad \text{Equação 18}$$

4.4.2 - Fluido Reagente no trocador de calor *i*

$$\frac{d(T_{c_i})}{dt} = \frac{F c_i}{V c_i} (T_i - T_{c_i}) - \left(\frac{U A_i}{V c_i \rho_i C_{p_i}} \right) LMDT_i \quad \text{Equação 19}$$

Onde:

$$LMDT_i = \frac{(T_i - T_{j_i}) - (T_{c_i} - T_{j_e})}{\ln \left(\frac{(T_i - T_{j_i})}{(T_{c_i} - T_{j_e})} \right)}$$

4.4.3 - Fluido de Resfriamento no trocador de calor *i*

$$\frac{d(T_{j_i})}{dt} = \frac{F j_i}{V j_i} (T_{j_e} - T_{j_i}) + \left(\frac{U A_i}{V j_i \rho_j C_{p_j}} \right) LMDT_i \quad \text{Equação 20}$$

4.5 - Balanços de massa aplicados ao sistema

A Fig. 7 mostra o esquema do processo fermentativo com as principais variáveis. Aplicando balanço de massa em partes do sistema obtém-se as equações abaixo:

$$F_w = \frac{F_0}{(1 - RR)} \quad \text{Equação 21}$$

$$F_R = F_w - F_0 \quad \text{Equação 22}$$

$$F_{l1} = \frac{F_R X_R}{X_l} \quad \text{Equação 23}$$

$$F_a = F_R - F_{l1} \quad \text{Equação 24}$$

$$F_V = F_w \frac{X_l - X_f}{X_l - X_v} \quad \text{Equação 25}$$

$$F_l = F_w - F_V \quad \text{Equação 26}$$

$$F_S = F_l - F_{l1} \quad \text{Equação 27}$$

$$S_R = \frac{F_{l1} S_f}{F_R} \quad \text{Equação 28}$$

$$S_w = \frac{F_R S_R + F S_0}{F_w} \quad \text{Equação 29}$$

$$P_R = \frac{F_{l1} P_f}{F_R} \quad \text{Equação 30}$$

$$P_w = \frac{F_R P_R}{F_w} \quad \text{Equação 31}$$

Assim o processo descrito na Figura 6 é constituído de 7 equações diferenciais para cada reator e mais 11 equações obtidas do balanço de massa no sistema. As equações são resolvidas pelo método de Runge-Kutta de 4ª ordem.

4.6 - Rendimento e Produtividade

As variáveis apresentadas neste capítulo até o momento são observadas diretamente no processo. No entanto, existem variáveis indiretas que são importantes para análise do processo, destacando rendimento do processo e produtividade de etanol.

$$REND = \frac{F_v \cdot P(4)}{F0.S0} \cdot \frac{100}{0,511} \quad \text{Equação 32}$$

$$PROD = \frac{F_v \cdot P(4)}{V} \quad \text{Equação 33}$$

Os valores de REND e PROD, nos perfis dinâmicos, são valores médios referentes a uma hora de processo. DECHECHI (1998) utilizou o artifício de colocar os valores da última hora passada para que os valores fossem o mais representativo possível dinamicamente.

4.7 - Conclusão

Este capítulo apresentou a modelagem matemática do bioprocessamento estudado. O modelo foi validado industrialmente mostrando-se adequado apesar das simplificações utilizadas.

Com o desenvolvimento do modelo matemático (determinístico) pode-se dar início a um estudo das variáveis do processo como será mostrado no capítulo seguinte. O estudo das variáveis auxiliou a modelagem do processo por RNAs (modelo tipo “caixa preta”) e na modelagem híbrida. Foi estudado também a sensibilidade e o grau de linearidade das variáveis envolvidas.

Capítulo 5

Estudo das Variáveis do Bioprocesso

5.1 – Introdução

Este capítulo apresenta, primeiramente, a dinâmica do processo quando este recebe perturbação degrau nas variáveis de entrada. Pretende-se analisar as respostas (variáveis independentes) a fim de identificar quais alcançam o estado estacionário mais rapidamente e em qual valor, e se há ou não ocorrência de resposta inversa. Em seguida é feita uma análise de sensibilidade de algumas variáveis selecionadas e aplicado um teste de verificação do seu grau de linearidade.

5.2- Dinâmica do Bioprocesso

Cada variável de entrada sofreu perturbações de $\pm 10\%$ no seu valor inicial mostrado na tabela 6, exceto F_0 que variou $\pm 20\%$. Alterações no valor das temperaturas nos reatores não foram realizadas, pois, o modelo determinístico do processo possui um controle proporcional para malhas de temperatura através de alterações nas vazões de fluido refrigerante em cada estágio, lembrando que as temperaturas devem estar em torno de

33,5°C. As Figuras 9 a 12 mostram alguns perfis que possibilitam análise do comportamento das variáveis.

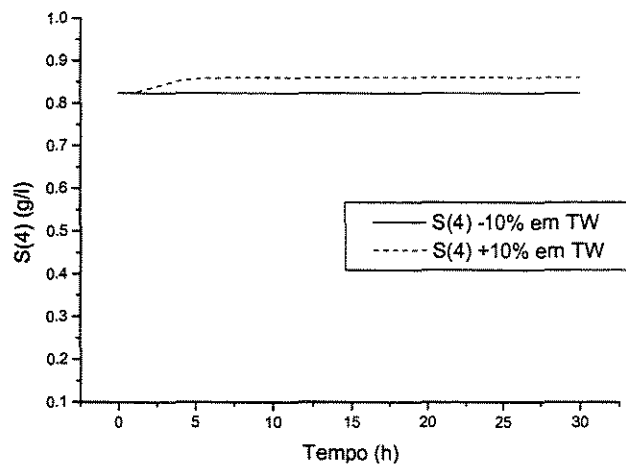


Figura 9: Perfil de S(4) para perturbação degrau em Tw.

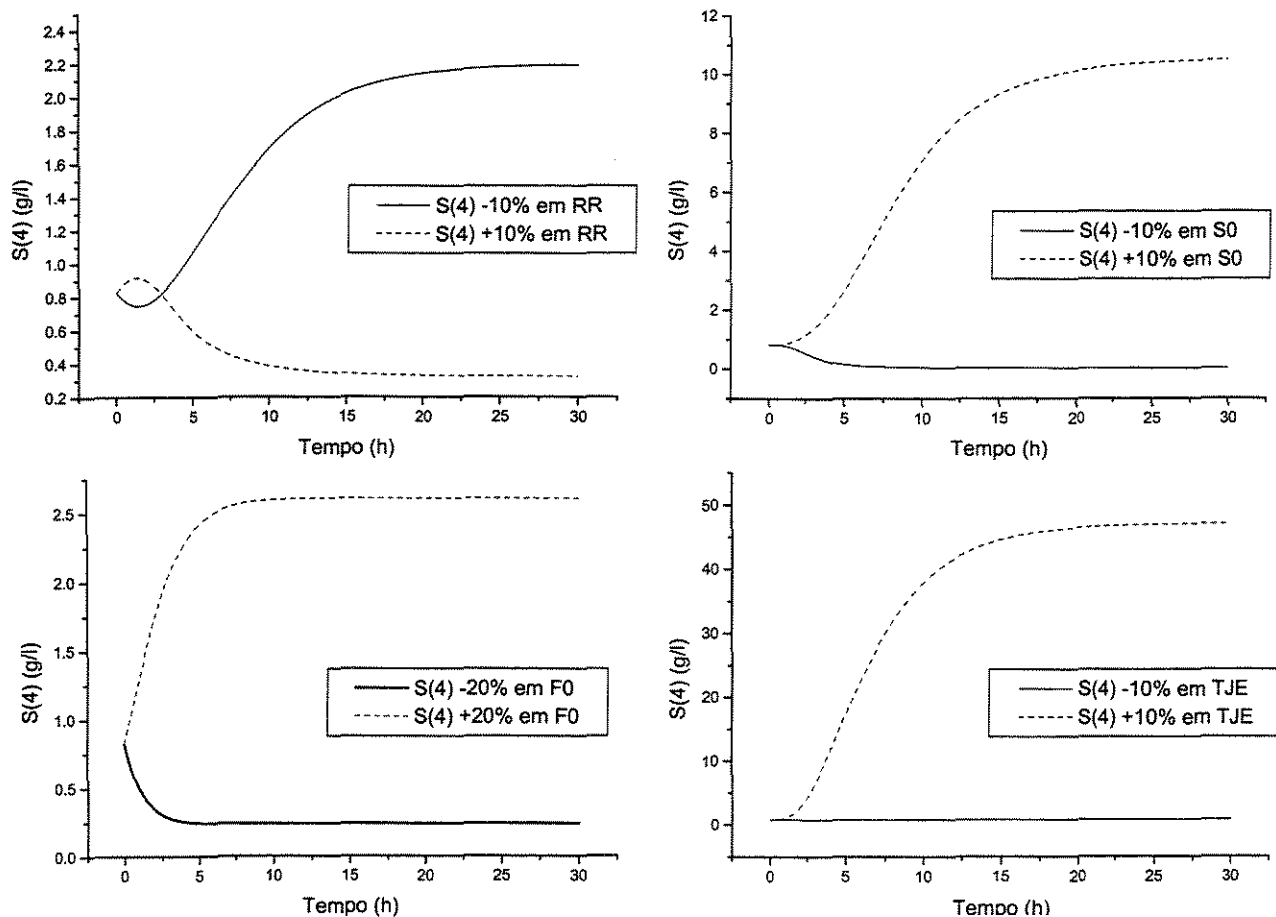


Figura 10: Perfil de S(4) para perturbação degrau em RR, S0, F0 e TJE

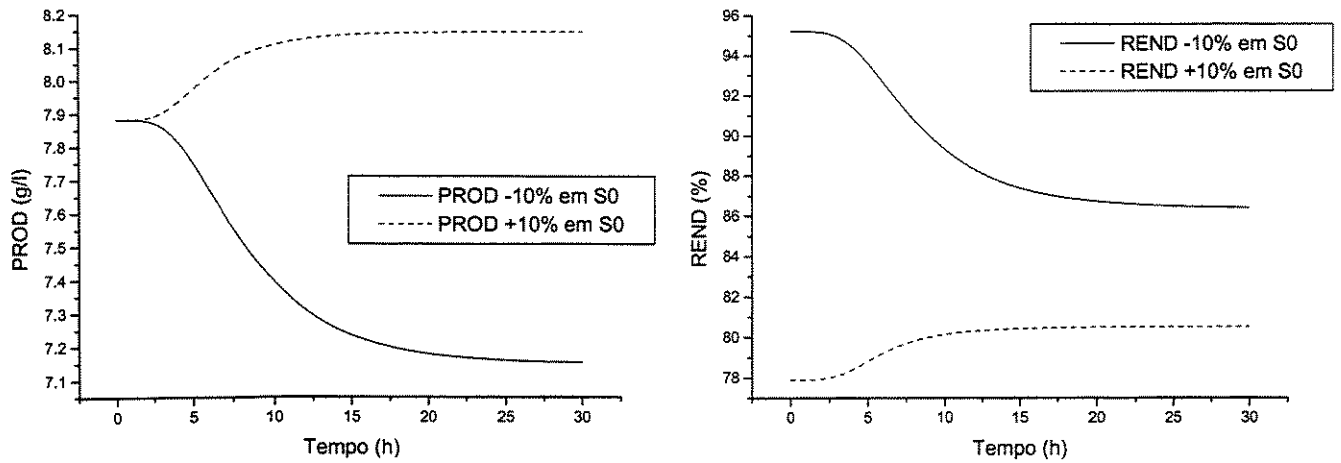


Figura 11: Perfil de Produtividade e Rendimento para perturbação degrau em S0.

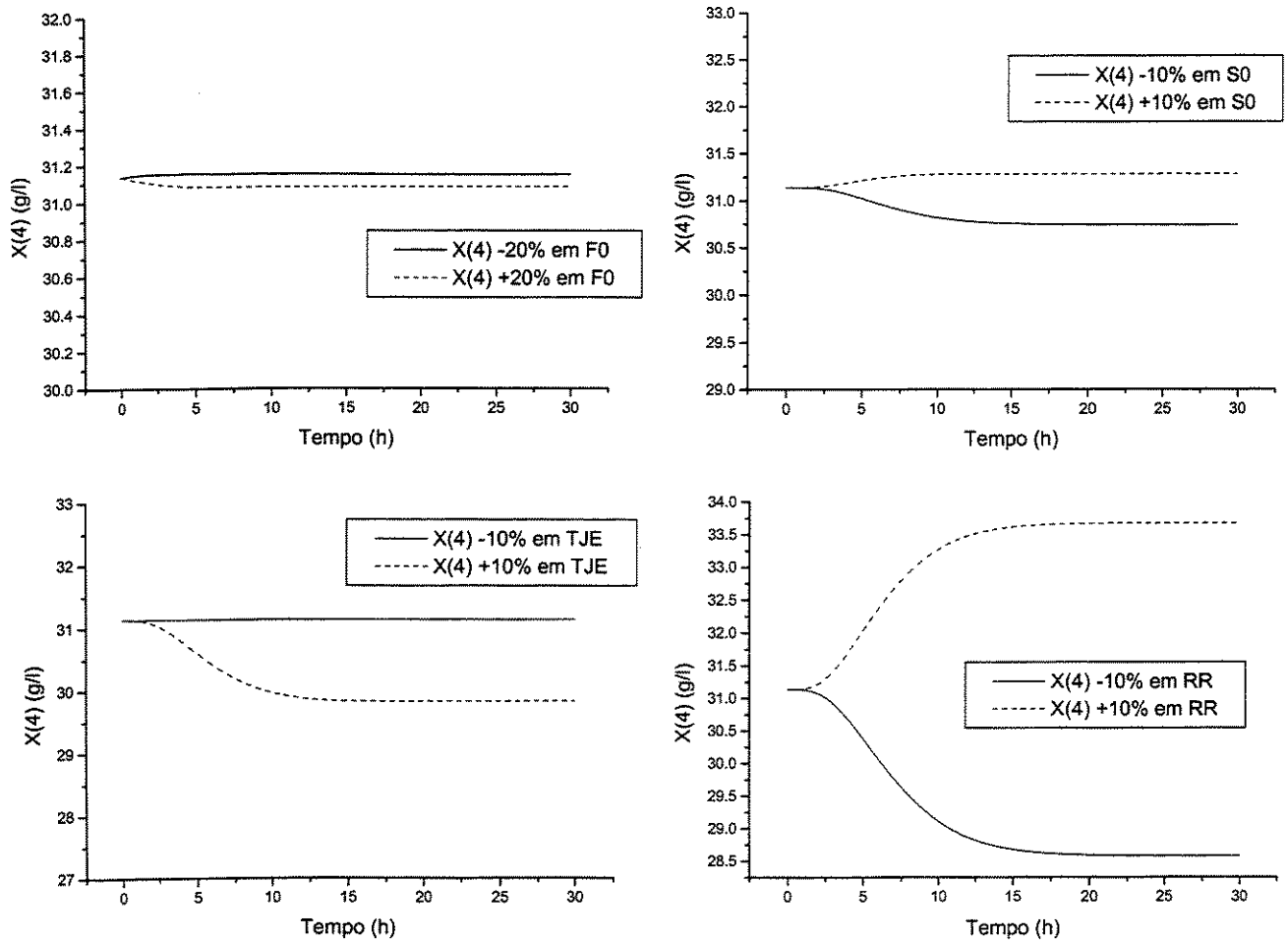


Figura 12: Perfil de X(4) para perturbação degrau em F0, S0, TJE, e RR.

Como pode-se observar nas Figuras 9 a 12, o processo fermentativo apresentado possui o comportamento dinâmico muitas vezes não-linear e ocorrendo até casos de resposta inversa (Fig. 10). No perfil de Rendimento para perturbação degrau em S_0 (Fig. 11) nota-se que o valor do rendimento no início da simulação é dependente de S_0 como visto na seção 4.6. Podem também ocorrer mudanças nos parâmetros cinéticos durante o processo visto que se trata de um processo biológico. Estes fatores fazem com que o controle seja difícil tornando-se de extrema importância a utilização de estratégias de controle. Também é importante notar tais características para a análise do poder de predição de modelos mais simplificados como os baseados em Redes Neurais Artificiais.

5.3 - Análise de Sensibilidade

A análise da sensibilidade das variáveis permite, de uma maneira rápida, a verificação da influência de cada variável do processo sobre as variáveis de saída. Neste trabalho a sensibilidade de uma variável de saída referente a uma dada entrada é definida como a porcentagem da variação do valor estacionário da saída em relação ao valor inicial da mesma:

$$S_{VS} = 100 * \left(\frac{VS_{Depois} - VS_{Antes}}{VS_{Antes}} \right)$$

Equação 34

Onde:

VS = variável de saída analisada

S_{VS} = sensibilidade de VS .

VS_{Depois} = valor estacionário de VS após a aplicação de perturbação degrau de 10% na variável de entrada.

VS_{Antes} = valor estacionário de VS antes da aplicação de perturbação degrau de 10% na variável de entrada.

A escolha das variáveis de entrada e de saída foi baseada na análise dos trabalhos de ANDRIETTA (1994) e DECHECHI (1997). A tabela 9 mostra os resultados obtidos com as devidas perturbações degrau.

Tabela 9: Sensibilidade das variáveis do processo sobre as variáveis de saída

Entrada/saída		S(4)	REND	PROD
F0	+10%	76.434	-0.416	9.543
	-10%	-44.475	0.241	-9.782
RR	+10%	-60.870	0.407	0.408
	-10%	165.760	-0.945	-0.944
S0	+10%	1172.820	-6.038	3.359
	-10%	-95.860	0.846	-9.238
Tw	+10%	4.379	-0.024	-0.023
	-10%	0.001	0.000	0.001
TJE	+10%	5605.450	-30.745	-30.744
	-10%	0.001	0.000	0.001

Antes de se analisar os resultados deve-se saber que o objetivo é conseguir os maiores valores possíveis para rendimento e produtividade e menores valores para S(4). Analisando a tabela 9 pode-se chegar a algumas conclusões. De modo geral, a variável TJE contribui de forma mais significativa para variações nas variáveis de saída, quando acrescida no valor. Isto mostra a importância do controle de temperatura no processo. Vale lembrar que o modelo matemático do processo implementado em FORTRAN já possui um controlador PI para malha de temperatura nos reatores, simulando o que de fato ocorre na operação industrial. O segundo fator em importância é a concentração de substrato no início do processo, S0. A variável S(4) é muito mais sensível às variações nas variáveis de entrada se comparada ao rendimento e à produtividade.

Alguns fatores práticos devem ser analisados. Por exemplo, sabe-se que as variáveis Tw e TJE são dependentes da variação da temperatura do rio do qual se coleta a água (ANDRIETTA,1994). Indústrias maiores possuem sistemas de resfriamento. A temperatura do rio, em média, não sofre mudanças bruscas. A variável S0 é dependente da matéria prima processada (tipo de cana, época de colheita, tipo de armazenamento, etc), sua variação é amenizada pois há uma correção da concentração de substrato inicial utilizando-se melão .

O objetivo principal do sistema de fermentação alcoólica é obter no final do processo altas concentrações de etanol e, conseqüentemente, baixas concentrações de substrato (ART). A baixa concentração de ART deve ser mantida constante pois ela está ligada ao valor do rendimento, influenciando diretamente no custo total do processo. A concentração celular e de ART são medidas em laboratório. Analisando-se todos os fatores

foram escolhidas as variáveis S(4), X(4), REND e PROD para serem as variáveis de saída e F0 e S0 para variáveis de entrada no treinamento da Rede Neural Artificial, sem modelagem híbrida. O desenvolvimento destes modelos tipo “caixa preta” será descrito no próximo capítulo.

5.4 - Análise da Linearidade entre as Variáveis de Entrada e Saída

Esta análise teve como objetivo a verificação do grau de linearidade das variáveis que foram estudadas no desenvolver do trabalho.

Para se consolidar a linearidade de uma variável de saída em relação à duas variáveis de entrada devem ser validados os seguintes princípios:

- **Princípio da superposição:** Seja VS a variável de saída, quaisquer que sejam as variáveis de entrada, VE_1 e VE_2

$$VS(VE_1 + VE_2) = VS(VE_1) + VS(VE_2)$$

Ou seja, o valor de VS após perturbação nas duas variáveis de entrada deve ser igual a soma das VS após perturbação individual em cada VE .

- **Princípio do Fator de escala:** Seja VS a variável de saída, qualquer que seja $K \in \mathbb{R}$ e a variável de entrada VE_1

$$VS(K \cdot VE_1) = K \cdot VS(VE_1)$$

As variáveis de saída analisadas são: S(4), X(4), REND e PROD; e as de entrada F0 e S0.

5.4.1 - Princípio da Superposição

Como um exemplo para maior entendimento será demonstrado a análise do princípio da superposição para a variável S(4) em relação as variáveis F0 e S0. Sendo $VS=S(4)$, $VE_1=F0$ e $VE_2=S0$.

- **Cálculo de S(4)*:** Aplicação de perturbação degrau de 5% em F0. Subtração do valor de S(4) no estado estacionário em cada tempo.

$$S(4)^* = S(4)_t - S(4)_{\text{Estacionário}}$$

Onde:

$S(4)_t$ = valor de $S(4)$ no tempo t .

$S(4)_{\text{Estacionário}}$ = valor de $S(4)$ no estado estacionário.

- Aplicação de perturbação degrau de 5% em S_0 . Subtração do valor de $S(4)$ no estado estacionário em cada tempo.
- **Cálculo de $VS(VE_1) + VS(VE_2)$** : Soma dos valores de $S(4)^*$ obtidos nas perturbações anteriores.
- **Cálculo de $VS(VE_1 + VE_2)$** : Aplicação de perturbação degrau de 5% em S_0 e F_0 , simultaneamente, e subtração do valor de $S(4)$ no estado estacionário em cada tempo.

Alguns dos resultados obtidos podem ser visualizados nas Figuras 12 e 13

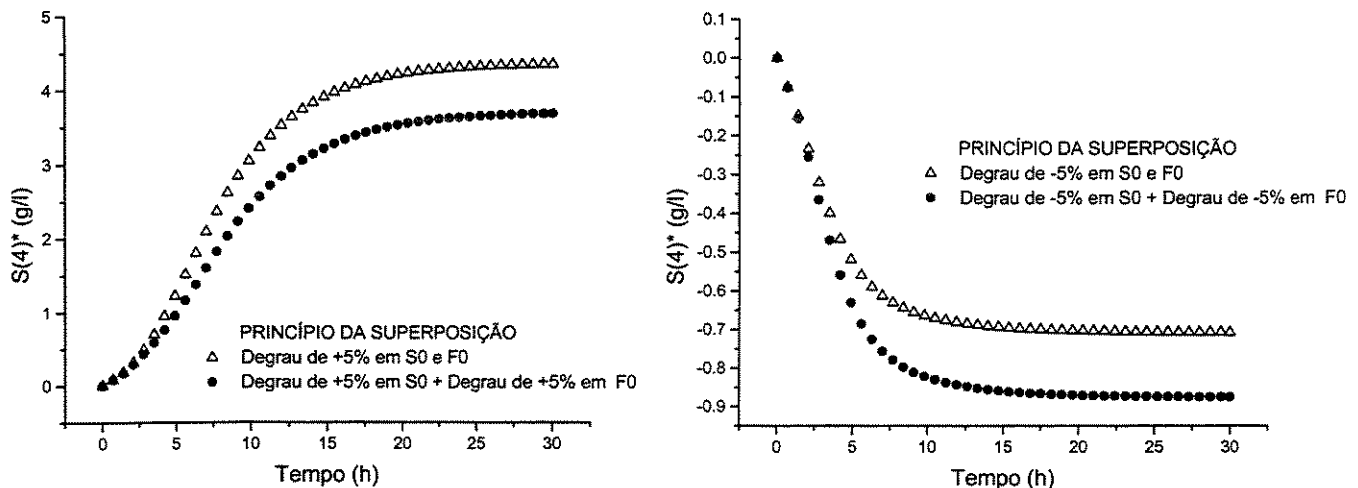


Figura 13: Teste do princípio da superposição de $S(4)$ para as variáveis de entrada S_0 e F_0 .

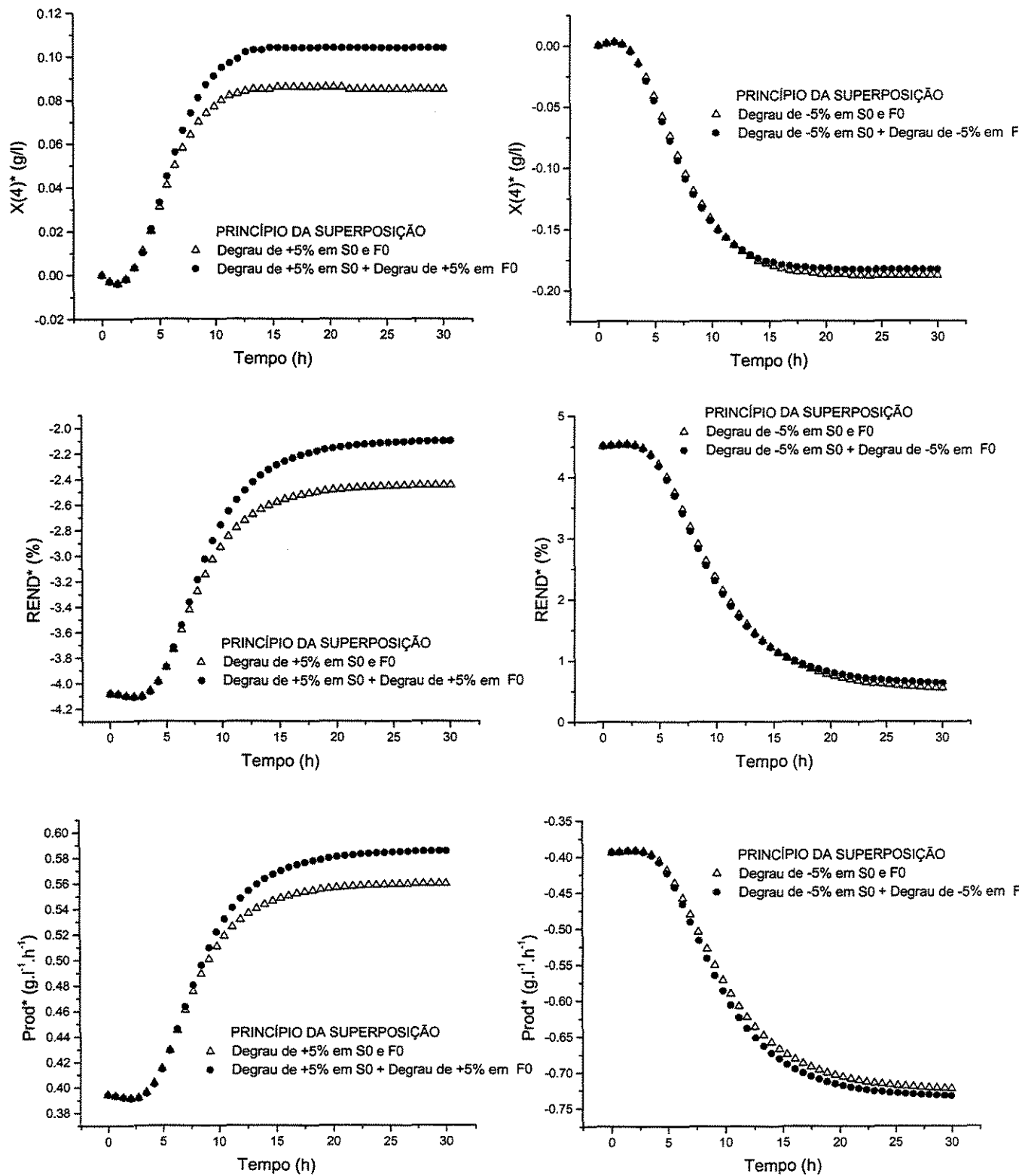


Figura 14: Teste do princípio da superposição de $S(4)$, REND e PROD para as variáveis de entrada S0 e RR

5.4.2 - Fator de Escala

A metodologia para o teste deste princípio é parecida com a utilizada no princípio da superposição.

- **Cálculo de $S(4)^*$:** Aplicação de perturbação degrau de 5% em F_0 . Subtração do valor de $S(4)$ no estado estacionário em cada tempo.
- **Cálculo de $K.VS(VE)$:** Multiplicação dos valores de $S(4)^*$ obtidos na perturbação anterior por um fator igual a 2.
- **Cálculo de $VS(K.VE)$:** Aplicação de perturbação degrau de 10% em F_0 . Subtração do valor de $S(4)$ no estado estacionário em cada tempo.

As Figuras 14 a 17 possuem os resultados obtidos.

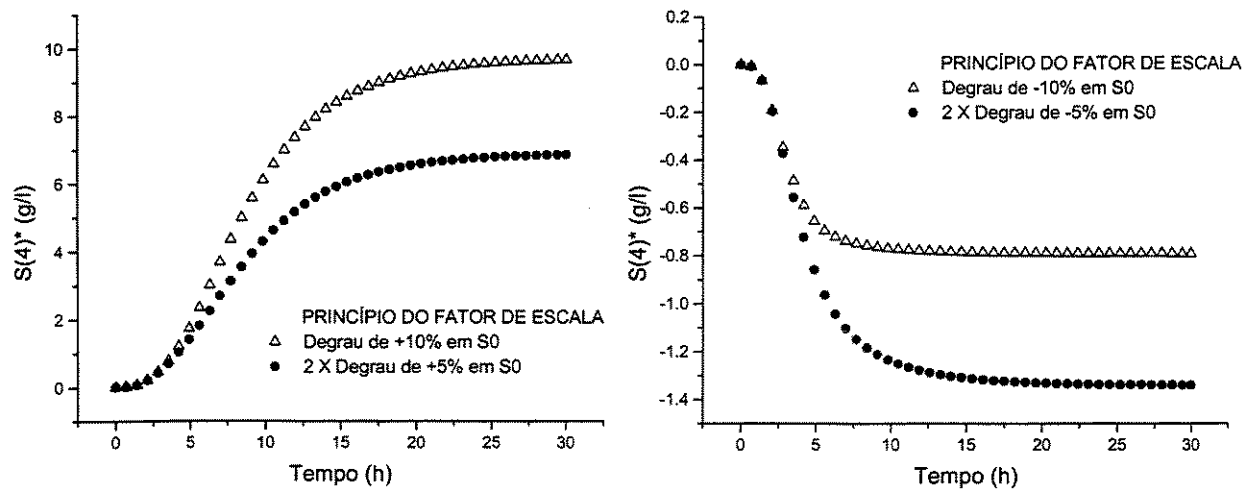


Figura 15: Teste do princípio do fator de escala de $S(4)$ para a variável de entrada S_0

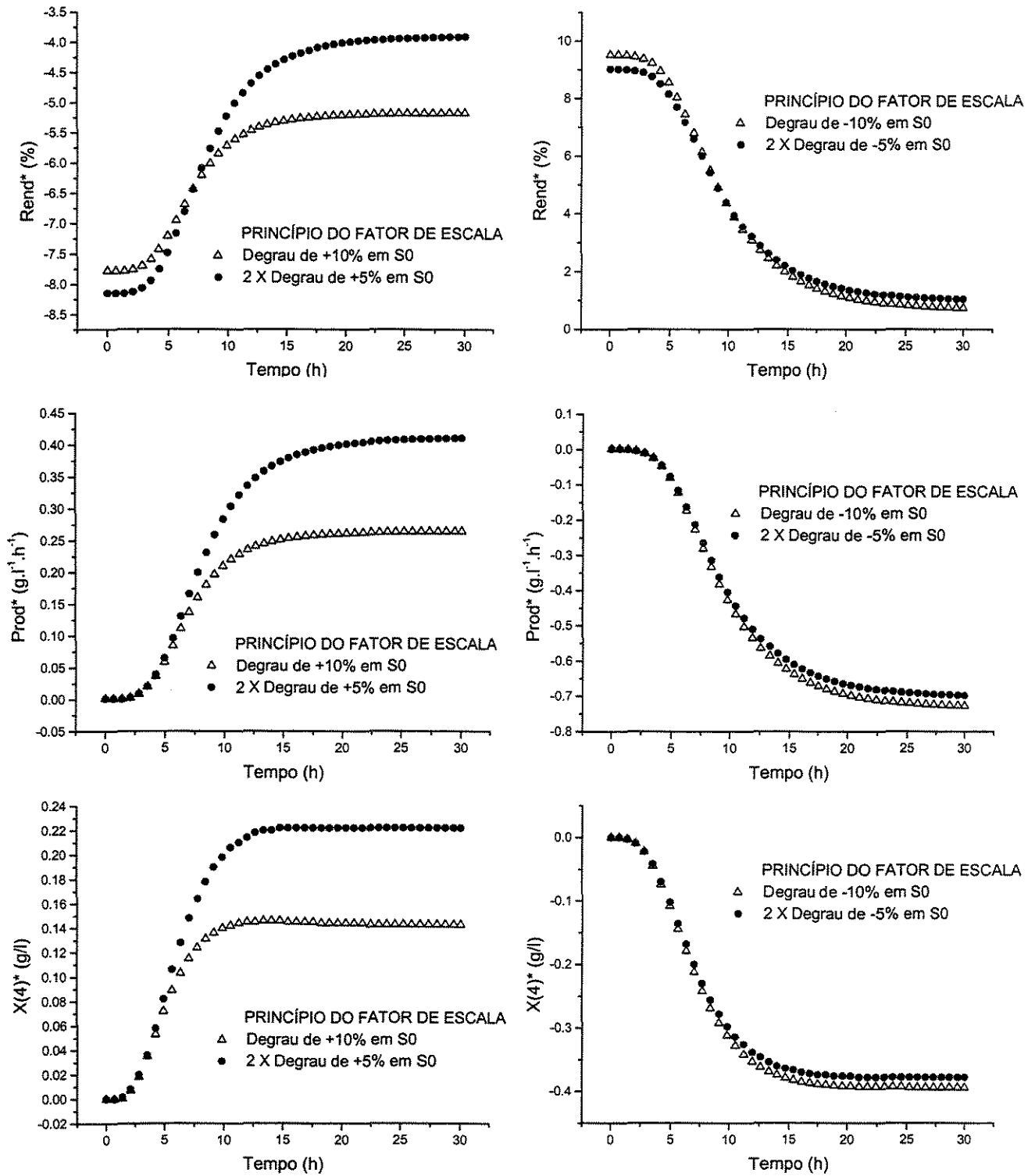


Figura 16: Teste do princípio do fator de escala de REND, PROD e X(4), para a variável de entrada S0.

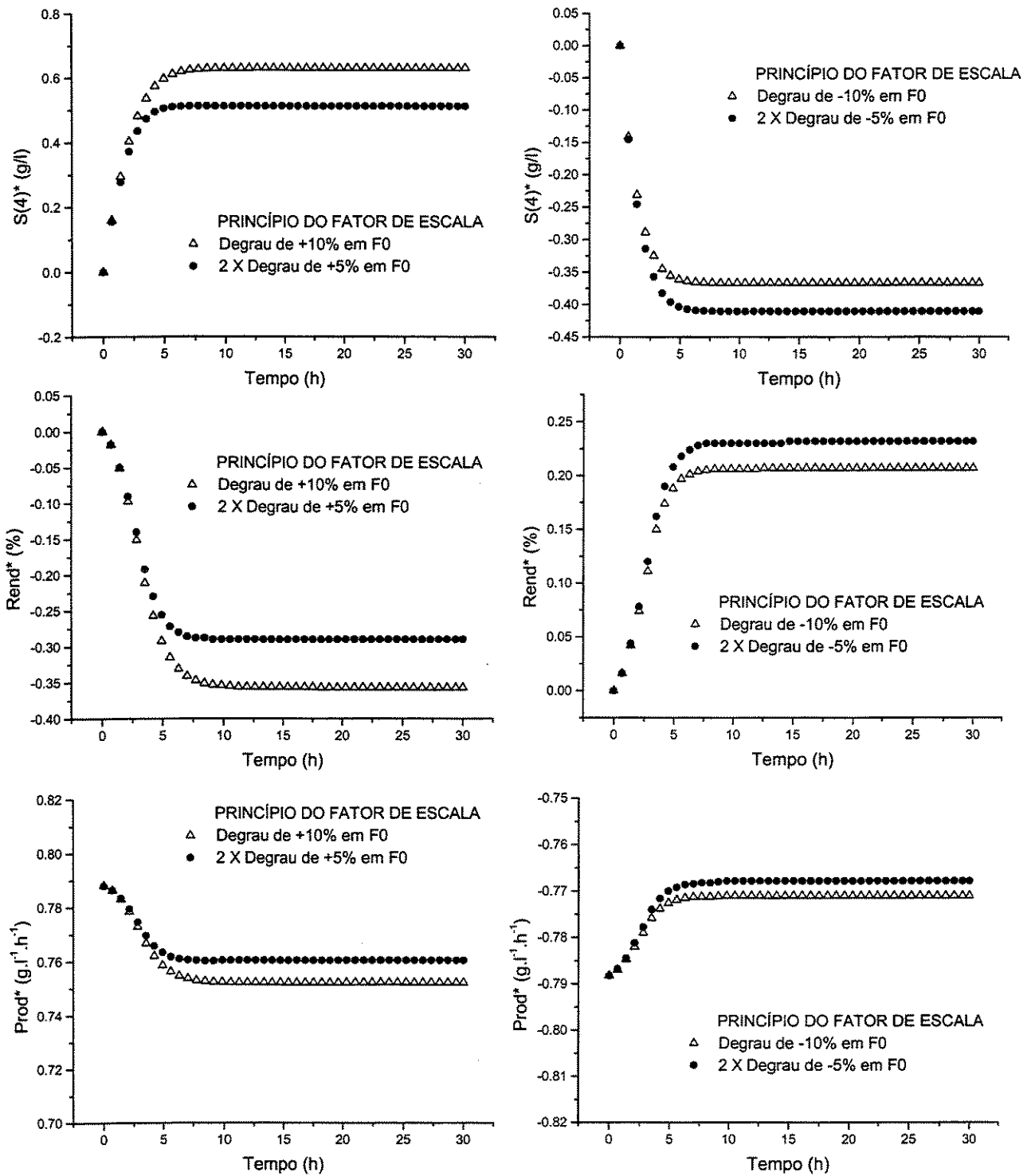


Figura 17: Teste do princípio do fator de escala de $S(4)$, REND e PROD, para a variável de entrada F_0 .

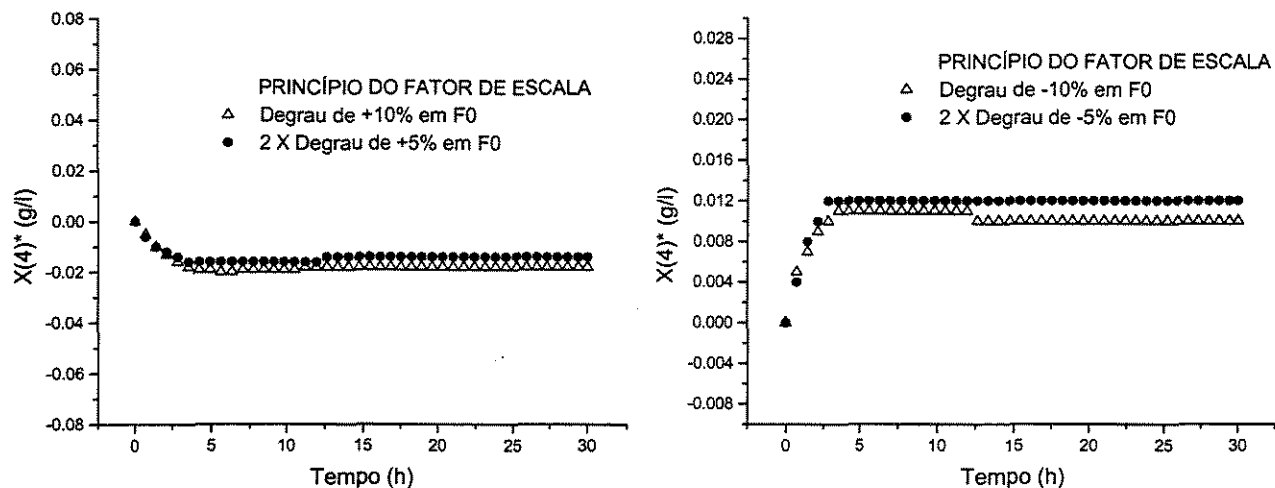


Figura 18: Teste do princípio do fator de escala de $X(4)$ para a variável de entrada F_0

5.5 - Conclusão

A dinâmica do processo apresenta não-linearidade e resposta inversa (Fig.10) comportamento característico de um bioprocesso. Após análise da sensibilidade do processo e análise das restrições praticas para a manipulação das variáveis foram escolhidas as variáveis $S(4)$, $X(4)$, $REND$ e $PROD$ para serem as variáveis de saída e F_0 e S_0 para variáveis de entrada no treinamento da Rede Neural Artificial, sem modelagem híbrida.

A análise da linearidade entre as variáveis de entrada e saída demonstrou que $S(4)$ é a variável que possui menor linearidade. No teste do princípio do fator de escala os valores de $S(4)^*$ para perturbação degrau de +10% em S_0 e duas vezes degrau de 5% em S_0 (Fig.15) diferem em 1,5 após 30h de simulação. As demais variáveis $X(4)$, $REND$ e $PROD$ possuem uma leve não-linearidade para o intervalo de perturbação estudado. Vale lembrar que os gráficos contidos nas Figuras 13 a 18 estão em escalas diferentes. Um aumento positivo tanto em F_0 como S_0 contribuem para o aumento da não linearidade das variáveis estudadas.

O treinamento de Redes Neurais Artificiais para construção de modelos tipo “caixa preta” é apresentado no capítulo seguinte.

Capítulo 6

Modelagem do Processo Fermentativo por Rede Neural Artificial

6.1 - Introdução

A finalidade deste capítulo é o desenvolvimento de dois modelos tipo “caixa preta” utilizando Redes Neurais Artificiais.

O primeiro modelo fornece os valores de $S(4)$, $X(4)$, $REND$ e $PROD$ depois de 30h de processo após perturbação nas variáveis de entrada $F0$ e $S0$. O segundo modelo representará o perfil dinâmico das variáveis $S(4)$, $X(4)$ e $P(4)$ após perturbação em $F0$. Para a obtenção deste perfil dinâmico será utilizado o conceito de treinamento para predição um passo adiante.

6.2 - Características das Redes Neurais Artificiais Utilizadas

Todas as Redes Neurais Artificiais apresentadas neste trabalho possuem as características citadas a seguir. A estrutura da rede escolhida é a “*Feedforward*”, totalmente conectada, constituída de camada de entrada, intermediária(s) e de saída. A camada intermediária possibilita a criação de representações próprias, internas da rede, do problema em questão. O paradigma de aprendizado utilizado é o supervisionado e são apresentados à rede pares de entrada e saída. Os pesos são ajustados durante a aprendizagem em função do estímulo recebido e do sinal de erro (diferença entre a resposta fornecida pela rede e a resposta do modelo matemático). O algoritmo utilizado na aprendizagem é o de “*Backpropagation*” .

Para treinamento da Rede Neural Artificial utilizou-se um programa implementado em linguagem FORTRAN. O programa possui um número variável de camadas e de neurônios. Os pesos são atualizados a cada par entrada/saída apresentado à rede. As camadas intermediárias e de saída possuem neurônios “Bias” que são atualizados pelo programa. É calculado também o erro quadrático global ao conjunto de dados e o programa termina o cálculo se este erro atingir um valor menor que o valor fornecido pelo usuário. Todos os pesos são inicializados automaticamente pelo programa. A taxa de aprendizado é fixa e estabelecida pelo usuário. A rede possui também o termo de momento para evitar oscilações.

Os treinamentos foram realizados em um PC com processador Pentium II Intel MMX, 128Mb de RAM. A Fig. 19 mostra um fluxograma do programa de treinamento da rede. A finalidade do treinamento é ajustar um conjunto de pesos.

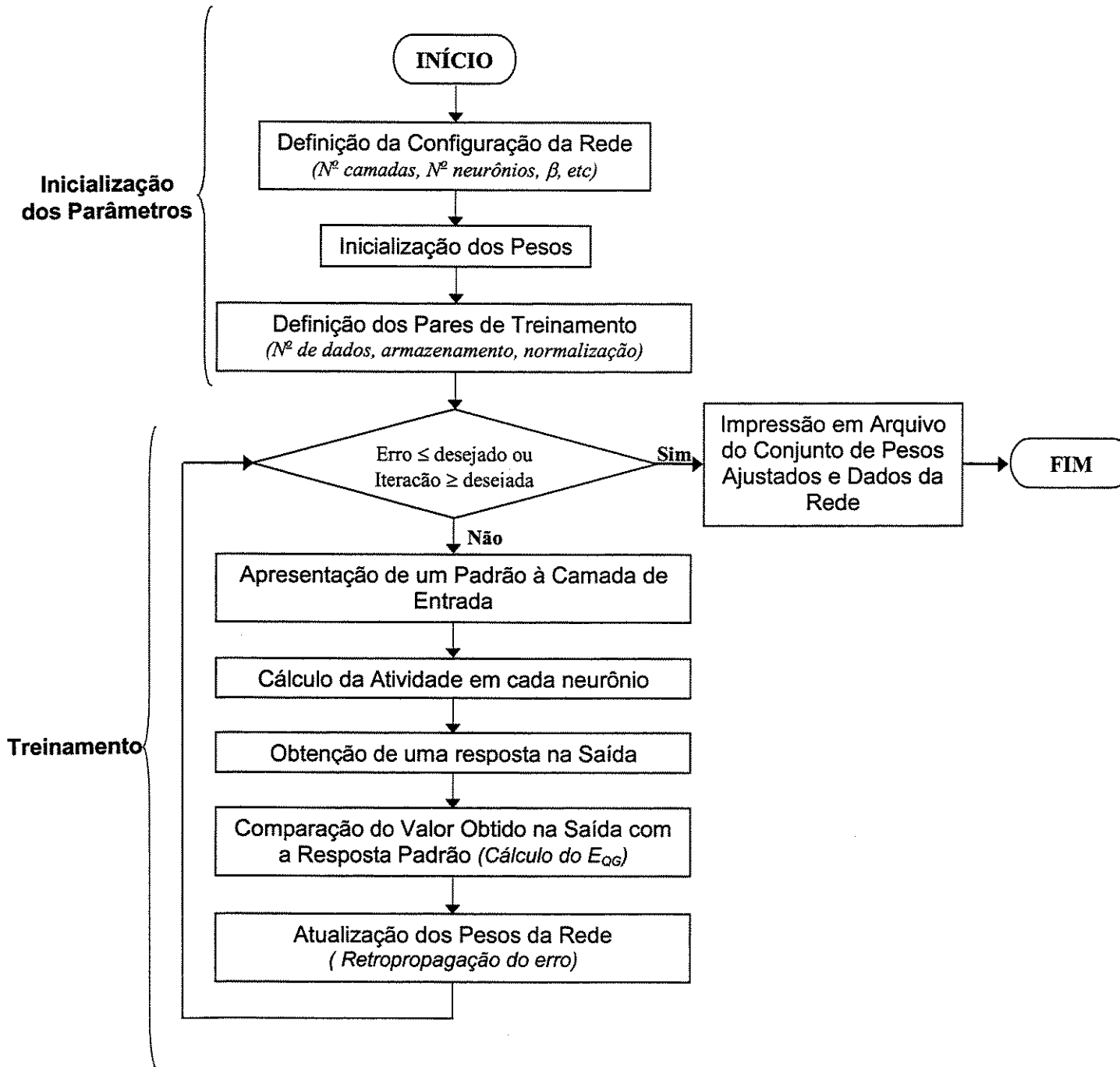


Figura 19: Fluxograma do programa de treinamento das Redes Neurais Artificiais.

6.3 - Treinamento das Redes Neurais Artificiais

6.3.1 - Caso A

A Rede Neural Artificial representada na Fig. 20 possui duas variáveis de entrada: F0 e S0. As variáveis de saída são S(4), X(4), rendimento e produtividade, após 30h de processo.

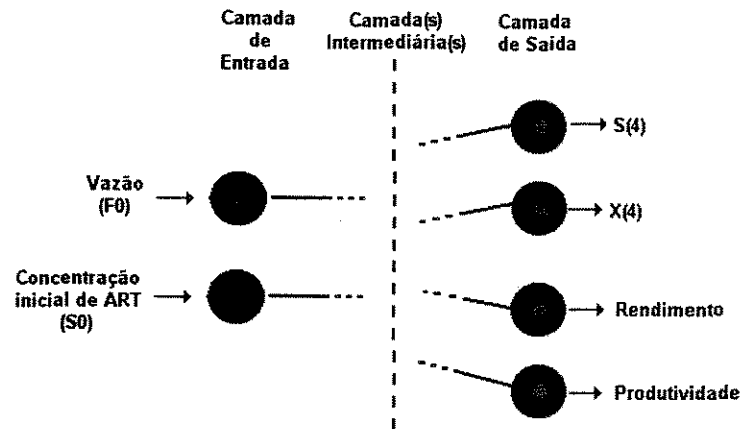


Figura 20: Configuração geral da Rede Neural Artificial (Caso A)

O treinamento da Rede Neural Artificial é realizado conforme demonstrado na seção 2.4 e no Apêndice. O conjunto de dados de entrada é apresentado à rede e o “sinal” se propaga pela rede até a camada de saída gerando as variáveis de saída. Estas variáveis são comparadas com os valores de treinamento (neste caso os valores obtidos pelo modelo determinístico). O erro é calculado e utilizado para gerar um novo conjunto de pesos. Cada par de entrada/saída vai sendo apresentado à rede até que o erro ou o número de iteração desejado seja atingido. Neste ponto tem-se um conjunto de pesos ajustados para o conjunto de dados usados no treinamento. Este conjunto de pesos ajustado é então utilizado para predição dos valores de $S(4)$, $X(4)$, $REND$ e $PROD$ após o usuário fornecer os valores de $F0$ e $S0$. Esta Rede Neural Artificial sofreu um treinamento “generalizado” que consiste na apresentação dos dados à rede de modo a cobrir um intervalo pré-determinado de cada variável de entrada.

Para cada configuração de rede foi calculado o Erro Relativo Médio (E_{Rel}) comparando-se o resultado da Rede Neural Artificial e o modelo. O treinamento é feito de modo a se minimizar o Erro Quadrático Global (E_{QG}). O erro E_{Rel} foi utilizado para se verificar a melhor configuração de rede.

$$E_{rel} = \frac{\sum |VS_{rede} - VS_{modelo}|}{N}$$

Equação 35

$$E_{QG} = \frac{1}{2} \cdot \sum_m [(o_m - d_m)^2]$$

Equação 36

onde VS_{rede} e VS_{modelo} são os valores da variável de saída VS provindos da

rede e do modelo, respectivamente, ao se utilizar o conjunto de pesos ajustados. N é o número de pares utilizados na representação. O_m e d_m são os valores obtidos pela rede e o desejado (modelo) durante o processo de treinamento. A variável m representa o número de variáveis de saída da rede.

A primeira e importante etapa do treinamento de Rede Neural Artificial é a obtenção do conjunto de dados para treinamento. Neste trabalho o conjunto de dados de treinamento foi obtido do modelo determinístico implementado em linguagem FORTRAN utilizando-se as equações descritas no capítulo 4. O modelo matemático permite gerar uma quantidade enorme de dados. No entanto os dados para o treinamento devem ser representativos e abranger todo o intervalo estudado. Ao se trabalhar com dados experimentais o número de pares de treinamento pode ser limitado requerendo cuidado na definição do intervalo de valores das variáveis.

Inicialmente, foram gerados 1517 pares (entrada/saída) a partir do modelo determinístico. Estes dados foram obtidos medindo-se o valor das variáveis após 30h de processo, ao sofrer perturbação de $\pm 20\%$ no valor inicial da variável F_0 e $\pm 10\%$ em S_0 . A tabela 10 mostra os intervalos das variáveis envolvidas. Como não se recomenda a extrapolação na utilização de Redes Neurais Artificiais o sistema deve ser simulado no intervalo em que as variáveis foram treinadas.

Tabela 10: Intervalo das variáveis de treinamento da rede (caso A).

Variável	Intervalo
F_0	80 à 120 (m^3/h)
S_0	162 à 198 (g/l)
$S(4)$	1,59E-02 à 15,28 (g/l)
$X(4)$	30,74 à 31,40 (g/l)
REND	78,08 à 86,47 (%)
PROD	5,93 à 9,48 ($g.l^{-1}.h^{-1}$)

Obs: Estes valores são normalizados (0,1 a 0,9) pelo programa de treinamento da Rede Neural Artificial.

O conjunto de dados gerados foi dividido em dois subconjuntos: um para o treinamento da rede (1062 pares) e outro para a validação (455 pares). É sugerido que o número de dados para validação seja em torno de 30% do total utilizado no treinamento. Os dados foram colocados em ordem aleatória para se evitar tendências no processo de treinamento. Para as duas fases (treinamento/validação) foi calculado o E_{Rel} em %.

Não existe, até o momento, nenhuma técnica ou teoria plenamente confiável que permita estimar a melhor configuração de rede de modo claro e objetivo. As escolhas são feitas, normalmente, de forma empírica sendo algumas sugestões sugeridas pela literatura. A tabela 11 apresenta o resultado de treinamentos iniciais. Para a validação das redes utilizou-se todo o conjunto de 455 pares que não participaram do treinamento.

Tabela 11: Configurações de rede iniciais e respectivos erros (caso A).

	Neu	β	Nº pares	E_{QG}	E_{rel} em % (Treinamento)				E_{rel} em % (Validação)			
					$S(4)$	$X(4)$	$Rend$	$Prod$	$S(4)$	$X(4)$	$Rend$	$Prod$
1	3	0,5	50	3,00E-03	27,07	0,03	0,05	0,56	23,31	0,03	0,07	0,62
2	5	0,5	50	2,00E-03	38,16	0,02	0,04	0,52	26,67	0,03	0,06	0,67
3	10	0,5	50	2,00E-03	16,30	0,02	0,04	0,50	16,20	0,03	0,07	0,60
4	3	0,5	200	8,34E-02	26,69	0,03	0,05	0,57	30,52	0,03	0,05	0,56
5	5	0,5	200	3,63E-02	25,61	0,02	0,05	0,28	32,14	0,02	0,05	0,29
6	10	0,5	200	2,00E-03	28,49	0,01	0,04	0,10	44,77	0,01	0,05	0,11
7	3	0,5	800	2,52E+01	0,08	0,36	5,22	5,26	19039	149	88	2229
8	5	0,5	800	2,33E+01	0,04	0,10	0,95	5,20	19507	148	92	2252
9	10	0,5	800	2,29E+01	0,01	0,06	1,25	5,20	19047	148	88	2486

Outras configurações de rede foram analisadas variando-se o número de pares de treinamento, número de neurônios na camada intermediária, dentre outros fatores. Alguns resultados sugerem a ocorrência de mínimo local. Ao ser encontrado um mínimo local o erro para o conjunto de treinamento pára de diminuir e estaciona em um valor maior que o aceitável.

Dentre as configurações analisadas, as características da rede que melhor a ajustou são:

- Taxa de aprendizado = 0,5
- Tolerância = 0,0009
- Uma camada intermediária com 10 neurônios,
- 75 pares (entrada/saída) utilizados no treinamento,
- Número de iterações: 361.360
- Tempo de treinamento: 22 min e 41 Seg,

Os erros para esta configuração são mostrados nas tabelas 12 e 13. É possível reduzir ainda mais estes erros através de novos treinamentos utilizando-se outras configurações e conjunto de dados. No entanto, considerou-se esta configuração satisfatória para o bioprocessamento em questão.

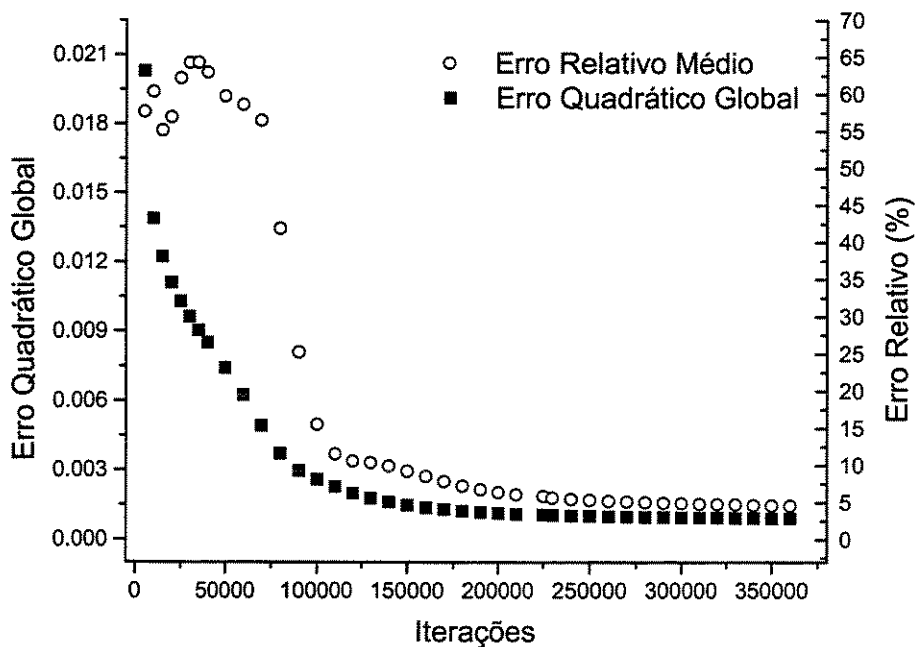
Tabela 12: Erros para os pares de treinamento

Entradas	Erro relativo médio	Erro relativo médio em %
S(4)	2,06E-02	4,57%
X(4)	1,12E-03	0,0038%
Rendimento	1,03E-02	0,012%
Produtividade	7,21E-03	0,0947%

Tabela 13: Erros para os pares de validação

Valor de S(4)	Erro relativo médio	Erro relativo médio em %
S(4)	2,74E-02	7,603%
X(4)	1,33E-03	0,004%
Rendimento	1,38E-02	0,016%
Produtividade	8,98E-03	0,118%

A Fig. 21 mostra a evolução do E_{QG} do treinamento e o E_{Rel} , em %, para a variável S(4).

Figura 21: Evolução do E_{QG} e E_{Rel} , para S(4), durante o treinamento.

A Fig. 22 apresenta os gráficos de comparação das variáveis de saída da rede com o modelo. Os coeficientes de correlação calculados e a proximidade dos pontos à reta $Y=X$ possibilitam verificar a eficácia do treinamento da Rede Neural.

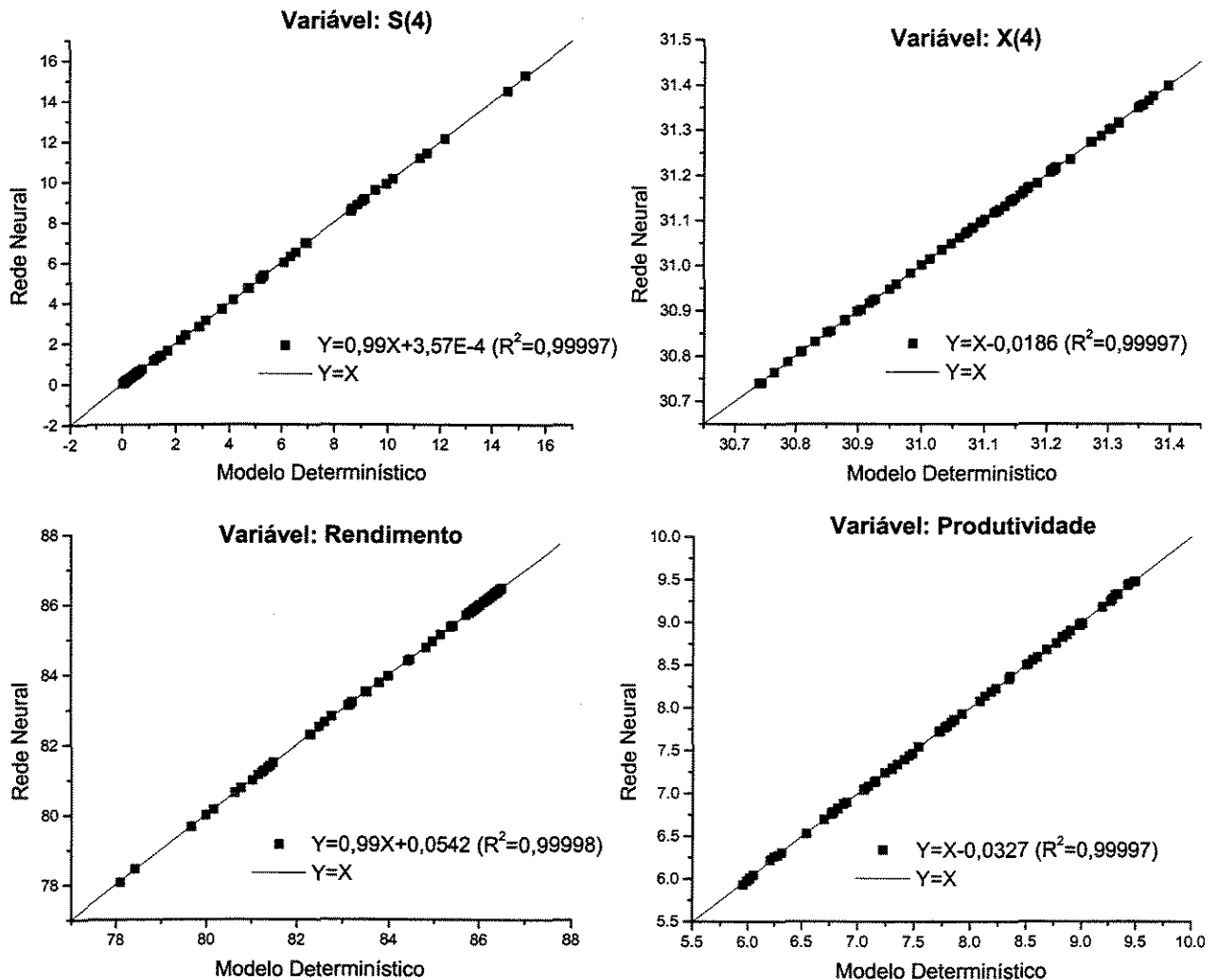


Figura 22: Gráficos de comparação das variáveis S(4), X(4), REND e PROD do modelo determinístico e da Rede Neural Artificial.

Pelos resultados obtidos pode-se notar uma maior dificuldade de se fazer o ajuste para a variável S(4). Isto se deve à sua alta não-linearidade demonstrada no capítulo 5. Este modelo foi criado para demonstração de como é feito o treinamento da rede e sua utilização. A seguir será apresentado um modelo dinâmico do processo no qual pode-se obter maiores informações sobre o mesmo.

6.3.2 - Caso B

A Fig. 23 mostra a configuração de uma Rede Neural Artificial que representa o perfil dinâmico das variáveis $S(4)$, $X(4)$ e $P(4)$ após perturbação degrau na variável F_0 para o processo fermentativo de produção de etanol.

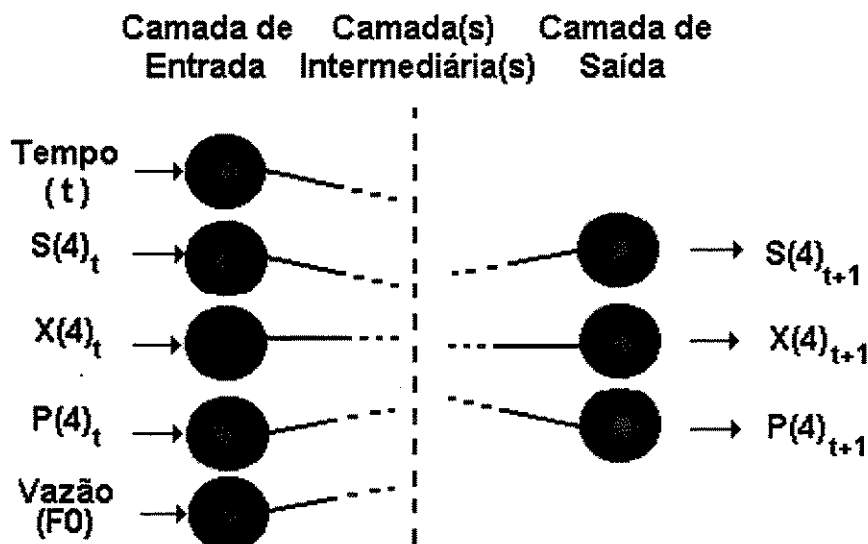


Figura 23: Configuração geral da Rede Neural Artificial (Caso B).

As variáveis de entrada são: F_0 , t (tempo), $S(4)$, $X(4)$ e $P(4)$ no tempo t . As variáveis de saída são: $S(4)$, $X(4)$ e $P(4)$ no tempo $t+1$.

Para o treinamento desta rede foram utilizados 200 pares entrada/saída extraídos de um conjunto de 4141 pares de dados. Este conjunto de 4141 dados foi obtido pela perturbação da variável F_0 no intervalo de $\pm 20\%$ do seu valor central. Para cada perturbação, obteve-se o valor de $S(4)$, $X(4)$ e $P(4)$ até 20h de simulação. No arquivo de treinamento são colocadas as entradas seguidas das saídas que serão utilizadas para o ajuste de pesos. A tabela 14 contém uma amostra do conjunto de dados de treinamento.

O treinamento com números de dados maiores (até 1200) demonstraram que acima de 200 dados não havia melhora significativa para a rede. Como já dito, as Redes Neurais Artificiais não têm capacidade de extrapolação. Sendo assim, ao se utilizar a rede treinada para um processo, deve-se respeitar os intervalos para os quais ela foi treinada. A tabela 15 contém os intervalos das variáveis desta rede.

Tabela 14: Amostra do conjunto de dados para treinamento da Rede Neural Artificial Dinâmica.

t	F0	S(4) _t	X(4) _t	P(4) _t	S(4) _{t+0,2}	X(4) _{t+0,2}	P(4) _{t+0,2}
9.4	104	1.0247	31.1289	65.5087	1.0247	31.1289	65.5087
5.8	98	0.7373	31.1376	65.6367	0.7371	31.1376	65.6367
7.0	87	0.3797	31.1487	65.7958	0.3795	31.1487	65.7959
9.0	110	1.4533	31.1161	65.3180	1.4535	31.1161	65.3179
15.0	96	0.6565	31.1396	65.6726	0.6565	31.1396	65.6726
0.2	111	0.8802	31.1331	65.5732	0.9300	31.1314	65.5510
9.4	80	0.2391	31.1526	65.8584	0.2391	31.1526	65.8584
4.6	86	0.3635	31.1497	65.8031	0.3620	31.1498	65.8037
3.8	104	1.0135	31.1286	65.5138	1.0156	31.1286	65.5129
15.8	92	0.5175	31.1436	65.7344	0.5175	31.1436	65.7344
16.6	90	0.4577	31.1453	65.7610	0.4577	31.1452	65.7610
7.8	87	0.3793	31.1486	65.7960	0.3792	31.1485	65.7960
15.8	105	1.0803	31.1276	65.4840	1.0803	31.1277	65.4840
7.8	98	0.7367	31.1375	65.6369	0.7367	31.1375	65.6369
17.8	82	0.2735	31.1505	65.8430	0.2735	31.1505	65.8430
17.8	96	0.6565	31.1396	65.6726	0.6565	31.1396	65.6726
17.0	109	1.3720	31.1195	65.3542	1.3720	31.1195	65.3542
11.0	111	1.5403	31.1141	65.2793	1.5403	31.1141	65.2793
17.8	84	0.3122	31.1494	65.8258	0.3122	31.1494	65.8258
16.2	118	2.2980	31.0933	64.9421	2.2980	31.0933	64.9421

Tabela 15: Intervalo das variáveis de treinamento da rede (caso B).

Variável	Intervalo
Tempo (t)	0.2 a 19.8 (h)
F0	80 a 120 (m ³ /h)
S(4)	0.239 a 2.606 (g/l)
X(4)	31.08 a 31.15 (g/l)
P(4)	64.80 a 65.866 (g/l)

A tabela 16 mostra algumas configurações testadas e os respectivos resultados para erros. Em algumas foi utilizado o número de iterações como critério de parada e em outras foi utilizado o valor do E_{OG} . Foi calculado o E_{Rel} em % de S(4) para os 200 dados de treinamento e E_{Rel} em % de S(4) para os valores obtidos da dinâmica do processo pelo modelo determinístico e pela Rede Neural Artificial (F0 variou de 80 a 120 m³/h). A variável S(4) foi utilizada como parâmetro pois é a mais difícil de se ajustar. A rede ajustou as demais variáveis (X(4) e P(4)) muito bem como será mostrado mais adiante.

Tabela 16: Algumas configurações de rede iniciais e respectivos erros (caso B).

	Neu	Iteração	β	E_{QG}	E_{Rel} (%) (Treinamento)	Dinâmica	
						E_{Rel} (%)*	Maior E_{Rel} (%)**
1	1	30.000	0,9	3,42E-01	11,03	54,62	178.66
2	2	30.000	0,9	2,19E-02	1,60	7,24	30.37
3	9	2.000	0,3	1,47E-02	1,38	11,20	49.79
4	10	5.000	0,9	3,46E-03	0,66	7,03	33.40
5	6	11.000	0,6	1,69E-03	0,41	3,02	21.59
6	5	125.000	0,6	1,40E-03	0,36	2,32	16.13
7	10	30.000	0,3	1,39E-03	0,34	3,41	24.36
8	5	30.000	1,0	8,80E-04	0,32	2,27	10.44
9	8	30.000	0,6	6,84E-04	0,29	1,69	12.15
10	10	30.000	0,9	5,43E-04	0,29	2,03	11.94
11	9	30.000	0,9	5,04E-04	0,26	1,83	11.88
12	7	106.128	0,9	4,00E-04	9,86E-04	1,41	5.58

* E_{Rel} (%) = Erro relativo médio em % de várias simulações para perturbações em F_0 .

** Maior E_{Rel} (%) = Maior valor para E_{Rel} (%) após simulações com perturbações em F_0 .

A rede que possibilitou melhor ajuste foi:

- Taxa de aprendizado = 0,9
- Tolerância = 0,0004
- Uma camada intermediária com 7 neurônios,
- 200 pares (entrada/saída) utilizados no treinamento,
- Número de iterações: 106.128
- Tempo de treinamento: 11 min e 21 Seg.

Todas as Redes Neurais Artificiais definidas como a melhor no caso A e B necessitam de pouco tempo para o treinamento. Nas tentativas de se obter a melhor configuração obteve-se redes que requereram até 24h de treinamento. O tipo de processador e quantidade de memória RAM influenciam, e muito, o tempo necessário para atingir o treinamento desejado.

O modelo tipo “caixa preta” baseado em Rede Neural para a dinâmica do processo foi desenvolvido utilizando-se o conjunto de pesos ajustados na rede anterior. O fluxograma do programa que utiliza este conjunto de pesos ajustados é apresentado na Fig. 24. O programa contém os valores iniciais de $S(4)$, $X(4)$ e $P(4)$ idênticos aos fornecidos para o modelo determinístico. O usuário entra com um valor de F_0 (de 80 a 120 m^3/h) e o programa retorna os valores de $S(4)$, $X(4)$ e $P(4)$ para cada tempo variando de 0 a 20h (com passo de 0,2).

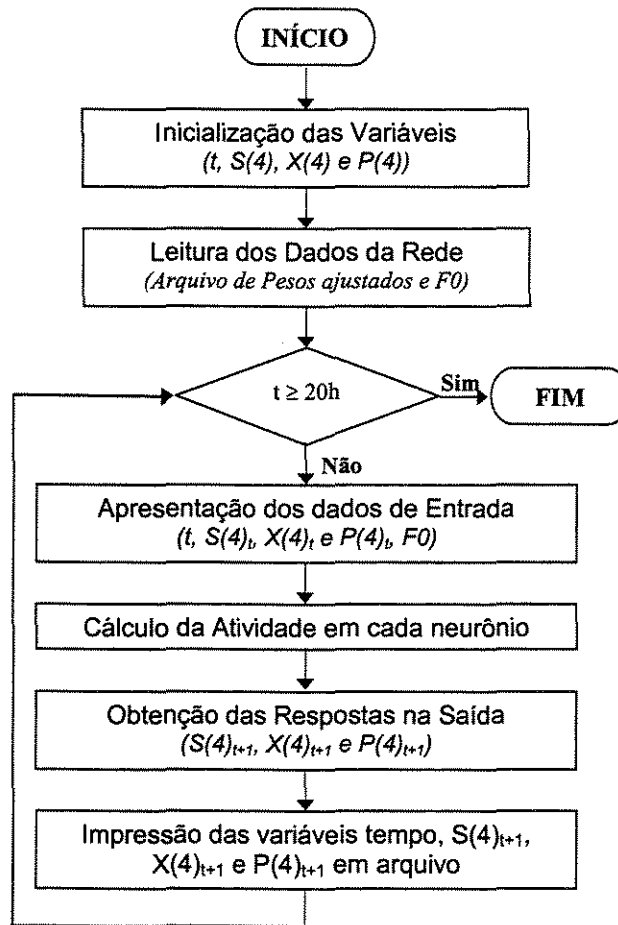


Figura 24: Fluxograma do programa de utilização dos pesos ajustados (Modelo “caixa preta”).

A Fig. 25 mostra os perfis dinâmicos das variáveis $S(4)$, $X(4)$ e $P(4)$ após perturbação de $\pm 20\%$ em $F0$ para valores obtidos do modelo determinístico e pela Rede Neural. Note que quase não há variação para $X(4)$ e $P(4)$ o que facilitou seu ajuste na fase de treinamento da rede.

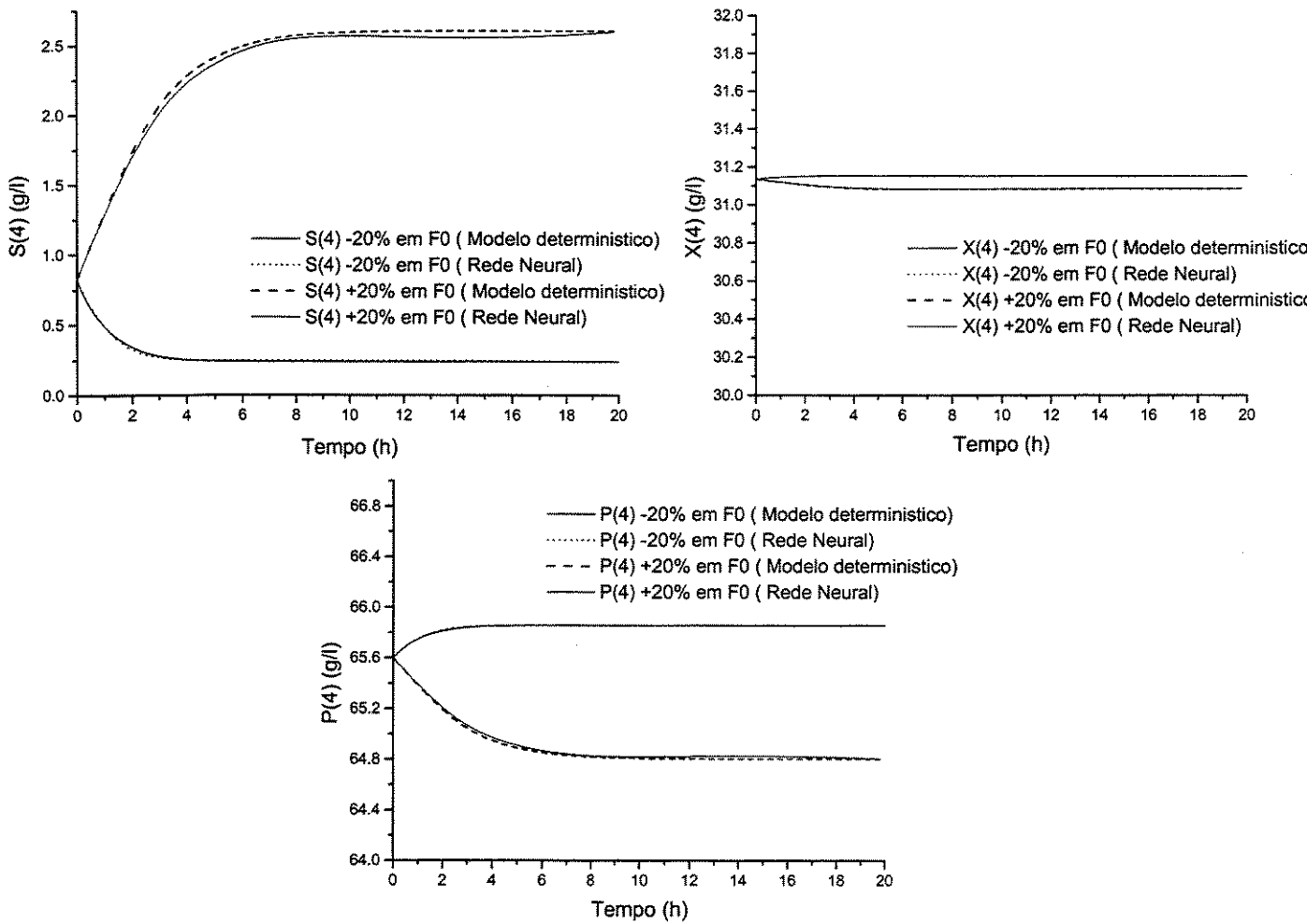


Figura 25: Perfil de $S(4)$, $X(4)$ e $P(4)$ para perturbação degrau em F_0 (Modelo determinístico e Redes Neurais Artificiais).

O modelo “caixa preta” foi testado através de várias simulações para perturbações em F_0 . O erro E_{Rel} (%) de $S(4)$ (comparando-se o valor do modelo determinístico e da rede) para cada dinâmica de perturbação é mostrado na Fig. 26. Pode-se notar que o modelo se aplica a todo o intervalo. Para demonstrar a baixa capacidade de extrapolação deste modelo foram realizadas simulações fora do intervalo de treinamento.

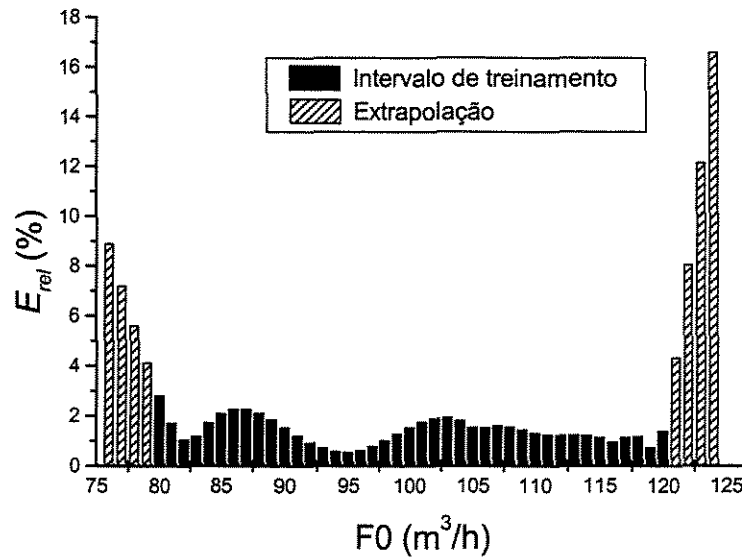


Figura 26: E_{rel} para cada perturbação em F_0 .

A utilização do algoritmo da Fig. 24 causa um acúmulo de erros, uma vez que os dados estimados na saída da rede no tempo t são utilizados como dados de entrada no tempo $t+0,2$. Como o ajuste dos pesos não é perfeito, as variáveis estimadas sempre contêm erros embutidos, em relação ao modelo determinístico. Os resultados obtidos nas simulações não revelaram um acúmulo de erro significativo. Geralmente esta predição chamada de “um passo a diante” é utilizada em controladores em que se deseja apenas saber o valor das variáveis no tempo $t+1$ ao serem fornecidos os valores das variáveis (medida diretamente no processo) no tempo t . Assim não há necessidade da obtenção de um perfil dinâmico não ocorrendo o acúmulo de erro explicado anteriormente. Este trabalho trouxe a inovação de se utilizar a variável tempo no treinamento da rede possibilitando a criação do perfil dinâmico. O trabalho desenvolvido por LINKO e ZHU (1991) aborda de maneira bastante clara, a utilização do conceito de predição um passo a diante.

Sabe-se que, para obtenção do modelo determinístico, várias simplificações são necessárias. A Rede Neural Artificial obtida do modelo determinístico incorpora suas simplificações. O treinamento de redes com dados provindos do processo pode levar à obtenção de resultados mais realistas sendo esta uma das razões para o aumento do estudo de Redes Neurais Artificiais.

Os modelos desenvolvidos neste capítulo podem ser incorporados a um

controlador e também utilizados para otimização do processo.

6.4 - Conclusão

Neste capítulo foram desenvolvidos dois modelos baseados em Redes Neurais Artificiais (tipo “caixa preta”). O primeiro modelo fornece os valores de $S(4)$, $X(4)$, $REND$ e $PROD$ após 30h de processo (estado estacionário). O desenvolvimento deste modelo serviu também como base para o aperfeiçoamento nas técnicas de treinamento de RNAs e sua validação. O segundo modelo introduziu a técnica de treinamento com predição um passo a diante. Esta técnica permitiu a obtenção de perfis dinâmicos das variáveis $S(4)$, $X(4)$ e $P(4)$ após perturbação em $F0$. Ambos os modelos apresentaram bons resultados.

O próximo capítulo apresenta a modelagem híbrida do processo fermentativo. Uma Rede Neural Artificial que calcula a velocidade específica de crescimento da levedura foi treinada e acoplada ao modelo determinístico.

Capítulo 7

Modelagem Híbrida

7.1 - Introdução

O modelo híbrido desenvolvido neste capítulo é composto do modelo determinístico mostrado no capítulo 4 e uma Rede Neural Artificial que calcula o valor da velocidade específica de crescimento da *Saccharomyces cerevisiae*. Primeiramente foi realizado o treinamento e validação da Rede Neural Artificial e em seguida realizada a análise da modelagem híbrida.

7.2 - Cálculo da Velocidade Específica de Crescimento da *Saccharomyces cerevisiae* com Rede Neural Artificial

Na modelagem determinística, o cálculo de μ é realizado através de uma subrotina do programa. Esta subrotina será substituída por um algoritmo de Rede Neural Artificial devidamente treinada e validada constituindo, assim, o modelo híbrido. Neste capítulo,

serão mostrados os passos para a obtenção da Rede Neural Artificial que fará o cálculo de μ .

Foi desenvolvido um programa em FORTRAN para geração dos dados utilizados no treinamento e validação da rede. Este programa é similar à subrotina que faz o cálculo de μ durante a simulação usando o modelo determinístico. Para o cálculo do erro da Rede Neural Artificial são comparados os valores de μ obtidos pela rede e pelo modelo determinístico.

Analisando o cálculo de μ (Equação 1), pode-se visualizar que a Rede Neural Artificial que representará este cálculo conterá 4 neurônios na camada de entrada (S, P, X e T) e 1 na camada de saída (μ). Como o modelo determinístico foi obtido com dados de processo que possui um controlador PI para temperatura, esta variável não foi considerada no cálculo de μ . Sendo assim a rede terá a configuração geral demonstrada na Fig. 27.

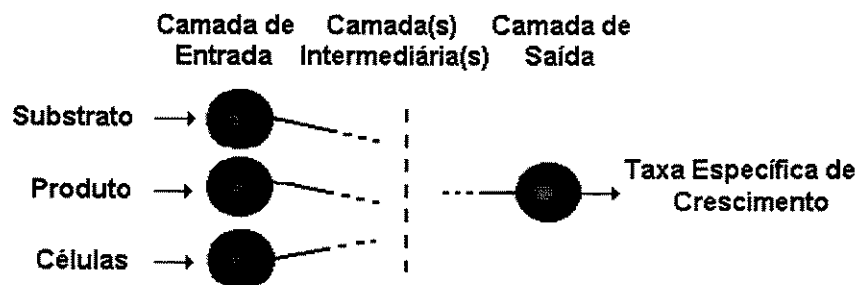


Figura 27: Configuração geral da Rede Neural Artificial para o cálculo de μ

A tabela 17 mostra o intervalo de valores de cada variável de entrada usada na fase de treinamento. Estas faixas foram definidas analisando-se os valores com os quais o sistema a ser simulado trabalha. Na última linha da tabela tem-se a faixa de valores da velocidade específica de crescimento obtida com estas variáveis de entrada.

Tabela 17: Intervalo das variáveis independentes e dependente.

Variável	Intervalo
Substrato (S)	0.1376 a 58.553 (g/l)
Produto (P)	39.909 a 65.904 (g/l)
Células (X)	29.233 a 31.157 (g/l)
Velocidade específica de crescimento (μ .)	6.605E-004 a 6.2261E-002 (1/h)

A rede sofreu um “treinamento orientado”, ou seja, é utilizado um conjunto de dados de entrada (S,P,X) provindos da simulação do modelo determinístico após sofrer perturbação de $\pm 20\%$ na variável F0.

Segundo PAPES (1999), a utilização de dados (combinações) “conhecidos” pelo reator simulado, durante o treinamento, colabora para a adequação da simulação à realidade do sistema (maior precisão), prejudicando, no entanto, a capacidade de generalização da rede (aplicável a uma faixa menor das variáveis operacionais e de projeto).

O conjunto de dados gerados foi dividido em dois subconjuntos: um para o treinamento da rede, onde a rede “aprende” a relação de μ com S,P,X e outro para a sua validação. Em testes iniciais foi definido o número de 1150 pares de dados (entrada/saída) para o treinamento. Utilizou-se para validação 400 pares de dados que não foram usados no treinamento. A tabela 18 apresenta alguns resultados obtidos.

Tabela 18: Configurações de algumas redes e respectivos E_{rel} em % do valor de μ para dados de treinamento e validação.

Rede	Neu	β	E_{QG}	Iteração	E_{rel} em %	
					Treinamento	Validação
1	7	0,80	2,37E-03	233.334	1,150	0,968
2	9	0,80	2,56E-03	233.334	1,180	1,023
3	8	0,65	2,40E-03	166.668	1,285	1,145
4	6	0,95	2,52E-03	166.668	1,124	0,946
5	8	0,95	2,43E-03	166.668	1,096	0,922
6	7	0,80	2,44E-03	200.000	1,170	1,010
7	6	0,65	1,26E-03	300.001	0,943	0,859
8	5	1,10	1,23E-03	300.000	1,003	0,935
9	8	0,95	2,48E-03	150.001	1,118	0,962
10	7	0,90	2,19E-03	287.501	1,152	0,909
11	6	0,95	2,29E-03	300.001	1,139	0,904
12	8	0,65	2,09E-03	250.001	1,026	0,859
13	7	0,90	2,50E-03	162.501	1,204	1,055
14	5	0,80	1,29E-03	233.334	1,012	0,961
15	5	1,10	2,47E-03	100.000	1,150	0,999
16	7	1,20	2,22E-03	225.001	1,123	0,898
17	7	0,50	2,42E-03	233.334	1,339	1,193
18	7	1,00	2,29E-03	225.001	1,129	0,908
19	9	1,10	3,42E-03	100.000	1,748	1,588
20	8	0,95	2,26E-03	250.001	1,126	0,900

Como se pode observar, a rede apresenta baixo E_{rel} para todos os conjuntos de dados usados. O ajuste desta rede deve ser muito bom pois ela entra no modelo determinístico como uma subrotina que fornece os valores de μ que serão posteriormente utilizados para gerar novas variáveis de saída (S, X e P). Estas variáveis são utilizadas novamente na rede para o cálculo de uma nova velocidade. Assim, o erro da rede se torna acumulativo na modelagem híbrida. Na escolha da melhor configuração analisou-se os E_{rel} exemplificados na tabela 18 e também os E_{rel} das variáveis S(4), X(4), P(4), REND e PROD, comparando-se o modelo híbrido (obtido com os pesos ajustados para cada configuração) com o determinístico. Esta comparação é demonstrada com maior detalhe na seção 7.3.

Após análise de todos os resultados pôde-se notar que a utilização de uma taxa maior de aprendizado (β) melhorou a capacidade de generalização e o treinamento da rede de modo geral, para o conjunto de 1150 dados. Ao se utilizar um outro conjunto de dados esta afirmação foi inversa, ou seja, obteve-se melhores resultados com β menores. Isto comprova que a escolha da configuração e tipo de rede depende de vários fatores não havendo relações claras que possam ajudar na escolha da melhor configuração. Para algumas redes, ocorreram aumentos dos erros após várias iterações. Isto acontece quando o erro caminha até o mínimo na superfície passando por ele e voltando a crescer (Fig. 28). Este mínimo é maior que o estipulado pelo usuário como critério de parada.

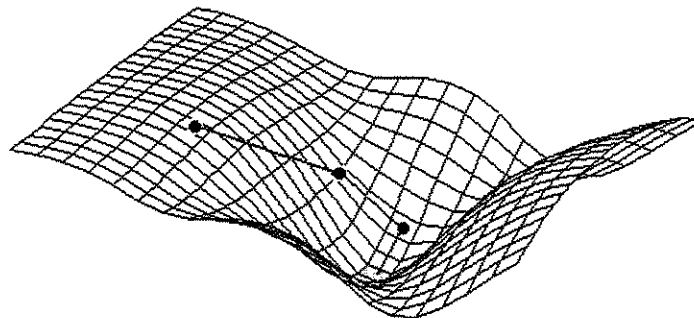


Figura 28: Exemplo de uma superfície de erro.

A melhor configuração obtida foi:

- Taxa de aprendizado = 0,9
- Tolerância = 2.189E-3

- Duas camadas intermediárias com 7 neurônios cada,
- 1150 pares (entrada/saída) utilizados no treinamento,
- Número de iterações: 287.501
- Tempo de treinamento: 17 min e 56 Seg,
- E_{rel} em % do valor de μ para treinamento, 1,152% e validação, 0,909%

A Fig. 29 mostra os resultados plotados em gráficos para os dados de treinamento e validação utilizando-se os pesos ajustados para a rede descrita anteriormente. Na Fig. 30 está representada a evolução do E_{QG} durante o treinamento. Foi representado o logaritmo do erro para facilitar a visualização uma vez que há uma queda brusca nas primeiras iterações.

No desenvolvimento de um modelo determinístico para o cálculo de μ são feitas simplificações. As Redes Neurais Artificiais correlacionam um conjunto de dados de entrada a valores correspondentes na saída. Uma rede treinada com dados experimentais ou mesmo industriais possui melhor representatividade do sistema, apesar de só trabalhar na faixa em que foi treinada.

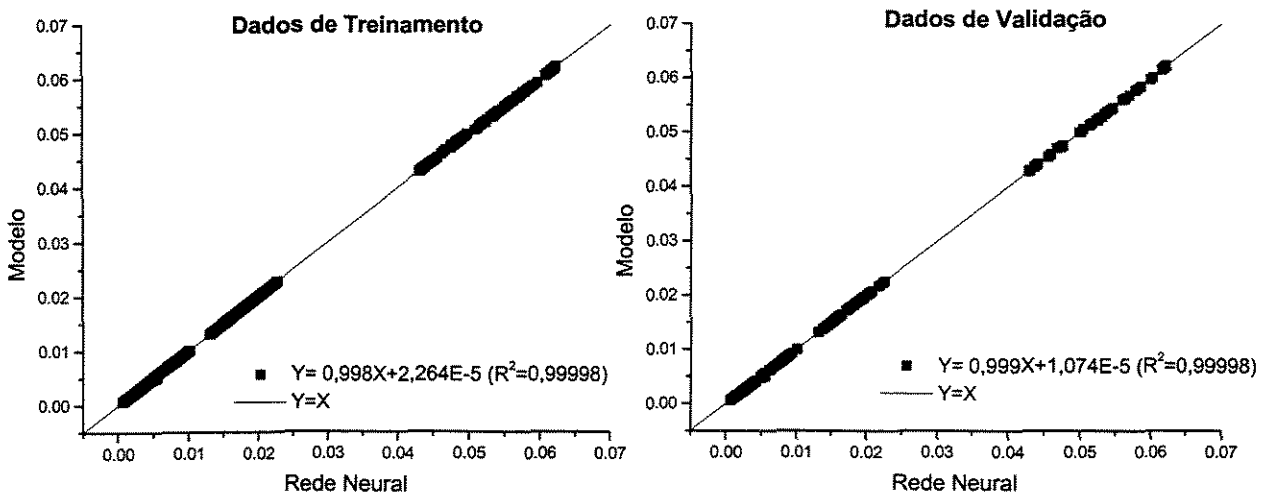


Figura 29: Valores de μ obtidos pela Rede Neural Artificial e pelo modelo.

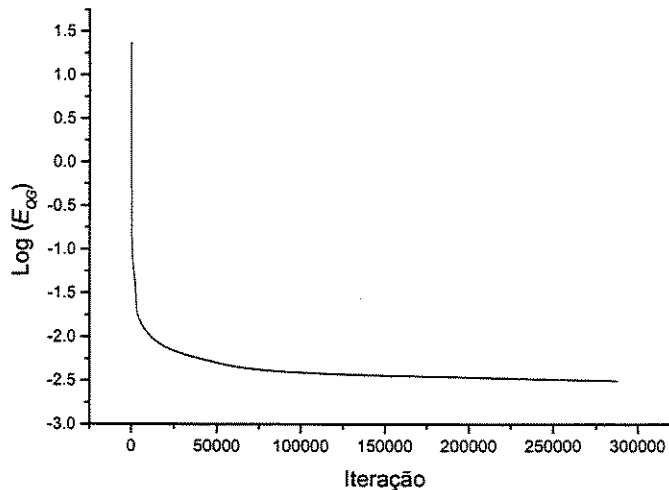


Figura 30: Evolução do E_{QG} durante o treinamento.

7.3 - Modelo Híbrido

O modelo híbrido do processo em estudo consiste do acoplamento de um algoritmo de Rede Neural Artificial ao algoritmo do modelo determinístico. A velocidade específica de crescimento do microrganismo é calculada a cada passo de integração através do algoritmo de Redes Neurais. No item anterior foi descrito como obter o conjunto de pesos, devidamente ajustados, que será utilizado para o cálculo de μ na modelagem híbrida.

Com a modelagem híbrida é possível se obter o perfil dinâmico das variáveis dependentes do processo fermentativo para obtenção de etanol. Todos os balanços de massa e energia do sistema continuam o mesmo, apenas o cálculo de μ é diferente. A velocidade específica de crescimento é particular para cada microrganismo, no caso em estudo *Saccharomyces cerevisiae*. Caso houvesse uma mudança na cinética deste processo uma nova rede poderia ser treinada (com dados do processo, por exemplo) e seus pesos utilizados no modelo híbrido.

No início do trabalho com o modelo híbrido ocorreram muitos problemas de “overflow” e “floating point error” na execução do programa ao se utilizar alguns conjuntos de pesos ajustados. O problema foi resolvido mas para utilização de modelos híbridos em controle em tempo real deve-se deixar disponível no sistema um dispositivo que detecte e alerte sobre este tipo de problema.

Como dito no item anterior, na escolha da melhor Rede Neural Artificial analisou-se o desempenho de seus conjuntos de pesos ajustados ao serem utilizados no modelo híbrido. Para isto foram realizadas simulações variando-se o valor de F0 com o modelo determinístico e híbrido e comparando-se os valores das variáveis S, X e P (para os 4 reatores) e REND e PROD obtidas. A variável S(4) foi a que teve maior dificuldade para ajuste sendo utilizada como ponto de referência para o cálculo de todos os E_{rel} dos resultados. A tabela 19 mostra os valores de E_{rel} em % para simulação com F0 variando em $\pm 20\%$ e o valor médio de 42 pares simulações (determinístico/rede) para F0 variando de 80 a 120 m³/h. Estes erros foram calculados para cada configuração de rede testada. A tabela 19 mostra alguns resultados e as configurações estão na tabela 18. A rede escolhida foi a N^o 10 por apresentar menor média e melhor relação nos erros.

Os valores do E_{rel} em % de S(4) para cada simulação após perturbação no valor de F0 é mostrado na Fig. 31. Nota-se que o modelo híbrido é satisfatório para todo o intervalo estudado.

Tabela 19: Valor de E_{rel} em % para S(4).

Rede	E_{rel} em %		
	80 (m ³ /h)	120 (m ³ /h)	Média*
1	3,32	8,27	1,94
2	1,68	9,59	2,38
3	4,33	9,24	2,61
4	3,84	8,21	1,91
5	2,41	8,98	1,89
6	3,45	8,55	2,19
7	31,89	5,93	9,52
8	41,60	5,07	11,81
9	2,47	9,15	2,09
10	3,19	7,57	1,64
11	4,35	7,25	1,84
12	4,92	7,59	1,96
13	3,62	8,79	2,35
14	32,34	3,73	9,62
15	5,43	9,43	2,16
16	3,62	7,74	1,80
17	4,72	8,30	2,77
18	3,35	7,96	1,69
19	24,26	8,76	6,38
20	2,54	8,37	1,59

Média* = Valor médio do E_{rel} para 42 simulações (F0 variando de 80 à 120m³/h)

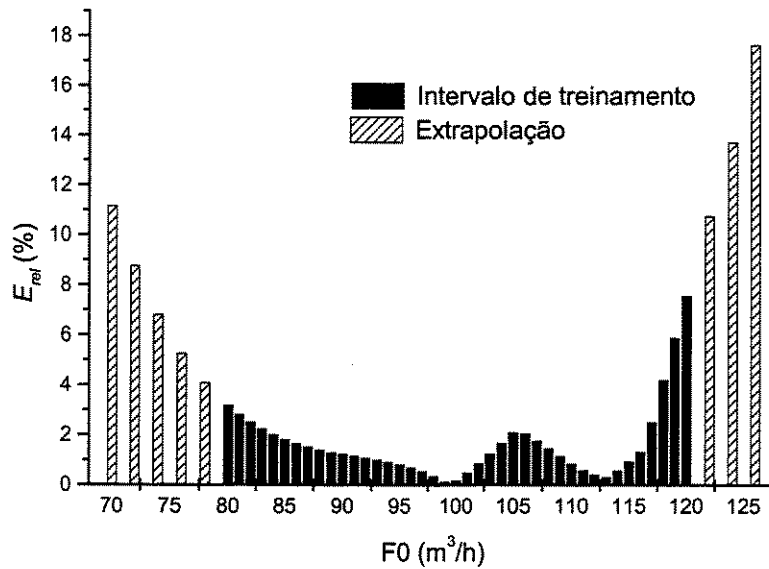


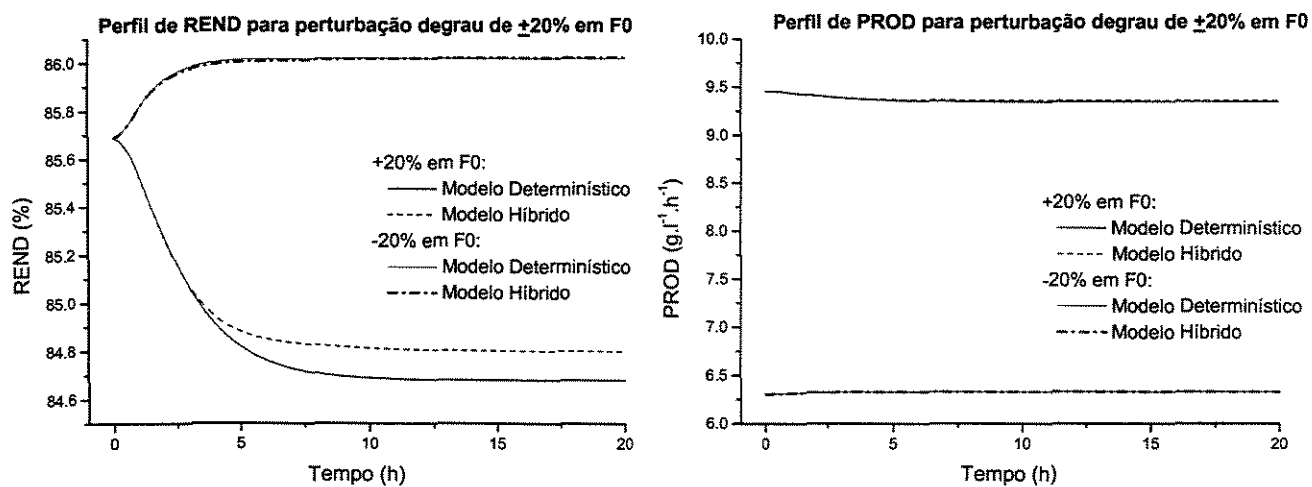
Figura 31: E_{rel} para cada perturbação em F_0 .

Utilizando o conjunto de pesos da rede N^o 10 foram realizadas várias simulações sendo alguns resultados para o E_{rel} mostrados na tabela 20. As Figuras 32 a 34 contém gráficos que mostram os perfis das variáveis estudadas para os dois modelos. As variáveis X e $REND$ praticamente não foram afetadas pela perturbação em F_0 durante as 30h de simulação. Seus perfis são praticamente constantes e o modelo híbrido respondeu bem, mesmo com pequenas variações. O tempo gasto na simulação do modelo híbrido e do determinístico são praticamente os mesmos. Deve-se tomar o cuidado ao se fazer o algoritmo do modelo híbrido para que o conjunto de pesos seja lido somente uma vez e seus valores guardados na memória pois a ação de leitura aumenta muito o tempo de simulação. Ao ser utilizado todo o sistema deve ser monitorado para indicar se há, ou não, necessidade de retreinamento.

Tabela 20: Valor de E_{rel} em % para S, X, P, REND e PROD.

Variáveis	E_{rel} em %		
	80 (m ³ /h)	120 (m ³ /h)	Média*
S(1)	0.102	0.088	0.036
S(2)	0.207	0.303	0.151
S(3)	3.688	0.181	1.125
S(4)	3.194	7.567	1.638
X(1)	0.006	0.009	0.003
X(2)	0.003	0.011	0.003
X(3)	0.009	0.002	0.004
X(4)	0.002	0.017	0.002
P(1)	0.049	0.058	0.020
P(2)	0.024	0.064	0.024
P(3)	0.057	0.011	0.029
P(4)	0.005	0.123	0.013
REND	0.005	0.118	0.013
PROD	0.005	0.118	0.013

Média* = Valor médio do E_{rel} para 42 simulações (F0 variando de 80 a 120m³/h)

Figura 32: Perfis de REND e PROD para perturbação degrau de $\pm 20\%$ em F0.

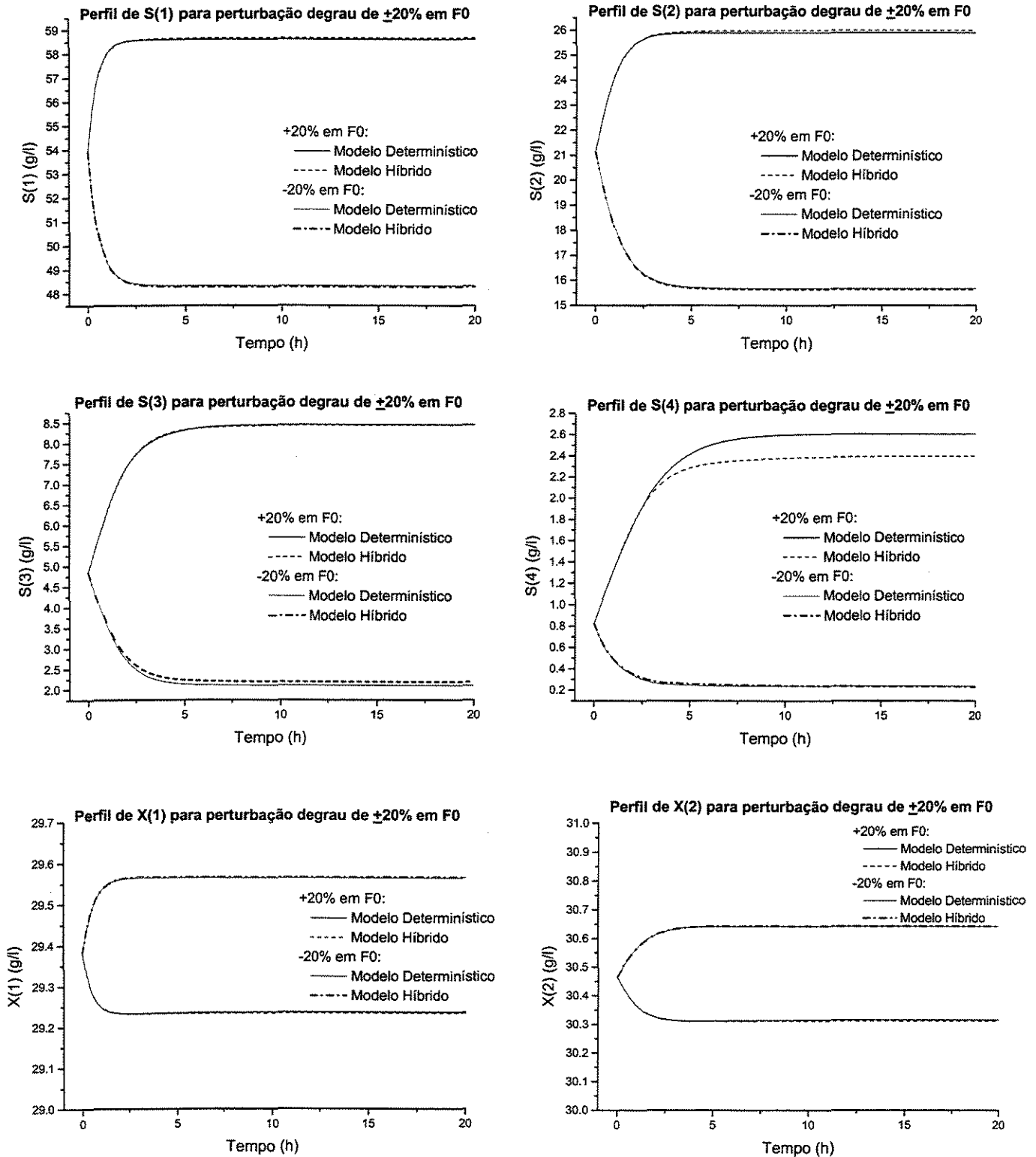


Figura 33: Perfis S(1), S(2), S(3), S(4), X(1) e X(2) para perturbação degrau de $\pm 20\%$ em F0.

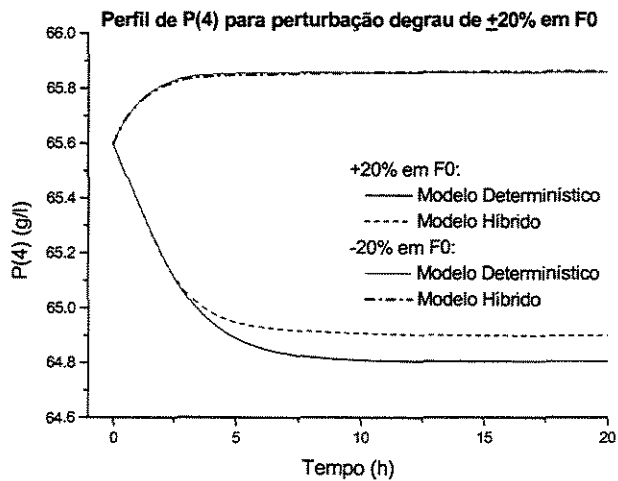
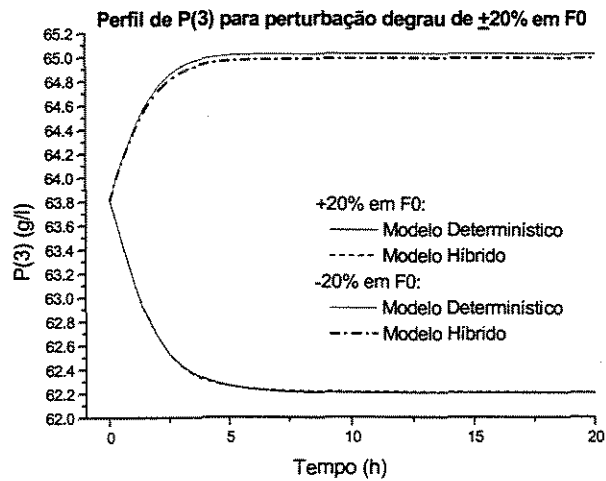
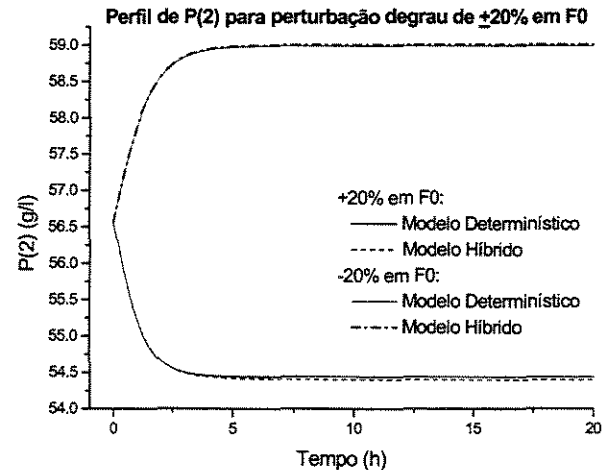
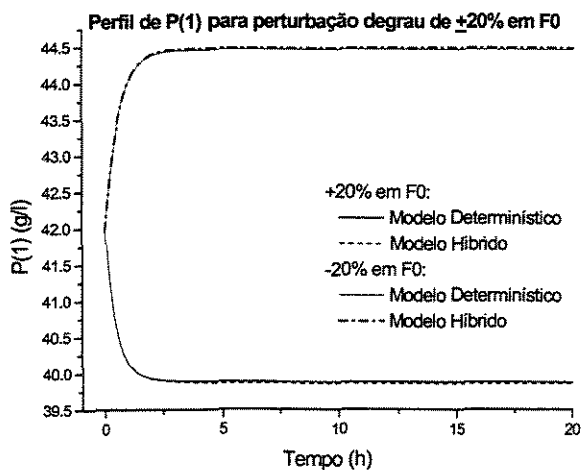
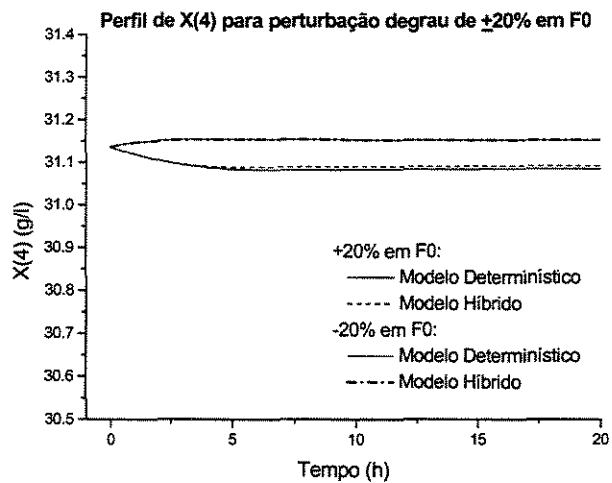
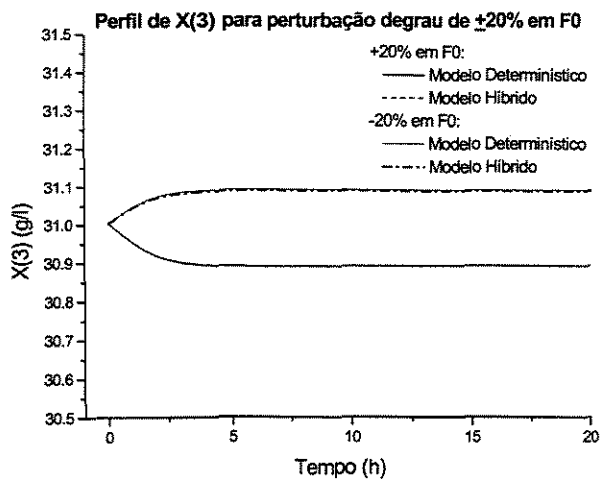


Figura 34: Perfis X(3), X(4), P(1), P(2), P(3) e P(4) para perturbação degrau de $\pm 20\%$ em F0.

7.4 - Conclusão

A modelagem híbrida apresentada mostrou-se satisfatória para o intervalo estudado. O treinamento da Rede Neural Artificial que calcula o valor de μ deve ser muito bom pois há acúmulo de erros durante a simulação.

O modelo híbrido trabalha somente dentro da faixa de valores mostrados na tabela 17. Esta é a grande desvantagem do modelo híbrido em relação ao determinístico. Por outro lado se a Rede Neural Artificial tivesse sido treinada com valores experimentais ela iria representar muito melhor o processo uma vez que estaria livre de simplificações.

Capítulo 8

Conclusões Gerais

O desenvolvimento deste trabalho permitiu uma avaliação das vantagens e possíveis problemas ao se utilizar modelagem híbrida em bioprocessos. Apesar da metodologia ainda empírica para treinamento de Redes Neurais Artificiais pode-se notar seu grande poder na modelagem de processos (modelos “caixa Preta”) e possível utilização em controle. Para fazer um bom uso das RNAs é preciso ter um bom conhecimento do processo e escolher bem as variáveis com as quais se deseja trabalhar. O modelo determinístico do processo fermentativo permitiu a análise das variáveis importantes para o sistema além de fornecer os dados utilizados no treinamento das Redes Neurais Artificiais. Na análise das variáveis concluiu-se que:

- O bioprocesso possui o comportamento dinâmico muitas vezes não-linear ocorrendo também resposta inversa.
- A variável TJE contribui de forma mais significativa para variações em $S(4)$, para rendimento e produtividade, mostrando a necessidade do controle de temperatura

no processo. A variável $S(4)$ é muito mais sensível às variações nas entradas se comparada ao rendimento e à produtividade.

- A análise de sensibilidade mostrou de forma mais clara as características não lineares do sistema, principalmente para a variável $S(4)$. Foi verificado também que um aumento positivo tanto em F_0 como S_0 contribui para o aumento da não linearidade das variáveis estudadas.

O desenvolvimento de modelos tipo “Caixa Preta” permitiu um melhor conhecimento da metodologia de treinamento de RNAs. O treinamento de redes que forneçam os valores das variáveis dependentes após entrarem em um novo estado estacionário é interessante do ponto de vista de controle e simulação. A aplicação do conceito de treinamento um passo adiante para a obtenção de perfis dinâmicos revelou-se apropriada com a utilização da variável tempo como entrada. Este artifício tende a diminuir o acúmulo de erros na simulação.

O treinamento da Rede Neural Artificial que calcula a velocidade específica de crescimento do microrganismo é o ponto chave para a modelagem híbrida. Para a escolha da melhor configuração de RNA e melhor conjunto de dados de treinamento para o cálculo de μ deve-se analisar os seguintes fatores:

- E_{QG} durante o treinamento;
- E_{rel} para os pares de treinamento e validação;
- E_{rel} dos dados de simulação ao se utilizar os pesos ajustados no modelo híbrido.

O modelo híbrido obtido é viável em todo o intervalo estudado, sendo a variável $S(4)$ utilizada como referência para o cálculo de erros. Uma desvantagem do modelo híbrido em relação ao determinístico é que ele trabalha em um intervalo definido não permitindo extrapolações. No caso estudado, este intervalo foi definido para variações de $\pm 20\%$ em F_0 pois os valores de S , P e X para cálculo de μ foram obtidos aplicando esta variação.

Capítulo 9

Referências Bibliográficas

ACKLEY, D. H., HINTON, G. E., SEJNOWSKI, T. J. A learning algorithm for Boltzmann machines. *Cognitive Science*, v.9, p.147-169, 1985.

ALEKSANDER, I., THOMAS, W. V., BOWDEN, P. A. WISARD. A radical step forward in image recognition. *Sensor Review*, v.4, p.120-124, 1984.

ANDERSON, J. A. A simple neural network generating an interactive memory. *Mathematical Biosciences*, v.14, p.197-220, 1972.

ANDERSON, J. A. *An Introduction to Neural Networks*. 1997.

ANDRIETTA, S. R. Fermentação alcoólica – Evolução dos processos industriais. *STAB*, v.17, n.1, p. 32, Set-Out, 1998.

ANDRIETTA, S. R., MAUGERI, F. Optimum design of a continuous fermentation unit of an industrial plant for alcohol production. *Advances in bioprocess engineering*, P.47-52, 1994.

ANDRIETTA, Silvio Roberto. *Modelagem. simulação e controle de fermentação alcoólica em escala industrial*. Campinas: Faculdade de Engenharia de Alimentos, UNICAMP, 1994. Tese (Doutorado).

ARBIB, M. A. *The Handbook of Brain Theory and Neural Networks*. 1995.

BATISTA, L. M. F. L. *Desenvolvimento de Software usando Modelos Determinísticos e redes neurais para o processo de Craqueamento Catalítico*. Campinas: Faculdade de Eng. Química, UNICAMP, 1996. Tese (Doutorado)

BHAGAT, P. An Introduction to Neural Nets. *Chemical Engineering Progress*, n.86, p.55-60, 1990.

CARPENTER, G., GROSSBERG, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer vision, graphics and image processing*, v.37, p.54-115, 1987a.

CARPENTER, G., GROSSBERG, S. ART2:Self-organization of stable category recognition codes for analog input patterns. *Applied optics*, v.26, p.4919-4930, 1987b.

CHITRA, S. P., BULSON, R. J., MORREL, D. G. Computer programs thata from experience. *Chemetech*, p.18-26, Feb, 1995.

DECHECHI, Eduardo Cesar. *Controle avançado preditivo adaptativo"DMC multivariavel adaptativo"*. Campinas: Faculdade de Eng. Química, UNICAMP, 1998. Tese (Doutorado)

ENDER, L., MACIEL FILHO, R. Design of multivariable controller based on neural networks. *Computers and Chemical Engineering*, v. 24, Issue: 2-7, p. 937 – 943, July 15, 2000.

FEIGENBAUM, B. A. E. *The Handbook of Artificial Intelligence*. Addison-Wesley, v.I, II, III, 1981.

FUKUSHIMA, K. Cognitron: a self-organizing multilayered neural network. *Biol.*

Cybernetics, v.20, p.121-136, 1970.

HEBB, D.O. *The organization of behaviour*. New York: John Wiley and Sons, 1949.

HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*. v.79, p.2554-2558, 1982.

KALIL, S. J. *Modelagem. simulação e otimização de um processo de fermentação alcoólica contínua. com reciclo de células. em reator cascata*. Campinas: Faculdade de Eng. de Alimentos, UNICAMP, 1997. Tese (Mestrado)

KANERVA, P. *Self-propagating search: a unified theory of memory*. University of Stanford, 1984. PhD dissertation

KARIM, M. N., YOSHIDA, T., RIVERA, S. L. Global and Local Neural Network Models in Biotchnology: Application to Different Cultivation Processes. *Review. J. of fermentation and Bioengineering*, v.3, n.1, p.1-11, 1997.

KOHONEN, T. Correlation matrix memories. *IEEE Transactions on Computers*. v.21, p.353-359, 1972.

KOVÁCS, Z. L. *Redes Neurais Artificiais: Fundamentos e Aplicações*. Edição Acadêmica. São Paulo, 1996

LEE, J. M., PALLARD, J. F., COULMAN, G. A. Ethanol fermentation with cell recycling: computer simulation. *Biotechnology and Bioengineering*, New york, v.25, n.2, p.497-511, Feb., 1983.

LIMA, U. A., AQUARONE, E., BORZANI, W. *Biotecnologia: Tecnologia das fermentações*. Ed. Edgard Blucher. 1982.

LINKO P., ZHU, Y. Neural networks programming in bioprocess variable estimation and state prediction. *J. of Biotechnology*, v.21, p. 253-270, 1991.

MACIEL FILHO, R., ZORZETTO, L. F. Modelling a batch beer production system, trough artificial neural networks anf hybrid models, using distinct levels of i. *The First*

European Congress on Chemical Engineering (ECCE-1), Florence, Italy, May 4-7, 1997, vol. 4, p. 2535-2538.

MCCULLOCK, W. S., PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. v.5, p.115-133, 1943.

MELEIRO, L.A.C., MACIEL FILHO, R. A self-tuning adaptive control applied to an industrial large scale ethanol production. *Computers and chemical engineering*, v.24, Issue: 2-7, pp. 925 – 930, july 15, 2000.

MENDES, E. F. F. *Projeto Evolucionario de Arquiteturas Neurais para Avaliacao de Credito Financeiro*. Sao Carlos: ICMSC, USP, 1997. Tese (Mestrado)

MINSKY, M., PAPERT, S. *Perceptrons*. Cambridge:MA: MIT Press, 1969.

NOVA, L. C. M. *Classificação de Imagens por Redes Neurais*. Rio de Janeiro: Instituto Militar de Engenharia, 1992. Tese (Mestrado)

PAPES, C. A. *Projeto e operação de reatores auxiliados por computador*. Campinas: Faculdade de Eng. Química, UNICAMP, 1999. Tese (Mestrado)

ROSENBLATT, F. The perceptrons: a probabilistic model for information storage and organization in the brain. *Psychological Review*. v.65, p.386-408, 1958.

RUMELHART, D. E., HINTON, G. E. WILLIAMS. R.J. Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*. MIT Press: Cambridge:MA. 1986a.

RUMELHART, D. E., HINTON, G. E. WILLIAMS. R.J. Learning representations by back- propagation of errors. *Nature*. v.323, p.533-536, 1986b.

SHULER, M.L., KARGI, F. *Bioprocess Engineering: Basic concepts*. Publication englewood cliffs:prentice-hall, 479p, 1992.

SILVA, F. L. H. *Modelagem. Simulação e controle de fermentação alcoólica continua extrativa*. Campinas: Faculdade de Eng. de Alimentos, UNICAMP, 1998. Tese

(Doutorado)

WIDROW, B. , HOFF, M. E. Adaptive switching circuits. *Proceedings of IRE WESCON Convention Record*, p.96-104, New York: IRE Press, 1960.

WILLSHAW, D. J., BUNEMAN, O. P. LONGUET-HIGGINS. H.C. Non-Holographic Associative Memory. *Nature*. v.222, p.960-962, 1969.

YAMAMOTO, C. I. , BABA, E. H., NASCIMENTO, C. A. O. Previsão de Comportamento de Reator de Síntese de Amônia via Rede Neural. *Em I Simpósio Brasileiro de Automação Inteligente (ISBAI)*, 1993, Rio Claro/SP, p. 48-57.

ZORZETTO, L.F.M., MACIEL FILHO, R., WOLF-MACIEL, M.R. Process modelling development through artificial neural networks and hybrid models. *Computers and chemical engineering*, v. 24, Issue: 2-7, p. 1355 – 1360, july 15, 2000.

Apêndice

Fundamentos das Redes Neurais artificiais

<i>A.1. COMO FUNCIONAM</i>	80
<i>A.2. ARQUITETURA</i>	82
<i>A.3. TOPOLOGIA</i>	83
<i>A.4. TREINAMENTO</i>	83
<i>A.5. ADALINE (ADAPTATIVE LINEAR ELEMENT)</i>	84
<i>A.6. REDES NEURAS MULTICAMADAS</i>	85
<i>A.7. BACKPROPAGATION</i>	86
<i>A.8. REGRA DELTA GENERALIZADA</i>	87
<i>A.9. PRINCIPAIS LIMITAÇÕES</i>	91

A.1 - Como Funcionam

Uma Rede Neural Artificial é composta de várias unidades de processamento simples que normalmente não fazem mais nada além de somar entradas de forma ponderada. Cada unidade (neurônio) trabalha com dados provindos de suas conexões.

Em 1943, Warren McCullock e Walter Pitts explicaram o funcionamento de um neurônio artificial, mais tarde denominado neurônio de McCullock-Pitts (Fig. 1). Este neurônio mostra, de forma simplificada, uma noção do funcionamento das Redes Neurais Artificiais .

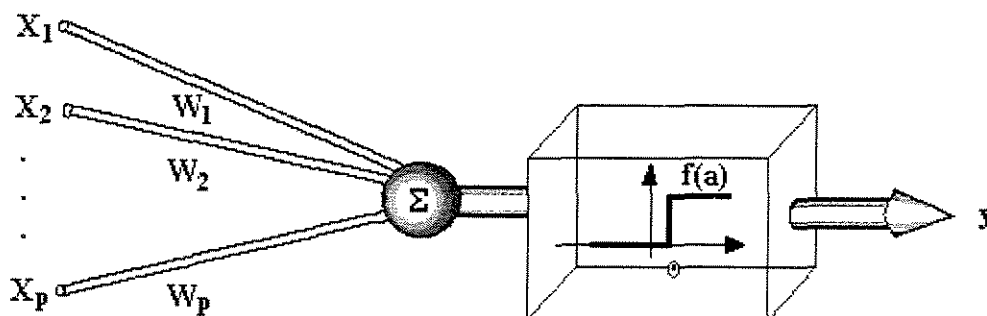


Figura 1 :Esquema de unidade McCullock – Pitts (MENDES, 1997).

Primeiramente os dados são apresentados à entrada (X_1, X_2, \dots, X_p). Cada sinal é, então, multiplicado por um número, ou peso (W_1, W_2, \dots, W_p), que indica sua influência na saída da unidade. A soma ponderada dos sinais é feita produzindo um nível de atividade. Ou seja:

$$a = W_1 * X_1 + W_2 * X_2 + \dots + W_p * X_p \quad \text{Equação 1}$$

onde a é a atividade.

Caso este nível de atividade exceda um determinado limite (“*threshold*”), a unidade de processamento (neurônio) produz uma certa resposta de saída.

MacCullock e Pitts trabalharam com sinais que assumem valores booleanos (0 ou 1) e θ como um limitador, assim a saída pode ser dada por:

$$\begin{cases} y = 1, & \text{se } a \geq \theta \text{ ou} \\ y = 0, & \text{se } a < \theta \end{cases}$$

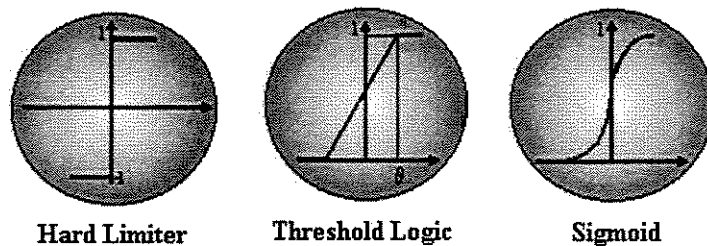
Esta é uma explicação simplificada. A Fig. 2 mostra um neurônio de forma mais detalhada.



Figura 2: Neurônio artificial (NOVA, 1992).

A função Soma (S) estabelece a forma de como combinar os valores de entrada de um neurônio B (Ex: Eq.1) com os pesos das conexões deste neurônio em um instante t gerando um valor, por exemplo, $S_B(t)$.

A função de Ativação estabelece como o valor $S_B(t)$ é combinado com o estado atual de ativação deste neurônio $a_B(t)$ para produzir um novo estado de ativação $a_B(t+1)$. As principais funções de ativação são:



A função sigmoide é a mais utilizada, sua expressão é:

$$f(a) = \frac{1}{(1 + e^{-a})}$$

Equação 2

O valor de saída do neurônio é obtido através da função Transferencia $T(a_B(t))$. A função de transferência mais utilizada é a função identidade, ou seja, o valor de saída do neurônio é igual ao valor de sua ativação.

As Redes Neurais Artificiais não se resumem a entrada e saída de dados, é preciso treinar a rede de modo que os pesos sejam ajustados. Uma rede neural é especificada, principalmente, pela sua topologia, pelas características dos nós e pelas regras de treinamento. A seguir, serão apresentados maiores detalhes.

A.2 - Arquitetura

A escolha da arquitetura da rede leva em conta o número de camadas, o número de neurônios, etc. A Fig. 3 mostra um tipo de arquitetura.

No início dos trabalhos com as redes neurais elas apresentavam uma única camada não sendo possível a formação de uma representação interna. O modelo fazia uma associação entre padrões de entrada e saída distintos à rede e, apesar das limitações, serviu como base para o desenvolvimento de outros modelos. Foi MINSKY, em 1969, que descobriu a limitação do PERCEPTRON (redes de uma camada) para problemas envolvendo classes linearmente separáveis. Para se resolver o problema foram criadas camadas intermediárias de neurônios.

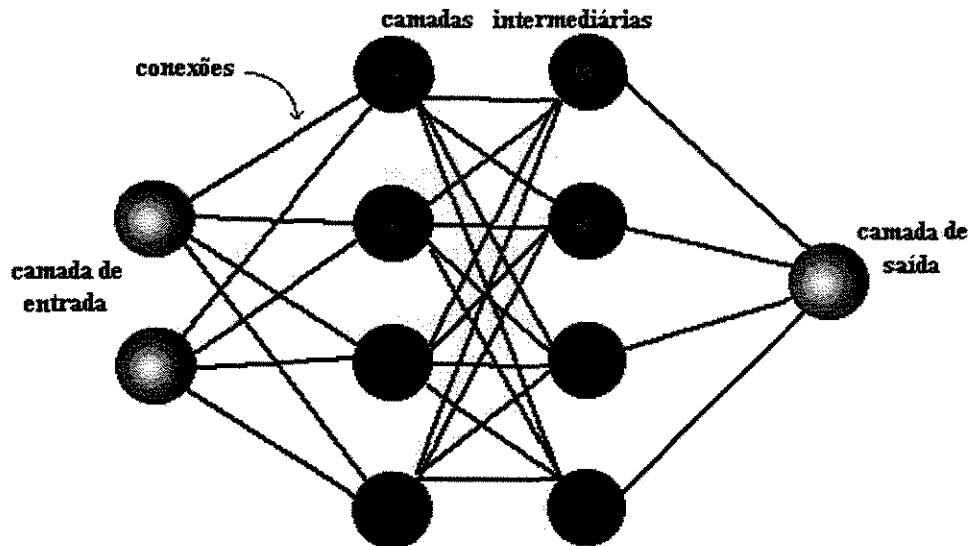


Figura 3: Organização em camadas (MENDES, 1997)

As camadas são classificadas em três tipos:

- **Camada de Entrada:** onde os padrões são apresentados à rede (“*Input*”);
- **Camadas Intermediárias ou Escondidas:** onde é feita a maior parte do processamento, através das conexões ponderadas. Podem ser consideradas como extratoras de características (“*Hidden*”);
- **Camada de Saída:** onde o resultado final é concluído e apresentado (“*Output*”).

Os pesos estão associadas às conexões. Um peso positivo pode representar uma entrada excitatória, enquanto que um peso negativo representa uma entrada inibitória.

A.3 - Topologia

A topologia se refere a como os neurônios são interligados. Os modos principais são :

- **Redes “FEED-FORWARD”**: os dados fluem da entrada (“*input*”) para a saída. Não existem conexões na direção contrária.
- **Redes RECORRENTES (“FEEDBACK”)**: existe um fluxo contrário entrada/saída de dados. Ao contrário da “*Feedforward*”, as propriedades dinâmicas da rede são importantes.

A.4 - Treinamento

Após escolhidas a arquitetura da rede, sua topologia e as funções utilizadas nos cálculos, pode-se dar início ao treinamento desta rede.

A rede “aprende” alterando os pesos associados às entradas de cada neurônio. A regra de aprendizagem descreve exatamente como os pesos devem ser alterados e é determinada pelo modelo da rede escolhido. O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

O aprendizado pode ser:

- **Supervisionado**: é apresentado à rede pares padrões (entrada/saída) de dados que funcionam como professores da rede.
- **Não-supervisionado (auto-organização)**: um conceito é incorporado à rede sem o conhecimento prévio da saída.

Dentre as regras de aprendizagem mais comuns podemos citar a Regra Delta, a regra Delta generalizada e a Regra Hebbiana. A regra delta simples funciona em redes de uma camada sendo seu entendimento um passo para se aplicar a regra delta generalizada, que é utilizada em redes multicamadas. A seguir uma explicação sobre a regra delta.

Ao se apresentar um padrão de entrada à rede com os pesos inicializados aleatoriamente ela produzirá uma saída (saída obtida). A regra delta prevê a modificação dos pesos da seguinte maneira:

$$W_{novo} = W_{anterior} + \beta * (d - O) * X$$

Equação 3

Onde d = resposta desejada (saída)

O = resposta Obtida (saída)

β = taxa de aprendizado. Constante positiva que influi na velocidade de aprendizado.

X = valor de entrada.

Modo de correção dos pesos na aprendizagem:

Denomina-se ciclo uma apresentação de todos os N pares (entrada/saída) do conjunto de treinamento no processo de aprendizado. A correção dos pesos num ciclo pode ser executado de dois modos:

- **Modo Padrão:** A correção dos pesos acontece a cada apresentação à rede de um exemplo do conjunto de treinamento. Cada correção de pesos baseia-se somente no erro do exemplo apresentado naquela iteração. Assim, em cada ciclo ocorrem N correções.
- **Modo Batch:** Apenas uma correção é feita por ciclo. Todos os exemplos do conjunto de treinamento são apresentados à rede, seu erro médio é calculado e a partir deste erro fazem-se as correções dos pesos.

A.5 - ADALINE (ADAPtative LINear Element)

Um exemplo clássico de redes “Feedforward” é a ADALINE. ADALINE é um modelo linear proposto por WIDROW em 1960.

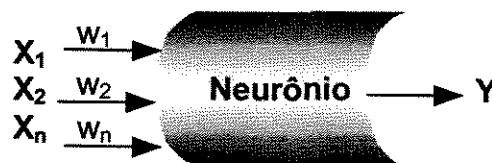


Figura 4: Esquema de neurônio simples.

A saída y do neurônio (Fig. 4) é uma simples combinação linear dos pesos e

valores de entrada:

$$y = \sum_{i=1}^n W_i * X_i$$

Equação 4

Deseja-se determinar um conjunto de pesos w^* tal que o erro quadrático E seja o mínimo possível. WIDROW propôs uma regra para se obter o ponto de mínimo conhecido como método do gradiente ou regra delta. A Equação 3 mostra a regra delta para correção pelo modo padrão, ou seja, a correção dos pesos será a cada apresentação à rede de um par de treinamento. O trabalho de WIDROW tem, segundo KOVÁCS (1996), relevância acadêmica e foi mais tarde generalizado para redes neurais com modelos mais elaborados.

A.6 - Redes Neurais Multicamadas:

A Fig. 3 mostra uma rede multicamada. Todas as rede utilizadas atualmente são multicamadas, variando apenas a forma das ligações (conexões), o número de camadas e a quantidade de neurônios existentes em cada uma.

Cada camada da rede, segundo MENDES (1997), tem uma função específica: a camada de saída recebe os estímulos da camada intermediária e constrói a resposta. As camadas intermediárias são extratoras de características, cujos pesos são uma codificação das características apresentadas nos padrões de entrada permitindo que a rede crie uma representação mais rica e complexa do problema.

Foi provado que são necessárias no máximo duas camadas intermediárias (com número suficiente de neurônios por camada) para se fazer qualquer mapeamento. É importante saber que o número de neurônios das camadas de entrada e saída são iguais ao número de variáveis de entrada e saída. O número de neurônios da camada intermediária influencia muito o funcionamento da rede. Poucos neurônios retardam o treinamento e muitos podem diminuir a sua capacidade de generalização

O algoritmo mais utilizado para treinamento de redes multicamadas é o de retropropagação ou “*backpropagation*”. O modelo de rede mais utilizado é o das redes “*Perceptron Multi-camadas*” (MLP) treinadas com o algoritmo de “*backpropagation*”.

A.7 – “Backpropagation”

O algoritmo de “backpropagation” utiliza uma modificação da regra delta: a regra delta generalizada. O algoritmo funciona da seguinte maneira:

1. Inicialização de todos os parâmetros (pesos, taxa de aprendizados, etc.);
2. Apresentação de um padrão à camada de entrada;
3. Cálculo da atividade em cada neurônio (a rede flui até a saída);
4. Obtenção de uma resposta na saída;
5. Comparação do valor obtido na saída com a resposta padrão calculando o erro;
6. Atualização de todos os pesos de forma a minimizar o erro (retropropagação do erro);

Novos padrões são apresentados à rede até se conseguir um erro mínimo sem perda da sua capacidade de generalização. A Fig. 5 mostra o esquema do algoritmo de “backpropagation”.

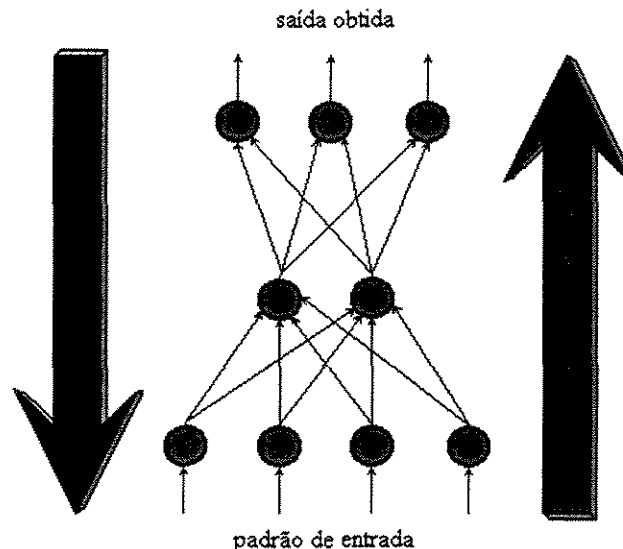


Figura 5: Esquema de treinamento com o algoritmo de “backpropagation”.

Após a rede estar devidamente treinada, ou seja, com um erro mínimo e uma capacidade de generalização máxima, ela pode ser utilizada como na Fig. 6. A rede é então integrada a um sistema do ambiente operacional da aplicação. Vale lembrar que o sistema

deve ser periodicamente monitorado para avaliar a performance da rede e fazer sua manutenção. Pode ocorrer, com o tempo, a necessidade de retreinamento da rede.

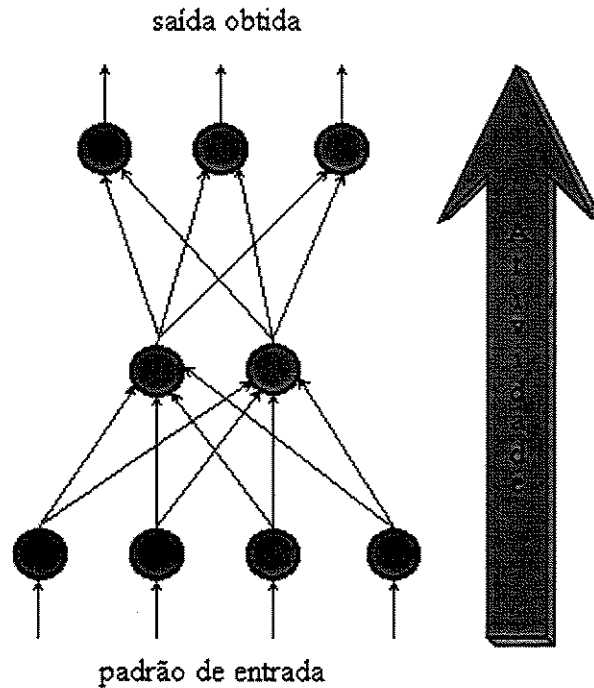


Figura 6: Utilização da MLP

A seguir será explicada a obtenção e o funcionamento da regra delta generalizada.

A.8 - Regra Delta Generalizada

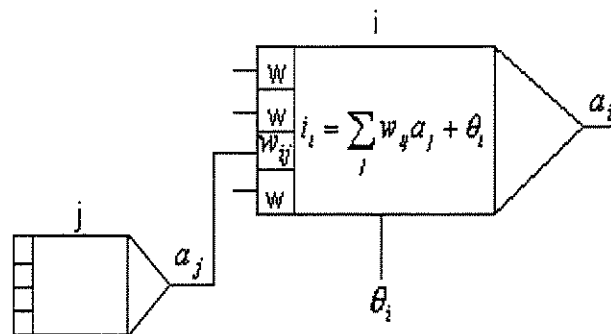


Figura 7: Conexão dos neurônios i e j

Consideremos dois neurônios i e j situados em camadas adjacentes como mostrado na Fig. 7. Estes neurônios fazem parte de uma rede sendo N_o a camada de saída da rede e p o vetor de entrada dos padrões. A saída do neurônio j a_j , é a entrada do neurônio i, cuja saída é

a_i . Sendo assim o peso relativo à conexão entre i e j é w_{ij} . θ_i é conhecido como BIAS e funciona como um neurônio extra acoplado a cada neurônio da rede com o objetivo de permitir uma convergência mais rápida do processo de treinamento. A fonte de treinamento deste “neurônio” é sempre 1, ao invés de um padrão de entrada.

A atividade do neurônio i num tempo t é dado por:

$$i_i^p(t) = \sum_j w_{ij}(t) a_j^p(t) + \theta_i(t) \quad \text{Equação 5}$$

Onde: $i_i^p(t)$ = atividade do neurônio i no tempo t para vetor de entrada de padrão P .

w_{ij} = peso relativo à conexão entre i e j .

$a_i^p(t)$ = saída do neurônio i no tempo t para padrão P .

$\theta_i(t)$ = valor de BIAS para neurônio i no tempo t .

A saída do neurônio i no tempo $t+1$ é dada por uma função da atividade no tempo t :

$$a_i^p(t+1) = f_i \left(\sum_j w_{ij}(t) a_j^p(t) + \theta_i(t) \right) \quad \text{Equação 6}$$

Como visto anteriormente, a função mais utilizada nestes casos é a sigmoideal (diferenciável e não decrescente), assim:

$$a_i^p = \frac{1}{1 + e^{-i_i^p}} \quad \text{Equação 7}$$

É deste modo que a rede multicamada calcula a atividade para todos os neurônios até chegar à camada de saída N_o , onde se tem a resposta obtida. A correção dos pesos é feita retropropagando o erro. A regra para correção é:

$$w_{novo} = w_{anterior} - \Delta w \quad \text{Equação 8}$$

Utilizando-se o método gradiente para minimizar o erro em relação ao peso o gradiente do erro deve ser calculado. O erro E^P é definido como o erro quadrático para o padrão p no neurônio de saída, e é dado por:

$$E^p = \frac{1}{2} \sum_{i=1}^{N_o} (d_i^p - a_i^p)^2 \quad \text{Equação 9}$$

Onde d_i^p é a resposta desejada (padrão) e a_i^p a resposta obtida da rede. O erro total é $E = \sum_p E^p$, e deverá ser minimizado. O Δw , para os neurônios i e j , é dado pela relação:

$$\Delta_p w_{ij} = -\beta \frac{\partial E^p}{\partial w_{ij}} \quad \text{Equação 10}$$

A relação mostrada na equação 10 origina-se do método do gradiente, onde se parte de um ponto arbitrário e caminha-se na superfície de erro em direção ao ponto de mínimo. A evolução é sempre no sentido oposto ao gradiente naquele ponto. A Fig. 8 mostra uma superfície de erro. β é uma constante e determina o tamanho do passo de correção.

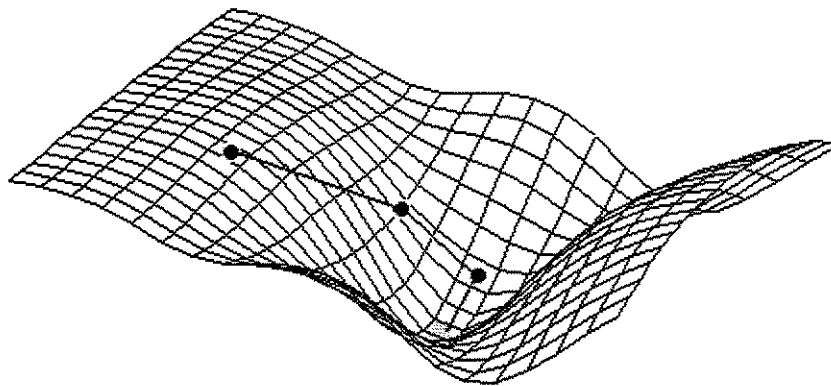


Figura 8: Superfície de erro

Queremos obter uma expressão para $\frac{\partial E^p}{\partial w_{ij}}$. Podemos aplicar a regra da

cadeia e obter:

$$\frac{\partial E^p}{\partial w_{ij}} = \frac{\partial E^p}{\partial i_i^p} \cdot \frac{\partial i_i^p}{\partial w_{ij}} \quad \text{Equação 11}$$

Derivando a equação 5 podemos ver que:

$$\frac{\partial i_i^p}{\partial w_{ij}} = a_j^p \quad \text{Equação 12}$$

Definindo: $\delta_i^p = -\frac{\partial E^p}{\partial i_i^p}$ e substituindo tudo na EQ.10, temos que:

$$\Delta_p w_{ij} = \beta \delta_i^p a_j^p \quad \text{Equação 13}$$

Agora deve-se obter uma relação para δ_i^p . Podemos escrever que:

$$\delta_i^p = -\frac{\partial E^p}{\partial i_i^p} = -\frac{\partial E^p}{\partial a_i^p} \cdot \frac{\partial a_i^p}{\partial i_i^p} \quad \text{Equação 14}$$

Mas pela equação 6 podemos ver que : $\frac{\partial a_i^p}{\partial i_i^p} = f'(i_i^p)$ sendo então uma simples derivada da função sigmoideal (se for a função utilizada). Para se determinar o primeiro termo da Eq. 14 $\frac{\partial E^p}{\partial a_i^p}$ deve-se considerar dois casos.

O primeiro caso considera que o neurônio i faz parte da camada de saída, assim segundo a definição de E^p (Eq.9), temos:

$$\frac{\partial E^p}{\partial a_i^p} = -(d_i^p - a_i^p) \quad \text{Equação 15}$$

Temos então que δ_i^p é dado por:

$$\delta_i^p = (d_i^p - a_i^p) f'(i_i^p) \quad \text{Equação 16}$$

No segundo caso o neurônio não faz parte da camada de saída e sim de uma intermediária qualquer. Assim, usando a regra da cadeia e h representando as camadas:

$$\frac{\partial E^p}{\partial a_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial a_h^p} \frac{\partial a_h^p}{\partial a_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial a_h^p} \frac{\partial}{\partial a_i^p} \sum_{k=1}^{N_h} w_{hk} a_k^p = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial a_h^p} w_{hi} = -\sum_{h=1}^{N_o} \delta_h^p w_{hi}$$

Com esta relação obtemos uma outra equação para δ_i^p :

$$\delta_i^p = f'(i_i^p) \sum_{h=1}^{N_o} \delta_h^p w_{hi} \quad \text{Equação 17}$$

As equações 8, 16 e 17 permite obter um processo recursivo para a atualização dos pesos, constituindo a regra delta generalizada para uma rede "feedforward" com função não linear.

A derivada da função sigmoïdal a ser substituída nas equações é dada por:

$$\begin{aligned} f'(i_i^p) &= \frac{\partial}{\partial i_i^p} \frac{1}{1 + e^{-i}} \\ &= \frac{1}{(1 + e^{-i})^2} (-e^{-i}) \\ &= \frac{1}{(1 + e^{-i})(1 + e^{-i})} e^{-i} \end{aligned}$$

$$f'(i_i^p) = a_i^p (1 - e_i^p)$$

Equação 18

A equação 18 deve ser substituída na equação onde a derivada aparece.

A.9 - PRINCIPAIS LIMITAÇÕES

As redes neurais são vistas como "caixas pretas", não se sabendo como a rede chega ao resultado final. Não há justificativas para respostas dos modelos.

O treinamento de redes neurais que utilizam "*backpropagation*" tende a ser muito lento. Deve-se utilizar computadores mais potentes para redes muito grandes. A arquitetura ideal da rede é difícil de se definir. Ela deve ser grande o suficiente para conseguir obter as representações necessárias, ao mesmo tempo pequena o suficiente para se ter um treinamento mais rápido.