UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA QUÍMICA

ÁREA DE CONCENTRAÇÃO

DESENVOLVIMENTO DE PROCESSOS QUÍMICOS

ESTRATÉGIA DE CONTROLE NÃO LINEAR BASEADA EM REDES NEURAIS ARTIFICIAIS COM APRENDIZAGEM "ON-LINE"

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

Aluna: Elis Regina Duarte

Orientador: Prof. Dr. Rubens Maciel Filho

Co-Orientador: Prof. Dr. Laércio Ender

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Campinas - SP

Março de 2004

UNIDADE BC
Nº CHAMADA
TUNICAMP
D85e
V
EX
TOMBO BC 59529
PROC. 16-0-117104
C
D X
PREÇO 11,00
DATA 14/09/2004
Nº CPD
BIB ID 321365

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

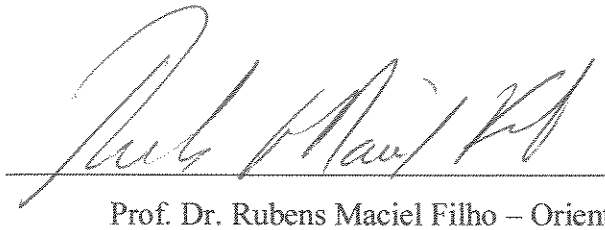
D85e

Duarte, Elis Regina
Estratégia de controle não linear baseada em redes neurais artificiais com aprendizagem "on-line" / Elis Regina Duarte. --Campinas, SP: [s.n.], 2004.

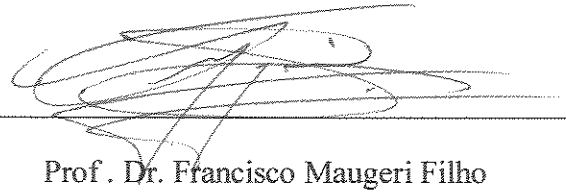
Orientadores: Rubens Maciel Filho e Laércio Ender.
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Teoria do controle não-linear. 2. Redes neurais (Computação). 3. Controle de processos químicos. I. Maciel Filho, Rubens. II. Ender, Laércio. III. Universidade Estadual de Campinas. Faculdade de Engenharia Química. IV. Título.

Dissertação de Mestrado defendida por Elis Regina Duarte e aprovada em 26 de março 2004 pela banca examinadora constituída pelos doutores:



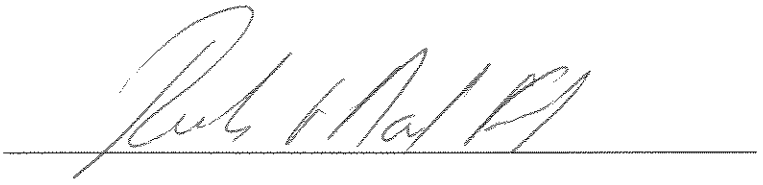
Prof. Dr. Rubens Maciel Filho – Orientador



Prof . Dr. Francisco Maugeri Filho

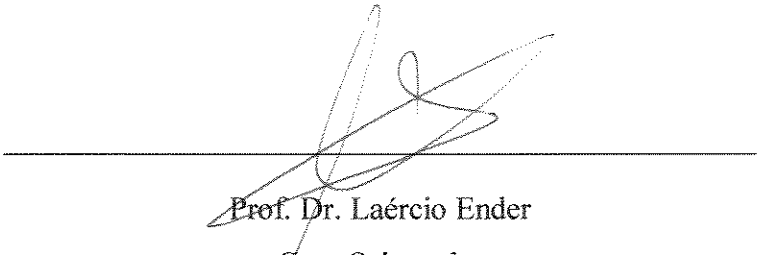
Prof. Dr. Roberto de Campos Giordano

Este exemplar corresponde à versão final da Dissertação de Mestrado em Engenharia Química



Prof. Dr. Rubens Maciel Filho

Orientador



Prof. Dr. Laércio Ender

Co – Orientador

Dedico esta dissertação de mestrado aos meus pais Alicio e Catarina e ao meu irmão Edielson.

AGRADECIMENTOS

Inicialmente agradeço a Deus por esta oportunidade de evolução contínua e principalmente pela presença constante da luz angelical de meu anjo protetor e guia Chavakiah.

Aos meus pais e meu irmão por todo carinho, amor e compreensão nos momentos de alegria e também de dificuldades.

Ao Raphael por todo carinho e companheirismo.

Aos meus orientadores Rubens Maciel Filho e Laércio Ender por toda orientação e incentivo.

Em especial a duas pessoas que me mostraram esse caminho de conhecimento científico, Karen e Chivanga por toda amizade e incentivo.

A todos meus antigos e novos amigos por toda amizade, em especial aos colegas da UNICAMP, LOPCA, FURB e LFC.

E por fim a todas pessoas que acreditaram em meu potencial e sempre me incentivaram nessa incessante busca de conhecimento.

A CAPES pelo apoio financeiro.

Devemos ser magos buscando incessantemente o oculto, desvendando os mistérios e transformando a ignorância em sabedoria.

RESUMO

Nos dias atuais com a globalização, competitividade do mercado e preocupações ambientais se faz necessários produtos com alta qualidade, baixo custo e baixa emissão de poluentes, sendo o controle de processos a principal maneira de se obter estes resultados. Por isso o grande avanço de pesquisas em busca de novas metodologias e concepções sobre desenvolvimento e controle para processos de todas as áreas.

O objetivo deste trabalho é propor o aperfeiçoamento da estratégia de controle CoNDEG, (Controle Neural Direto Baseado no Erro Global), desenvolvida por ENDER (2002). A estratégia está baseada em redes neurais artificiais, com aprendizagem em tempo real, de acordo com as alterações que ocorrem no processo. Desta forma o controlador apresentará uma performance mais apropriada, quando comparado com o controlador que não apresenta esta atualização ao longo do tempo.

Para desenvolvimento desta dissertação estudou-se o CoNDEG e foram propostas modificações na forma de ajuste do controlador e análise de diferentes configurações para aprendizagem *on-line*. A estratégia proposta CoNDEG Modificado foi implementada em linguagem de programação Fortran 90.

Foram utilizados dois processos não lineares para avaliar a estratégia proposta, um reator catalítico de leito fixo (TOLEDO, 1999) e um processo fermentativo extrativo para produção de etanol (COSTA, 2001).

As simulações em malha fechada realizadas utilizando o programa CoNDEG Modificado para ambos casos de estudo mostraram que o algoritmo de controle não linear proposto é eficiente e robusto, pois apresentaram bons resultados em problemas dos tipos servo e regulador.

Palavras – Chaves: Estratégia de Controle, Redes Neurais Artificiais, Aprendizagem "On-line"

ABSTRACT

In now days with the globalization, market competitivity and environmental preoccupation, there is a real need for products with high quality, low cost and low emission of pollutants. The process control is the main form to reach such objectives. That is the reason for the great advance on researches in the investigation of new methodologies and conceptions about process development and process control in several areas.

The objective of the present work is to propose the improvement of CoNDEG strategy control (Direct Neural Control based on Global Error), developed by ENDER (2002). The strategy is based in artificial neural networks, with on line learning, according to modifications that occur in the process. Hence the controller based on such approach will present better performance than a controller without on-line upgrade.

In order to develop the present work the CoNDEG was studied and it was proposed some modifications on the way the controller is adjusted as well as it was carried out the analysis of different configurations for the on-line learning. The proposed modified strategy CoNDEG was implemented in Fortran 90 programming language.

It was used two non-linear processes to evaluate the proposed strategy, to know, a fixed bed catalytic reactor (TOLEDO, 1999) and an extractive fermentative process for the production of ethanol (COSTA, 2001).

The results obtained from closed-loop simulations using the modified CoNDEG program for both studied cases showed that the proposed non-linear control algorithm is efficient and robust for servo and regulatory applications.

Key - Words: Control Strategy, Artificial Neural Network, On-line Learning.

SUMÁRIO

AGRADECIMENTOS ..... *IX*
RESUMO ..... *XIII*
ABSTRACT ..... *XV*
SUMÁRIO ..... *XVII*
LISTA DE FIGURAS ..... *XIX*
NOMENCLATURA ..... *XXIII*
1. INTRODUÇÃO *1*
2. REDES NEURAIS ARTIFICIAIS *5*
2.1 Introdução 5
2.2 Inteligência Artificial 5
2.3 História da arte 7
2.4 Redes Neurais 9
2.4.1 Neurônio Biológico 9
2.4.2 Neurônio Artificial 12
2.5 Arquitetura das redes neurais artificiais 16
2.6 Classificação das redes 17
2.6.1 Aplicações na Área de Controle 17
2.6.2 Fluxo de Informações 18
2.6.3 Método de Identificação 19
2.7 Processo de aprendizagem 19
2.7.1 Aprendizado Supervisionado 19
2.7.2 Aprendizado Não Supervisionado 20
2.8 Algoritmo de Aprendizagem *Backpropagation* 21
2.9 Conclusões 27
3. ESTRATÉGIA DE CONTROLE *29*
3.1 Introdução 29
3.2 Controle de Processos 29
3.3 Controlador Neural – Dinâmica Inversa 30
3.4 Modelo Neural do Processo – Identificação 32
3.5 Estratégia de Controle – CoNDEG Modificado 33
3.5.1 Ajuste do Controlador 34
3.6 Aprendizagem *On-Line* e *Off-Line* das *RNA's* 37
3.7 Configurações de Aprendizagem *On-line* 40
3.7.1 Aprendizagem *On-line* - Configuração 1 40

3.7.2 Aprendizagem *On-line* - Configuração 2 41
3.7.3 Aprendizagem *On-line* - Configuração 3 42
3.8 Conclusões 43
4. CASO 1 - PROCESSO EXTRATATIVO DE FERMENTAÇÃO ALCOÓLICA CONTÍNUA 45
4. 1 Introdução 45
4.2 Processo 45
4.3 Modelagem Matemática 47
4.4 Comportamento Estacionário 54
4.5 Comportamento Dinâmico 54
4.6 Análise de Sensibilidade dos Parâmetros de Processo – Planejamento Fatorial 58
4.7 Arquitetura e Dados Para Treinamento Das *RNA's* 59
4.8 Resultados Para o Caso de Estudo 1 – Processo Fermentativo 60
4.8.1 Escolha da Estrutura da Rede Usada na Estratégia 63
4.8.2 Resultados Comparativos das Três Configurações de Aprendizagem 70
4.6.2.1 Problema Servo para Variável S 72
4.6.2.2 Problema Servo para Variável P 76
4.6.2.3 Problema Servo para Variável X 80
7.8.3 Análise da Estratégia Proposta 85
4.9 Conclusões 88
5. CASO 2 – REATOR CATALÍTICO DE LEITO FIXO 88
5. 1 Introdução 89
5.2 Processo 89
5.3 Modelagem Matemática 91
5.4 Comportamento Estacionário 93
5.5 Comportamento Dinâmico 95
5.6 Análise de Sensibilidade dos Parâmetros de Processo – Planejamento Fatorial 98
5.7 Arquitetura e Dados Para Treinamento Das *RNA's* 104
5.8 Resultados Para o Caso de Estudo 2 – Reator 105
5.9 Conclusões 109
6. CONCLUSÕES E SUGESTÕES 111
REFERÊNCIAS BIBLIOGRÁFICAS 115

LISTA DE FIGURAS

Figura 2.1 Neurônio biológico (IANET,2003) 11
Figura 2.2 Neurônio Artificial 13
Figura 2.3 Rede Neural Artificial (Tipo Feedforward) 13
Figura 2.4 Anatomia do i-ésimo neurônio artificial da k-ésima camada 14
Figura 2.5 Funções de ativação: (a) Função Sigmoidal; (b) Função Tangente Hiperbólica 15
Figura 2.6 Estruturas de Rede Feedforward (a,b,c) e Feedback (d,e) 19
Figura 2. 7 Aprendizado Supervisionado 20
Figura 2.8 Aprendizado Não Supervisionado 20
Figura 2.9 Algoritmo Backpropagation 21
Figura 3.1 Arquitetura feedforward multivariável (Modelo Inverso) 32
Figura 3.2 Estratégia de Controle CoNDEG Modificado 34
Figura 3.3 Etapa 1 - Aprendizagem Off-line (Ajuste de Pesos com Dados Padrões) 38
Figura 3.4 Etapa 2 – Aprendizagem On-line (Ajuste de Peso em Tempo Real) 39
Figura 3.5 Esquema de Aprendizagem On-line para Configuração 1 41
Figura 3.6 Esquema de Aprendizagem On-line para Configuração 2 42
Figura 3.7 Esquema de Aprendizagem On-line para Configuração 3 43
Figura 4.1 Representação esquemática da planta de fermentação alcoólica extrativa 46
Figura 4.2 Perfil da Concentração de Substrato (S) para Perturbações Positivas 55
Figura 4.3 Perfil da Concentração de Produto(P) para Perturbações Positivas 56
Figura 4.4 Perfil da Concentração de Microorganismos (X) para Perturbações Positivas .56
Figura 4.5 Perfil da Concentração de Substrato (S) para Perturbações Negativas 57
Figura 4.6 Perfil da Concentração de Produto(P) para Perturbações Negativas 57
Figura 4.7 Perfil da Concentração de Microorganismos (X) para Perturbações Negativas . 58
Figura 4.8 Perturbações usadas para Treinamento das Redes 60
Figura 4.9. Perturbações No Processo Fermentativo para Variável So 61
Figura 4.10. Perturbações No Processo Fermentativo para Variável To 62
Figura 4.11.Resposta para Variável X para Perturbações em So e To (Malha Aberta) 62
Figura 4.12 Redes Treinadas com Passado Três 64
Figura 4.13 Redes Treinadas com Passado Cinco 64

Figura 4.14 Comparação das Redes com Passados Três e Cinco 65
Figura 4.15 Análise da Função de Ativação para Redes do Controlador e Processo 66
Figura 4.16 Teste para Memória das Redes para Variável S 67
Figura 4.17 Teste para Memória das Redes para Variável P 67
Figura 4.18 Teste para Memória das Redes para Variável X 68
Figura 4.19 Resultado Comparativo Resposta Real e Rede Neural para variável S 69
Figura 4.20 Resultado Comparativo Resposta Real e Rede Neural para variável P 69
Figura 4.21 Resultado Comparativo Resposta Real e Rede Neural para variável X 70
Figura 4.22 Variáveis Controladas para Problema Servo em S 72
Figura 4.23 Variáveis Manipuladas Para Problema Servo em S 73
Figura 4.24 Redes do Controlador Para Configurações 1, 2 e 3 74
Figura 4.25 Redes do Processo Para Configurações 1, 2 e 3 75
Figura 4.26 Variáveis Controladas Para Problema Servo em P 76
Figura 4.27 Variáveis Manipuladas Para Problema Servo em P 77
Figura 4.28 Redes do Controlador Para Configurações 1, 2 e 3 78
Figura 4.29 Redes do Processo Para Configurações 1, 2 e 3 79
Figura 4.30 Variáveis Controladas Para Problema Servo em X 80
Figura 4.31 Variáveis Manipuladas Para Problema Servo em X 81
Figura 4.32 Redes do Controlador Para Configurações 1, 2 e 3 82
Figura 4.33 Redes do Processo Para Configurações 1, 2 e 3 83
Figura 4.34 Comparativo Configuração 1 (25-25) e Configuração3 (25-200) 84
Figura 4.35 Comparação On-line e Off-line para problema Servo em S 85
Figura 4.36 Comparação On-line e Off-line para problema Servo em X 86
Figura 4.37 Comparação On-line e Off-line para problema Servo em P 86
Figura 4.38 Resultado com Perturbação Estocástica para Problema Servo em S 87
Figura 4.39 Resultado com Perturbação Estocástica para Problema Servo em P 87
Figura 4.40 Resultado com Perturbação Estocástica para Problema Servo em X 88
Figura 5.1 Esquema do Caso de Estudo 1 – Reator Catalítico de Leito Fixo 90
Figura 5.2 Perfil da Concentração de Acetaldeído ao Longo do Reator 94
Figura 5.3 Perfil da Temperatura ao Longo do Reator 95
Figura 5.4 Perfil da Temperatura (T1) para as Perturbações Positivas 96
Figura 5.5 Perfil da Temperatura (T3) para as Perturbações Positivas 97
Figura 5.6 Perfil da Temperatura (T1) para as Perturbações Negativas 97

Figura 5.7 Perfil da Temperatura (T3) para as Perturbações Negativas..... 98
Figura 5.8 Superfícies de resposta e diagramas de contorno para a temperatura (T1)..... 101
Figura 5.9 Superfícies de resposta e diagramas de contorno para a temperatura (T3)..... 103
Figura 5.10 Perturbações nas Variáveis Manipuladas para Treinamento das RNA's..... 104
Figura 5.11 Resultado das Variáveis T(1) e T(3) para Problema Servo..... 106
Figura 5.12 Resultado das Variáveis T(1) e T(3) com Perturbação Estocástica 106
Figura 5.13 Resultado para Variável R com Perturbação Estocástica 107
Figura 5.14 Resultado para Variável Gmo com Perturbação Estocástica 107
Figura 5.15 Resultado das Variáveis T(1) e T(3) para Problema Regulador em Tfo..... 108
Figura 5.16 Resultado das Variáveis T(1) para Problema Regulador em Tro 109
Figura 5.17 Resultado das Variáveis T(3) para Problema Regulador em Tro 109

NOMENCLATURA

REDES NEURAIS ARTIFICIAIS

Letras Latinas

d_j Resposta alvo de um neurônio j da camada de saída

N_p Número de padrões

n_k Número de neurônios da camada k

n Número de camadas ocultas

$S_i^{(k)}$ Função que indica a soma ponderada de todas as entradas de um neurônio i da camada k

$X_i^{(k)}$ Saída (ou ativação) de um neurônio i da camada k

$w_{i,j}^{(k)}$ Peso da conexão entre o neurônio j da camada $(k\text{-}1)$ e o neurônio i da camada k

Letras Gregas

ε Função objetiva para treinamento

$\delta_i^{(k)}$ Erro equivalente na aprendizagem

η Taxa de aprendizagem

β Conjunto de padrões passados usados na rede neural

ESTRATÉGIA DE CONTROLE

Letras Latinas

$\hat{e}(k)$ Vetor de erro global estimado da malha fechada, para o instante de tempo k

$u(i,j)$ Ação de controle da malha i no tempo j

$U(k)$ Vetor de ações de controle

$Y(k)$ Vetor de respostas da malha fechada

$\hat{Y}(k)$ Vetor de predições da malha fechada

$y_r(i)$ Referência desejada para a malha de controle i

$Y_r(k)$ Vetor de referências da malha fechada

$y_w(i,j)$ Trajetória de referência da malha i no tempo j

$Y_w(k)$ Vetor de trajetória de referência da malha fechada

Letras Gregas

α_i Parâmetro de ajuste da trajetória de referência para a malha i

$\Delta u(i,j)$ Incremento na ação de controle i no tempo j

δu_j Erro estimado na ação de controle da malha fechada j

δy_j Erro estimado na resposta da malha fechada j;

$\delta U\ (k)$ Vetor de erros estimados nas ações de controle da malha fechada

$\delta Y\ (k)$	Vetor de erros estimados nas respostas da malha fechada
$\lambda(i)$	Penalização na ação de controle da malha *i*

REATOR CATALÍTICO DE LEITO FIXO

Letras Latinas

Bih	Número de Biot
C_{pf},	Capacidade calorífica para o gás (Kcal/KgK)
C_{ps}	Capacidade calorífica para o catalisador (Kcal/KgK)
C_{PR}	Capacidade calorífica para o fluido refrigerante (Kcal/KgK)
D_{ef}	Difusividade radial efetiva (m/h)
f	Fator de fricção
GM_o	Fluxo mássico de alimentação (Kg/m^2h)
h_W	Coeficiente de filme das paredes internas ($Kcal/hm^2K$)
L	Comprimento do reator (m)
P	Pressão Total do reator
P_{ref}	Pressão de referência (atm)
PM	Peso molecular médio (Kg/Kmol)
r	Comprimento radial do reator (m)
R	Relação ar/etanol
T_{fo}	Temperatura de alimentação

T_W	Temperatura da parede do reator
T_R	Temperatura do fluido refrigerante
T_{ref}	Temperatura de referência
u_R	Velocidade do fluido refrigerante (m/h)
x	Conversão
z	Comprimento axial do reator (m)

Letras Gregas

ε	Porosidade
λ_{ef}	Condutividade radial efetiva (Kcal/m h K)
ΔH_R	Entalpia molar (Kcal/Kmol)
ρ_f	Massa específica do fluido reagente (Kg/m^3)
ρ_s	Massa específica do catalisador (Kg/m^3)
ρ_R	Massa específica do fluido refrigerante (Kg/m^3)

PROCESSO EXTRATIVO DE FERMENTAÇÃO ALCOÓLICA

Letras Latinas

F_W	Vazão de água (m^3/h)
F_p	Vazão de purga (m^3/h)
F_{pl}	Vazão de alimentação da purga (m^3/h)

F	Vazão de alimentação do fermentador (m^3/h)
F_w	Vazão de água de diluição (m^3/h)
F_R	Vazão de reciclo de células (m^3/h)
K_{eq}	Equilíbrio líquido - vapor
P	Concentração de etanol (kg/m^3)
$R_{microrg}$	Razão de reciclo de microorganismos
R_{flash}	Razão de reciclo do tanque flash
S	Concentração de substrato no fermentador (kg/m^3)
S_0	Concentração de substrato na corrente de alimentação (kg/m^3)
V	Volume do fermentador (m^3)
X	Concentração de microorganismos (kg/m^3)
X_v	Concentração de células viáveis (kg/m^3)
X_d	Concentração de células mortas (kg/m^3)
X_t	Concentração de células totais (kg/m^3)
P	Concentração de produto (kg/m^3)
ΔH	Calor de reação ($2{,}167.10^3$ J/kg de ART)
Cp	Calor específico (J/kgK)
T	Temperatura (°C)

Letras Gregas

τ	Tempo de residência do fermentador

ρ_m	Densidade do meio fermentativo (1000 kg/m^3)

Subscritos

E	Corrente que segue para tanque flash
C	Corrente que segue para cuba
R	Corrente de reciclo
F	Corrente de alimentação

INTRODUÇÃO

Os processos químicos na sua maioria são caracterizados por apresentarem comportamento dinâmico complexo, resultado da presença de não linearidades, decorrentes de termos não-lineares e interações entre os vários fenômenos que normalmente ocorrem de forma simultânea. Alguns exemplos típicos são os reatores químicos e bioquímicos do tipo tanque agitado, tubulares (inclusive os multifásicos), de leito fixo e colunas de destilação convencional e azeotrópica. É importante ressaltar que mesmo que um processo seja linear, a existência de um sistema de controle (válvulas de controle, saturação de atuadores, entre outros) pode levar o processo a ter comportamento não linear.

Uma etapa importante no desenvolvimento, implementação e análise de controladores se refere a representação adequada do comportamento do sistema. Dentro deste contexto, os modelos dinâmicos de processos são ferramentas básicas para diversos propósitos tais como simulação, controle, predição de estados, identificação de sistemas, entre outros. Os modelos dinâmicos lineares, devido à sua versatilidade e simplicidade, têm sido muito utilizados, porém não são adequados para descrever processos com multiplicidade de estados estacionários e as não linearidades inerentes à maioria dos processos da Engenharia Química.

Como efeito da globalização surge a necessidade de produtos com menor custo e maior rendimento, promovendo um avanço tecnológico significativo em diversas áreas. O controle e otimização de processos são ferramentas que auxiliam nessa nova concepção, trazendo muitas melhorias, como produtos com maior qualidade, menor custo e um melhor aproveitamento dos recursos.

A literatura tem mostrado que é necessário o desenvolvimento de algoritmos de controle baseados nos conceitos de controle avançado para se obter a operação do sistema a altos níveis de desempenho, conciliado a condições seguras. Com esta crescente necessidade de se desenvolver controladores de processos com desempenho cada vez melhores, emerge a necessidade do desenvolvimento de estratégias não lineares para atuarem de forma mais efetiva nos processos não lineares, as quais deverão ter a capacidade de adaptação em função das perturbações externas que podem ocorrer.

Dentro deste contexto, emergiram as estratégias de controle usando otimizações em tempo real, onde diversas ferramentas são empregadas objetivando desenvolver técnicas para o controle de processos considerados com dinâmica complexa.

As Redes Neurais Artificiais (*RNA's*) tem se tornado recentemente foco de muita atenção principalmente devido à sua capacidade de aprendizado, à ampla gama de aplicabilidade e à facilidade com que armazena conhecimento experimental e o torna disponível para o uso em aplicações em engenharia de processos.

Muitas pesquisas foram desenvolvidas na última década na identificação e controle de processos aplicando as redes neurais artificiais ou *RNA's* por estas características e principalmente pela grande capacidade de representação de modelos não lineares.

Uma *RNA* é um sistema com muitos elementos de processamento operando em paralelo, tendo como função definida pela topologia da rede, peso das conexões e operações efetuadas nos elementos de processamento. Representam um novo paradigma metodológico no campo da inteligência artificial, ou seja, no desenvolvimento de sistemas computacionais capazes de imitar tarefas intelectuais complexas, tais como a resolução de problemas, o reconhecimento e classificação de padrões.

Estes modelos são simplificados, e às vezes não plausíveis no que concerne aos seus correspondentes biológicos. Alguns tipos de redes apresentam características da cognição humana que as tornam bastante populares, tais como a aprendizagem através de exemplos e a capacidade de abstrair características relevantes a partir de dados abundantes. Isto é, quando dados de entrada e saída de um certo sistema são fornecidos a uma rede específica, ela é capaz de reproduzir o comportamento deste sistema, depois de ajustar as conexões entre os neurônios artificiais que a compõem (DE SOUZA JR, 1993).

Por isso as *RNA's* estão sendo empregadas no desenvolvimento de técnicas de controle para processos com dinâmica complexa e otimizações em tempo real.

As *RNA's* apresentam as seguintes características:

- Capacidade de processamento paralelo;

- Capacidade de representação distribuída, aumentando a tolerância a falhas de componentes, ou seja, mesmo que alguns neurônios sejam removidos o desempenho global da rede pode ser pouco alterado.
- Possibilidade de implementação utilizando "hardware" analógico;
- Grande capacidade de adaptação;
- Flexibilidade no atendimento a critérios de generalização;
- Possibilidade de incorporação de poderosas ferramentas algébricas, facilitando a análise e interpretação de resultados.

No entanto estas características estão presentes também em outros modelos, inclusive os lineares. Mas quando projetado de maneira adequada um modelo não linear é potencialmente melhor que um linear no seu desempenho (VAN ZUBEN, 1996).

Uma das áreas em que essas características das redes podem ser exploradas é a da Engenharia Química. Na indústria de processos químicos, muitas vezes a mesma empresa possui diversas plantas com dinâmica complexa, devido a características tais como as não linearidades, tempos mortos e ruídos.

Segundo ÅSTRÖM e McAVOY (1992), a modelagem usando *RNA's* pode ser mais genérica e economicamente vantajosa que aquela baseada em princípios básicos de conservação, uma vez que as redes empregam apenas dados históricos da planta e são, portanto, "facilmente transportadas de uma planta para outra". Ainda segundo os mesmos autores, alguns engenheiros de processo consideram as redes neurais mais fáceis de usar que os métodos tradicionais. Particularmente, para o controle desses processos, as *RNA's* vêm ao encontro às exigências de um controle mais rigoroso, na presença de incertezas quanto ao modelo da planta.

O objetivo geral desta dissertação é o aperfeiçoamento da estratégia desenvolvida por ENDER (2002), avaliando a forma de aprendizagem e teste das redes utilizadas, que possuem aprendizagem *on-line* e implementação da estratégia modificada. Para avaliar a estratégia foram considerados dois casos de estudos representados por processos não lineares; a saber, um processo de fermentação alcoólica extrativa para o caso de estudo 2

(COSTA *et al.*, 2001) e um reator catalítico de leito fixo para o caso de estudo 1 (TOLEDO, 1999).

A estratégia aborda a aplicação de redes neurais como modelo dinâmico e controlador. Em todos os casos utilizou-se aprendizagem ao longo do tempo (*on-line*) das redes neurais, a fim de evitar problemas de execução em tempo real, pois este tipo de treinamento permite a adaptação das redes, treinando-as em situações não abordadas durante a etapa de aprendizagem a partir de dados históricos (*off-line*).

O Capítulo 2 apresenta a história da arte das redes neurais, trazendo uma revisão da literatura do assunto e sua aplicação na área de controle de processos. Após esta abordagem inicial, se enfatiza as características das redes, como classificação, arquitetura, formas de aprendizagem e os métodos de treinamento e a definição da estrutura das redes utilizadas no desenvolvimento desta dissertação.

O Capítulo 3 traz a descrição da estratégia utilizada e um comparativo das alterações propostas para estratégia e seus respectivos resultados.

O Capítulo 4 explora o caso de estudo 1, representado por um processo de produção de etanol por via fermentativa (COSTA *et al*, 2001) trazendo uma breve descrição do processo, suas variáveis de interesse e faixa de operação e para finalizar os resultados obtidos com a estratégia proposta.

No Capítulo 5 apresenta-se o caso de estudo 2, representado por um reator de leito fixo (TOLEDO, 1999) trazendo uma breve descrição do processo, a definição das variáveis de interesse e suas faixas de operação e os resultados obtidos para este processo com a estratégia proposta.

As conclusões gerais sobre os resultados obtidos na dissertação e algumas sugestões para o desenvolvimento de trabalhos futuros são apresentadas no capítulo 6.

REDES NEURAIS ARTIFICIAIS

2.1 INTRODUÇÃO

O capítulo 2 apresenta uma introdução sobre redes neurais artificiais (RNA's) partindo da sua evolução dentro do ramo da inteligência artificial até sua aplicação na área de controle de processos. O objetivo principal deste capítulo é apresentar uma base teórica necessária para o entendimento sobre RNA's além de definir a topologia e algoritmo de aprendizagem das redes que serão utilizadas na estratégia de controle.

2.2 INTELIGÊNCIA ARTIFICIAL

A Inteligência artificial (IA) é um ramo de pesquisa multidisciplinar, resultado de intensa cooperação entre engenharia, matemática, física, neurologia, psicologia e filosofia. Reune técnicas que buscam integrar aos sistemas de processamento características típicas do comportamento inteligente. Segundo STOCK (1989), Inteligência artificial é uma extensão da computação que trata da solução de problemas através de técnicas não algorítmicas.

Um estado comum ao ser humano é o de, repentinamente e a partir de um estímulo incidental, ter a sua mente invadida por memórias de situações vividas e/ou aprendidas. A memória é uma das habilidades do cérebro humano. A busca do entendimento dessa e de outras habilidades, de sua influência no comportamento humano e da própria estrutura do cérebro tem atraído a atenção não só da neurobiologia e da psicologia, mas também da matemática aplicada e da engenharia.

Aos engenheiros interessa criar/aperfeiçoar tecnologia a partir desse entendimento. Construir computadores que apresentem características mais próximas às inerentes ao cérebro humano, proporcionando maior interação entre a máquina e o seu usuário, substituindo o homem na realização de tarefas repetitivas, preservando o conhecimento especialista (DE SOUZA JR, 1993).

A arquitetura clássica de computadores valoriza a manipulação rápida e eficiente de símbolos numéricos e lógicos e boa parte das pesquisas de IA concentra-se no estudo da lógica e de estrutura complexas de dados capazes de representar o conhecimento em uma forma tratável pelo computador. A manipulação conveniente de fórmulas da lógica de predicados de primeira ordem obteve sucesso na questão de demonstração automática de teoremas e permitiu a definição das chamadas linguagens lógicas. Existem duas linhas de abordagens: abordagem simbólica e a abordagem conexionista, que surgiu em 1980 (ROMARIZ, 1995).

De acordo com DE SOUZA JR (1993), nem todas as informações que a mentes humana capta, no entanto, são facilmente representadas por regras. Muitas vezes as pessoas resolvem problemas – por exemplo, conseguem reconhecer padrões de características relevantes em um conjunto de dados, assim como criar modelos que são adaptativos e separar dados do ruído (PRETT e GARCIA, 1988) – como resolvem estes problemas, as pessoas podem não saber explicar e simplesmente dizer que a solução vem naturalmente (GIARRATANO *et al.*, 1990).

De modo a reproduzir esses mecanismos de entendimento ou cognição, modelos matemáticos simplificados que tentam imitar as redes altamente interconectadas de neurônios que compõem o cérebro vem sendo empregados e classificados como processos sub-simbólicos, que são modelos matemáticos de sistemas biológicos neurais teorizados, que fazem parte da abordagem conexionista.

A idéia desta abordagem é a possibilidade de simulação de comportamentos inteligentes através de modelos baseados na estrutura e funcionamento do cérebro humano. Computações complexas podem ser obtidas pela combinação de muitos processadores simples altamente interconectados. Este paradigma segue uma inspiração biológica, ao aproveitar idéias do conhecimento que se tem da estrutura de cérebro e de modelos para os neurônios (ROMARIZ, 1995). Esses modelos são chamados de redes neural artificiais (*RNA's*) ou *Artificial Neural Networks (ANN's)*, modelos conexionistas, sistemas adaptativos, redes adaptativas, neurocomputadores ou processadores de distribuição paralela (SIMPSON, 1990).

As *RNA's* são apresentadas como modelos matemáticos cujas características gerais em termos estruturais (neurônios artificiais interligados entre si formando uma rede) e de comportamento (aprendizagem, reconhecimento de características relevantes em ambientes com ruído, operações em paralelo) guardam uma vaga similaridade com o cérebro humano. Dessa maneira, as reais potencialidades das *RNA's* podem ser evidenciadas, ao mesmo tempo em que expectativas errôneas são refutadas (DE SOUZA JR,1993).

A operação de uma rede neural artificial é constituída por três etapas:

- Etapa de Aprendizagem – Ajuste dos pesos sinápticos;
- Etapa de Representação – Estimativa para valores de saída;
- Etapa de Generalização – Resposta para novos padrões, não apresentados na etapa de aprendizagem.

2.3 HISTÓRIA DA ARTE

O trabalho pioneiro sobre redes neurais artificiais foi realizado no ano de 1943, por McCulloch e Pitts. Eles desenvolveram um estudo sobre o comportamento do neurônio biológico, com o objetivo de criar um modelo matemático para este. As conclusões de suas pesquisas geraram um modelo de elemento de rede neurológica e seu mecanismo de cálculo que foram uma inspiração para diversos trabalhos posteriores (McCULLOCH e PITTS, 1943).

Em 1949, Donald Hebb no livro *The Organization of Behavior* (A Organização do Comportamento), apresentou pela primeira vez uma explicação filosófica para o aprendizado através da modificação da junção sináptica dos neurônios.

O primeiro modelo de rede neural implementado foi o *Perceptron*, por Frank Rosenblatt em 1958. Em seu livro *Principles of Neurodynamics* mostra o modelo dos *Perceptrons*, onde os neurônios eram organizados em camada de entrada e saída, onde os pesos das conexões adaptados a fim de se atingir a eficiência sináptica. Este trabalho

estabeleceu a base para os algoritmos de treinamento de modelos não supervisionados como o de *Kohonen* e para modelos supervisionados como o *backpropagation*.

Neste período também surgiu o primeiro neurocomputador a obter sucesso (Mark I *Perceptron*), criado por Rosenblatt, Charles Wightman e outros (EBERHART *et al*, 1990).

Vários trabalhos foram desenvolvidos durante este período; parecia que as *RNA's* (Redes Neurais Artificiais) poderiam fazer qualquer coisa, e este entusiasmo exagerado de muitos pesquisadores que imaginavam máquinas tão poderosas quanto ao cérebro humano tirou quase toda credibilidade dos estudos desta área e causou grande aborrecimento a teóricos de outras áreas. Em 1969, Minsky e Papert escreveram o livro *Perceptrons* fazendo uma crítica ao modelo e demonstrando matematicamente as suas limitações. Após a publicação desta obra iniciaram-se os anos negros das *RNA's*.

O renascimento da área ocorreu na década de 80, com o trabalho de John Hopfield que criou o modelo de Hopfield em 1982, caracterizado pelo tipo *feedback*, isto é, existe uma conexão das entradas com as saídas. A rede de Hopfield pode ser comparada a um modelo físico, onde a cada troca de estado da rede a energia da mesma diminui, portanto, a fase de aprendizagem chegará ao fim no momento em que a rede tiver sua energia minimizada (HOPFIELD, 1982).

Em 1986 foi publicado o livro Parallel Distributed Processing de David Rumelhart e James McClelland que trouxe grandes contribuições para área, também nesta época foi desenvolvido o algoritmo de aprendizado para redes de múltiplas camadas denominado *Backpropagation* por Rumelhart, Hilton e Willians, 1986 (RUMELHART *et al*,1986).

A partir destes trabalhos as *RNA's* tiveram uma grande redescoberta no mundo científico, sendo hoje considerada multidisciplinar, pois existem pesquisas nas áreas de engenharia, matemática, administração, biomédica entre outras. É importante reconhecer as capacidades e limitações de cada modelo e não criar expectativas alarmantes como no passado.

HECHT-NIELSEN (1990) destaca as aplicações de *RNA's* aos seguintes casos:

- Processamento de sensores, tais como reconhecimento de padrões (por exemplo, reconhecimento de letras impressas ou manuscritas) e compreensão de imagens;
- Extração de informações a partir de dados brutos;
- Controle de processos.

Na engenharia química são muito utilizadas devido à natureza não linear dos processos químicos – são encontradas nos três domínios citados, e muitas vezes essas categorias se sobrepõem, como no caso em que sinais de sensores são processados para uso em controle.

No controle de processos aplicando *RNA's* pode-se citar HOSKIN e HIMMELBLAU (1988), HUNT *et al.* (1992), PSICHOGIOS e UNGAR (1991), DE SOUZA JR (1993), LIEW *et al.* (1996), HUSSAIN (1999), RAMIREZ-BELTRAN e JACKSON (1999), BRODERICK *et al.* (1999), HANG *et al.* (1999), NARENDRA *e CHEN* (2001), ENDER (2002) e MELEIRO (2002).

2.4 REDES NEURAIS

Uma rede neural é especificada, principalmente através de sua topologia, pelas características dos neurônios, e pelo processo de aprendizado.

2.4.1 Neurônio Biológico

O cérebro humano é considerado o mais fascinante processador baseado em carbono existente, sendo composto por aproximadamente 10 bilhões de neurônios. Todas as funções e movimento do organismo estão relacionados ao funcionamento destas pequenas células. Os neurônios estão conectados uns aos outros através de sinapses e juntos formam uma grande rede, chamada de Rede Neural. As sinapses transmitem estímulos através de diferentes concentrações de Na^+ (Sódio) e K^+ (Potássio) e o resultado disto é estendido para todo o corpo humano. Esta grande rede proporciona a capacidade de processamento e armazenamento das informações (IANET, 2003).

O sistema nervoso detecta estímulos internos e externos, tanto físicos quanto químicos e desencadeia as respostas musculares e glandulares. Assim, é responsável pela integração do organismo com o seu meio ambiente. Ele é formado basicamente, por células nervosas, que se interconectam de forma específica e precisa, formando os chamados circuitos neurais. Através desses circuitos, o organismo é capaz de produzir respostas estereotipadas que constituem os comportamentos fixos e invariantes (reflexos) ou então produzir comportamentos variados em maior ou menor grau.

Todo ser vivo dotado de um sistema nervoso é capaz de modificar o seu comportamento em função de experiências passadas. Essa modificação comportamental é chamada de aprendizado e ocorre no sistema nervoso através da propriedade chamada plasticidade cerebral.

A célula nervosa ou simplesmente neurônio é o principal componente do sistema nervoso. Considerada sua unidade anatomo - fisiológica, estima-se que no cérebro humano existam aproximadamente 10 bilhões destas células responsáveis por todas as funções do sistema.

Existem diversos tipos de neurônios, com diferentes funções dependendo da sua localização e estrutura morfológica, mas em geral constituem-se dos mesmos componentes básicos (Figura 2.1).

- Os dentritos, que tem por função receber os estímulos transmitidos por outros neurônios;
- O corpo do neurônio (Soma), que é responsável por coletar e combinar informações vindas de outros neurônios;
- O axônio, que é constituído de uma fibra tubular que pode alcançar até alguns metros, sendo responsável por transmitir os estímulos à outras células.

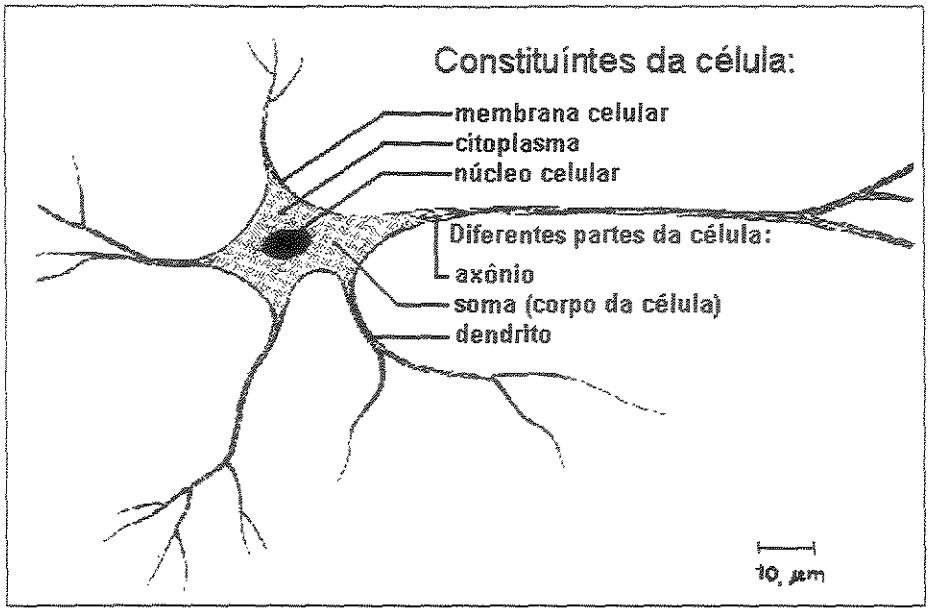


Figura 2.1 Neurônio biológico (IANET,2003)

A sinapse é a estrutura dos neurônios através da qual ocorrem os processos de comunicação entre os mesmos, ou seja, onde ocorre a passagem do sinal neural (transmissão sináptica) através de processos eletroquímicos específicos, isso graças a certas características particulares da sua constituição.

Em uma sinapse os neurônios não se tocam, permanecendo um espaço entre eles denominado de fenda sináptica, onde um neurônio pré-sináptico, liga-se a um outro denominado pós-sináptico. O sinal nervoso (impulso), que vem através do axônio da célula pré-sináptica chega em sua extremidade e provoca na fenda a liberação de neurotransmissores depositados em bolsas chamadas de vesículas sinápticas. Este elemento químico se liga quimicamente a receptores específicos no neurônio pós-sináptico, dando continuidade a propagação do sinal.

Um neurônio pode enviar e receber entre 1.000 a 100.000 conexões sinápticas em relação a outros neurônios, dependendo de seu tipo e localização no sistema nervoso. O número e a qualidade das sinapses em um neurônio pode variar, entre outros fatores, pela experiência e aprendizagem, demonstrando a capacidade plástica do sistema nervoso.

Embora cada neurônio seja um sistema de processamento de informação relativamente lento o paralelismo massivo do processamento de informações em muitas

sinapses simultaneamente conduz a um potencial no processamento efetivo que excede aos supercomputadores de hoje em dia (BISHOP, 1994).

Funcionalmente o sistema nervoso é composto por neurônios sensoriais, motores e de associação. As informações provenientes dos receptores sensoriais aferem ao sistema nervoso central, onde são integradas (codificação/comparação/armazenagem e decisão) por neurônios de associação ou interneurônios, e enviam uma resposta que atua em algum órgão (músculo, glândula) (EPUB, 2003).

O processo de aprendizagem do ser humano é complexo e o cérebro necessita de uma grande concentração de neurônios formando a rede de circuitos neurais conectando suas principais áreas sensoriais e motoras.

A capacidade destes neurônios armazenarem, interpretarem e emitirem respostas eficientes a qualquer estímulo, tendo a capacidade de, a todo instante, em decorrência de novas informações, provocar modificações e rearranjos em suas conexões sinápticas, possibilitando novas aprendizagens, fascina o ser humano, que procura a cada dia uma melhor forma de representar e utilizar o que a natureza tem a ensinar.

2.4.2 Neurônio Artificial

Da mesma maneira que as redes neurais biológicas, as redes neurais artificiais são organizadas na forma de elementos individuais simples que são os neurônios artificiais. Fazendo uma analogia com o neurônio biológico, o conjunto de dados de entrada (X_1, X_2, ..., X_p) correspondem aos dentritos e o conjunto de dados de saída (y) correspondem ao axônio (Figura 2.2).

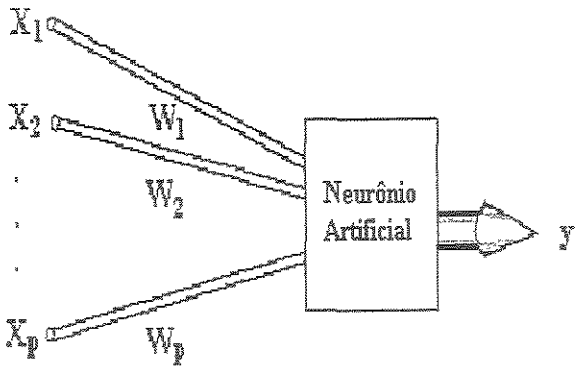


Figura 2.2 Neurônio Artificial

Os neurônios se interconectam uns aos outros através das sinapses formando as redes neurais artificiais (Figura 2.3). São os pesos das sinapses que determinam a magnitude do efeito do sinal recebido. Este sinal pode ser excitatório ou inibitório.

As redes neurais artificiais têm a capacidade de auto-organização, ou plasticidade, ou seja, através de um processo de aprendizado, é possível alterar os pesos das sinapses entre os neurônios artificiais como ocorre nos neurônios biológicos. Os pesos são os detentores da maior parte das informações que uma rede neural pode acumular. Isto implica que a fase de ajuste (aprendizagem) destes pesos tem uma influência direta na performance da rede.

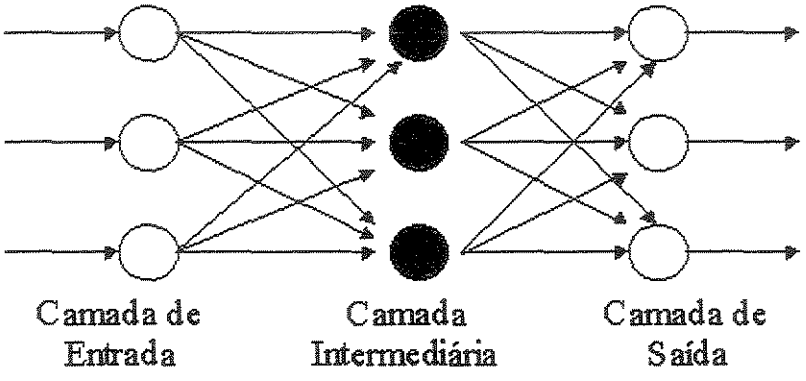


Figura 2.3 Rede Neural Artificial (Tipo Feedforward)

Segundo ENDER (2002) o processamento das informações da rede neural artificial seguem as equações apresentadas a seguir.

Considerando que um neurônio "i" qualquer da camada "k", recebe um conjunto de informações $X_j^{(k-1)}$ (j=1,..., n_{k-1}), correspondentes as saídas dos n_{k-1} neurônios da camada anterior, ponderadas, cada uma, pelo correspondente peso da conexão $w_{i,j}^{(k)}$. O neurônio soma essas entradas ponderadas e o valor resultante é por sua vez somado a um limite interno de ativação, um "bias" que pode ser descrito por $w_{i,0}^{(k)}$. A este sinal resultante, o neurônio "i", produz uma resposta $X_i^{(k)}$, de acordo com uma função limite ou função de ativação $f(\bullet)$, conforme mostrado na Figura 2.4, matematicamente pode-se expressar por:

$$S_i^{(k)} = w_{i,0}^{(k)} + \sum_{j=1}^{(n_{k-1})} w_{i,j}^{(k)} X_j^{(k-1)} \tag{2.1}$$

$$X_i^{(k)} = f\left(S_i^{(k)}\right) \tag{2.2}$$

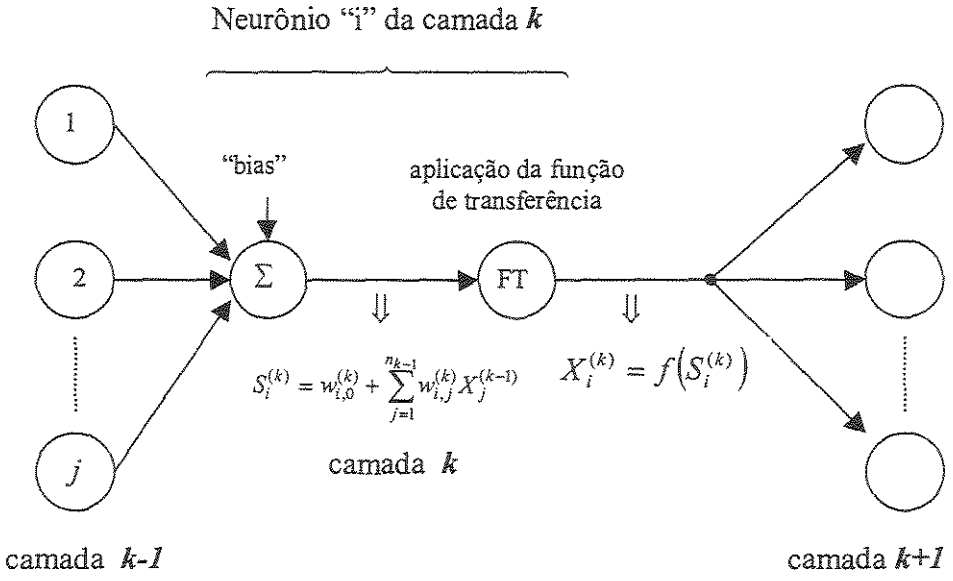


Figura 2.4 Anatomia do i-ésimo neurônio artificial da k-ésima camada

De acordo com BAKSHI e STEPHANOPOULOS (1993), para a ativação dos neurônios pode-se utilizar dois tipos de funções, a local e a global. A função de ativação

local é aquela que só apresenta uma resposta ativa para valores de entrada na vizinhança imediata do seu centro de ativação. Um exemplo é a função gaussiana.

A função de ativação global produz uma saída ativa para um largo intervalo de entradas. Por exemplo, as funções: linear, degrau limiar, sigmoidal e tangente hiperbólica.

Funções de ativação lineares praticamente não são utilizadas, pois as redes basicamente são usadas para modelagem de comportamentos não lineares e o uso de camadas internas com neurônios com ativação linear não aumenta o poder computacional da rede multicamada comparando-a com uma rede sem camadas internas (RUMELHART e McCLELLAND, 1986).

Na literatura observou-se o uso das funções sigmoidal e tangente hiperbólica com resultados satisfatórios utilizando o método de aprendizagem *backpropagation* entre eles, ENDER e MACIEL FILHO (2001a), KAUFMANN e ANDERSSON (1999), MORENO (2001). Nesta dissertação serão utilizadas como funções de ativação as sigmoidal e tangente hiperbólica. A Figura 2.5 mostra o comportamento das funções respectivas.

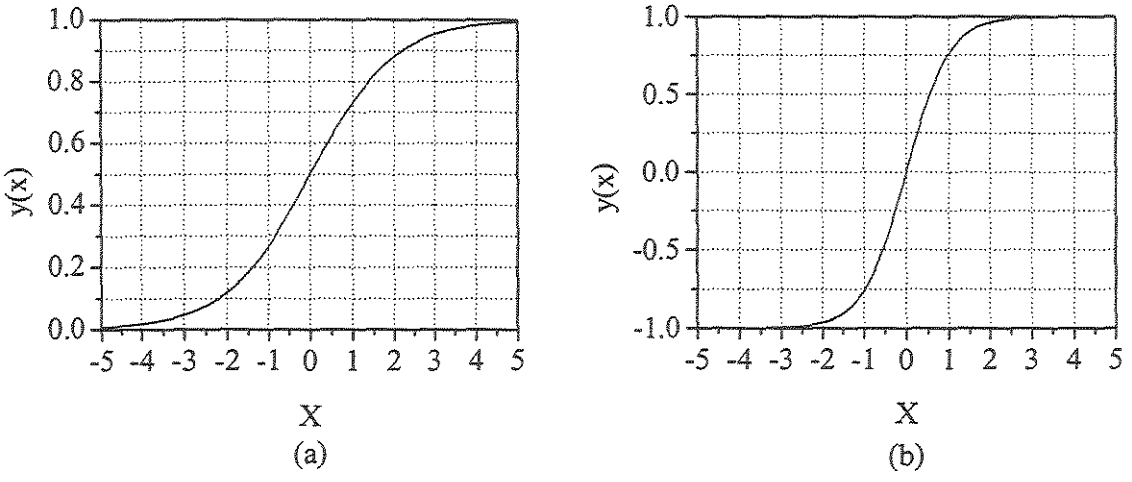


Figura 2.5 **Funções de ativação:** *(a) Função Sigmoidal; (b) Função Tangente Hiperbólica*

A diferença básica entre a função sigmoidal e a função tangente hiperbólica, é que a primeira é "positiva" (saída total variando de 0 a 1) e a segunda é "média-zero" (saída total variando de –1 a 1) (HUNT *et al.*, 1992).

As funções seguem as equações (2.3) e (2.4):

- função sigmoidal: $$f\left(S_i^{(k)}\right)=\frac{1}{1+e^{-S_i^{(k)}}} \tag{2.3}$$

- função tangente hiperbólica: $$f\left(S_i^{(k)}\right)=\tanh\left(S_i^{(k)}\right) \tag{2.4}$$

2.5 ARQUITETURA DAS REDES NEURAIS ARTIFICIAIS

Uma rede neural artificial é formada por diversas camadas como visto na Figura 2.3. Basicamente apresenta uma camada de entrada, camadas intermediárias (ocultas) e camada de saída. Os neurônios da camada de entrada são responsáveis pela distribuição dos dados de entrada para a próxima camada (camada intermediária). Nas camadas intermediárias as informações são processadas e propagadas até a última camada (camada de saída).

Segundo HENRIQUE *et al.* (2000) dois aspectos importantes a serem determinados são o número de camadas intermediárias e o número de neurônios das camadas intermediárias a serem utilizados. Por não existirem padrões, estes valores variam de acordo com a função que se deseja determinar. O que se observa é que aumentando a estrutura da rede (número de camadas e neurônios) ocorre uma maior precisão dos resultados; neste caso ocorre um aumento do tempo computacional requerido. Por isso para todas as redes utilizou-se apenas uma camada intermediária.

Segundo HECHT-NIELSEN (1989), uma função contínua pode ser aproximada para qualquer grau de precisão usando uma rede neural *backpropagation* com três camadas, desde que haja um número suficiente de neurônios ativos na camada oculta.

CYBENKO (1989) mostrou que qualquer representação de um conjunto de dados de $\Re^p$ para $\Re^q$ pode ser obtido com duas camadas internas. HORNIK *et al.* (1990)

também mostraram que qualquer mapeamento pode ser obtido, para um grau de precisão, usando uma rede neural com uma camada interna com um número suficiente de neurônios.

Segundo HECHT-NIELSEN (1989), o matemático russo KOLMOGOROV (1957), publicou um teorema relativo a representação de funções contínuas que pode ser diretamente adaptado às redes neurais multicamadas, reformulando este teorema da seguinte maneira (KOVÁCS, 1996):

Teorema – Kolmogorov-Nielsen: Dada uma função contínua arbitrária $f:[0,1]^n \rightarrow \Re^m$, $f(x)=y$, existe sempre para f, uma implementação exata com uma rede neural de três camadas, sendo a camada de entrada um vetor de dimensão n, a camada oculta composta por $(2n+1)$ neurônios, e a camada de saída com m neurônios representando as m componentes do vetor y.

De acordo com LOESCH *et. al* (1996), vários autores sugerem que em redes de 3 camadas com mais entradas que saídas o número ideal de elementos na camada oculta seja algo em torno da média geométrica entre as camadas de entrada e de saída.

2.6 CLASSIFICAÇÃO DAS REDES

2.6.1 Aplicações na Área de Controle

BAUGHMAN e LIU (1995), classificaram as redes em três categorias na aplicação em controle:

- Controle Direto: A rede neural é treinada para atuar como controlador e as ações de controle são determinadas diretamente;
- Controle Inverso: A rede neural é treinada como modelo inverso de processo, predizendo a ação de controle necessária para produzir a resposta desejada do processo;
- Controle Indireto: A rede neural é treinada para atuar como modelo do processo, ou seja, para determinar o *setpoint* para um controlador local, assim afetando indiretamente o processo controlado.

PSICHOGIOS e UNGAR (1991) classificaram a aplicação das redes neurais como método direto e método indireto. No método direto, a rede neural é treinada com dados de entrada / saída do sistema para representar a dinâmica inversa deste processo. Ou seja, dado o estado atual do sistema dinâmico e o estado desejado (*setpoint*) para o próximo tempo de amostragem, a rede neural é treinada a produzir a ação de controle que leva o sistema para o estado desejado. O modelo inverso resultante poderá ser usado como controlador em uma estrutura *feedforward* típica. No método indireto a rede neural é treinada com dados do sistema dinâmico para representar dinâmicas futuras, ou seja, conhecida o estado atual e a ação de controle atual, a rede aprende a predizer o novo estado do sistema.

2.6.2 Fluxo de Informações

Quanto ao fluxo das informações as *RNA's* podem ser da forma acíclica (*Feedforward*) ou da forma cíclica (*Feedback*) (BAUGHMAN e LIU, 1995). As conexões do tipo *feedback* são usualmente utilizadas em rede neurais recorrentes, como exemplo a rede de Hopfield (NARENDRA e PARTHASARATHY, 1990). A Figura 2.6 traz exemplos de estruturas *feedforward* (a, b, c) e *feedback* (d, e) (Fonte IANET).

Segundo BULSARI (1995) nas redes com conexões do tipo *feedforward*, todos os sinais se propagam através das camadas para "frente". Não há conexões laterais ou como o sinal retroceder. Este tipo de rede possui a capacidade de generalização, habilidade para aproximar funções complexas, dados com ruídos e são apropriadas para a modelagem dinâmica, onde se deseja mapear uma resposta de saída baseada em um sinal de entrada. As redes usadas na estratégia de controle propostas são constituídas por este tipo de arquitetura.

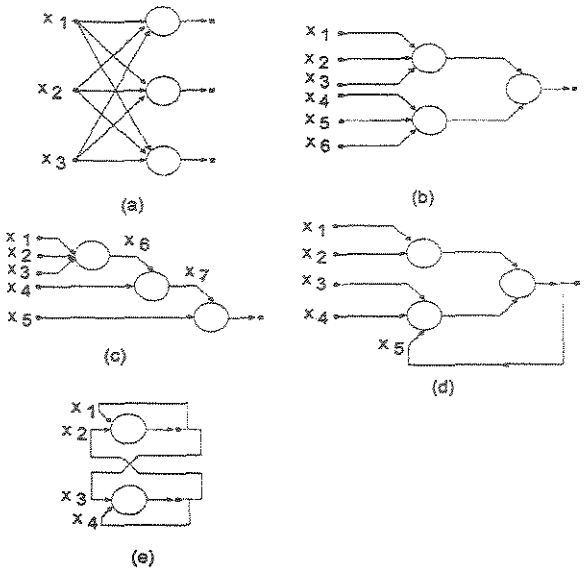


Figura 2.6 Estruturas de Rede Feedforward (a,b,c) e Feedback (d,e)

2.6.3 Método de Identificação

Paralelo: resposta predita pela rede é realimentada na mesma;

Série-Paralelo: saída do processo realimenta a rede.

2.7 PROCESSO DE APRENDIZAGEM

A aprendizagem ou treinamento é o processo onde os pesos e biases da rede são modificados (SIMPSON, 1990). O processo de aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

2.7.1 Aprendizado Supervisionado

Esta forma de aprendizagem utiliza um agente externo que indica à rede um comportamento bom ou ruim de acordo com o padrão de entrada (Figura 2.7). Com se existisse um professor, fornecendo a resposta correta para a rede a ser treinada. A rede aprende por experiência, ou seja, um conjunto de pares de exemplos de entrada e saída são apresentados a mesma assim obtém-se um sinal de erro com a resposta atual da rede e a saída desejada, com o qual ela pode adaptar seus pesos.

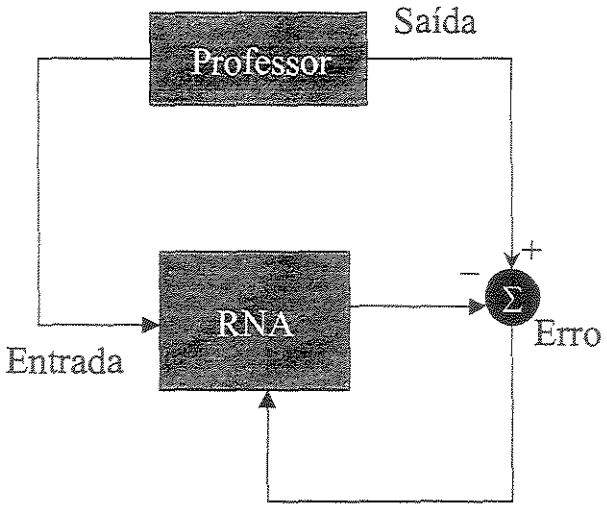


Figura 2. 7 Aprendizado Supervisionado

2.7.2 Aprendizado Não Supervisionado

Aprendizado não supervisionado ocorre de um modo auto-organizado não sendo necessário um professor para ordenar as mudanças sinápticas. Os parâmetros da rede são dirigidos pelos dados de entrada, de forma a criar classificações de acordo com os atributos percebidos (Figura 2.8). Essas classes são funções da distribuição estatística dos dados de entrada. Uma vez que a rede aprendeu os padrões estatísticos e formou representações internas dos atributos da entrada, ela é capaz de criar novas classes automaticamente (BRAGA *et al.* , 2000).

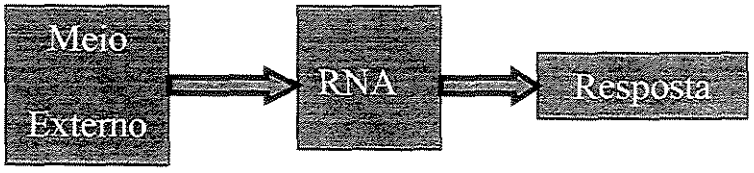


Figura 2.8 Aprendizado Não Supervisionado

Os métodos de treinamento mais empregados são os supervisionados, pois um modelo interno do processo que rege os dados fornecidos à rede é criado pela mesma, de modo que ela pode fazer predições para novas entradas e reconhecer padrões, mesmo com ruídos (RUMELHART e McCLELLAND (1986), GIARRATANO *et al.* (1990), DE SOUZA Jr (1993)).

2.8 ALGORITMO DE APRENDIZAGEM *BACKPROPAGATION*

As redes treinadas pelo método *backpropagation* (retropropagação) são multicamadas e não apresentam conexões laterais entre neurônios da mesma camada, seus sinais são do tipo *feedforward* (RUMELHART e McCLELLAND, 1986).

Durante o treinamento com o algoritmo *backpropagation*, a rede opera em uma seqüência de dois passos. No primeiro passo (Figura 2.9), um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se o valor da saída calculada não estiver correto, o erro é calculado. Este erro é então propagado a partir da camada de saída até a camada de entrada (daí o nome *backpropagation*), e os pesos das conexões dos neurônios da camada interna vão sendo modificados conforme o erro é retropropagado.

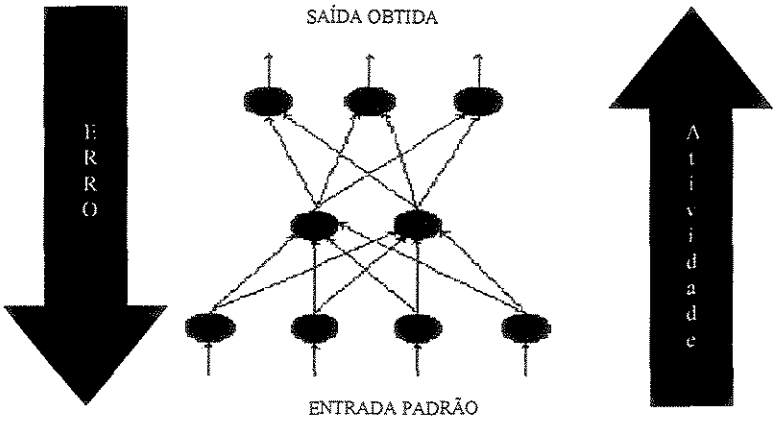


Figura 2.9 Algoritmo Backpropagation

Este tipo de algoritmo utiliza a regra delta generalizada, esta usa em essência a regra da cadeia do cálculo diferencial. Seu objetivo é reduzir o erro quadrático da rede, emprega o método do passo na direção oposta do gradiente, sobre a superfície de erros, onde se deseja minimizar um erro que é função de todos os pesos da rede (GOZDZIEJEWSKI Jr. e ENDER, 1998). O erro quadrático calculado na rede é dado por:

$$\varepsilon = \sum_{l=1}^{n_k} \left(d_l - X_l^{(k)} \right)^2 \tag{2.5}$$

Sendo que n_k o número de neurônios da camada de saída, $X_l^{(k)}$ a l-ésima saída da rede e d_l a saída desejada.

Considerando um neurônio "i" qualquer da camada de saída "k", em uma determinada iteração. O gradiente do erro para este elemento de processamento é:

$$\vec{\nabla}_{i,j}^{(k)} \varepsilon = \frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} \tag{2.6}$$

Onde $\vec{\nabla}_{i,j}^{(k)} \varepsilon$ denota o gradiente do erro quadrático do i-ésimo neurônio da k-ésima camada gerado pela j-ésima conexão sináptica.

Sabe-se que:

$$X_i^{(k)} = f\left(S_i^{(k)} \right) \tag{2.7}$$

$$S_i^{(k)} = \sum_{j=0}^{n_{k-1}} w_{i,j}^{(k)} X_j^{(k-1)} \tag{2.8}$$

Sendo que para $j = 0$, $X_j^{(k-1)} = +1$ que representa a conexão extra (*bias*) e n_{k-1} é o número de neurônios da camada $k-1$. Então:

$$\varepsilon = \sum_{l=1}^{n_k} \left(d_l - f\left(S_l^{(k)} \right) \right)^2 \tag{2.9}$$

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = \frac{\partial\left(\sum_{l=1}^{n_k}\left(d_l - f\left(S_l^{(k)}\right)\right)^2\right)}{\partial w_{i,j}^{(k)}} \tag{2.10}$$

Sabe-se que quando $i \neq l$ as derivadas serão nulas, portanto:

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = -2\left(d_i - f\left(S_i^{(k)}\right)\right)\frac{\partial f\left(S_i^{(k)}\right)}{\partial w_{i,j}^{(k)}} \tag{2.11}$$

Aplicando a regra da cadeia, chega-se a:

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = -2\left(d_i - f\left(S_i^{(k)}\right)\right)\frac{\partial f\left(S_i^{(k)}\right)}{\partial S_i^{(k)}}\frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}} \tag{2.12}$$

Pela equação (2.8) pode-se determinar:

$$\frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}} = \frac{\partial\left(\sum_{l=0}^{n_{k-1}} w_{i,l}^{(k)} X_l^{(k-1)}\right)}{\partial w_{i,j}^{(k)}} \tag{2.13}$$

A derivada em questão na equação anterior é nula quando $l \neq j$, desta forma:

$$\frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}} = X_j^{(k-1)} \tag{2.14}$$

A equação (2.12) pode ser reescrita:

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = -2\left(d_i - X_i^{(k)}\right)f'\left(S_i^{(k)}\right)X_j^{k-1} \tag{2.15}$$

Definindo:

$$\delta_i^{(k)} = \left(d_i - X_i^{(k)}\right) f'\left(S_i^{(k)}\right) \tag{2.16}$$

Segundo ENDER (2002), NGUYEN e WIDROW (1990) e CHEN (1990):

$\delta_i^{(k)}$ pode ser interpretado como *erro equivalente*, o qual será igual ao erro $d_i - X_i^{(k)}$ se $f\left(S_i^{(k)}\right)$ for a função identidade. Substituindo tem-se:

$$\vec{\nabla}_{i,j}^{(k)} \varepsilon = \frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = -2\delta_i^{(k)} X_j^{(k-1)} \tag{2.17}$$

Desta forma, os pesos podem ser reajustados na direção oposta ao gradiente do erro, da seguinte forma:

$$w_{i,j}^{(k)\,new} = w_{i,j}^{(k)\,old} + \eta\left(-\vec{\nabla}_{i,j}^{(k)} \varepsilon\right) \tag{2.18}$$

Ou substituindo a equação (2.17) na equação (2.18), tem-se:

$$w_{i,j}^{(k)\,new} = w_{i,j}^{(k)\,old} + 2\eta\delta_i^{(k)} X_j^{(k-1)} \tag{2.19}$$

Onde $0 \langle \eta \langle\langle 1$ é o termo de taxa de aprendizagem.

A forma de determinação de $\delta_i^{(k)}$, apresentada na equação (2.16), é valida para o ajuste dos pesos dos neurônios da camada de saída. Para ajustar os pesos das camadas ocultas, o erro ε deverá ser retropropagado até a camada oculta em análise, conforme exposto a seguir.

Considerando um neurônio "*i*" de uma camada oculta "*k*", em uma rede neural de 3 camadas, ou seja, uma camada de entrada (*k*-1), uma camada oculta (*k*) e uma camada de saída (*k*+1). Pela regra da cadeia, pode-se escrever:

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = \frac{\partial \varepsilon}{\partial S_i^{(k)}} \frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}} \tag{2.20}$$

Analisando primeiramente o termo $\frac{\partial \varepsilon}{\partial S_i^{(k)}}$,

$$\frac{\partial \varepsilon}{\partial S_i^{(k)}} = \frac{\partial\left(\sum_{l=1}^{n_{k+1}}\left(d_l - X_l^{(k+1)}\right)^2\right)}{\partial S_i^{(k)}} \tag{2.21}$$

$$\frac{\partial \varepsilon}{\partial S_i^{(k)}} = -2\sum_{l=1}^{n_{k+1}}\left(d_l - f\left(S_l^{(k+1)}\right)\right)\frac{\partial f\left(S_l^{(k+1)}\right)}{\partial S_l^{(k+1)}}\frac{\partial S_l^{(k+1)}}{\partial S_i^{(k)}} \tag{2.22}$$

Porém:

$$\frac{\partial S_l^{(k+1)}}{\partial S_i^{(k)}} = \frac{\partial\left(\sum_{j=0}^{n_k} w_{l,j}^{(k+1)} f\left(S_j^{(k)}\right)\right)}{\partial S_i^{(k)}} \tag{2.23}$$

Sabe-se que as derivadas serão nulas para $j \neq i$, portanto:

$$\frac{\partial S_l^{(k+1)}}{\partial S_i^{(k)}} = w_{l,j}^{(k+1)} f'\left(S_i^{(k)}\right) \tag{2.24}$$

Então:

$$\frac{\partial \varepsilon}{\partial S_i^{(k)}} = -2\sum_{l=1}^{n_{k+1}}\left(d_l - X_l^{(k+1)}\right) f'\left(S_l^{(k+1)}\right) w_{l,i}^{(k+1)}\ f'\left(S_i^{k}\right) \tag{2.25}$$

Pela equação (2.16), pode-se afirmar:

$$\delta_l^{(k+1)} = \left(d_l - X_l^{(k+1)}\right) f'\left(S_l^{(k+1)}\right) \tag{2.26}$$

Portanto:

$$\frac{\partial \varepsilon}{\partial S_i^{(k)}} = -2 f'\left(S_i^{(k)}\right)\sum_{l=1}^{n_{k+1}} \delta_l^{(k+1)}\ w_{l,i}^{(k+1)} \tag{2.27}$$

Analisando agora o termo $\frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}}$,

$$\frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}} = \frac{\partial\left(\sum_{m=1}^{n_{k-1}} w_{i,m}^{(k)}\ X_m^{(k-1)}\right)}{\partial w_{i,j}^{(k)}} \tag{2.28}$$

As derivadas para $m \neq j$ são nulas, então:

$$\frac{\partial S_i^{(k)}}{\partial w_{i,j}^{(k)}} = X_j^{(k-1)} \tag{2.29}$$

Portanto a derivada do erro em relação aos pesos da camada oculta pode ser determinada da seguinte forma:

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = -2 f'\left(S_i^{(k)}\right) \sum_{l=1}^{n_{k+1}} \left(\delta_l^{(k+1)}\ w_{l,i}^{(k+1)}\right) X_j^{(k-1)} \tag{2.30}$$

Então:

$$\frac{\partial \varepsilon}{\partial w_{i,j}^{(k)}} = -2\delta_i^{(k)} X_j^{(k-1)} \tag{2.31}$$

Onde tem-se para a camada intermediária:

$$\delta_i^{(k)} = f'\left(S_i^{(k)}\right) \sum_{l=1}^{n_{k+1}} \left(\delta_l^{(k+1)}\ w_{l,i}^{(k+1)}\right) \tag{2.32}$$

Esta forma para o cálculo de $\delta_i^{(k)}$ é usada para retropropagar o erro para as demais camadas ocultas, caso a rede neural apresentar $n \rangle 1$ camadas ocultas. A equação para o ajuste dos pesos fica da mesma forma que apresentada na equação (2.19), ou seja:

$$w_{i,j}^{(k)\,new} = w_{i,j}^{(k)\,old} + 2\eta\delta_i^{(k)} X_j^{(k-1)} \tag{2.33}$$

A inicialização dos pesos das interconexões e os *biases* é realizada com números aleatórios e pequenos, desta forma procura-se evitar os problemas de simetria (que haveria caso se começasse com pesos iguais) e saturação (produzidos por pesos grandes). No desenvolvimento desta dissertação adotaram-se pesos iniciais aleatórios no intervalo entre –1 e 1, conforme ENDER (2002), RUMELHART e McCLELLAND (1986).

Segundo ENDER (2002), para cada padrão p (p=1, 2,..., N_p; onde N_p é o número de padrões), tem-se um conjunto de entradas e um conjunto de saídas desejadas, d_i, que são usadas no critério de erro. Na implementação da aprendizagem *backpropagation*, pode-se atualizar os pesos após a apresentação de cada padrão (aprendizagem por padrão) ou após a apresentação de todos os padrões (aprendizagem batelada) (RUMELHART e McCLELLAND (1986). A forma de implementação adotada foi a aprendizagem por padrão.

Depois do treinamento, pode-se usar a rede, onde tem-se apenas a fase em que as entradas passam através da rede para frente (*feedforward*), correspondendo às equações (2.1) e (2.2).

2.9 CONCLUSÕES

No desenvolvimento desta dissertação, adotou-se como controle direto quando a rede neural atua como controlador e controle indireto quando as redes neurais são usadas em modelos de predição de estados ou dinâmicas futuras.

De acordo com a revisão da literatura apresentada neste capítulo determinou-se a arquitetura das redes utilizadas na estratégia de controle proposta com aprendizagem *on-line* das redes, estas serão compostas por 3 camadas, estrutura *feedforward* e como funções de ativação serão utilizadas as funções sigmoidal e tangente hiperbólica. O algoritmo de aprendizagem utilizado será do tipo supervisionado, *backpropagation* apropriado para o controle de processos devido à capacidade de representar processos dinâmicos e generalizar mesmo com ruídos.

ESTRATÉGIA DE CONTROLE

3.1 INTRODUÇÃO

O capítulo 3 apresenta a aplicação de redes neurais artificiais em estratégia de controle não linear, abordando aspectos relativos à aprendizagem *off-line* e *on-line* das redes. Para aprendizagem *on-line* é utilizada uma estratégia com três redes paralelas para suprir a limitação deste tipo de aprendizagem, além disto são propostas diferentes configurações de aprendizagens diferenciadas nos vetores de aprendizagem e teste das redes.

3.2 CONTROLE DE PROCESSOS

Controlar um processo significa atuar sobre ele, ou sobre as condições a que o processo está sujeito, de modo a atingir algum objetivo, por exemplo, manter o processo próximo de um determinado estado estacionário, mesmo que os efeitos externos tentem desvia-lo desta condição (STEPANOPOULOS, 1984).

Com a globalização, nos dias atuais as industrias buscam aumentar seus lucros e se tornarem mais competitivas, oferecendo produtos com maior qualidade e menor custo. Uma forma de se obter estes produtos é através do controle do processo de fabricação dos mesmos.

A maioria das plantas químicas dificilmente operam em estado estacionário, apesar deste ser uma condição de operação desejável, nem sempre é atingido ou mantido por muito tempo. Por isso existe a necessidade de monitoramento da operação destas plantas e intervir para garantir a satisfação dos objetivos operacionais.

O controle é realizado principalmente para:

- Garantir que restrições de segurança e ambientais não sejam violadas;
- Garantir certas condições operacionais;

- Garantir produtos finais com maior qualidade e menor custo.

Em geral, os parâmetros do controlador são determinados de maneira *off–line,* o que significa que a ação de controle é baseada em uma análise que pode não refletir as variações do processo e neste caso gerar um erro.

A estratégia desenvolvida neste trabalho busca uma alternativa para esta limitação, ou seja, atualiza estes parâmetros ao longo do tempo (*on-line*), através de uma etapa de otimização do controlador, utilizando redes neurais artificiais. Todas as redes utilizadas estão baseadas na arquitetura *feedforward* com aprendizagem *on-line* usando o algoritmo *backpropagation* conforme descrito no Capítulo 2.

3.3 CONTROLADOR NEURAL – DINÂMICA INVERSA

Um processo pode ser descrito na seguinte forma discreta (NARENDRA e PARTHASARATHY (1990), CHEN e BILLINGS (1992), NAHAS *et al.* (1992), SU *et al.*(1992), BULSARI(1995) e ENDER (2002)):

$$Y(k+1) = f(Y(k),...,Y(k-n+1),U(k),...,U(k-m+1)) \tag{3.1}$$

com:

$$\begin{aligned} Y(k) &= [y(1,k),\ldots y(i,k)]^T \\ U(k) &= [u(1,k),\ldots u(i,k)]^T \qquad i = \textit{número de malhas de controle} \end{aligned} \tag{3.2}$$

onde $U(k)$ e $Y(k)(k=0,1,...n)$ são os vetores de perturbações e respostas no instante de tempo k, respectivamente, e $f(\bullet)$ é uma função desconhecida. Considerando o desenvolvimento de uma função aproximada para $f(\bullet)$ pelo uso de redes neurais, onde $Y(k+1)$ e $Y(k),...,Y(k-n+1),U(k),...,U(k-m+1)$ são respostas e perturbações da rede neural respectivamente, e W são os pesos das conexões e os limiares da rede neural. Esta rede neural pode gerar um modelo do processo, pelo ajuste de W, denominado modelo de controle indireto:

$$\hat{Y}(k+1) = \hat{f}(Y(k),...,Y(k-n+1),U(k),...,U(k-m+1);W) \tag{3.3}$$

Resolvendo a equação (3.1) em relação a $U(k)$, obtém-se:

$$U(k) = \phi(Y(k+1),Y(k),...,Y(k-n+1),U(k-1),...,U(k-m+1)) \tag{3.4}$$

onde $\phi(\bullet)$ é uma função desconhecida, que representa a dinâmica inversa do processo. Como a função $\phi(\bullet)$ é desconhecida, bem como a função $f(\bullet)$, não é possível avaliar a equação (3.4). Neste caso utiliza-se uma rede neural para gerar uma aproximação da função ϕ, da seguinte maneira:

$$\hat{U}(k) = \hat{\phi}(Y(k+1),Y(k),...,Y(k-n+1),U(k-1),...,U(k-m+1);W) \tag{3.5}$$

Considere um problema de controle em que a resposta $Y(k)$ procura acompanhar as referências da malha fechada, $Y_r(k)$. Então para o próximo tempo de amostragem:

$$Y(k+1) = Y_r(k+1), \tag{3.6}$$

Substituindo a Equação (3.6) na Equação (3.5), a ação de controle desejada para o processo é dada por:

$$\hat{U}(k) = \hat{\phi}(Y_r(k+1),Y(k),...,Y(k-n+1),U(k-1),...,U(k-m+1);W) \tag{3.7}$$

Este modelo representa o mapeamento de perturbações/respostas não lineares através de redes neurais, denominado de modelo de controle direto, que terá suficiente exatidão se um grande número de dados estiver disponível (PSICHOGIOS e UNGAR (1991), BULSARI, (1995)).

Conforme a Figura 3.1, as entradas desta rede são formadas pelas referências das variáveis manipuladas para o próximo tempo de amostragem das variáveis manipuladas e controladas passadas. Como resposta da rede obtém-se as ações de controle para o próximo tempo de amostragem, onde β representa o número de informações passadas do processo usadas como entrada na rede.

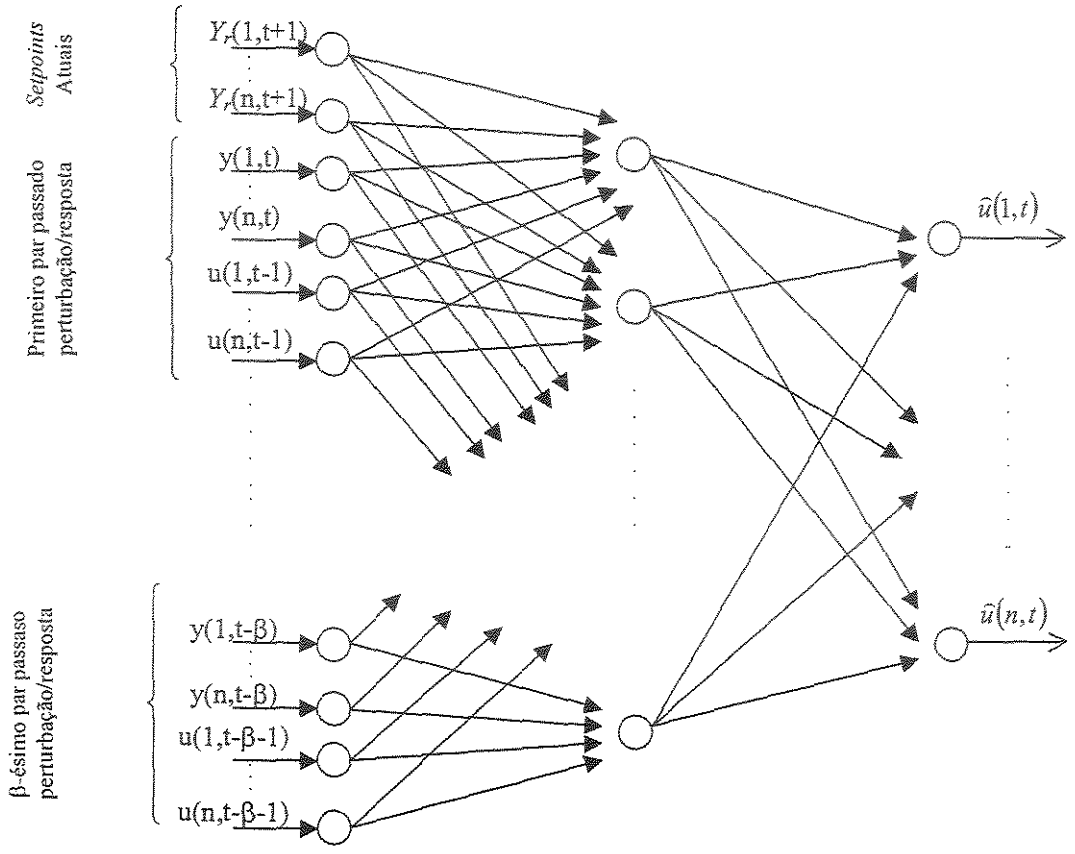


Figura 3.1 Arquitetura feedforward multivariável (Modelo Inverso)

O controle neural baseado em modelo inverso tem como objetivo conduzir a variável controlada para a referência desejada do processo. Segundo ENDER (2002), este tipo de modelo não apresenta resultados eficientes para comportamentos dinâmicos, podendo gerar *off-set.*

3.4 MODELO NEURAL DO PROCESSO – IDENTIFICAÇÃO

As redes neurais artificiais são amplamente empregadas na identificação de processos, gerando modelos a serem usados no projeto de controladores não lineares. Na

estratégia de controle proposta denominou-se esta rede como Modelo Neural do Processo. Utilizou-se o método de identificação série - paralelo, descrito no Capítulo 2.

3.5 ESTRATÉGIA DE CONTROLE – CONDEG MODIFICADO

O objetivo do controle adaptativo é ajustar os parâmetros do controlador de forma a compensar possíveis perturbações dinâmicas no sistema, causadas por variações nas condições operacionais do processo.

A estratégia proposta foi denominada Controle Neural Direto baseado no Erro Global (CoNDEG) Modificado, pois foram realizadas modificações no programa CoNDEG desenvolvido por ENDER (2002). A presente estratégia está baseada no modelo inverso e no controle adaptativo, apresentando um mecanismo diferenciado para ajustar o controlador que é baseado em redes neurais artificiais. A principal vantagem do uso da estratégia é a atualização dos parâmetros do controlador ao longo do tempo (*on-line*), que permite um controle mais efetivo, pois acompanha as variações dinâmicas do processo.

A Figura 3.2 mostra o esquema da estratégia, onde são utilizadas redes neurais artificiais atuando como controlador neural (modelo direto), e modelo neural do processo (modelo indireto). O qual é utilizado na rotina de otimização do controlador. A rede neural do controlador é otimizada *on-line*, a cada tempo de amostragem, usando o erro global estimado para malha fechada, obtido através do modelo neural do processo.

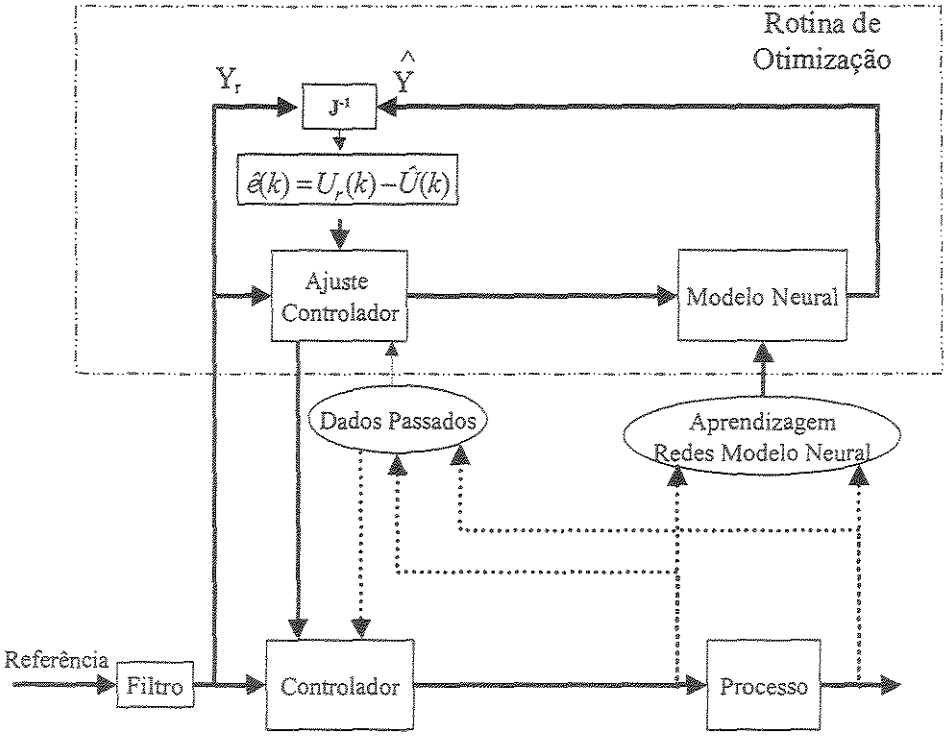


Figura 3.2 Estratégia de Controle CoNDEG Modificado

3.5.1 Ajuste do Controlador

Na rotina de otimização, o projeto do controlador utiliza as mesmas entradas que serão utilizadas no controlador neural para cada instante de amostragem. Esta rotina ajusta os pesos da rede neural do controlador de tal forma a minimizar o erro global estimado, onde a resposta da malha fechada é proveniente do modelo neural do processo.

O erro global não pode ser diretamente aplicado na aprendizagem das redes neurais do controlador, pois estes estão associados às respostas do processo e não às ações de controle, as quais são as respostas da rede neural a ser treinada. Por isso são retropropagados através de matriz Jacobiana do processo.

Para o ajuste do controlador, ENDER (2002), utilizou para ativação dos neurônios a função tangente hiperbólica e atingiu bons resultados. Avaliando a estratégia com redes treinadas com funções de ativação sigmoidal e tangente hiperbólica, observou-se que a estratégia não estava apta à redes treinadas com a função sigmoidal. Isto ocorre devido a função de ativação sigmoidal apresentar respostas somente positivas (0-1), fazendo com

que o erro global estimado $(Y_r - \hat{Y})$ quando retropropagado através da matriz jacobiana seja normalizado, gerando somente valores positivos, impedindo assim o ajuste dos pesos. Alterou-se a forma do cálculo deste erro, devido a esta limitação. Primeiramente foram retropropagados e normalizados individualmente os valores de referência (Y_r) e valores preditos da rede neural para a malha fechada $(\hat{Y})$ e posteriormente gerados os erros. Desta forma o erro utilizado na aprendizagem pode assumir valores positivos e negativos e pode-se utilizar as funções de ativação sigmoidal e tangente hiperbólica, para treinamento das redes, conforme descrito a seguir.

A matriz Jacobiana pode ser determinada a partir do modelo neural pela variação de cada variável manipulada, em torno do seu ponto de operação e medindo a variação nas respostas do processo. Considerando a matriz jacobiana:

$$J = \begin{pmatrix} \frac{\partial y_1}{\partial u_1} & \cdots & \frac{\partial y_1}{\partial u_i} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_i}{\partial u_1} & \cdots & \frac{\partial y_i}{\partial u_i} \end{pmatrix} \tag{3.8}$$

Considerando que as variáveis y_j e u_j são as respostas estimadas da malha fechada e a ação de controle que gerou esta resposta estimada da malha fechada, respectivamente, pode-se escrever:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_i \end{pmatrix} = \begin{pmatrix} \frac{\partial y_1}{\partial u_1} & \cdots & \frac{\partial y_1}{\partial u_i} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_i}{\partial u_1} & \cdots & \frac{\partial y_i}{\partial u_i} \end{pmatrix} * \begin{pmatrix} u_1 \\ \vdots \\ u_i \end{pmatrix} \tag{3.9}$$

Então os valores encontrados com a retropropagação são:

$$\hat{U} = J^{-1} * \hat{Y}$$

$$U_r = J^{-1} * Y_r \tag{3.10}$$

Assim o erro global $(\hat{e}(k))$ é calculado de acordo com a expressão:

$$\hat{e}(k) = U_r(k) - \hat{U}(k) \tag{3.11}$$

com:

$$\begin{aligned} U_r(k) &= [u_r(1),\ldots,u_r(i)]^T \\ \hat{U}(k) &= [\hat{u}(1,k),\ldots,\hat{u}(i,k)]^T \end{aligned} \tag{3.12}$$

onde:

- $U_r(k)$ é o vetor de ações de controle desejado para malha fechada;
- $\hat{U}(k)$ é o vetor de ações de controle estimado para malha fechada.
- i : refere-se as malhas de controle;
- $u_r(i,k)$: vetor que contém as ações de controle de referências do processo no instante de tempo k;
- $\hat{u}(i,k)$: vetor que contém as ações de controle preditas a partir do modelo neural do processo.

Este erro global é gerado inúmeras vezes, até ser atingida a tolerância permitida para o erro quadrático da rede neural ou alcançar-se o número máximo de iterações preestabelecidas, para atender as limitações de tempo disponível para processamento das informações, devido à aprendizagem *on-line*. Para garantir a boa representação dinâmica do processo através de redes neurais, adotou-se uma estratégia formada por três redes neurais que atuam paralelamente (ENDER e MACIEL FILHO, 2001[a,b]). Considerando que o ajuste da rede neural do controlador está baseado em predições, torna-se fundamental usar-se um modelo que represente com fidelidade o comportamento dinâmico do processo. Por esta razão, sempre que o erro quadrático, de pelo menos uma das redes neurais do modelo do processo, não atingir a tolerância desejada, o procedimento de aprendizagem das redes neurais do controlador não será executado. Nestas circunstâncias, o modelo do processo a ser usado, será aquele que apresentar menor erro quadrático na representação do vetor de padrões e a rede do controlador a ser usada na geração das ações de controle será aquela

que apresentar menor erro quadrático para a situação atual. O modelo dinâmico do processo é utilizado para a geração do erro quadrático da rede neural do controlador.

Para evitar variações bruscas nas referências do processo um filtro de primeira ordem foi utilizado de tal forma a produzir-se uma trajetória de referências, da seguinte forma:

$$y_w(i,k+1) = \alpha_i y_w(i,k) + (1-\alpha_i) y_r(i) \tag{3.13}$$

onde $y_r(i)$ é o valor desejado para a malha de controle i. Portanto, o erro estimado apresentado na equação (3.8), assumirá a seguinte forma:

$$\hat{e}(k) = Y_w(k) - \hat{Y}(k) \tag{3.14}$$

onde $Y_w(k)$ contém os vetores de trajetória de referência para cada malha de controle.

O controlador baseado em redes neurais representando a dinâmica inversa, apresenta a característica de gerar ações de controle enérgicas. Por esta razão utiliza-se um filtro na saída do controlador amenizando os incrementos nas ações de controle gerados. O filtro a ser utilizado para penalizar as ações de controle apresenta a seguinte forma:

$$u(i,k) = \lambda_i u(i,k-1) + (1-\lambda_i) u(i,k) \tag{3.15}$$

3.6 APRENDIZAGEM *ON-LINE* E *OFF-LINE* DAS *RNA'S*

Aprendizagem *off-line* é aquela realizada com um conjunto de informações passadas do processo (dados padrões), ou seja, pares de entrada e saída (Perturbações/Repostas) obtidos através de perturbações realizadas no processo. Estes dados são apresentados várias vezes à rede neural, até a representação apropriada destes padrões.

Aprendizagem *on-line* é aquela realizada com dados obtidos do processo em tempo real e estes são alimentados ao conjunto de dados de aprendizagem. Esta

aprendizagem permite a rede neural "aprender" o comportamento do processo num momento de não-linearidades fortes ou mapear uma região ainda não apresentada à rede neural, durante a sua utilização.

A aprendizagem *on-line* apresenta limitações devido ao número de iterações necessárias para atender o critério de tolerância adotado e à necessidade de aprendizagem em tempo real. Para contornar estas dificuldades, limitou-se o número máximo de iterações, usando um vetor de padrões formado por um número pré-estabelecido das últimas informações passadas de perturbações/respostas do sistema em questão. Além disto, para garantir a boa representação dinâmica do processo através de redes neurais, adotou-se uma estratégia formada por três redes neurais que atuam paralelamente (ENDER e MACIEL FILHO, 2001[a,b]). Essa estratégia foi utilizada para o controlador e modelo do processo, pois todas as redes foram submetidas à aprendizagem *on-line*.

As aprendizagens *on-line* e *off-line* são essencialmente idênticas do ponto de vista matemático, diferindo apenas na forma de obtenção dos padrões de aprendizagem. Na aprendizagem *off-line* o conjunto de padrões é fixo e na *on-line* é flutuante.

Na aprendizagem *on-line,* numa primeira etapa as redes devem ser treinadas com os dados padrões da aprendizagem (*off-line)*, conforme a Figura 3.3.

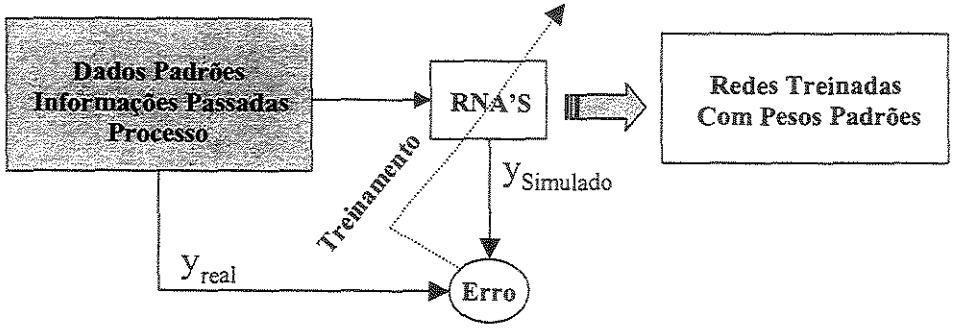


Figura 3.3 Etapa 1 - Aprendizagem Off-line (Ajuste de Pesos com Dados Padrões)

Na segunda etapa de aprendizagem, a primeira rede neural (*RNA 1*) é denominada de Padrão, pois não terá seus pesos atualizados, ou seja, seus pesos são os obtidos através

da aprendizagem *off-line* (pesos padrões). A segunda rede neural (*RNA 2*) é denominada de Base–Padrão, pois inicia a aprendizagem com os pesos padrões (obtidos na primeira etapa da aprendizagem (*off-line*) e tem os mesmos sendo atualizados ao longo do tempo através da aprendizagem *on-line*. Sempre que a *RNA 1* apresentar o menor erro quadrático em relação as outras redes, a *RNA 2* terá seus pesos substituídos pelos pesos padrões e inicia novamente o processo de ajuste de pesos (aprendizagem *on-line*). A terceira rede (*RNA 3*) é denominada de Nova, pois inicia o processo de aprendizagem com os pesos padrões e tem seus pesos sendo atualizados ao longo do tempo (Figura 3.4).

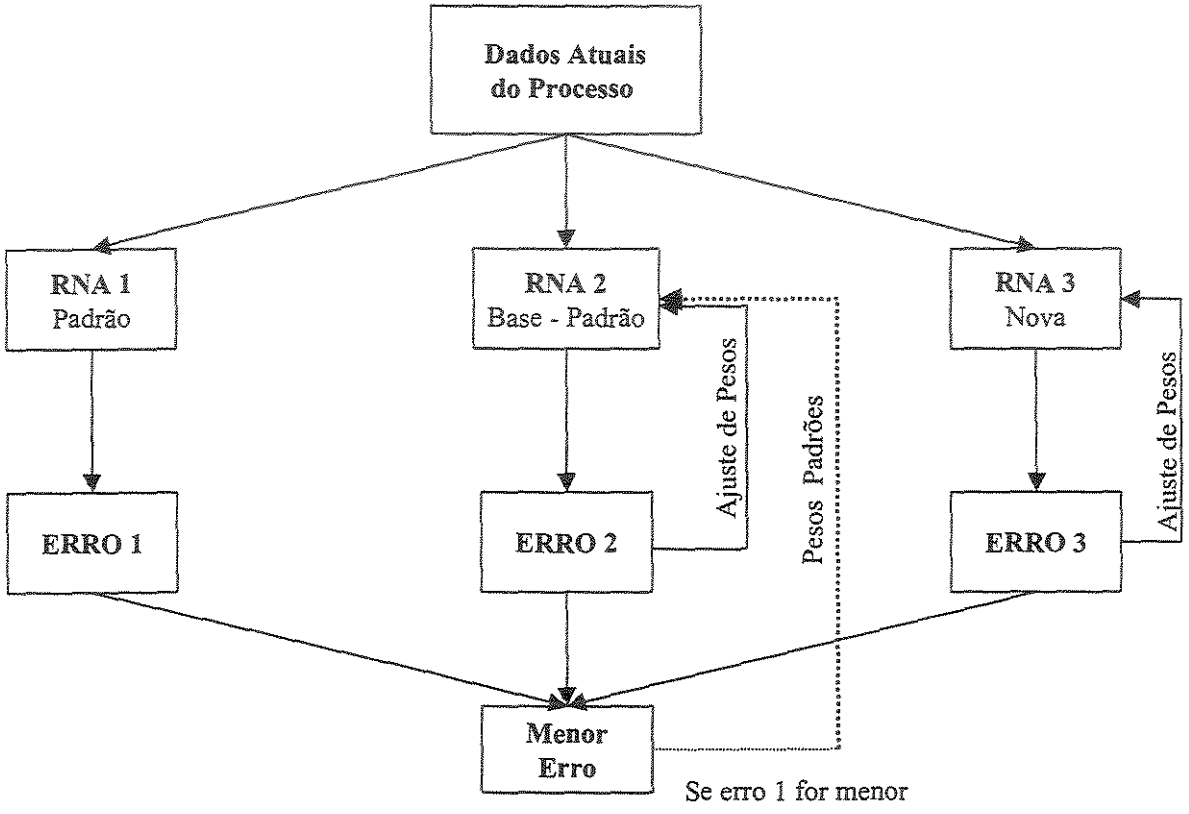


Figura 3.4 Etapa 2 – Aprendizagem On-line (Ajuste de Peso em Tempo Real)

A diferença na representação destas redes é que a rede treinada apenas com aprendizagem *off-line* (*RNA* 1) por ter em seu vetor de treinamento um conjunto de dados de várias perturbações e maior número de dados consegue representar bem o perfil do processo, pode –se dizer que tem uma maior "memória" do processo. Já a rede treinada com dados *on-line (RNA's 2* e *3)* possuem uma pequena "memória" ou seja, conseguem representar muito bem as últimas perturbações mas não guardam um histórico do processo.

Isto pode ser considerado uma limitação se o processo apresentar uma mesma perturbação, pois a rede não conseguirá lembrar desta perturbação, não irá representar o processo se não for novamente treinada. Mas este é o objetivo da aprendizagem *on-line* permitir que a rede seja treinada para novas perturbações ao longo do tempo.

3.7 CONFIGURAÇÕES DE APRENDIZAGEM *ON-LINE*

Foram propostas três configurações para avaliar a aprendizagem *on-line* das redes neurais artificiais, para predição do processo, utilizadas na estratégia de controle (CoNDEG Modificado). As alterações foram realizadas na formação dos vetores de aprendizagem e teste das redes. Na fase de teste, a rede neural que apresentar o menor erro quadrático na representação do vetor teste é utilizada naquele instante de amostragem. O objetivo destas modificações é determinar qual configuração terá melhor performance para a estratégia proposta.

3.7.1 Aprendizagem *On-line* - Configuração 1

Nesta configuração de aprendizagem utiliza-se o mesmo vetor para as fases de aprendizagem e teste das redes, formado por informações do processo (entradas/saídas), correspondentes aos padrões do tempo (t-b) até (t), onde b é um parâmetro de projeto e t o tempo atual (Figura 3.5). Em todas simulações realizadas com esta configuração foram utilizados 25 pares de informações do processo para aprendizagem e teste das redes. Esta configuração foi proposta por ENDER (2002).

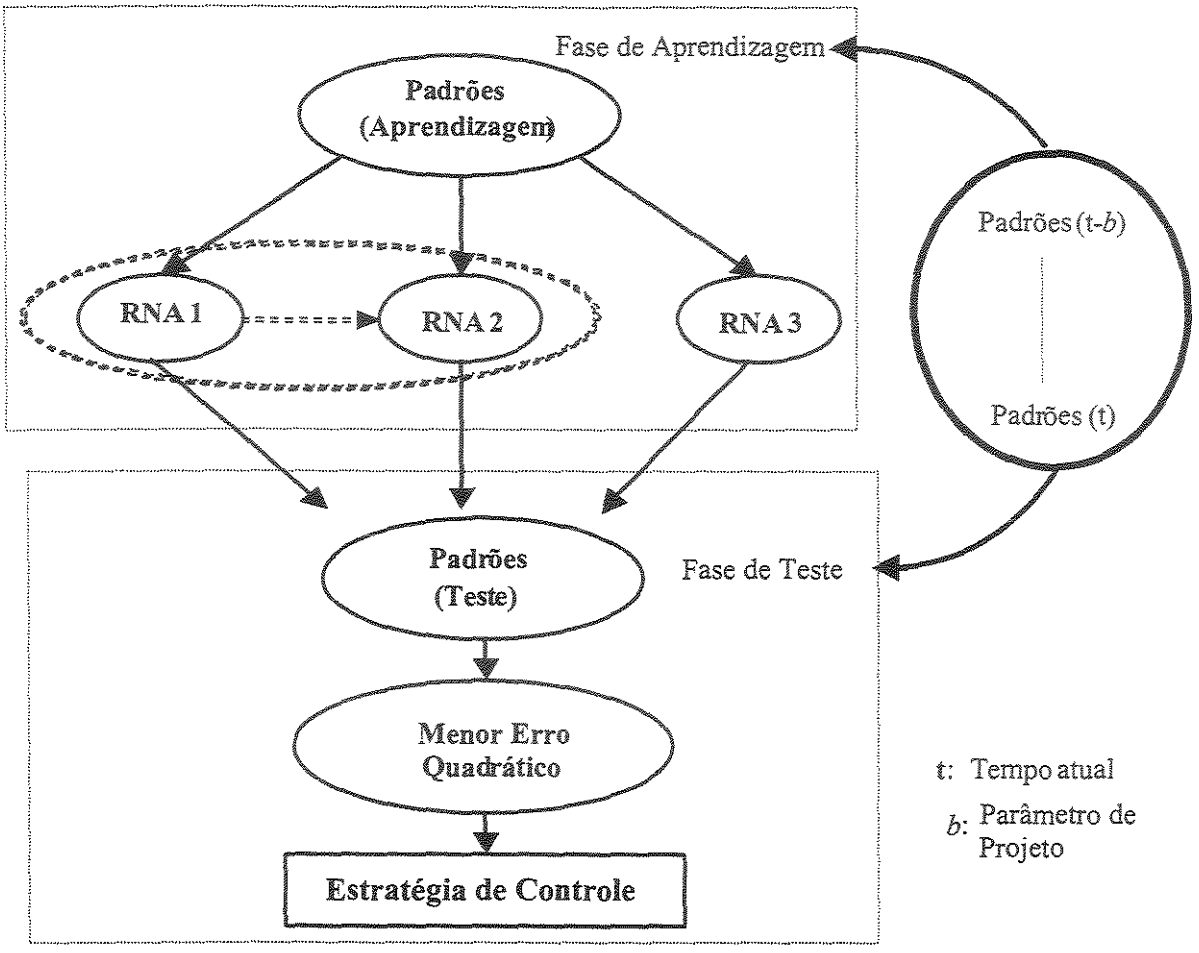


Figura 3.5 Esquema de Aprendizagem On-line para Configuração 1

3.7.2 Aprendizagem *On-line* - Configuração 2

Nesta configuração de aprendizagem são formados dois vetores para as fases de aprendizagem e teste das redes. O vetor da fase de aprendizagem é formado por um conjunto de informações do processo (entradas/saídas), correspondente aos padrões do tempo (*t-b*) até (*t-a*). O vetor da fase de teste é formado por um conjunto das ultimas informações do processo, correspondente aos padrões do tempo *(t-a)* até (*t*), onde *a* e *b* são parâmetros de projeto e *t* o tempo atual (Figura 3.6).

Em todas simulações realizadas com esta configuração foram utilizados 25 pares de informações do processo para fase de aprendizagem e outros 50 pares para fase de teste.

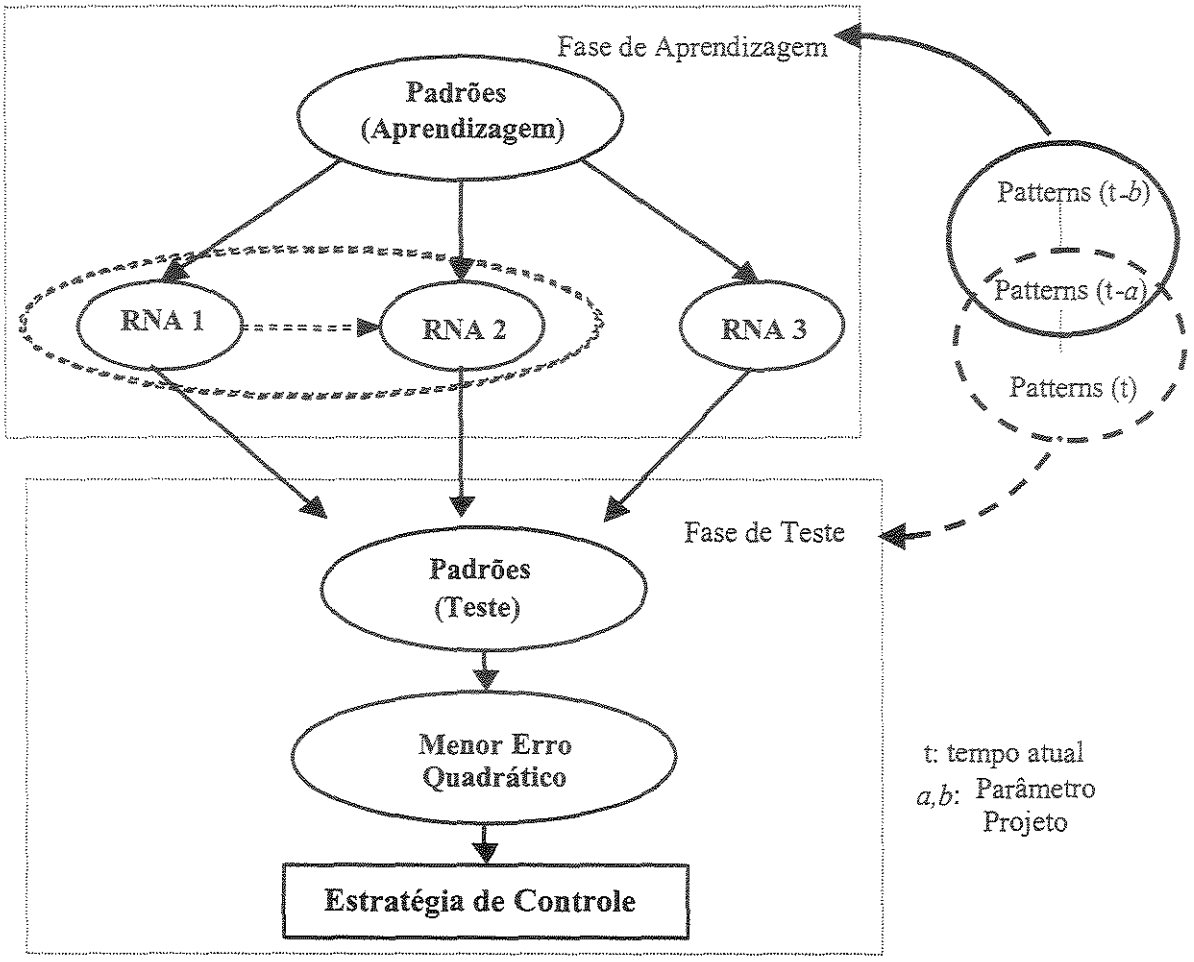


Figura 3.6 Esquema de Aprendizagem On-line para Configuração 2

3.7.3 Aprendizagem *On-line* - Configuração 3

Nesta configuração de aprendizagem o vetor teste é formado por todo conjunto de dados de informações do processo (padrões do tempo (*t-b*) até (*t*)) e o vetor de aprendizagem formado pelos últimos tempos de amostragem do processo (padrões do tempo (*t-a*) até (*t*)), conforme o esquema da Figura 3.7.

Em todas simulações realizadas com esta configuração foram utilizados 25 pares de informações do processo para fase de aprendizagem e 50 pares para fase de teste das redes, exceto quando determinado um valor diferenciado para o vetor de teste.

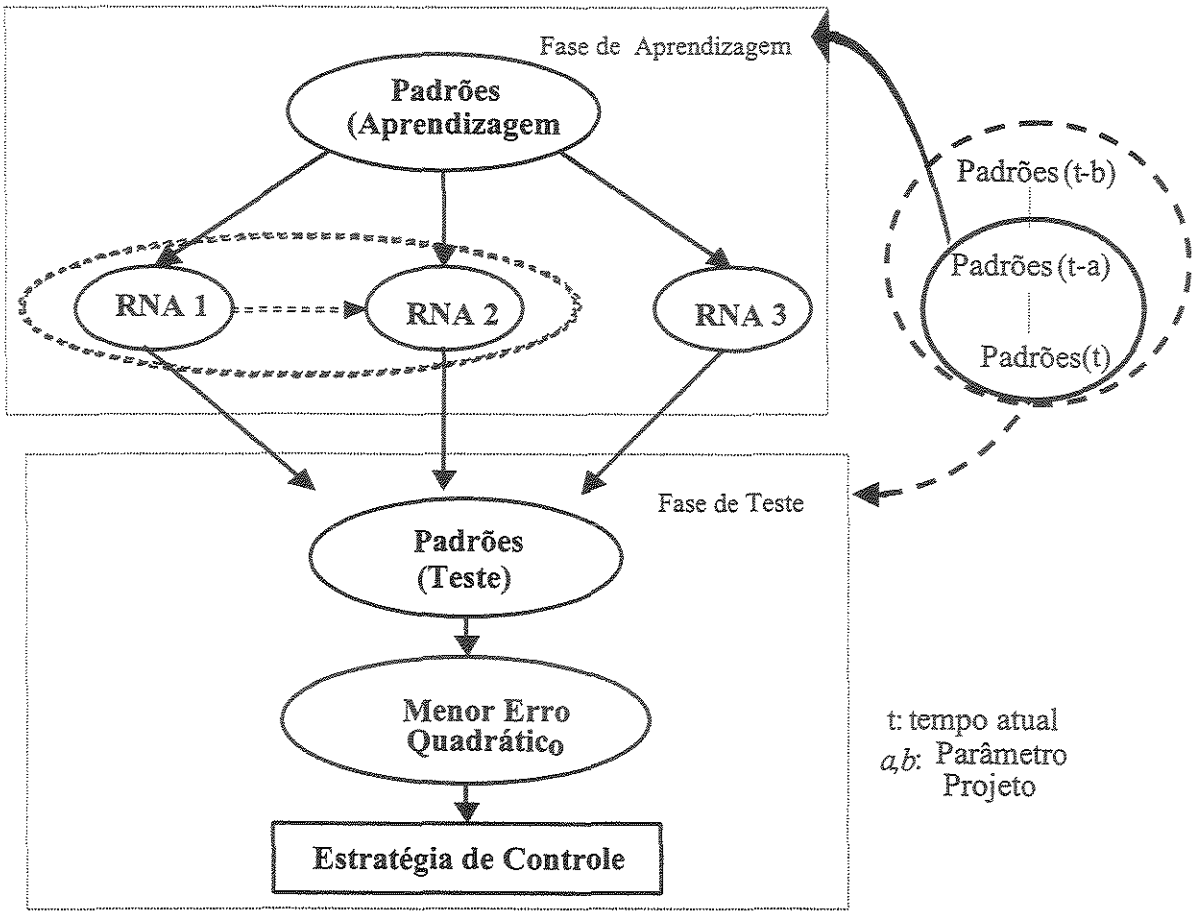


Figura 3.7 Esquema de Aprendizagem On-line para Configuração 3

3.8 CONCLUSÕES

Neste capítulo discutiu-se a estratégia de controle proposta neste trabalho, que utiliza aprendizagem *on-line* para as redes neurais e as diferentes configurações de aprendizagem propostas, diferenciadas na forma de apresentação dos padrões para as fases de aprendizagem e testes Objetivando-se uma melhor representatividade do processo estudado.

Para avaliar a estratégia utilizou-se dois processos não lineares, típicos na engenharia química, como casos de estudo: um reator de leito fixo (TOLEDO, 1999) e um processo fermentativo (COSTA, 2001), conforme descrito nos capítulos seguintes.

CASO 1 - PROCESSO EXTRATATIVO DE FERMENTAÇÃO ALCOÓLICA CONTÍNUA

4. 1 INTRODUÇÃO

O objetivo deste capítulo é apresentar o primeiro caso de estudo representado por um processo extrativo de fermentação alcoólica contínua. Primeiramente serão analisadas as características do processo e sua dinâmica para definição das propriedades das redes a serem utilizadas. Por fim serão apresentados os resultados obtidos pelo uso da estratégia de controle proposta para este processo.

4.2 PROCESSO

O etanol apresenta várias vantagens como combustível, principalmente na qualidade do ar, mas o seu uso como combustível viável substituindo a gasolina somente será possível se a sua produção for economicamente competitiva. Deste modo, há um grande interesse na otimização de todas as etapas do processo de produção do etanol (MELEIRO, 2002).

O processo fermentativo é inibido pelo etanol e uma forma de melhorar o processo é a extração seletiva deste produto, durante a fermentação, aumentando a produtividade. SILVA *et al.* (1999) propôs um processo extrativo de fermentação alcoólica com este princípio e segundo COSTA *et al.* (2001), apresentou alta produtividade comparado aos processos convencionais.

Para o desenvolvimento desta dissertação será utilizado o processo de fermentação alcoólica extrativa representado na Figura 4.1 (Meleiro, 2002).

A planta é formada por quatro unidades interligadas:

- *Fermentador* – unidade onde o etanol é produzido;

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

- *Centrífuga* – unidade responsável pela separação das fases sólida (microorganismo) e líquida (mosto e etanol);
- *Cuba* – unidade de tratamento dos microorganismos;
- *Tanque Flash* – unidade de separação da mistura etanol – mosto (fase aquosa).

O tanque flash é utilizado para extrair continuamente parte do etanol formado, aumentando desta forma a produtividade, como discutido anteriormente.

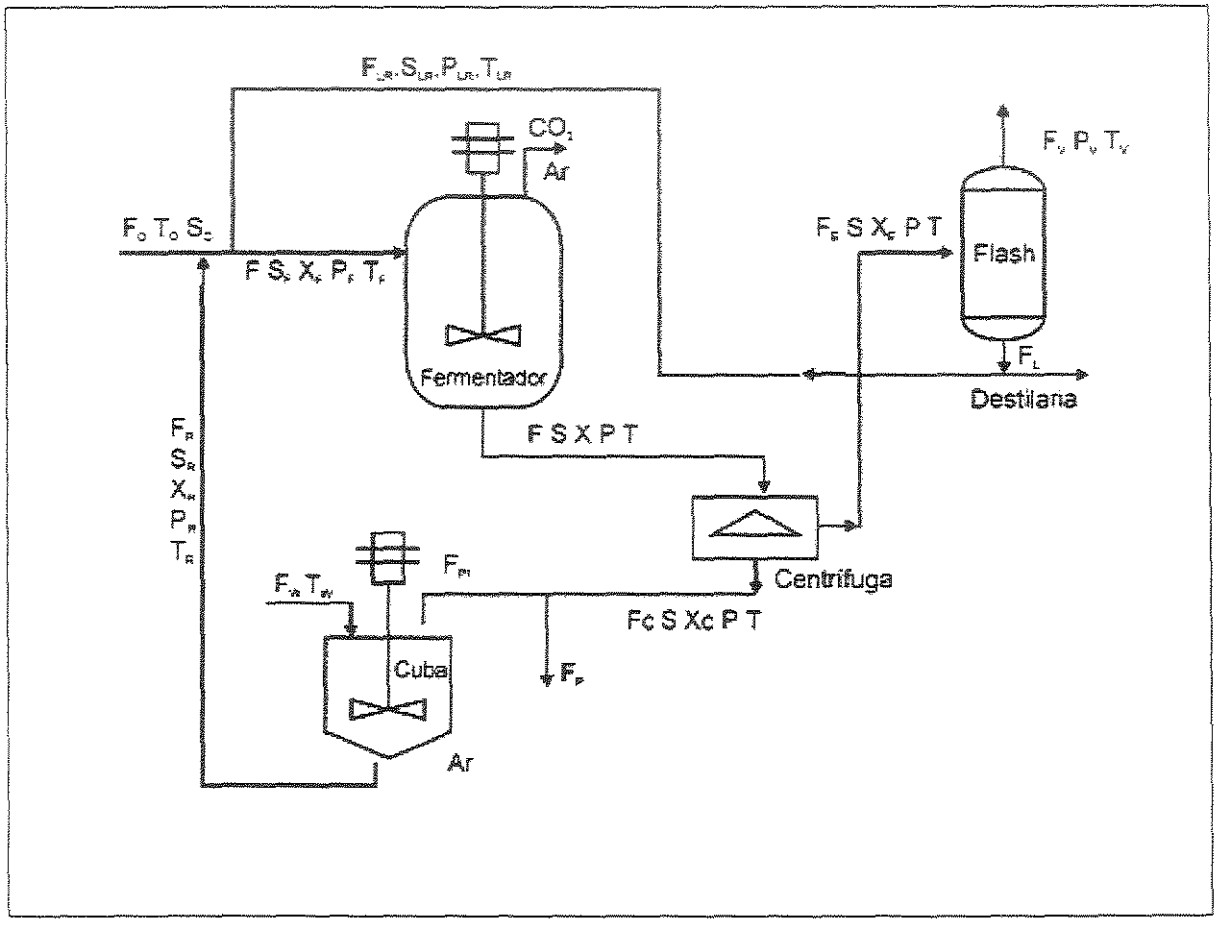


Figura 4.1 Representação esquemática da planta de fermentação alcoólica extrativa

Segundo MELEIRO (2002) o processo inicia-se com a alimentação de substrato no fermentador, que contém a concentração de microorganismos adequada para dar início

ao processo fermentativo. A corrente de saída do fermentador, contendo substrato, células e etanol, é enviada para centrífuga e separada em duas fases: fase pesada, contendo maior parte das células (microorganismos); e a fase leve, praticamente isenta de células. A fase pesada, após sofrer uma purga para eliminar parte das células é enviada para a cuba. Na cuba o creme de leveduras é diluído com água e acidulado com ácido sulfúrico, sob forte agitação, para diminuir a contaminação do processo por bactérias. A suspensão de células tratadas é enviada para o fermentador para um novo ciclo do processo.

Após o estado estacionário ser atingido o sistema de separação a vácuo é ativado e a fase leve que sai da centrífuga é enviada para o tanque flash, onde parte do etanol e separado da mistura aquosa.

A corrente vaporizada e parte da corrente líquida que deixam o flash são enviadas para uma coluna de retificação. A outra parte da corrente líquida oriunda do flash retorna ao fermentador com o objetivo de manter a concentração de etanol em um nível mínimo no qual ele apresente propriedades anti-sépticas.

4.3 Modelagem Matemática

A modelagem matemática do processo considerado consiste em balanços de massa e energia. Todos os equipamentos, com exceção do fermentador são modelados assumindo a condição de estado pseudo-estacionário. Segundo COSTA *et al.* (2002) esta aproximação é adequada porque a dinâmica do fermentador é muito lenta quando comparada com as dinâmicas das demais unidades.

O modelo matemático para o fermentador utilizado foi desenvolvido por COSTA *et al.* (2001), que considerou a queda de viabilidade celular ao longo da fermentação, de modo que a concentração de microorganismos foi dividida em duas partes: uma ativa (células viáveis) e outra inativa (células mortas). Os parâmetros cinéticos utilizados na modelagem do mesmo equipamento foram determinados experimentalmente por ATALA *et al.* (2001).

Segundo COSTA *et al.* (2001) para se determinarem as concentrações e as vazões de entrada para o fermentador, fez-se um balanço de massa global no processo assumindo as seguintes aproximações:

- As concentrações de substrato e de produto na saída da centrífuga, nas fases leve e pesada, são iguais as concentrações na saída do fermentador;
- A concentração de biomassa na corrente de reciclo é fixa. Para isto ajusta-se a vazão de água que dilui o creme de leveduras;
- A vazão de reciclo de células, cujo valor é determinado pela razão de reciclo de microorganismos é mantida constante através da variação da vazão de purga. A purga serve para retirar produtos secundários acumulados no fermentador e permite uma constante renovação celular.

De acordo com estas aproximações o modelo matemático do processo é descrito pelas equações descritas a seguir:

- Vazão de Alimentação do fermentador:

$$F = F_O + F_{LR} + F_R \tag{4.1}$$

- Vazão de Reciclo de células:

$$F_R = FR_{microrg} \tag{4.2}$$

Substituindo o valor de F_R (equação 5.2) na equação 5.1, tem-se:

$$F = \frac{F_O + F_{LR}}{(1 - R_{microrg})} \tag{4.3}$$

- Tempo de Residência do Fermentador:

$$\tau = \frac{V}{F} \tag{4.4}$$

onde V é o volume do reator.

- Balanço de Biomassa (Microorganismos) para Centrífuga:

$$FX = F_E X_E + F_C X_C \tag{4.5}$$

onde X representa a concentração de microorganismos e os índices $_E$ e $_C$ estão associados às correntes que seguem para o flash e para a cuba, respectivamente.

- Balanço Global Centrífuga:

$$F_C = F - F_E \tag{4.6}$$

- Vazão de Alimentação do Tanque Flash:

$$F_E = \frac{F(X - X_C)}{X_E - X_C} \tag{4.7}$$

- Vazão de Alimentação da Cuba:

$$F_{C1} = \frac{F_R X_R}{X_C} \tag{4.8}$$

- Vazão de Água de Diluição que Alimenta a Cuba:

$$F_W + F_R + F_{C1} \tag{4.9}$$

- Concentração de Substrato na Corrente de Reciclo:

$$S_R = \frac{F_{C1} S}{F_R} \tag{4.10}$$

- Concentração de Etanol na Corrente de Reciclo:

$$P_R = \frac{F_{C1} P}{F_R} \tag{4.11}$$

- Balanço de Massa no Tanque Flash:

$$F_E = F_V + F_L \tag{4.12}$$

$$F_E X_{Ei} = F_V y_i + F_L x_i \tag{4.13}$$

- Equilíbrio Líquido Vapor da Mistura Etanol e Água:

$$k_{eqi} = \frac{y_i}{x_i} = \gamma_i \frac{P_i^{sat}}{p} \tag{4.14}$$

onde o valor de P_i^{sat} é dado pela equação de Antoine e o valor de γ_i é calculado com ao modelo *NRTL* (SILVA, 1998).

As concentrações de substrato, etanol e microorganismos presentes na corrente que retorna para o fermentador, vindas do tanque flash seguem as equações 4.15, 4.16 e 4.17, respectivamente.

$$S_{LR} = \frac{F_E S}{F_L} \tag{4.15}$$

$$P_{LR} = \frac{F_E P - F_V P_V}{F_L} \tag{4.16}$$

$$X_{LR} = \frac{F_E X_E}{F_L} \tag{4.17}$$

- Vazão de Reciclo do tanque Flash;

$$F_{LR} = R_{flash} F_L \tag{4.18}$$

onde R_{flash} é a razão de reciclo do tanque Flash.

As concentrações de substrato, etanol e microorganismos na corrente de alimentação do fermentador são dadas pelas equações 4.19, 4.20 e 4.21, respectivamente.

$$S_F = \frac{F_O S_O + F_{LR} S_{LR} + F_R S_R}{F} \tag{4.19}$$

$$P_F = \frac{F_R P_R + F_{LR} P_{LR}}{F} \tag{4.20}$$

$$X_F = \frac{F_R X_R + F_{LR} X_{LR}}{F} \tag{4.21}$$

Os balanços de energia em torno da Cuba e do Fermentador fornecem as temperaturas nas correntes de reciclo da Cuba e de alimentação do Fermentador, respectivamente:

$$T_R = \frac{F_W T_W + F_{C1} T}{F_R} \tag{4.22}$$

$$T_F = \frac{F_O T_O + F_R T_R + F_{LR} T_{LR}}{F} \tag{4.23}$$

Assumindo que o volume do reator é constante, as equações de balanço de massa e energia para o fermentador, utilizando o modelo intrínseco, são escritas da seguinte forma:

$$\frac{dX_v}{dt} = r_x - r_d - \frac{F}{V}\left(X_v - X_{vF}\right) \tag{4.24}$$

$$\frac{dX_d}{dt} = r_d - \frac{F}{V}\left(X_d - X_{dF}\right) \tag{4.25}$$

$$\frac{d\left(1 - \frac{X_t}{\rho}\right)SV}{dt} = F(S_F - S) - r_S V \tag{4.26}$$

$$\frac{d\left[\left(1 - \frac{X_t}{\rho}\right)PV + \frac{X_t}{\rho}\Gamma PV\right]}{dt} = r_p V + F(P_F - P) \tag{4.27}$$

$$\frac{dT}{dt} = \frac{F}{V}(T_F - T) + \frac{\Delta Hr_s}{\rho_m Cp} \tag{4.28}$$

onde:

X_v é a concentração de células viáveis;

X_d é a concentração de células mortas;

X_t é a concentração de células totais;

S é a concentração de substrato;

Sendo que o indicador F indica na alimentação;

P é a concentração de produto;

T é a temperatura do fermentador;

ΔH é o calor de reação ($2{,}167.10^3$ J/kg de ART);

ρ_m é a densidade do meio fermentativo (1000 kg/m^3);

Cp é o calor específico ($4{,}183.10^3$ J/kgK);

F é a vazão de alimentação, V é o volume do fermentador, ρ é a relação de massa seca pelo volume de massa úmida e Γ é a relação entre as concentrações de etanol intra e extra-celular.

- Taxa Cinética de Crescimento

$$r_x = \mu_{\max} \frac{S}{K_S + S} \exp(-K_i S)\left(1 - \frac{X_t}{X_{\max}}\right)^m \left(1 - \frac{P_t}{P_{\max}}\right)^n X_v \tag{4.29}$$

- Taxa de Morte

$$r_d = K_{dT} X_v \exp(K_{dP} P) \tag{4.30}$$

- Consumo de Substrato

$$r_s = \frac{r_x}{y_x} + m_x X_v \tag{4.31}$$

- Formação de Produto

$$r_p = Y_{px} r_x + m_p X_v \tag{4.32}$$

Os parâmetros cinéticos utilizados foram determinados experimentalmente por ATALA *et al.* (2001) e são descritos na Tabela 4.1.

Tabela 4.1 Parâmetros Cinéticos em Função da Temperatura (em °C)

Parâmetros	***Expressão ou Valor***
μ_{max}	$1{,}57 \exp\left(\frac{-41{,}47}{T}\right) - 1{,}29.10^{4} \exp\left(\frac{-431{,}4}{T}\right)$
X_{max}	$-0{,}3279\,T^2 + 18{,}484\,T - 191{,}06$
P_{max}	$-0{,}4421\,T^2 + 26{,}41T - 279{,}75$
B	$2{,}704 \exp(-0{,}1225\,T)$
Y_{PX}	$0{,}2556 \exp(-0{,}1086\,T)$
K_s	4,1
K_i	$1{,}393.10^{-4} \exp(-0{,}1004T)$
m_p	0,1
m_x	0,2
m	1
n	1,5
K_{dP}	$7{,}421.10^{-3}\,T^2 - 0{,}4654\,T + 7{,}69$
K_{dT}	$4.10^{13} \exp\left(\frac{41947}{1{,}987\,(T + 273{,}15)}\right)$
ρ	390
Γ	0,78

4.4 COMPORTAMENTO ESTACIONÁRIO

Os parâmetros estacionários adotados para as simulações são apresentados na Tabela 4.2, que seguem MELEIRO (2002).

Tabela 4.2 Parâmetros Estacionários Adotados

Variável	*Valor Estacionário*
To	30 ° C
So	180 kg/m^3
Fo	100 kg/m^3
$R_{microrg}$	0,35
R_{flash}	0,4
P	5,49 kg/m^3
S	41,12 kg/m^3
X	34,81 kg/m^3

4.5 COMPORTAMENTO DINÂMICO

Para observar o comportamento dinâmico do fermentador foram feitas perturbações positivas e negativas nas variáveis de entrada: vazão de alimentação (Fo), razão de reciclo de microorganismo ($R_{microrg}$), razão de reciclo do tanque flash (R_{flash}), temperatura de alimentação (To) e Concentração de substrato na alimentação (So).

A Tabela 4.3 mostra o valor de cada variável após as perturbações positivas (+) e negativas (-) dos valores estacionários.

Tabela 4.3 Valores das Variáveis de Entrada após Perturbações Especificadas

Variável	*Perturbação (+)*	*Estacionário*	*Perturbação (-)*
To	35° C	30° C	25° C
So	200 kg/m^3	180 kg/m^3	160 kg/m^3
Fo	120 m^3/h	100 m^3/h	80 m^3/h
$R_{microrg}$	0,2	0,35	0,5
R_{flash}	0,2	0,4	0,5

Nas Figuras 4.2, 4.3 e 4.4, pode se observar os perfis de concentração do substrato (S), produto (P) e microorganismo (X) para as perturbações positivas dos valores estacionários, conforme descrito na Tabela 4.3, respectivamente.

Nas Figuras 4.5, 4.6 e 4.7 pode se observar os perfis de concentração do substrato (S), produto (P) e microorganismo (X) para as perturbações negativas dos valores estacionários, conforme descrito na Tabela 4.3, respectivamente.

Nestas Figuras pode-se observa o comportamento dinâmico do fermentador para as perturbações impostas nas variáveis de entrada.

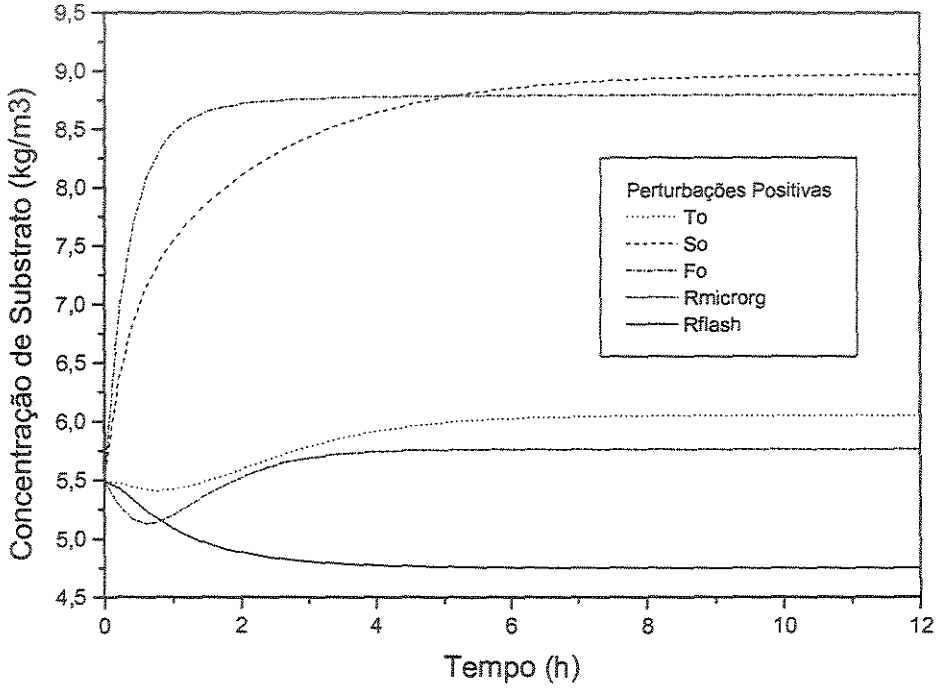


Figura 4.2 Perfil da Concentração de Substrato (S) para Perturbações Positivas

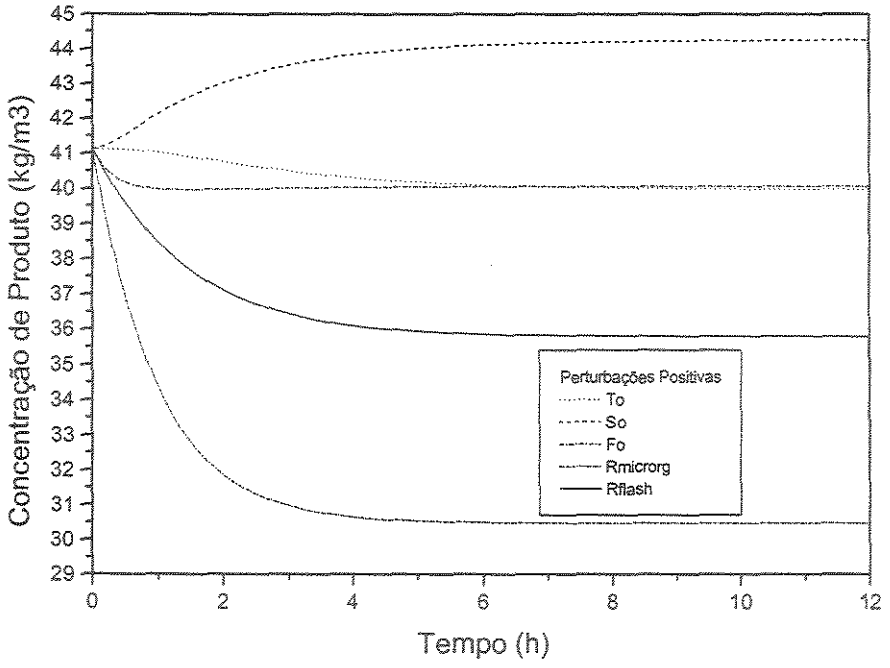


Figura 4.3 Perfil da Concentração de Produto(P) para Perturbações Positivas

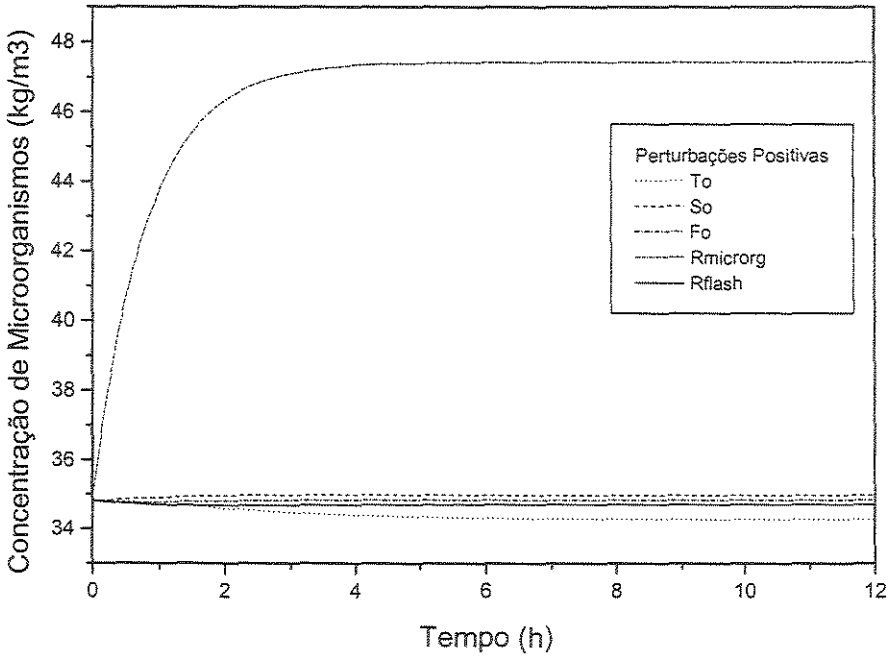


Figura 4.4 Perfil da Concentração de Microorganismos (X) para Perturbações Positivas

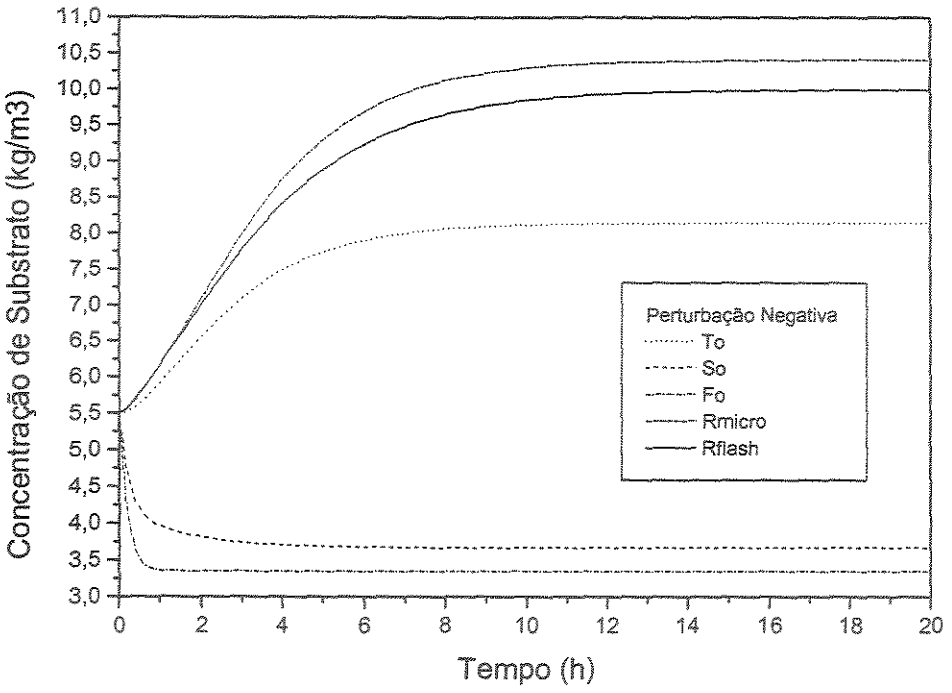


Figura 4.5 Perfil da Concentração de Substrato (S) para Perturbações Negativas

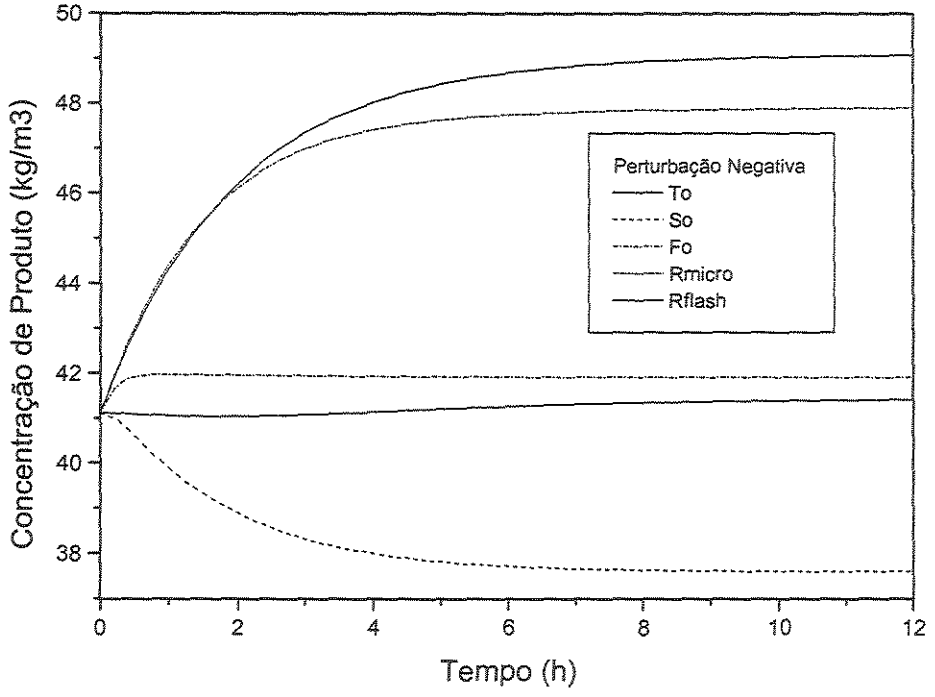


Figura 4.6 Perfil da Concentração de Produto(P) para Perturbações Negativas

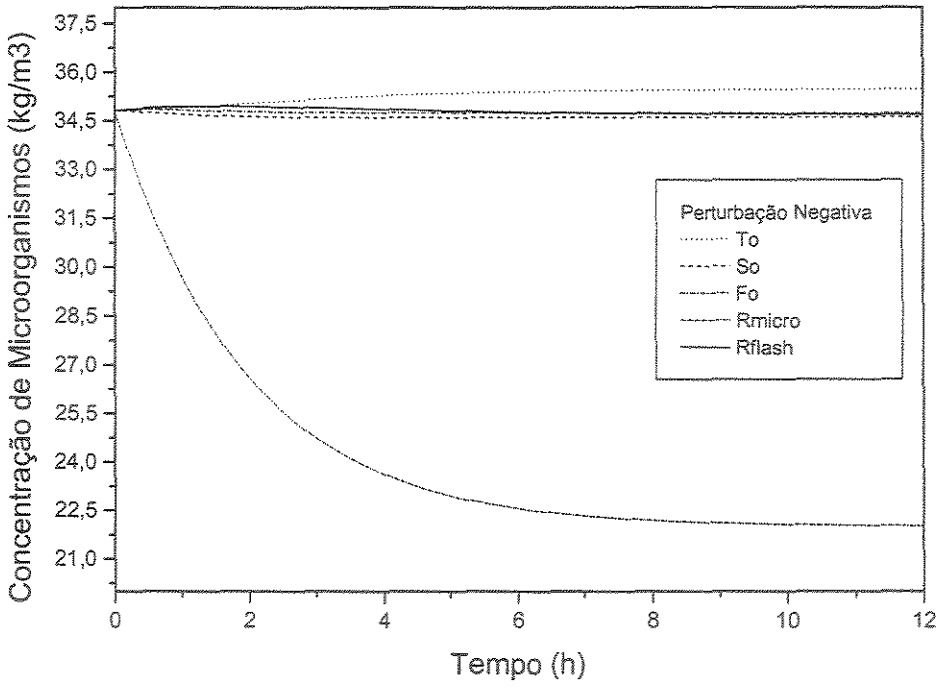


Figura 4.7 Perfil da Concentração de Microorganismos (X) para Perturbações Negativas

Nas Figuras 4.2 até 4.7 pode-se observar a dinâmica do processo estudado bem como a presença de não linearidades.

4.6 ANÁLISE DE SENSIBILIDADE DOS PARÂMETROS DE PROCESSO – PLANEJAMENTO FATORIAL

Para determinação das variáveis de entrada, saída e perturbação para este processo fermentativo SILVA *et al.* (1999) utilizaram os métodos de planejamento fatorial e análise da superfície de resposta gerada a partir de dados de simulação para determinar as variáveis de processo mais relevantes.

COSTA *et al.* (2001) também fez um estudo completo utilizando a mesma técnica combinada com a simulação dinâmica do processo para determinar como as variáveis de saída (controladas) variam no tempo de acordo com mudanças impostas nas variáveis de entrada, sendo estas variáveis manipuladas e possíveis perturbações.

De acordo com estes estudos foram escolhidas as principais variáveis de interesse para este caso de estudo, conforme descrito a seguir.

Variáveis de Entrada:

- Vazão da Corrente de Alimentação: Fo (m^3/h);
- Razão de Reciclo de Microorganismos: $R_{microrg}$ (adimensional);
- Razão de Reciclo do Tanque Flash: R_{flash} (adimensional).

Variáveis de Saída:

- Concentração de Etanol no Fermentador: P (kg/m^3);
- Concentração de Substrato no Fermentador: S (kg/m^3);
- Concentração de Microorganismos no Fermentador: X (kg/m^3).

Variável Perturbação:

- Concentração de Substrato na Corrente de Alimentação: So (kg/m^3);
- Temperatura na Corrente de Alimentação: To (°C).

O período de amostragem foi escolhido em função do tempo médio de análise de um cromatógrafo (equipamento utilizado para medir as concentrações de ART e etanol), pois o turbidímetro que mede a concentração celular opera em intervalos pequenos. Segundo ANDRIETA (1994) doze minutos é o tempo de amostragem adequado para o processo.

4.7 ARQUITETURA E DADOS PARA TREINAMENTO DAS *RNA'S*

Conforme descrito no Capítulo 2 a definição da arquitetura das redes é uma etapa importante, para este caso de estudo utilizou-se para todas as redes 3 camadas (entrada, intermediária,saída), sendo a camada intermediária composta por 10 neurônios, seguindo a

lei de LOESCH (1996). Esta arquitetura foi escolhida após algumas simulações pois apresentou resultados eficientes na representação dos resultados.

Para treinamento das redes utilizadas para este caso de estudo foram realizadas alterações nos valores das variáveis manipuladas (Fo, R_{micro} e R_{flash}) dentro de suas faixas de operação, variando individualmente e depois de forma conjunta cada variável, conforme mostra a Figura 4.8.

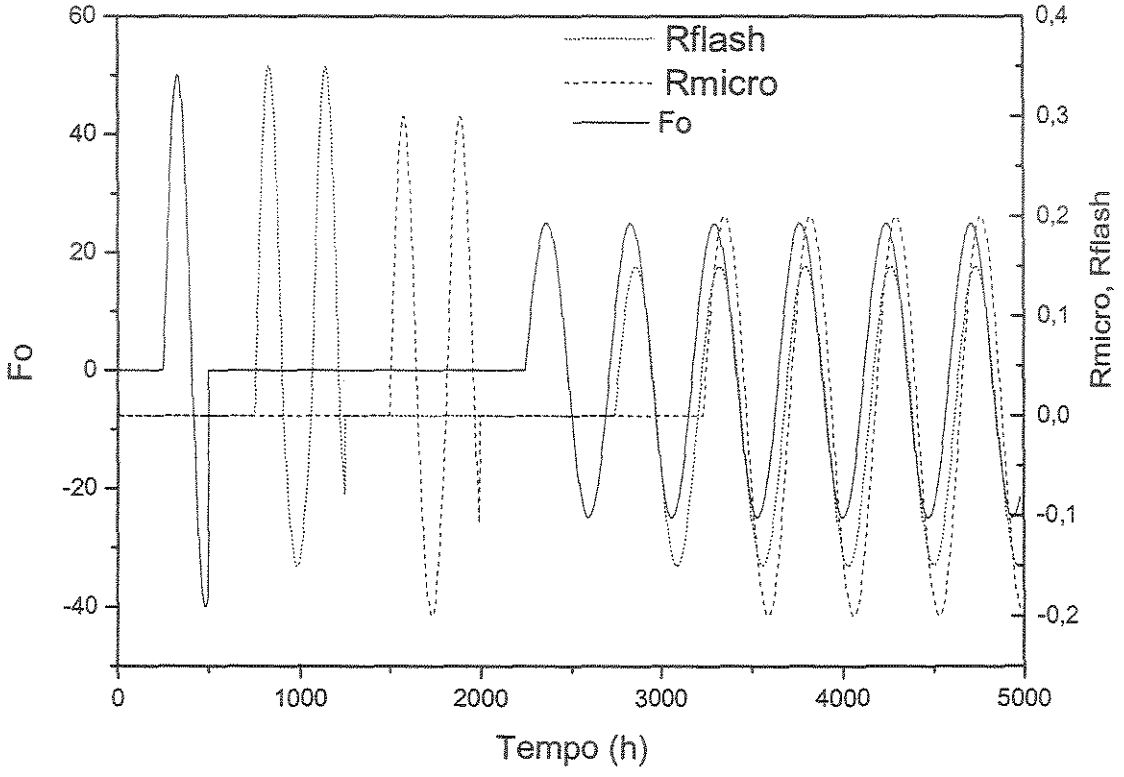


Figura 4.8 Perturbações usadas para Treinamento das Redes

4.8 Resultados Para o Caso de Estudo 1 – Processo Fermentativo

Os resultados do controle do processo fermentativo utilizando o CoNDEG Modificado, descrito no Capítulo 3, são apresentados neste Capítulo. Para verificar o

desempenho da estratégia, foram utilizados problemas dos tipos: servo, regulador e estocástico.

Como problema regulador foram usadas as variáveis perturbações So e To. Estas foram submetidas a oscilações dentro de suas faixas de operação assumindo valores diferentes ao longo do tempo em todas as simulações, gerando uma oscilação no processo e a necessidade de um controlador que seja robusto e efetivo. As Figuras 4.9 e 4.10 descrevem as perturbações cargas impostas em So e To ao longo do tempo, respectivamente.

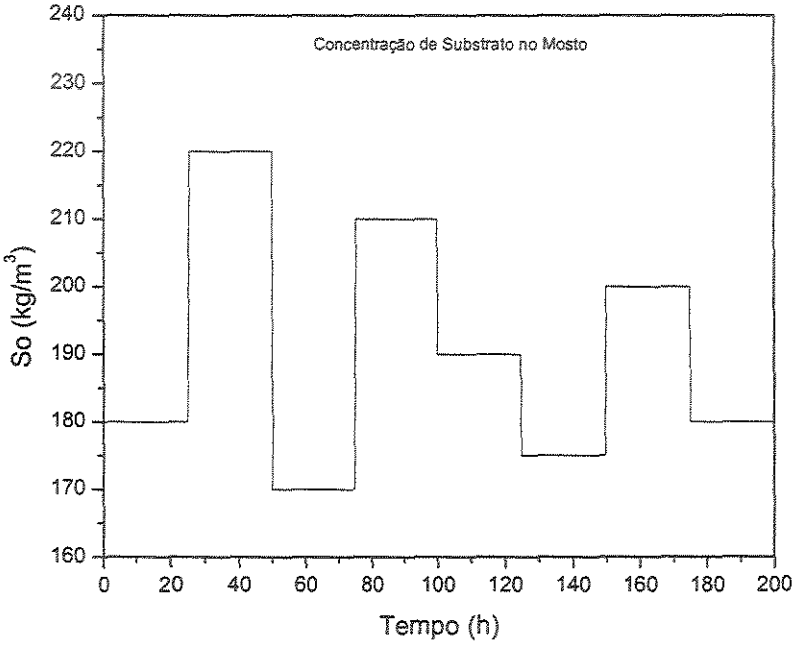


Figura 4.9. Perturbações No Processo Fermentativo para Variável So

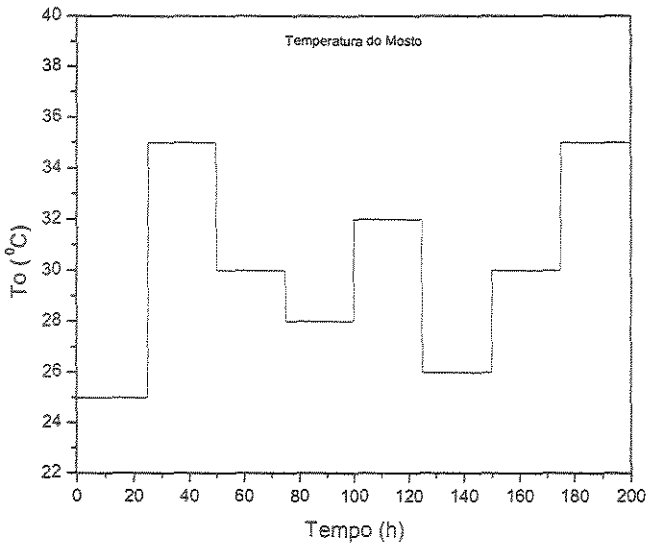


Figura 4.10. Perturbações No Processo Fermentativo para Variável To

A Figura 4.11 descreve o comportamento do sistema para as perturbações descritas em malha aberta gerando uma oscilação no processo e a necessidade de um controlador que seja robusto e efetivo.

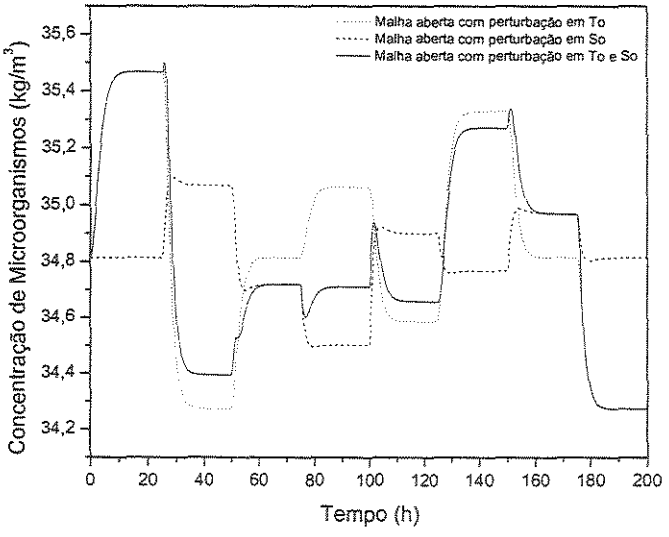


Figura 4.11.Resposta para Variável X para Perturbações em So e To (Malha Aberta)

Como problema servo, foram usadas as seguintes variáveis manipuladas: concentração de substrato (S), concentração de produto (P) e concentração de microorganismos (X). Os desvios nas respectivas referências estão apresentadas na Tabela 4.4.

Tabela 4.4 Desvios nas Referências do Processo Fermentativo

Tempo (h)	*S*	*P*	*X*
0-50	4.5	4.0	2.2
50-100	-2.5	-1.0	-2.8
100-150	2.5	2.0	1.2
150-200	0	1.0	0.0

4.8.1 Escolha da Estrutura da Rede Usada na Estratégia

Para determinar as melhores redes a serem usadas na estratégia, foram avaliadas redes treinadas com diferentes arquiteturas, foram analisados, seguindo a opção de aprendizagem com configuração 3:

- *Número de Passados* – Foram avaliados os passados 3 e 5;
- *Função de Ativação* – Foram avaliadas as funções de ativação Sigmoidal e tangente hiperbólica;
- *Parâmetros Mensuráveis e Não Mensuráveis.*

Uma rede treinada com parâmetros mensuráveis é aquela que possui em seus pares de treinamento algum histórico das possíveis perturbações do tipo carga que possam ocorrer ao longo do processo, neste caso de estudo as variáveis foram *So* e *To*. Uma rede treinada com parâmetros não mensuráveis não possui nenhum histórico das perturbações em seus pares de treinamento, e sim, somente as variáveis manipuladas e controladas.

Foram analisadas as respostas para todas as variáveis manipuladas: vazão de alimentação (*F*), razão de reciclo de microorganismos ($R_{microrg}$) e razão de reciclo do tanque flash (R_{flash}), como todas variáveis apresentaram respostas semelhantes escolheu-se a variável $R_{microrg}$, que teve uma maior variação com as modificações realizadas, para demonstrar as redes escolhidas.

As respostas obtidas da variável $R_{microrg}$ na estratégia para redes treinadas com as duas funções de ativação para o processo e função fixa (tangente hiperbólica) para o controlador, parâmetros mensuráveis e não mensuráveis com passados três e cinco são descritas nas Figuras 4.12 e 4.13, respectivamente.

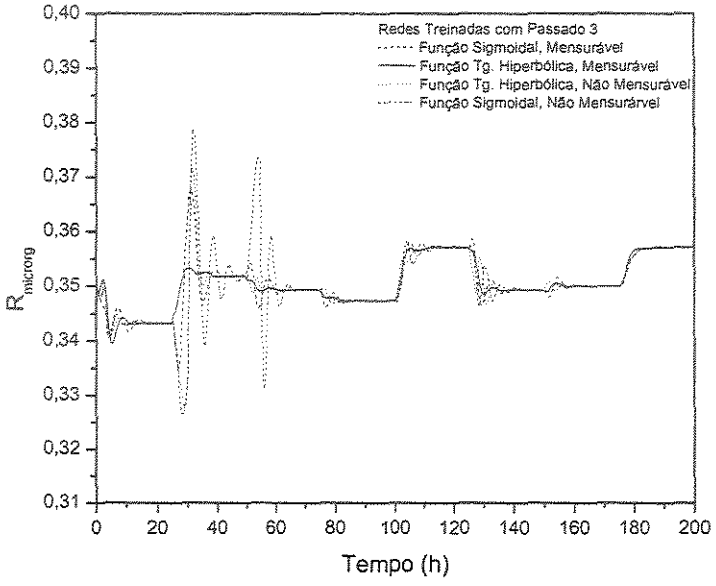


Figura 4.12 Redes Treinadas com Passado Três

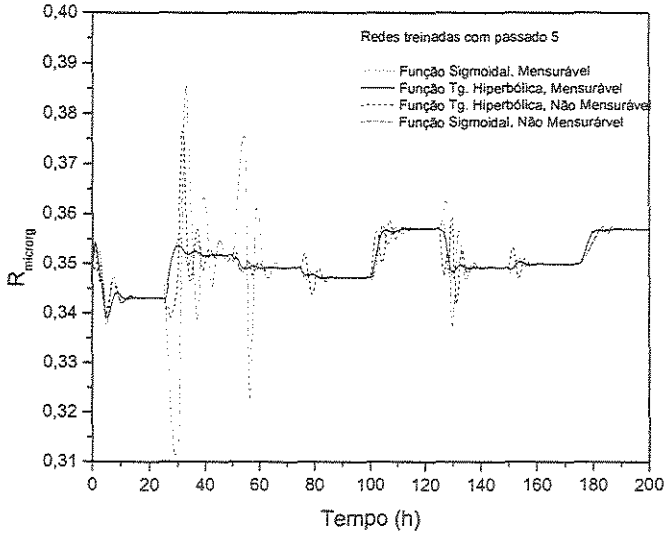


Figura 4.13 Redes Treinadas com Passado Cinco

Observa-se que para os dois passados utilizados as redes treinadas com parâmetros mensuráveis e função de ativação sigmoidal para o processo apresentaram os melhores resultados, pois não apresentaram grandes variações para a variável estudada.

A Figura 4.14 traz uma análise das redes treinadas com parâmetros mensuráveis e função de ativação sigmoidal com passados três e cinco, a rede treinada com passodo (β) três possui uma melhor performance comparada com a treinada com o passado cinco que mostrou-se mais oscilatório e o passado três possui a vantagem de requerer um menor tempo computacional, pois o número e pesos da rede neural é reduzido.

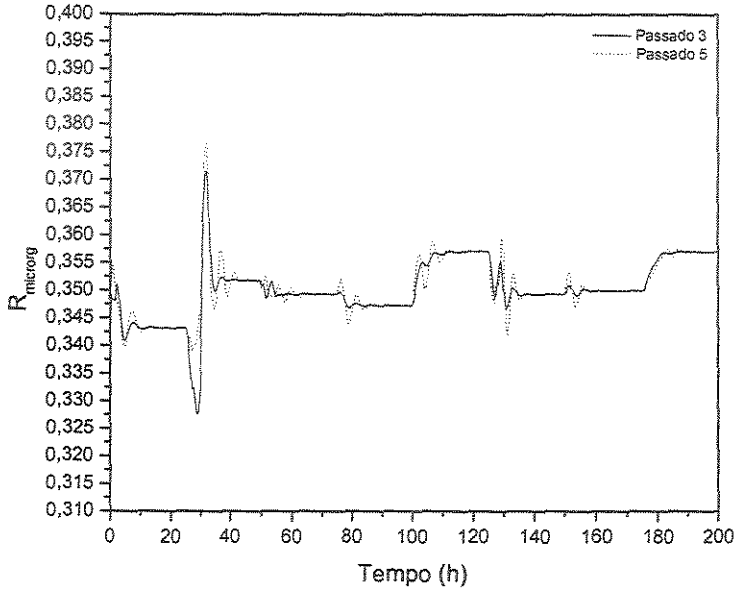


Figura 4.14 Comparação das Redes com Passados Três e Cinco

A Figura 4.15 traz a resposta da variável $R_{microrg}$ usando para o treinamento das redes do controlador e processo as funções de ativação tangente hiperbólica e sigmoidal, respectivamente, e passado β=3. Nesta análise da função de ativação observou-se que as redes usadas para representar o modelo do processo apresentam uma melhor performance quando é utilizada a função sigmoidal; para o controlador as redes treinadas com a função tangente hiperbólica apresentaram uma maior estabilidade.

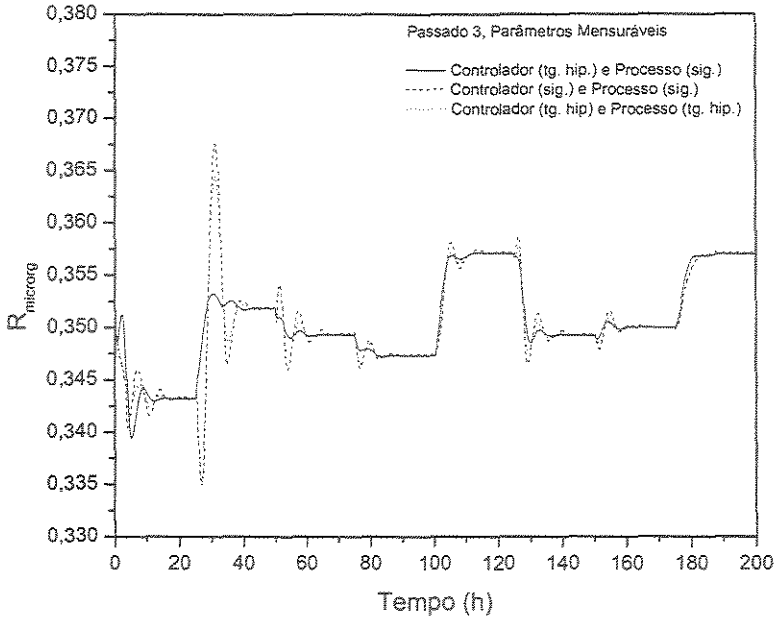


Figura 4.15 Análise da Função de Ativação para Redes do Controlador e Processo

Para verificar a "memória" das redes com aprendizagem *on-line* e *off-line*, compararam-se os resultados obtidos por duas redes na representação das concentrações de substrato, produto e microorganismos para problema regulador (perturbações em So e To) obtidos na predição da malha fechada na rotina de otimização do controlador. Considerando um horizonte passado β=5. Foram usadas: uma rede com os pesos da aprendizagem *off-line* e outra rede com os pesos da aprendizagem *on-line* até o tempo de 124 horas. Os resultados são apresentados nas Figuras 4.16 à 4.18, respectivamente.

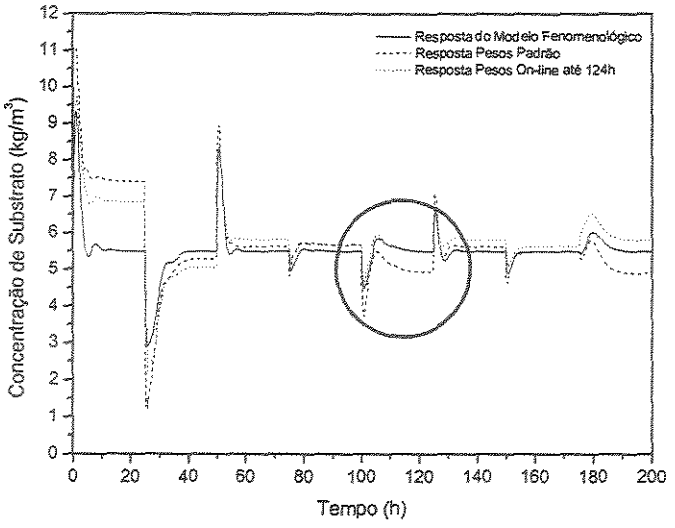


Figura 4.16 Teste para Memória das Redes para Variável S

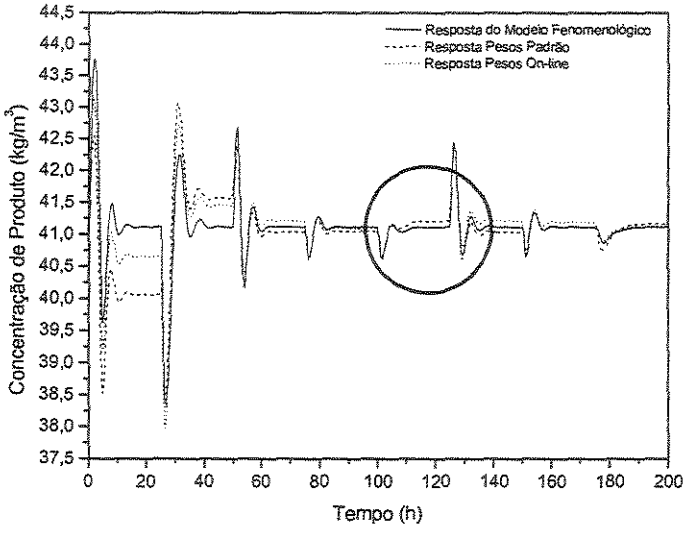


Figura 4.17 Teste para Memória das Redes para Variável P

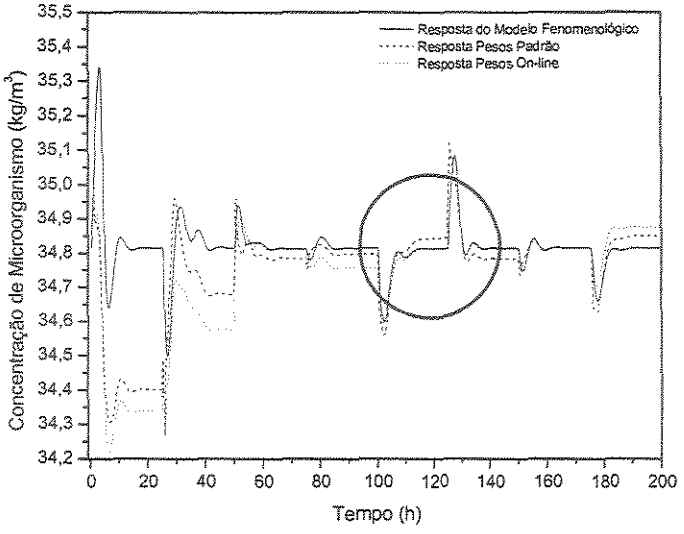


Figura 4.18 Teste para Memória das Redes para Variável X

Pode-se observar nas Figuras anteriores que as respostas geradas pela rede treinada com aprendizagem *on-line* apresentam os melhores resultados para a última faixa de tempo de treinamento (tempo de 124 horas), ou seja, esta rede consegue representar com maior eficiência o processo para o tempo que foi treinada, comparada à rede com aprendizagem *off-line*. Nas demais regiões é observado melhor resultado da rede neural padrão, devido ao melhor mapeamento das respostas obtidas na aprendizagem *off-line*.

As Figuras 4.19 à 4.20 mostram os resultados obtidos para rede treinada *on-line* comparada com a resposta real para o processo. Pode-se verificar que a rede conseguiu acompanhar as variações ao longo do processo.

Na Figura 4.19, pode-se verificar no tempo 27 horas um desvio considerável na concentração de substrato do seu valor real. Isto ocorreu devido à perturbação acentuada imposta no processo. Isto não pode ser considerado uma limitação para utilização da estratégia de controle proposta, pois nesta situação outra rede neural (Padrão ou Base-Padrão) assumiria a predição na rotina de ajuste dos pesos da rede neural do controlador.

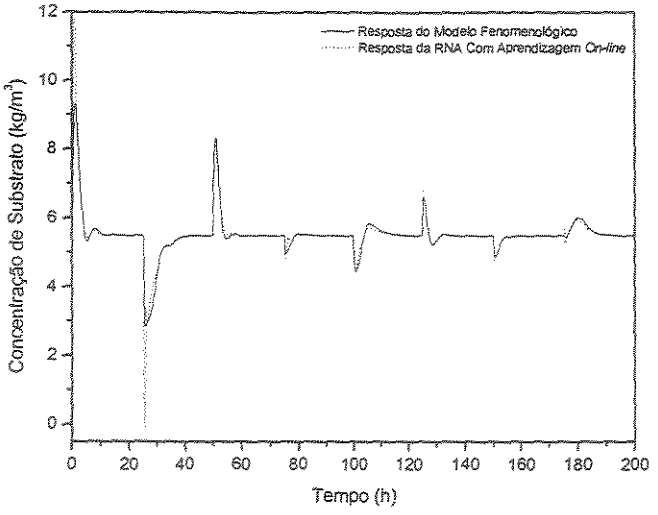


Figura 4.19 Resultado Comparativo Resposta Real e Rede Neural para variável S

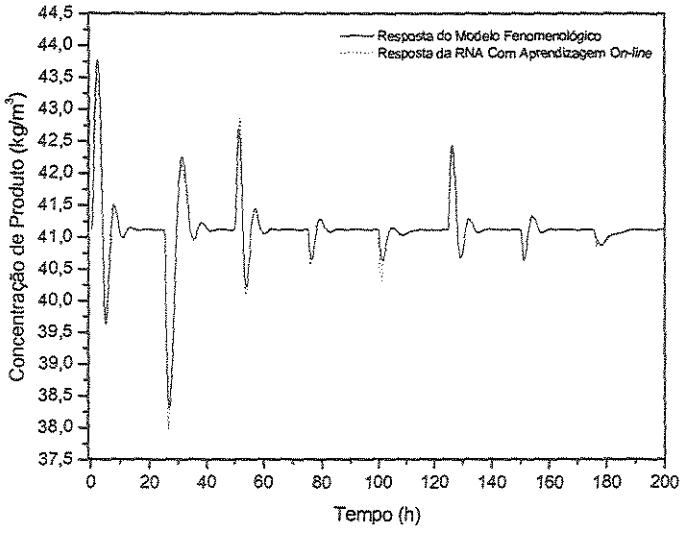


Figura 4.20 Resultado Comparativo Resposta Real e Rede Neural para variável P

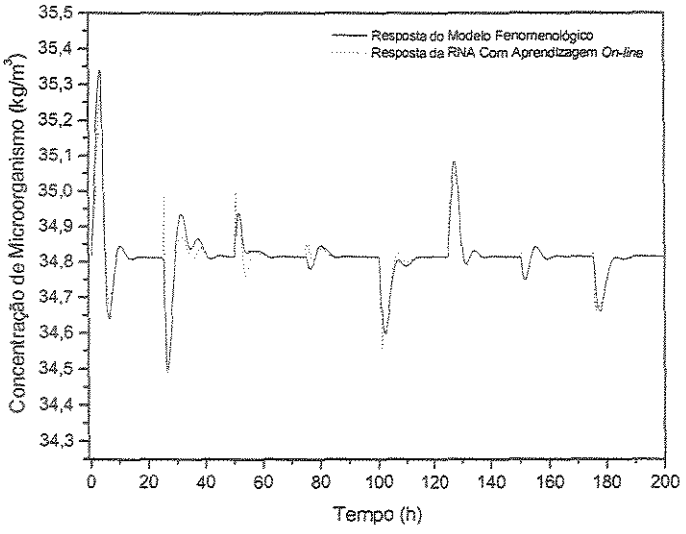


Figura 4.21 Resultado Comparativo Resposta Real e Rede Neural para variável X

4.8.2 Resultados Comparativos das Três Configurações de Aprendizagem

Conforme descrito no Capítulo 3, foram utilizadas três configurações para aprendizagem *on-line* objetivando definir qual configuração apresenta melhor resultado para a estratégia proposta.

Inicialmente estudou-se separadamente os problemas do tipo servo para cada variável (Tabela 4.4) e regulador para as variáveis perturbações, para posteriormente trabalhar com os problemas tipo servo e regulador de forma simultânea.

Para os problemas puramente servos não se obteve grandes variações entres as configurações propostas, mas dentre eles, pode-se observar uma pequena variação nas variáveis de controle P e X quando foi utilizado o problema do tipo servo em S para a configuração 2. Isto é devido a diferença dos vetores de aprendizagem e teste utilizados para aprendizagem das redes.

Os problemas do tipo reguladores (So e To) apresentaram comportamento semelhante entre as configurações propostas, mas pode-se observar uma maior variação

para a configuração 2, nestes problemas reguladores do que nos tipo servos propostos. Isto porque as perturbações foram bastante fortes para teste da estratégia.

Como o objetivo principal é avaliar qual configuração tem uma melhor aplicabilidade na estratégia proposta, apresentou-se somente os resultados na forma de Figura para as perturbações acopladas de problemas do tipo regulador (para So e To juntos) e servo para cada variável proposta, conforme detalhado nas seções seguintes.

Em cada seção serão apresentados os resultados das comparações entre as três configurações para o problema do tipo servo para as variáveis S, P e X, respectivamente.

Serão apresentadas as Figuras que descrevem os resultados para as variáveis controladas, variáveis manipuladas e para as redes que representam o controlador e processo, respectivamente para cada configuração proposta.

4.6.2.1 Problema Servo para Variável S

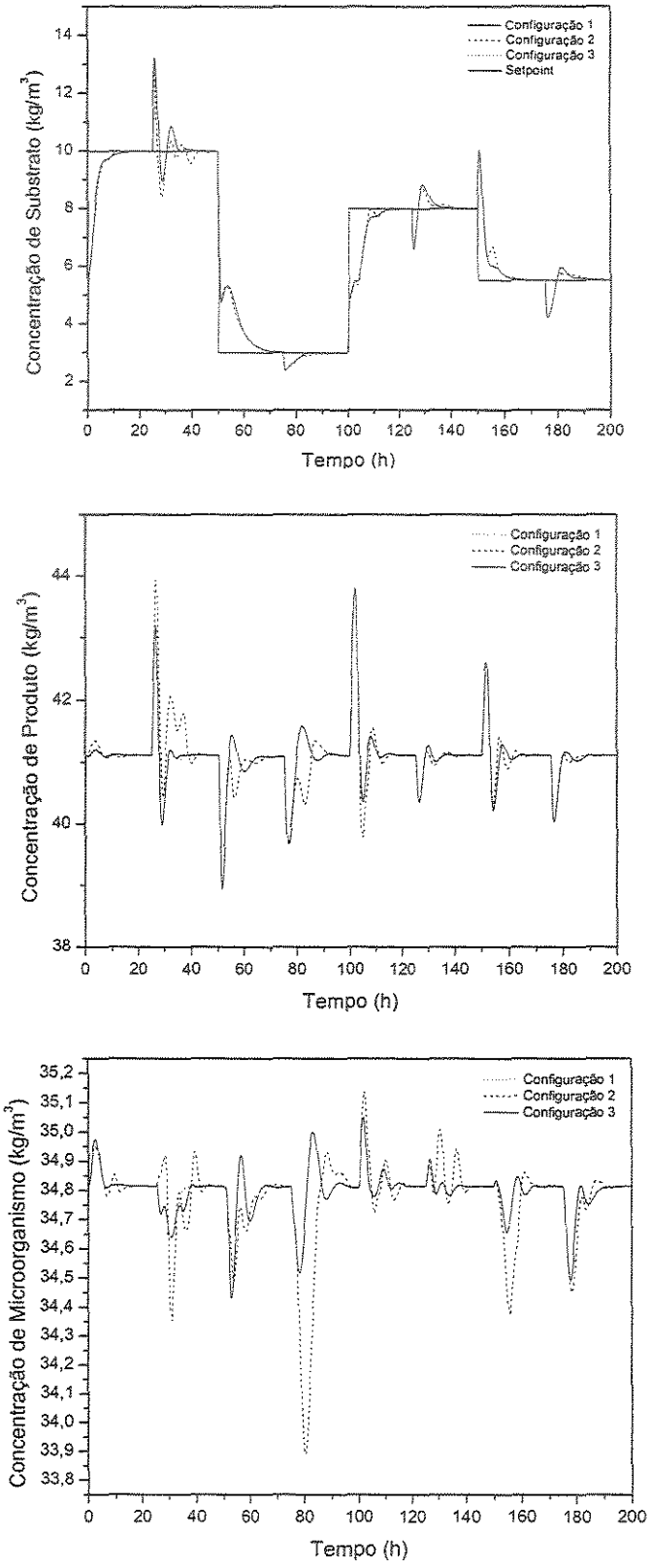


Figura 4.22 Variáveis Controladas para Problema Servo em S

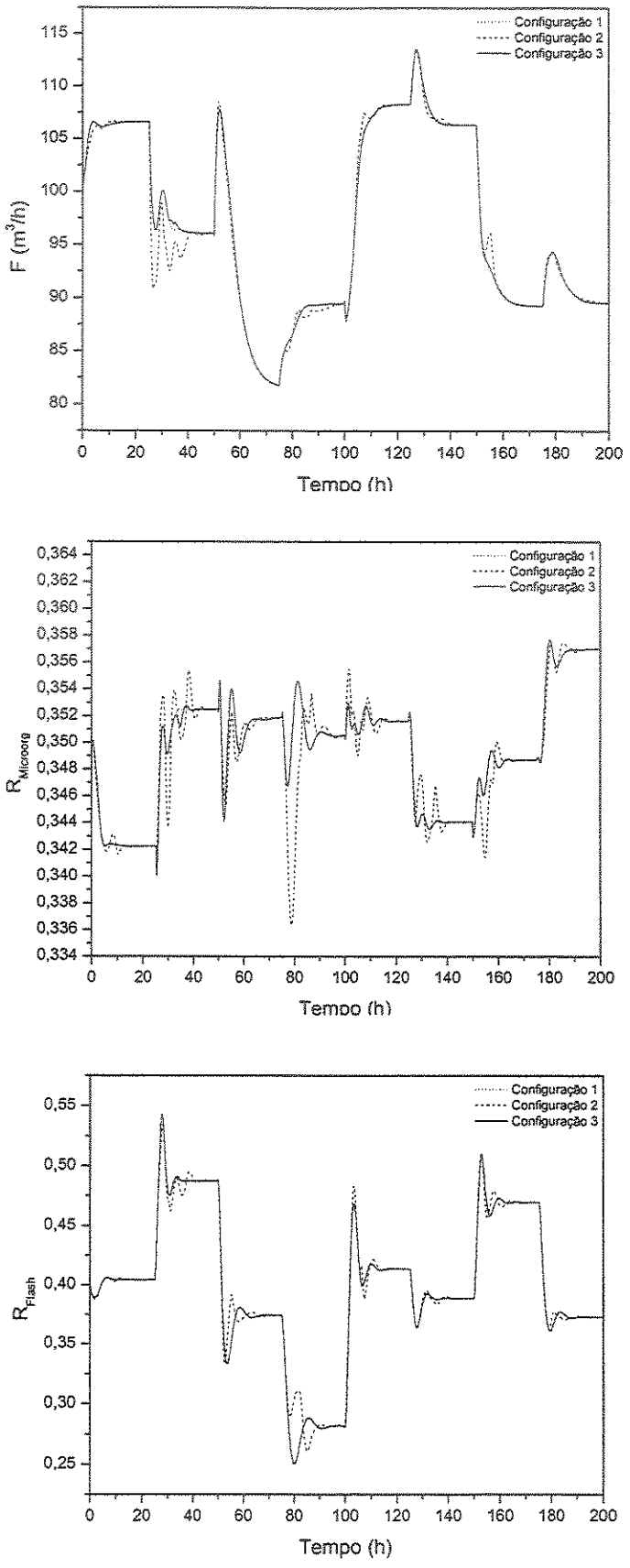
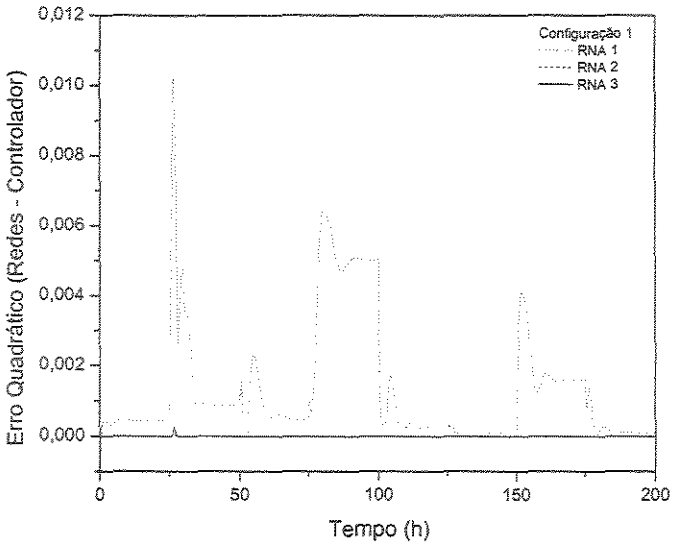
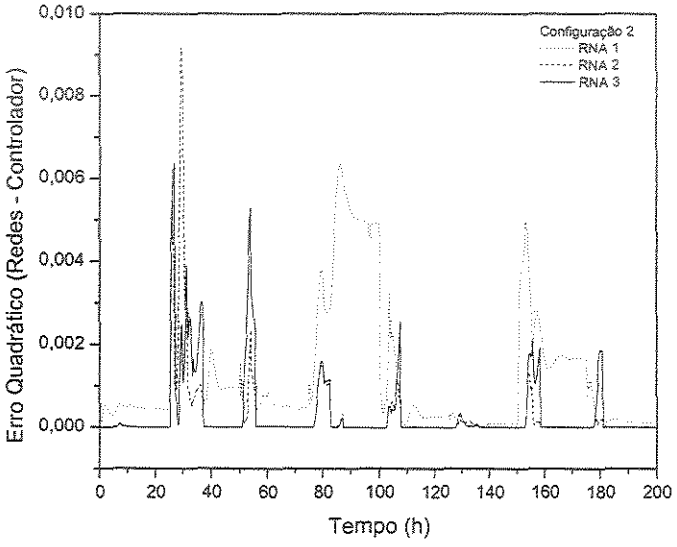


Figura 4.23 Variáveis Manipuladas Para Problema Servo em S





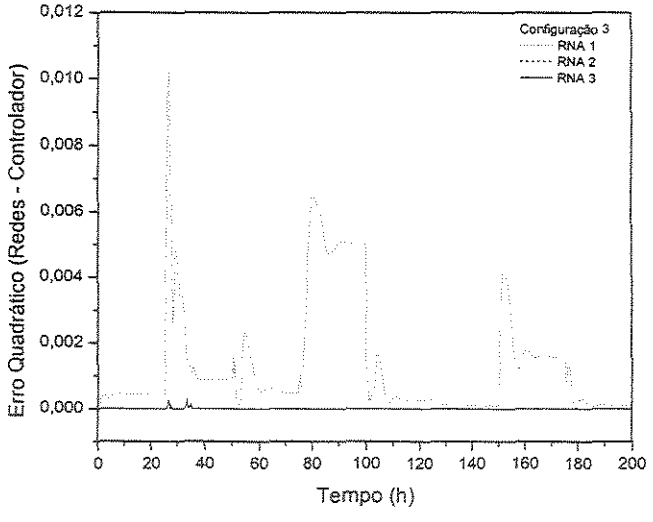


Figura 4.24 Redes do Controlador Para Configurações 1, 2 e 3

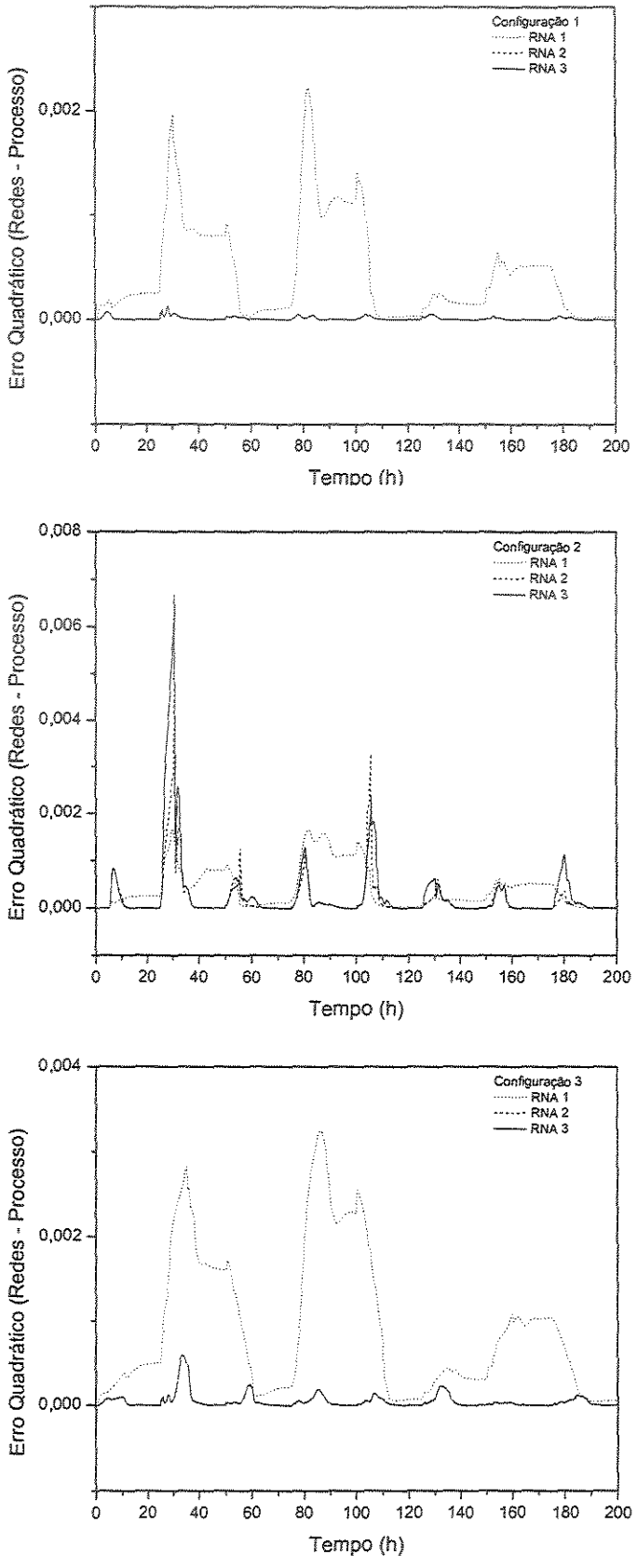


Figura 4.25 Redes do Processo Para Configurações 1, 2 e 3

4.6.2.2 Problema Servo para Variável P

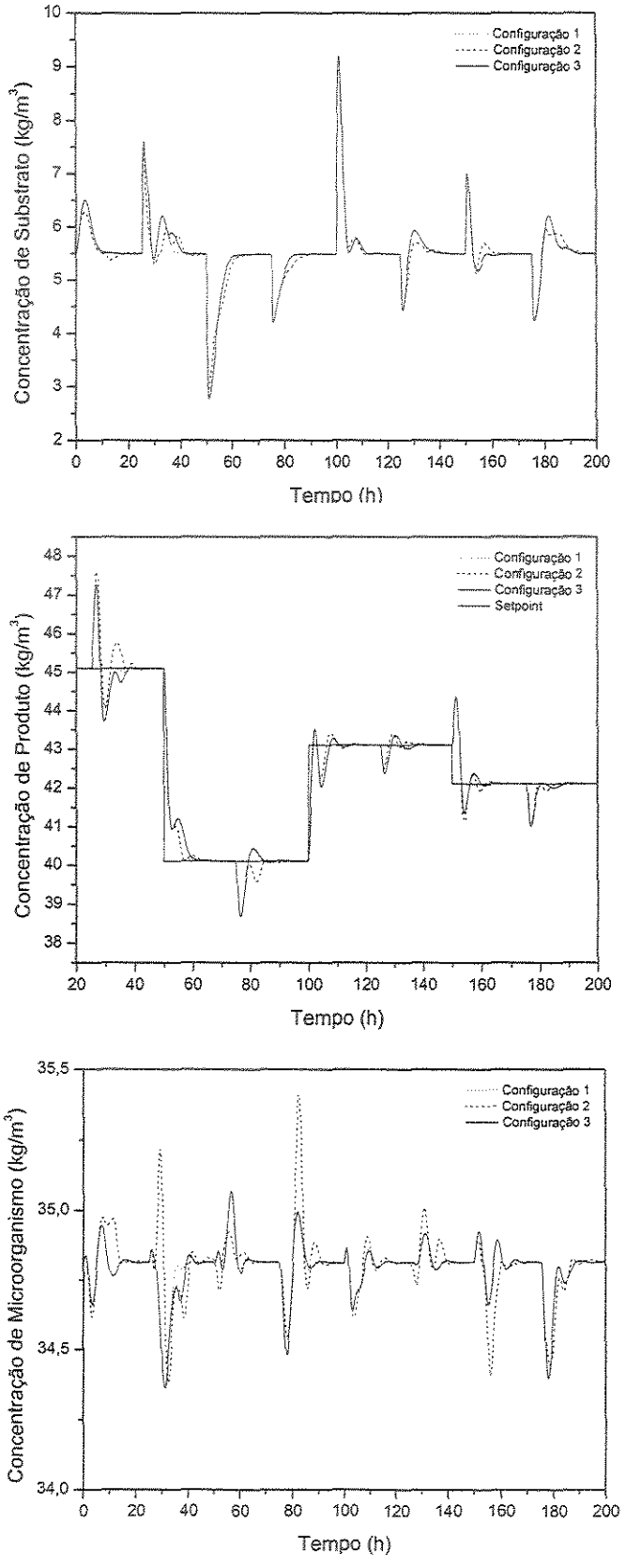


Figura 4.26 Variáveis Controladas Para Problema Servo em P

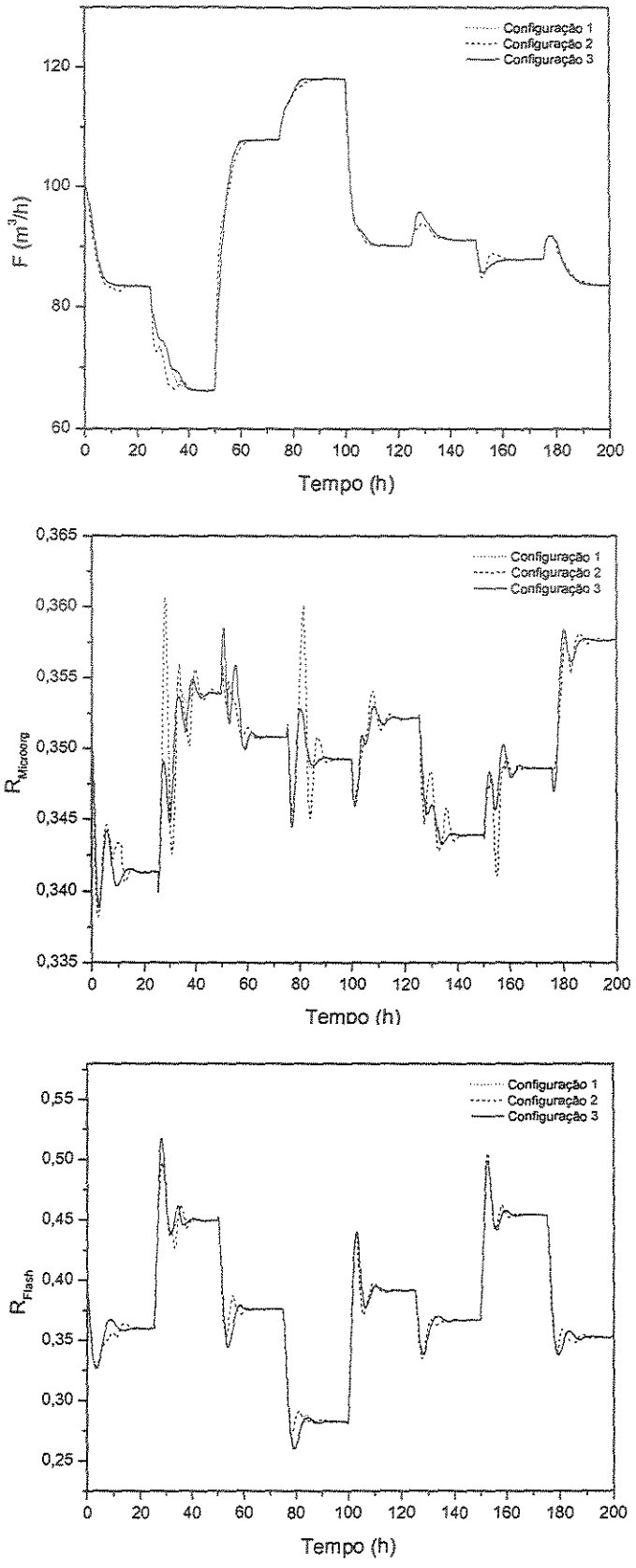


Figura 4.27 Variáveis Manipuladas Para Problema Servo em P

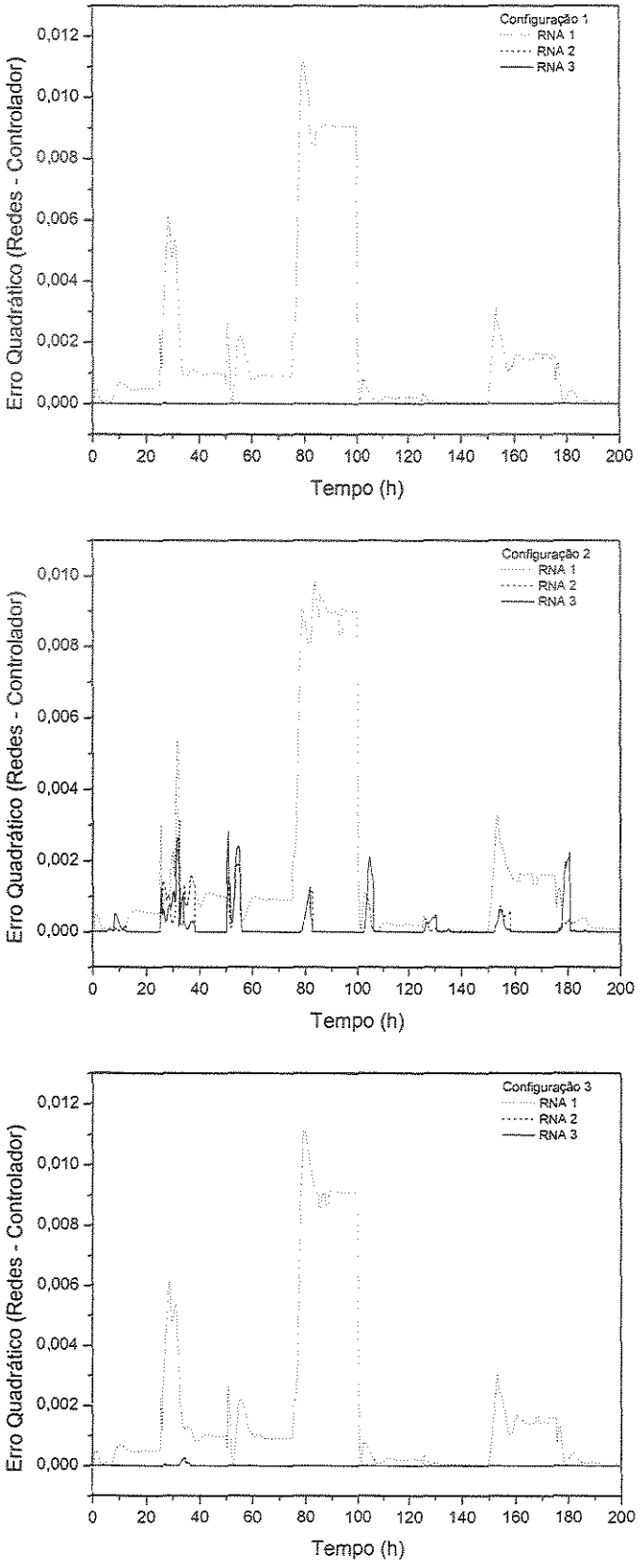


Figura 4.28 Redes do Controlador Para Configurações 1, 2 e 3

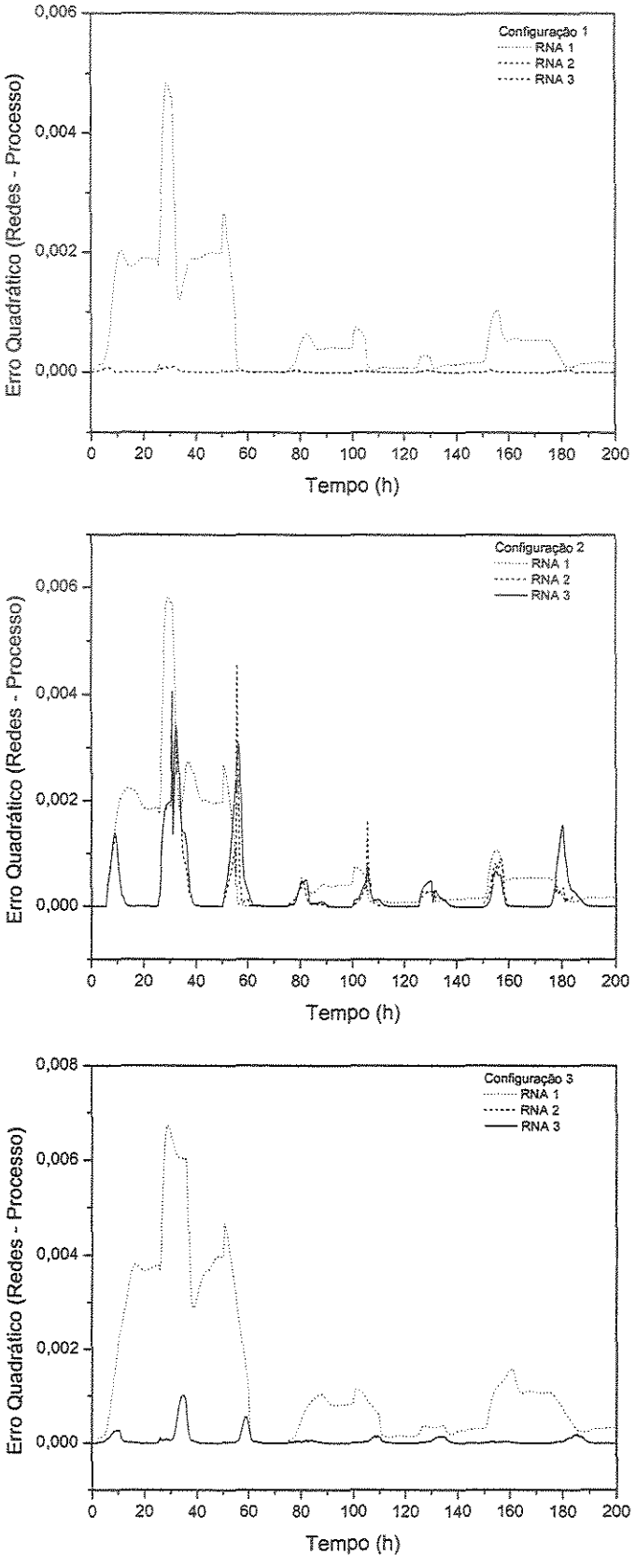


Figura 4.29 Redes do Processo Para Configurações 1, 2 e 3

4.6.2.3 Problema Servo para Variável X

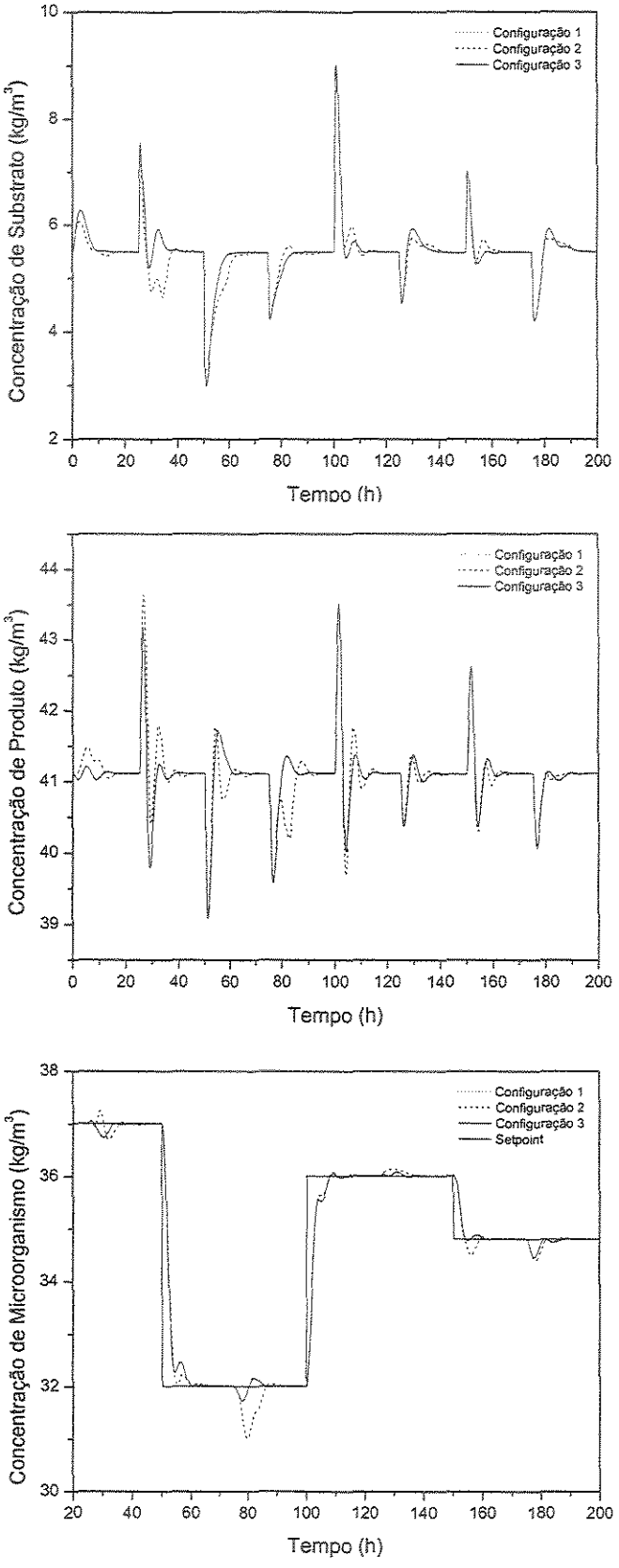
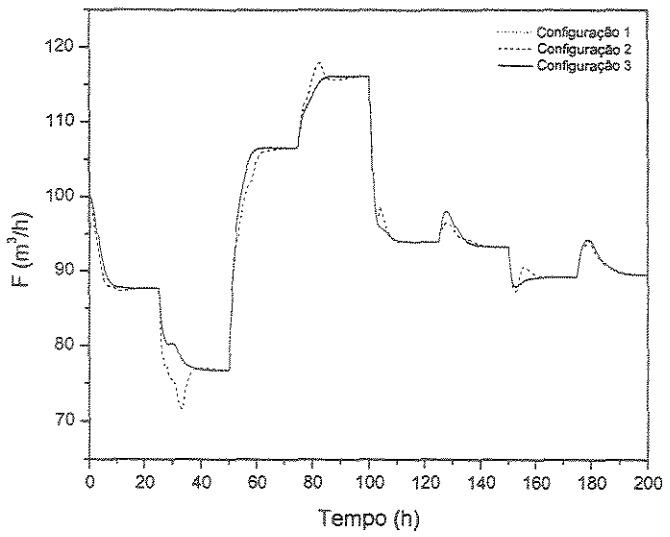
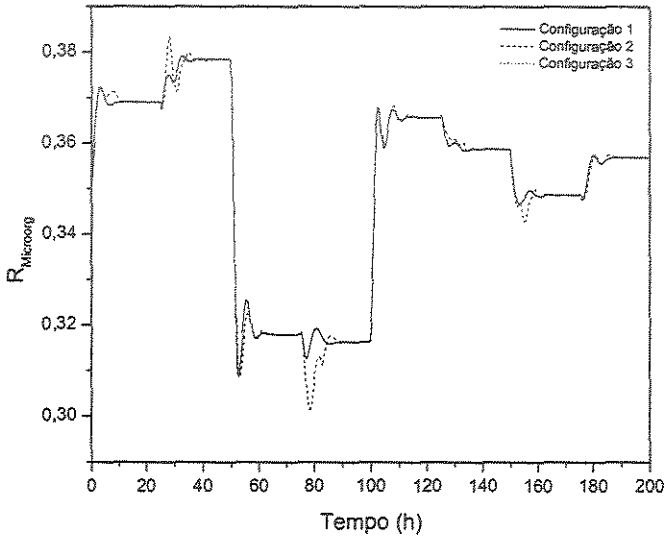


Figura 4.30 Variáveis Controladas Para Problema Servo em X





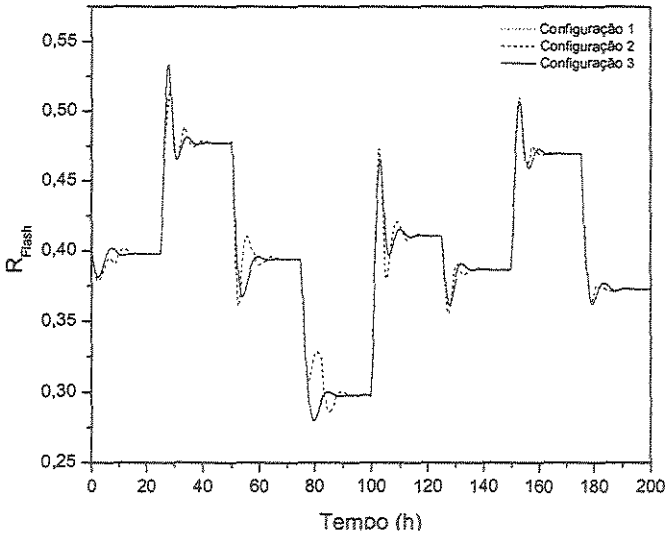


Figura 4.31 Variáveis Manipuladas Para Problema Servo em X

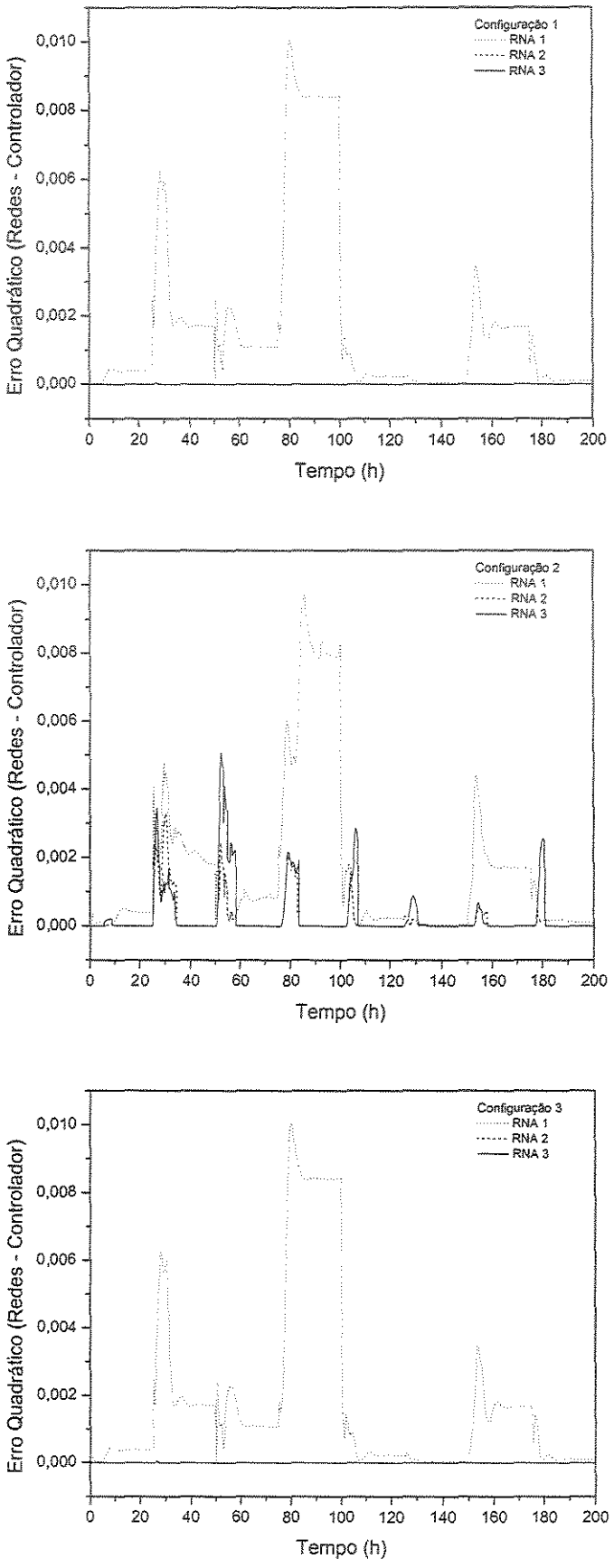


Figura 4.32 Redes do Controlador Para Configurações 1, 2 e 3

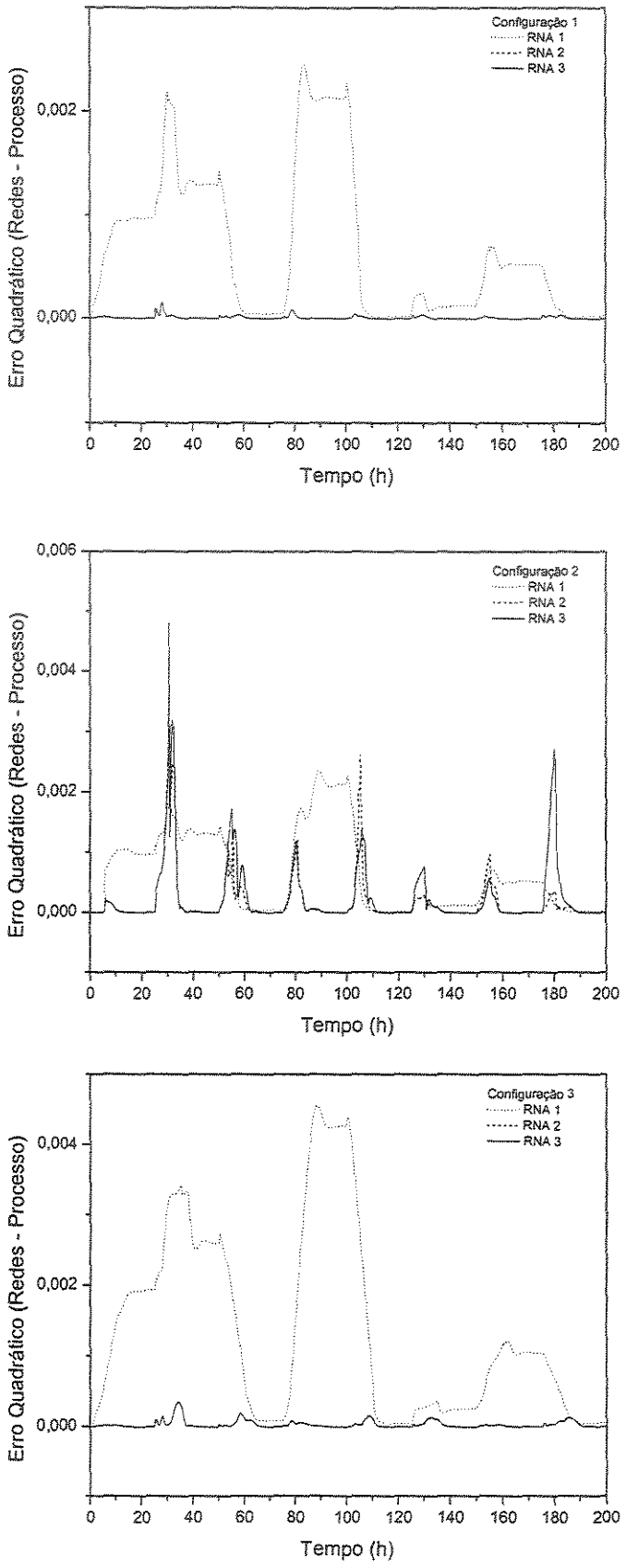


Figura 4.33 Redes do Processo Para Configurações 1, 2 e 3

Nas secções anteriores observou-se que para todas as respostas analisadas todas as configurações apresentaram o mesmo perfil.

As configurações 1 e 3 apresentaram as mesmas respostas para as variáveis manipuladas e controladas. Isto é devido a semelhança dos vetores utilizados para aprendizagem e teste. Para comprovar esta observação foi utilizado um vetor diferenciado para teste na configuração 3, ou seja foram utilizados 25 pares de aprendizagem e aumentou-se o número de pares utilizados para o teste das redes. Observa-se uma diferença significativa para as estratégias 1 e 3 apenas com o número de padrões de teste igual a 200, conforme a Figura 4.34.

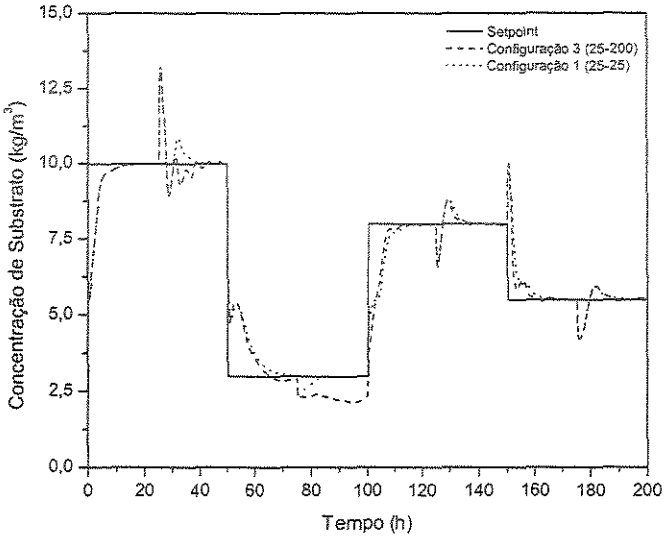


Figura 4.34 Comparativo Configuração 1 (25-25) e Configuração3 (25-200)

A configuração 2 apresenta pequena diferença (picos) comparada com as outras configurações. Esta região de picos é referente as maiores perturbações e a geração destes picos é devido aos vetores de aprendizagem e teste serem diferenciados.

Os resultados do erro quadrático das redes mostram que as redes que tiveram treinamento ao longo do tempo apresentaram os menores erros.

No resultado do erro quadrático para a configuração 2 pode-se observar a importância do uso de três redes paralelas pois nas regiões de maiores perturbações a aprendizagem *on-line* não foi eficiente (pois os vetores eram diferentes) mas utilizando a rede padrão e base-padrão as redes conseguiram acompanhar as variações do sistema.

As configurações 1 e 3 apresentaram os melhores resultados, mas a configuração 3 é a mais realística, pois o horizonte de teste é maior que o de aprendizagem, o que permite avaliar se a *RNA* ainda representa um passado próximo (ultimas variações do sistema).

7.8.3 Análise da Estratégia Proposta

Para analisar a estratégia de controle proposta foram comparados os resultados obtidos para malha fechada com e sem aprendizagem on-line, para as redes que representam o processo; e malha aberta para os problemas estudados do tipo servo e regulatório em So e To de forma conjunta.

As Figuras 4.35 até 4.37 mostram a potencialidade da estratégia on-line comparada com a off-line.

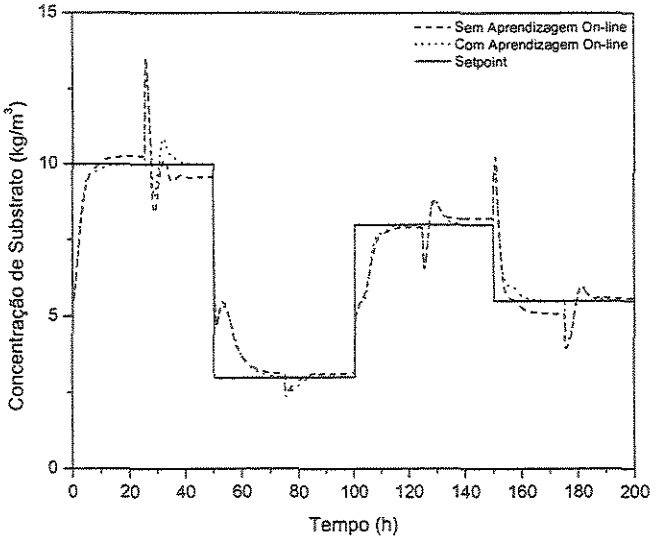


Figura 4.35 Comparação On-line e Off-line para problema Servo em S

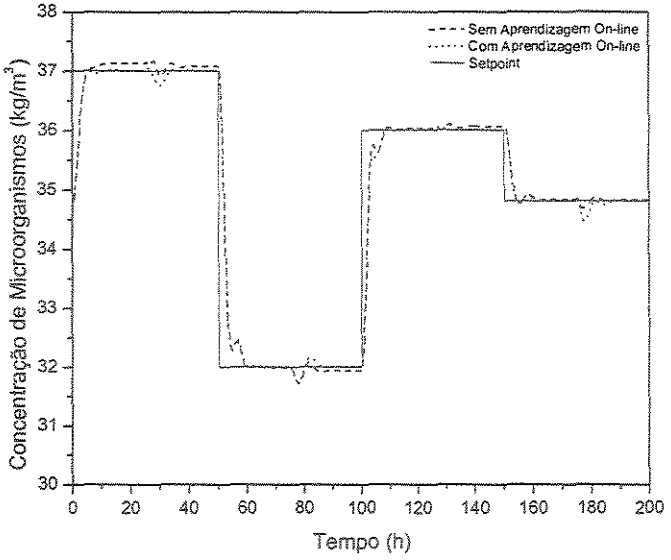


Figura 4.36 Comparação On-line e Off-line para problema Servo em X

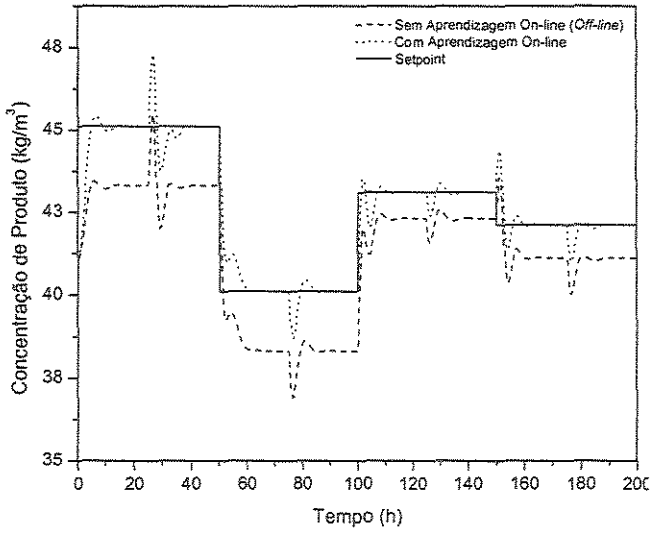


Figura 4.37 Comparação On-line e Off-line para problema Servo em P

As Figuras 4.38 à 4.40, mostram os resultados obtidos para a estratégia quando são aplicados ruídos estocásticos, (ruído colorido com perturbações nas variáveis So e To), o que permite verificar a eficiência mesmo com a presença de ruídos.

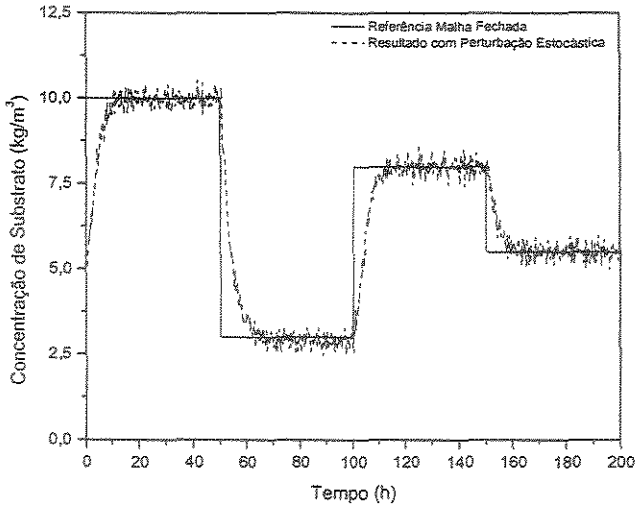


Figura 4.38 Resultado com Perturbação Estocástica para Problema Servo em S

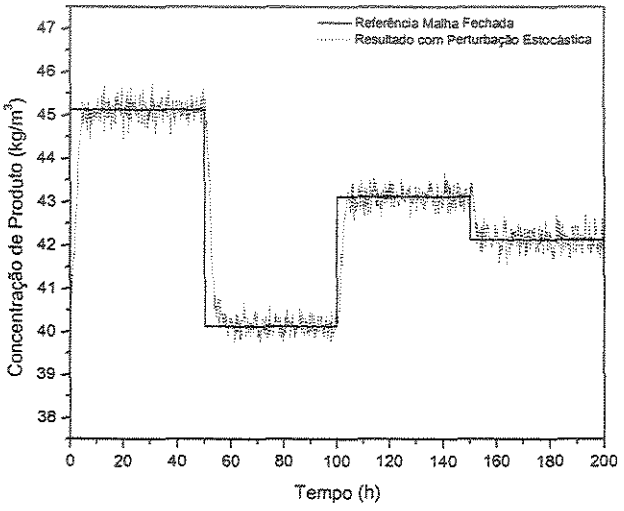


Figura 4.39 Resultado com Perturbação Estocástica para Problema Servo em P

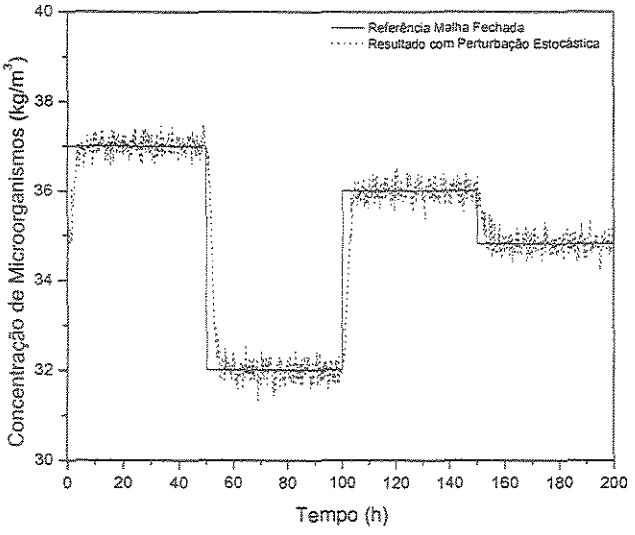


Figura 4.40 Resultado com Perturbação Estocástica para Problema Servo em X

4.9 CONCLUSÕES

O uso de três redes neurais permite que as redes apresentem diferentes "memórias" do processo, onde a rede Padrão (RNA 1) terá uma visão geral do processo e a rede Nova (RNA 3) uma visão mais pontual do processo, o uso dessa estratégia de aprendizagem permite as redes uma adaptação ao longo do tempo, mesmo quando o número máximo de iterações não for suficiente para a obtenção do critério de convergência com a aprendizagem *on-line*.

As configurações propostas foram adequadas para a aprendizagem *on-line,* mas a configuração 3 é a mais adequada, ou seja consegue determinar a melhor rede e se está ainda esta representando as últimas perturbações. A estratégia proposta apresentou-se robusta para os problemas do tipo servo e regulador de forma conjunta e na presença de ruídos estocásticos. A aprendizagem *on-line* permite ao controlador acompanhar as modificações do processo e apresentar um controle efetivo, pois consegue ter seus parâmetros atualizados em tempo real.

Buscando uma estratégia que possa ser aplicada a diferentes processos, com dinâmica e características distintas o próximo capítulo apresenta o segundo caso de estudo, representado por um reator de leito fixo.

CASO 2 – REATOR CATALÍTICO DE LEITO FIXO

5. 1 INTRODUÇÃO

Neste capítulo é considerado como caso de estudo um reator de catalítico de leito fixo, sistema que possui um comportamento dinâmico complexo (devido as não linearidades e presença de resposta inversa) além de possuir uma dinâmica mais rápida do que a do caso de estudo do processo extrativo de fermentação alcoólica contínua. De forma similar ao capítulo anterior apresentou-se o processo estudado e posteriormente os resultados aplicando-se a estratégia de controle proposta.

5.2 PROCESSO

Os reatores de leito fixo são equipamentos muito utilizados na indústria química, principalmente em reações catalíticas, por isso escolheu-se este processo para representar o Caso de Estudo 2.

O controle efetivo desses reatores é fundamental para se obter operações seguras, especialmente quando alta performance é desejada. Estes reatores exibem complicado comportamento estacionário e dinâmico resultante das não linearidades do sistema e da diferença de capacidades térmicas do sólido e do fluido, as quais entre outras coisas levam ao surgimento da resposta inversa e "*runaway*" da temperatura do reator (TOLEDO, 1999).

O fenômeno de resposta inversa pode ser explicado porque o catalisador tendo uma capacidade térmica *(rCp)s* maior que a do fluído refrigerante, *(rCp)f,* inicialmente absorve o calor introduzido no reator pelo aumento em *Tfo* e promove uma diminuição da temperatura do reator, comportamento não esperado. Entretanto, na seqüência devido a reação química ser extremamente exotérmica que ocorre dentro da partícula de catalisador há uma liberação de energia para o meio fluido e há o esperado aumento de temperatura no reator (MACIEL FILHO, 1989).

Foi considerada a reação de oxidação de etanol à acetaldeído, sendo esta reação altamente exotérmica e com uma cinética muito complexa o que torna o processo extremamente interessante do ponto de vista de simulação e controle. A taxa de reação utilizada considera como catalisador o Fe–Mo (TOLEDO,1999 e MACIEL FILHO, 1985).

A reação global para a oxidação do etanol a acetaldeído sobre catalisador de Fe-Mo é:

$$CH_3CH_2OH + \frac{1}{2}O_2 \rightarrow CH_3CHO + H_2O$$

Utiliza-se o ar como fonte de oxigênio considerando a seguinte composição:

Oxigênio (O2) = 21%

Nitrogênio (N2) = 79%

A Figura 5.1 mostra o esquema do processo, tem-se uma entrada de reagentes e do fluido refrigerando e estes passando ao longo do reator promovem a reação sobre os catalisadores, gerando como produto o acetaldeído.

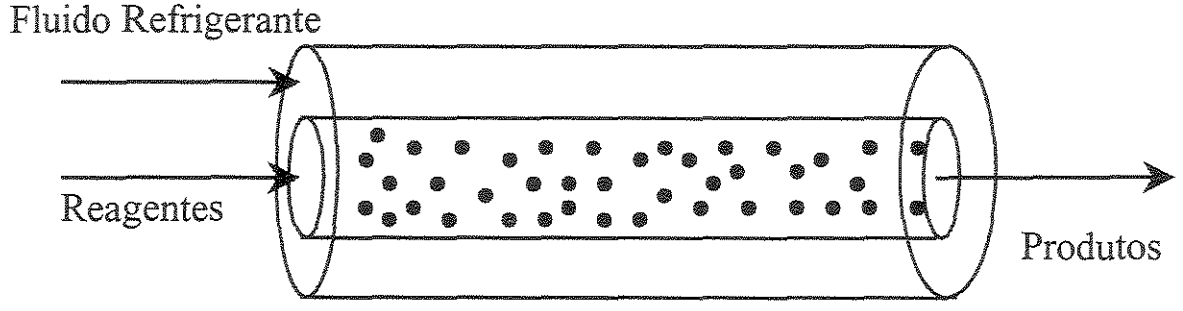


Figura 5.1 Esquema do Caso de Estudo 1 – Reator Catalítico de Leito Fixo

5.3 MODELAGEM MATEMÁTICA

A maior dificuldade na modelagem usualmente refere-se a transferência de calor, pois a taxa de liberação de energia ao longo do comprimento do reator não é uniforme e a maior parte das reações ocorrem nas proximidades da entrada do reator. Em reações exotérmicas a taxa de reação pode ser relativamente maior na entrada do reator devido a maior concentração de reagentes (TOLEDO, 1999).

As equações que descrevem o comportamento dinâmico do reator consideram um modelo pseudo-homogêneo (TOLEDO 1999).

Neste modelo as partículas de catalisador e fluido reagente são tratadas como um único meio e contínuo. Assume-se que o fluido dentro de um elemento de volume associado a uma partícula de catalisador pode ser caracterizado por uma certa temperatura, pressão e composição, e estas grandezas variam continuamente com a posição dentro do reator. Foram adotadas ainda as seguintes considerações:

- Perfil de velocidade " *Plug-flow*" ;
- Dispersão axial desprezível.

O modelo matemático segundo TOLEDO (1999) é descrito através das equações seguintes:

- ***BALANÇO DE MASSA***

$$\frac{\partial X}{\partial t} = \frac{D_{ef}}{R_t^2}\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial X}{\partial r}\right] - \frac{GM_o}{\rho_f L}\frac{\partial X}{\partial z} + \frac{(1-\varepsilon)PM\rho_B}{\varepsilon\rho_f}R_W \tag{5.1}$$

- ***BALANÇO DE ENERGIA***

$$\frac{\partial T}{\partial t} = \frac{\lambda_{ef}}{C_m R_t^2}\frac{1}{r}\frac{\partial}{\partial r}\left[r\frac{\partial T}{\partial r}\right] - \frac{\varepsilon GM_o C_{pf}}{C_m L}\frac{\partial T}{\partial z} + \frac{(1-\varepsilon)\rho_B(-\Delta H_R)}{C_m T_{ref}}\frac{R_W}{(R+1)} \tag{5.2}$$

onde:

$$C_m = \varepsilon\,(\rho_f\, C_{pf}) + (1 - \varepsilon)\,(\rho_s\, C_{ps})$$

- ***BALANÇO DE MOMENTO***

$$\frac{\partial p}{\partial t} = -\frac{GM_o}{\rho_f L}\left[\frac{\partial p}{\partial z} + f\right] \tag{5.3}$$

- ***EQUAÇÃO DA CONTINUIDADE***

$$\frac{\partial}{\partial z}\left(\rho_f V_f\right) = 0 \tag{5.4}$$

- ***EQUAÇÃO PARA O FLUIDO REFRIGERANTE***

$$\frac{\partial T_R}{\partial t} = \frac{-u_R}{L}\frac{\partial T_R}{\partial z} + \frac{h_w}{\rho_R C_{PR}}(T(1,z) - T_R) \tag{5.5}$$

- ***CONDIÇÕES DE CONTORNO***

$$r = 0 \qquad \frac{\partial X}{\partial t} = \frac{\partial T}{\partial r} = 0 \qquad \text{(Simetria)} \tag{5.6}$$

$$r = 1 \qquad \frac{\partial X}{\partial t} = 0, \frac{\partial T}{\partial r} = Bih(T_W - T_R) \tag{5.7}$$

$$z = 0 \qquad X = 0, T = {T_{fo}}/{T_{ref}}, p = {p_o}/{p_{ref}}, T_R = {T_{R0}}/{T_{ref}} \tag{5.8}$$

RW é a taxa de oxidação de etanol à acetaldeído sobre catalisador de Fe-Mo, (Kmol de mistura reagente/h Kg de catalisador), TOLEDO, 1999:

$$R_W = \frac{2K_1K_2P_{O2}P_{ET}}{K_3K_1P_{ET}P_{AC} + K_1P_{ET} + 2K_2P_{O2} + K_3K_4P_{AC}P_{H2O}} \tag{5.9}$$

$$P_{N2} = \frac{0.79R}{R + 1 + 0.5X} \tag{5.10}$$

$$P_{O2} = \frac{0.21R - 0.5X}{R + 1 + 0.5X} \tag{5.11}$$

$$P_{H2O} = \frac{X}{R + 1 + 0.5X} \tag{5.12}$$

$$P_{ET} = \frac{1-X}{R+1+0.5X} \tag{5.13}$$

$$P_{AC} = \frac{X}{R+1+0.5X} \tag{5.14}$$

onde P_{O2}, P_{ET}, P_{H2O}, P_{AC} são pressões parciais do oxigênio, etanol, água e acetaldeído respectivamente, P é a pressão do sistema e K_i constantes cinéticas na forma de Arrhenius. Estas equações são resolvidas pelo método das linhas, onde as coordenadas axial e radial foram discretizadas por colocação ortogonal e a integração em relação ao tempo foi resolvida pelo método de Gear (TOLEDO, 1999).

5.4 COMPORTAMENTO ESTACIONÁRIO

Os parâmetros estacionários adotados para as simulações são apresentados na Tabela 5.1, que seguem DOMINGUES e MACIEL FILHO (1992), TOLEDO (1999) e ENDER (1999).

Tabela 5.1 Parâmetros Estacionários Adotados

Variável	*Valor Estacionário*
Tfo	445,15K
Tro	445,15K
Gmo	4500 kg/h
R	25
Ur	3,0 m/h
Po	1 atm
Dt	0,017 m
Dp	0,002 m
L	1,0 m
T(1)	453,21K
T(3)	451,3 K

As variáveis controladas foram as temperaturas T(1) e T(3) nos pontos de colocação ortogonal z_1=0,0254 e z_3=0,297, respectivamente, pois estes pontos são os mais sensíveis as perturbações, e por esse motivo a reação ocorre mais fortemente no início do reator, indicando ser esta região que se deve ter um maior conhecimento para assegurar um melhor desempenho e controle do reator (TOLEDO, 1999 e ENDER 2002).

- T(1) = Temperatura no primeiro ponto de colocação axial ortogonal 453,21K;
- T(3) = Temperatura no terceiro ponto de colocação axial ortogonal 451,30K.

A Figura 5.2 representa o perfil para concentração de acetaldeído ao longo da posição do reator e a Figura 5.3 representa o perfil da temperatura ao longo do reator.

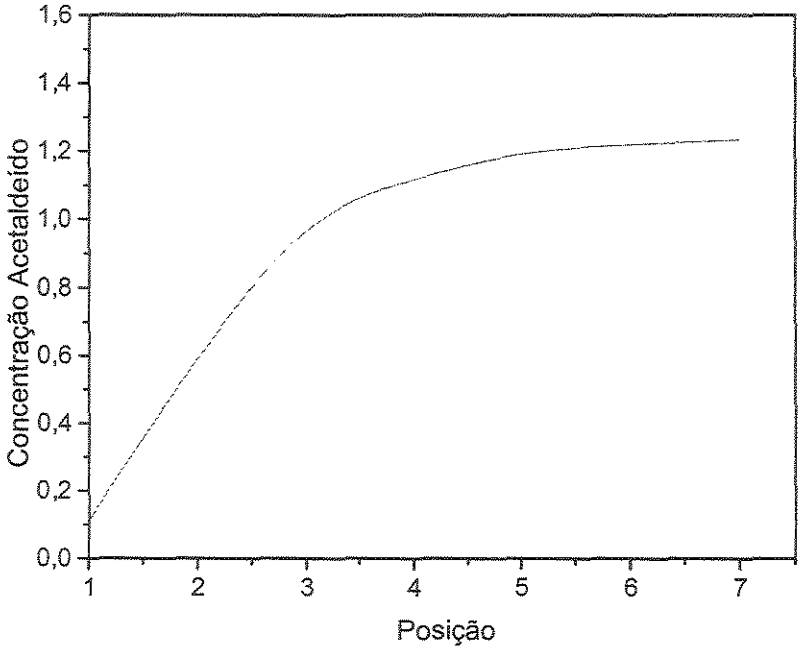


Figura 5.2 Perfil da Concentração de Acetaldeído ao Longo do Reator

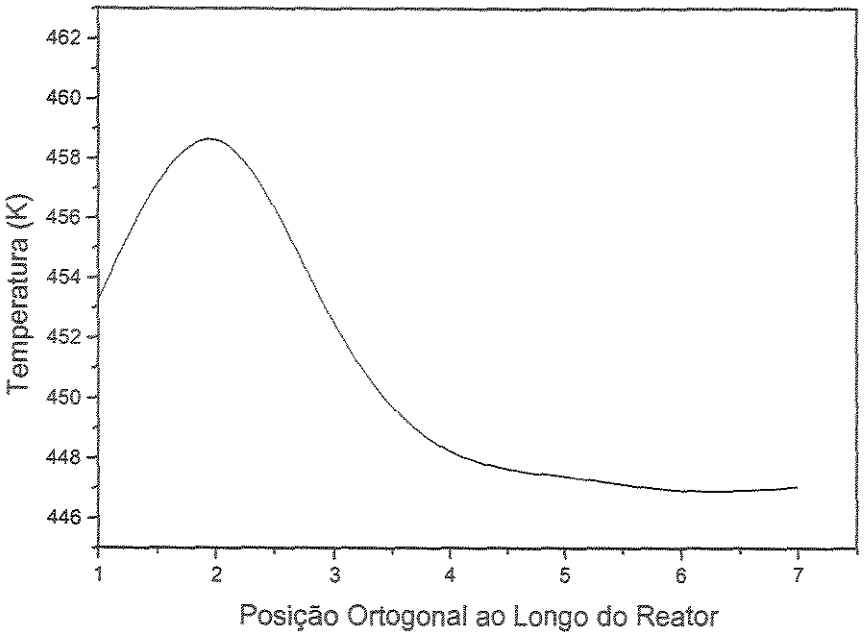


Figura 5.3 Perfil da Temperatura ao Longo do Reator

5.5 COMPORTAMENTO DINÂMICO

Para observar o comportamento dinâmico do reator foram feitas perturbações positivas e negativas nas variáveis de entrada: temperatura de alimentação (Tfo), temperatura do fluído refrigerante (Tro), razão de alimentação ar/etanol (R) e fluxo mássico de alimentação (GMo).

A Tabela 5.2 mostra o valor de cada variável após as perturbações positivas (+) e negativas (-). Estas perturbações foram de acréscimo e decréscimo nos valores estacionários de 40% em GMo, 20% em R e 5% Tfo e Tro, respectivamente.

Tabela 5.2 Valores das Variáveis de Entrada após Perturbações Especificadas

Variável	*Perturbação (+)*	*Estacionário*	*Perturbação (-)*
Tfo	422,9 K	445,15K	467,4 K
Tro	422,9 K	445,15K	467,4 K
Gmo	2700 kg/h	4500 kg/h	6300 kg/h
R	20	25	30

As Figuras 5.4 e 5.5 mostram os perfis das temperaturas do primeiro (T1) e terceiro (T3) ponto de colocação ortogonal, para as perturbações de acréscimo nos valores estacionários de 40% em GMo, 20% em R e 5% Tfo e Tro, respectivamente.

As Figuras 5.6 e 5.7 mostram os perfis das temperaturas (T1) e (T3) para as perturbações de decréscimo nos valores estacionários de 40% em GMo, 20% em R e 5% Tfo e Tro, respectivamente.

Nesta Figuras pode se observar o comportamento dinâmico do reator, com suas não linearidades e resposta inversa.

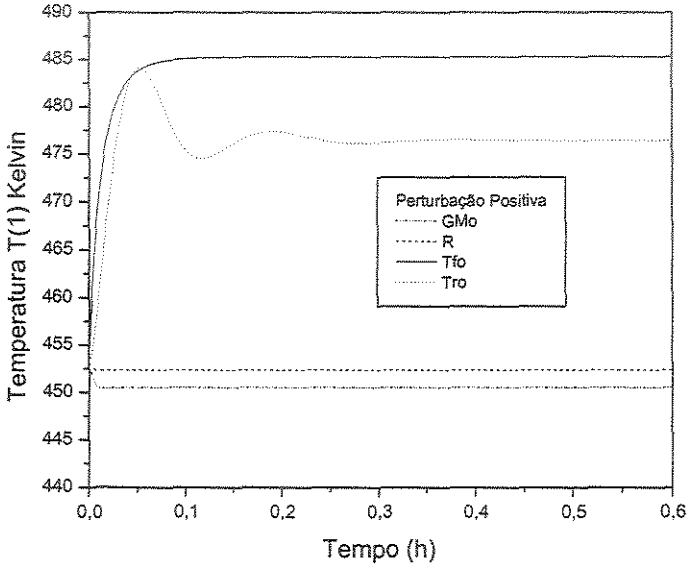


Figura 5.4 Perfil da Temperatura (T1) para as Perturbações Positivas

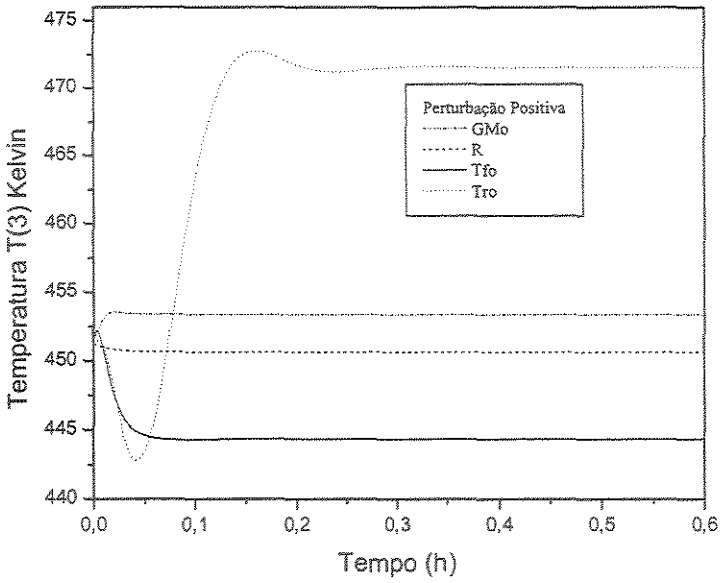


Figura 5.5 Perfil da Temperatura (T3) para as Perturbações Positivas

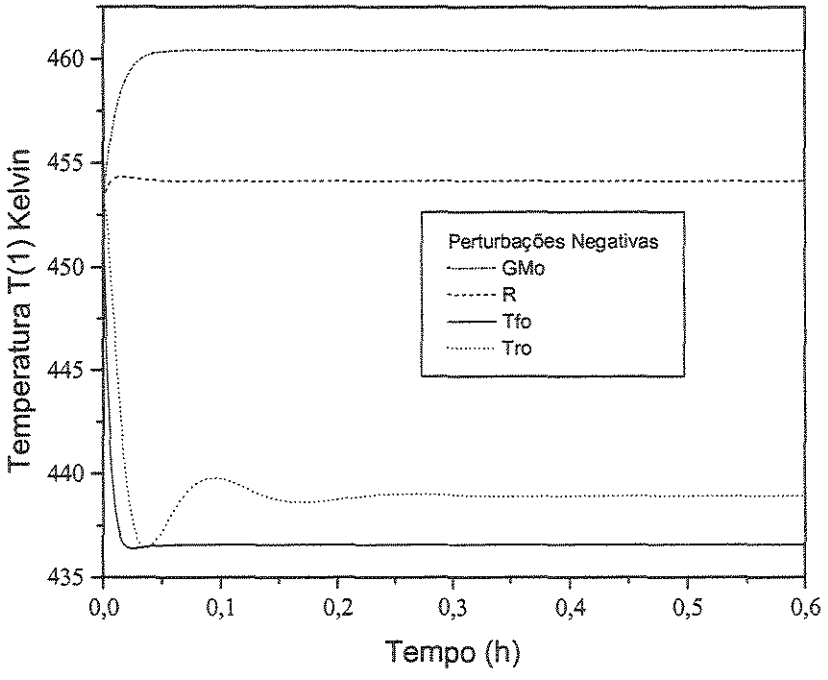


Figura 5.6 Perfil da Temperatura (T1) para as Perturbações Negativas

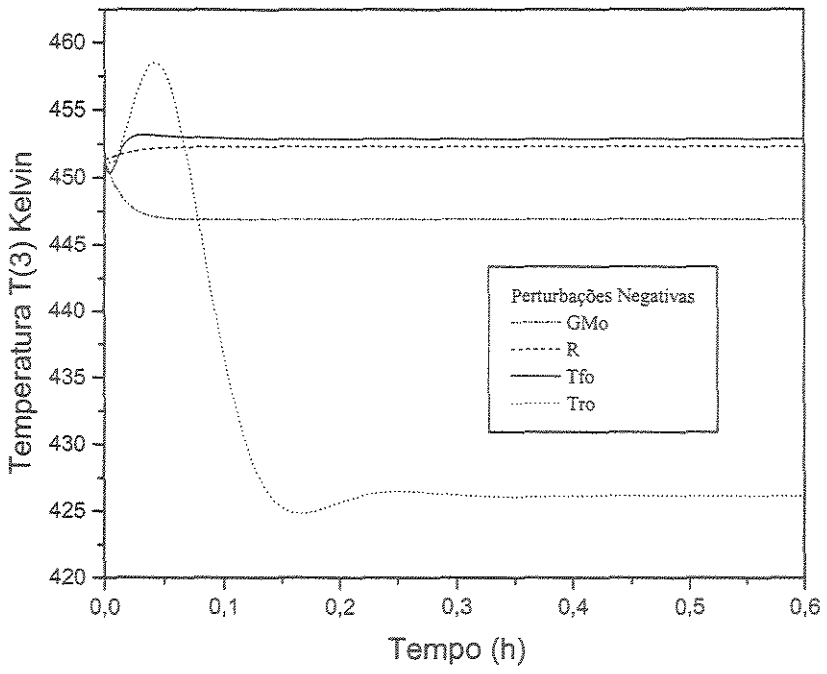


Figura 5.7 Perfil da Temperatura (T3) para as Perturbações Negativas

5.6 ANÁLISE DE SENSIBILIDADE DOS PARÂMETROS DE PROCESSO – PLANEJAMENTO FATORIAL

Uma das etapas mais importantes no desenvolvimento de uma estratégia de controle é definir as variáveis que serão manipuladas e controladas, ou seja, qual é o objetivo operacional do sistema.

Para monitorar o desempenho do processo é necessário medir os valores de certas variáveis, que podem ser medidas diretamente (medidas primárias) ou indiretamente (medidas secundárias). No caso de medidas secundárias utiliza-se relação matemática para determinação das mesmas, que podem resultar de considerações empíricas, experimentais ou teóricas.

Uma característica que deve ser observada na definição da variável manipulada é a influência desta sobre a variável controlada, ou seja, se a variável manipulada for muito sensível, qualquer variação, mesmo pequena desta variável produz um efeito grande na

variável controlada. Da mesma maneira se a variável manipulada for pouco sensível, são necessárias mudanças muito grandes na mesma para produzir efeitos na variável controlada.

De acordo com TOLEDO (1999) e ENDER (2002), para determinação da influência das variáveis do processo sobre as variáveis de interesse o planejamento fatorial mostra ser uma ferramenta interessante.

O método de planejamento fatorial consiste em selecionar um número fixo de níveis e então executar experimentos com todas as possíveis combinações destes. Nesta dissertação se utilizará um experimento fatorial com dois níveis (nível –1 e nível +1). Estes níveis são o inferior e o superior em relação ao nível central (nível 0) para cada variável. Para n variáveis envolvidas no estudo, o número de experimentos que será realizado é igual a 2^n. Com os resultados obtidos do planejamento fatorial é possível calcular os principais efeitos e iterações das variáveis de interesse sobre as respostas, determinando os efeitos mais significativos (BARROS NETO, *et al,* 1996).

Para análise de sensibilidade das variáveis de operação do reator catalítico de leito fixo sobre o perfil da temperatura (T(1) e T(3)), utilizou-se a metodologia de planejamento fatorial completo 2^n, cuja descrição é mostrada em BARROS NETO *et al.* (1996).

As variáveis utilizadas para o planejamento fatorial completo do reator foram relação ar/etanol (R), fluxo mássico de alimentação (GMo) e velocidade do fluido refrigerante (Ur), pois estas variáveis podem ser facilmente manipuladas num processo industrial. Não foi utilizada a variável temperatura do fluído refrigerante pois tem difícil implementação real. TOLEDO (1999) também faz uma análise completa das variáveis de operação do reator em estudo.

A Tabela 5.3 traz a descrição dos níveis das variáveis usadas na análise de sensibilidade do processo, sendo o ponto central os valores estacionários das variáveis e os níveis: superior (+1) os valores estacionários com um acréscimo de 10% e inferior (-1) os valores estacionários com um decréscimo de 10%.

A Tabela 5.4 descreve os ensaios usados no planejamento fatorial completo para determinar as temperaturas no primeiro e terceiro ponto de colocação ortogonal axial do reator T(1) e T(3), respectivamente.

Tabela 5.3 Níveis das Variáveis Usadas na Análise de Sensibilidade do Processo

Variável	Nível Inferior	Nível Central	Nível Superior
GMo	4050	4500	4950
R	22,5	25	27,5
Ur	2,7	3	3,3

Tabela 5.4 Ensaios Usados no Planejamento Fatorial Completo do Reator

Ensaio	*Gmo*	*R*	*Ur*	*T(1)*	*T(3)*
1	-1	-1	-1	455,04	450,78
2	1	-1	-1	452,65	452,6
3	-1	1	-1	453,83	450,18
4	1	1	-1	451,95	451,66
5	-1	-1	1	455,04	450,68
6	1	-1	1	452,63	452,52
7	-1	1	1	453,83	450,12
8	1	1	1	451,95	451,6

A Tabela 5.5 descreve os efeitos estimados para temperatura do primeiro ponto de colocação ortogonal (T(1)), obtidos através do planejamento fatorial completo. Foram utilizados níveis de confiança de 95 e 93% para definição dos valores estatisticamente significativos. Observa-se que:

- Quando o valor de GMo é aumentado do nível inferior para o nível superior, na faixa de valores estudada, ocorre uma diminuição do valor da temperatura T(1) em 2,14%, considerando o nível de confiança de 95%;

- Quando o valor de R é aumentado do nível inferior para o nível superior, na faixa de valores estudada, ocorre uma diminuição do valor da temperatura T(1) em 0,95%, considerando o nível de confiança de 95%;

- Quando são aumentados os valores de GMo e R ao mesmo tempo, do nível inferior para o nível superior, ocorre um aumento de 0,26% da temperatura T(1), considerando o nível de confiança de 93%;

- A variável Ur e as interações das variáveis 1 com 3 e 2 com 3 não apresentaram resultados estatisticamente significativos para os níveis de confiança de 93 e 95%.

Tabela 5.5 Efeitos Estimados para temperatura T(1)

	Efeito	*p*
Média	453,34811	5,81E-09
(1)Gmo*	-2,14075	0,0011705
(2)R*	-0,95075	0,0058924
(3)UR	-0,00575	0,9446209
1 by 2	0,26075	0,0707592
1 by 3	-0,00425	0,9590391
2 by 3	0,00575	0,9446209

Estatisticamente Significativo: * nível de confiança de 95%

**nível de confiança de 93%

A Figura 5.8 mostra os resultados para o planejamento fatorial, onde se pode observar a superfície de reposta e diagramas de contorno para temperatura T(1) expresso como função das variáveis GMo e R.

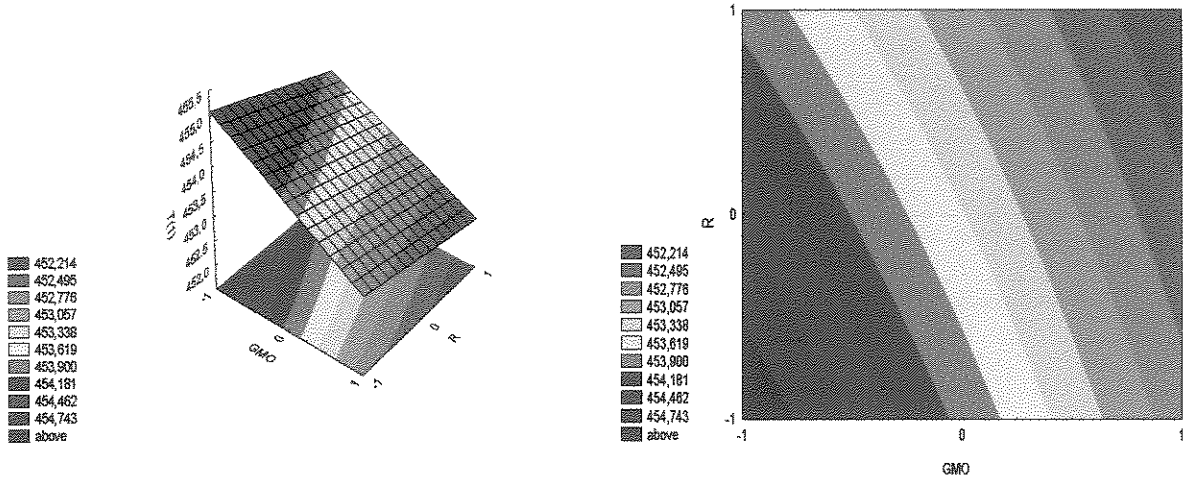


Figura 5.8 Superfícies de resposta e diagramas de contorno para a temperatura (T1) expresso como função de: GM0 (vazão de entrada) e R.

A Tabela 5.6 descreve os efeitos estimados para temperatura do terceiro ponto de colocação ortogonal (T(3)), obtidos através do planejamento fatorial completo. Foi utilizado nível de confiança de 95% para definição dos valores estatisticamente significativos. . Observa-se que:

- Quando o valor de GMo é aumentado do nível inferior para o nível superior, na faixa de valores estudada, ocorre um aumento do valor da temperatura T(3) em 1,65%;

- Quando o valor de R é aumentado do nível inferior para o nível superior, na faixa de valores estudada, ocorre uma diminuição do valor da temperatura T(3) em 0,73%;

- Quando são aumentados os valores de GMo e R ao mesmo tempo, do nível inferior para o nível superior, ocorre uma diminuição do valor da temperatura T(3) em 0,16%.

- A variável Ur e as interações das variáveis 1 com 3 e 2 com 3 não apresentaram resultados estatisticamente significativos para o nível de confiança de 95%.

Tabela 5.6 Efeitos Estimados para temperatura T(3)

	Efeito	*p*
Média	451,2756	3,06E-10
(1)Gmo*	1,675	1E-04
(2)R*	-0,735	0,000519
(3)UR	-0,055	0,08156
1 by 2*	-0,165	0,010148
1 by 3	0,015	0,465006
2 by 3	0,025	0,274099

Estatisticamente Significativo: * nível de confiança de 95%

A Figura 5.9 mostra os resultados para o planejamento fatorial, onde se pode observar a superfície de reposta e diagramas de contorno para temperatura T(3) expresso como função das variáveis GMo e R.

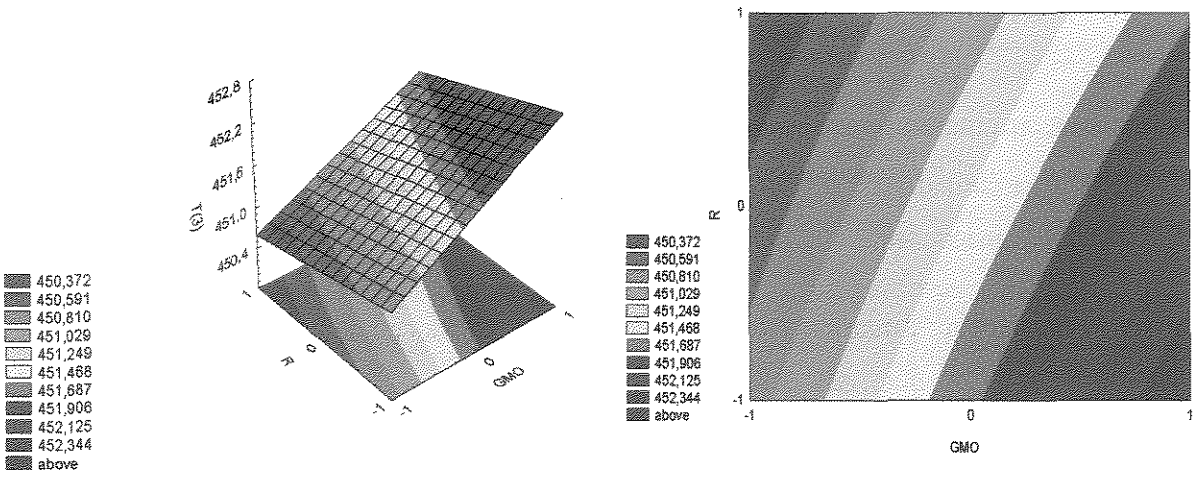


Figura 5.9 Superfícies de resposta e diagramas de contorno para a temperatura (T3) expresso como função de: GM0 (vazão de entrada) e R.

Através dos resultados analisados é possível concluir que:

- A variável GMo tem maior influência nas variáveis respostas analisadas. Apresenta uma relação inversa com T(1) e direta com T(3);
- A variável R tem uma relação inversa com as respostas analisadas T(1) e T(3);
- A variável Ur apresenta pequena influência nas respostas analisadas, T(1) e T(3), para faixa de valores estudados por isso não é uma variável manipulada apropriada para este caso;
- As variáveis GMo e R apresentam interações dificultando o controle, mas são as variáveis manipuladas mais apropriadas e portanto foram escolhidas para o desenvolvimento da estratégia de controle desta dissertação. É importante ressaltar que o conhecimento da existência das interações entre as variáveis permite o desenvolvimento de estratégias de controle usando estas variáveis porém com as devidas atenções.

5.7 ARQUITETURA E DADOS PARA TREINAMENTO DAS *RNA's*

A arquitetura das redes para o reator foi constituída de três camadas, com cinco neurônios na camada intermediária e passado β=5. Estes valores foram determinados de acordo com simulações.

Conforme discutido no Capítulo 3, as *RNA's* passam por um processo de aprendizagem para que seus pesos sejam modificados e tenham a capacidade de predizer respostas a partir de dados de entrada.

Ao escolher o conjunto de dados de treinamento da rede deve-se observar as possíveis alterações que as variáveis manipuladas podem sofrer ao longo do processo. Uma forma de obter um bom conjunto de dados para treinamento é promover perturbações das variáveis manipuladas.

Para o treinamento das redes utilizadas no caso de estudo 1, foram feitas perturbações nas variáveis manipuladas GMo e R, de forma individual e conjunta, conforme descreve a Figura 5.10.

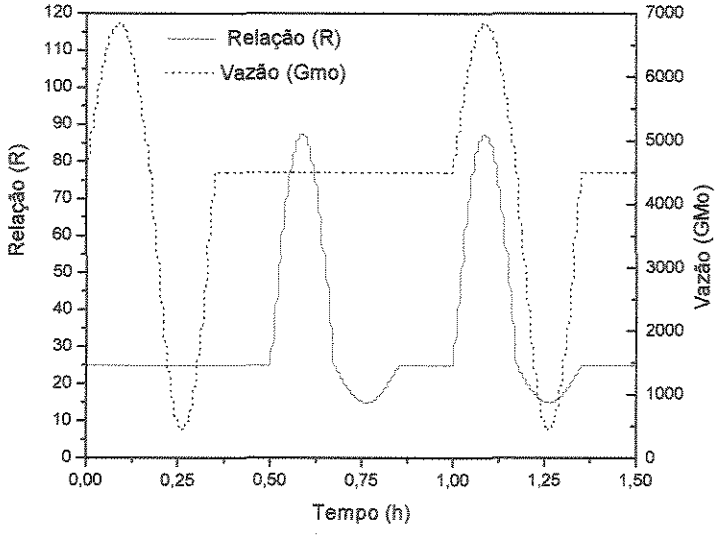


Figura 5.10 Perturbações nas Variáveis Manipuladas para Treinamento das RNA's

5.8 RESULTADOS PARA O CASO DE ESTUDO 2 – REATOR

Os resultados do controle do reator utilizando o CoNDEG Modificado, descrito no Capítulo 3, são apresentados neste Capítulo. Para comparar os resultados obtidos com a nova configuração a ser utilizada, definida como a configuração 3 (Capítulo 4), foram realizadas simulações com os problemas dos tipos: servo, regulador e estocástico.

Comparando as configurações 1 e 3 não foi observada diferença nos resultados pela similaridade dos vetores utilizados, onde pode-se concluir que a configuração 3 permite à estratégia determinar a melhor rede, e verificar se esta ainda está representando a realidade do processo para o intervalo testado e também pelo menor número de padrões utilizados para obter uma melhora no tempo computacional requerido. Por isso todos os resultados apresentados serão para a configuração 3.

A Tabela 5.7 descreve o problema do tipo servo utilizado, para o processo em estudo, para as variáveis controladas T(1) e T(3).

Tabela 5.7 Desvios nas Referências do Processo Reator Catalítico

Tempo (h)	*T(1)*	*T(3)*
0-0,2	-2,0	0,0
0,2-0,6	-3,5	-1,0
0,6-0,8	-2,0	0,0
0,8-1,0	0,0	-1,0

A Figura 5.11 apresenta os resultados para as variáveis controladas (T(1) e T(3)) utilizando a estratégia com aprendizagem *on-line,* aprendizagem *off*-line, aprendizagem *on-line,* para as redes que representam o processo e malha aberta.

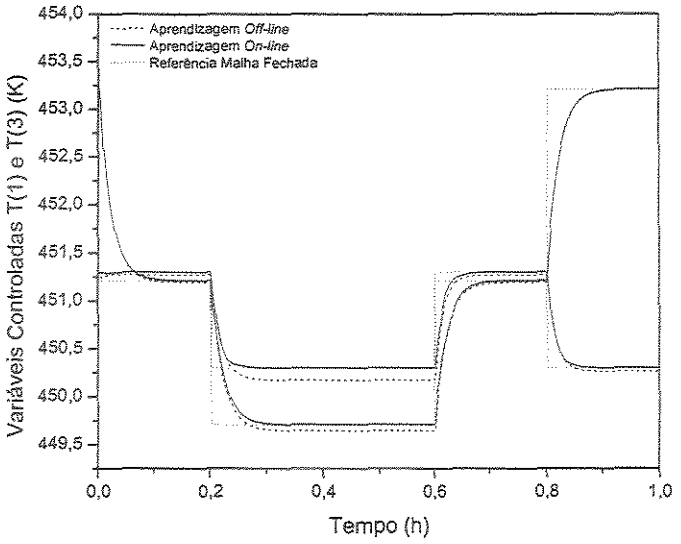


Figura 5.11 Resultado das Variáveis T(1) e T(3) para Problema Servo

A Figura 5.12 traz as respostas para as variáveis controladas para o problema do tipo servo associado a uma perturbação estocástica, ruído colorido para variáveis Tfo e Tro. Mesmo com a presença de ruído a estratégia mostrou-se robusta.

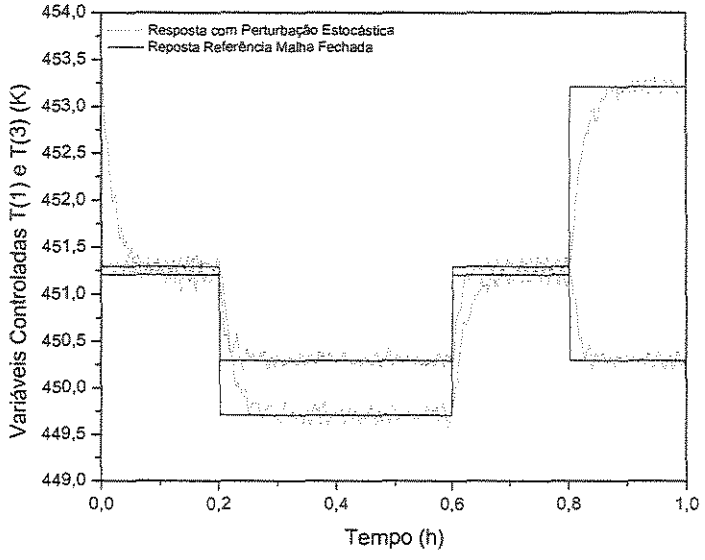


Figura 5.12 Resultado das Variáveis T(1) e T(3) com Perturbação Estocástica

As Figuras 5.13 e 5.14 apresentam os resultados para as variáveis razão de alimentação ar/etanol (R) e fluxo mássico de alimentação (GMo), respectivamente na presença de ruídos.

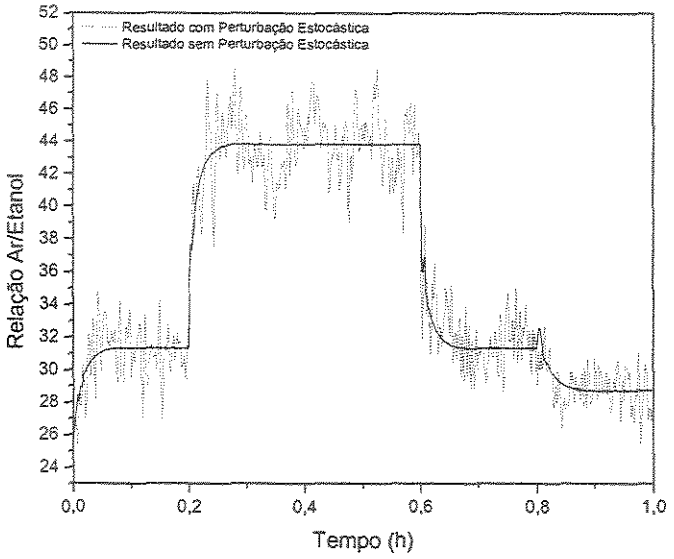


Figura 5.13 Resultado para Variável R com Perturbação Estocástica

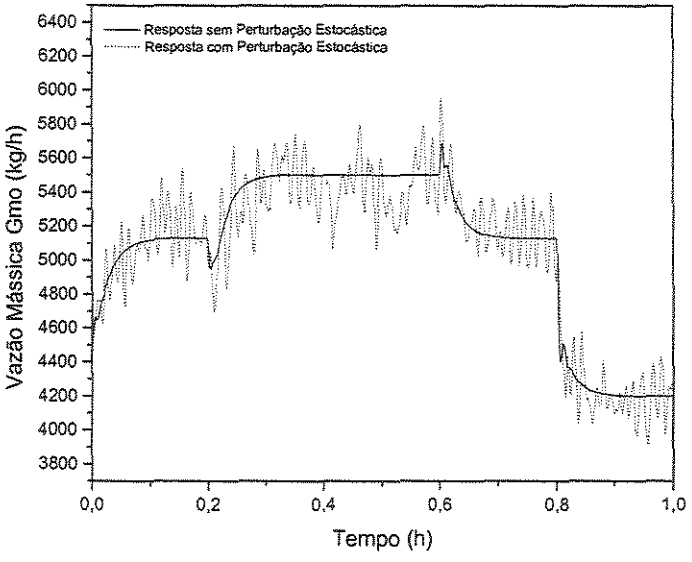


Figura 5.14 Resultado para Variável Gmo com Perturbação Estocástica

Para o problema do tipo regulador na variável temperatura de alimentação do reator (Tfo) foi acrescentado 4 Kelvin entre os tempos de 0,05 e 0,5 horas. Os resultados para este

tipo de problema são apresentados na Figura 5.15, onde é possível observar que a eficiência da estratégia e importância da aprendizagem on-line, que permite um controle efetivo e robusto para o problema proposto.

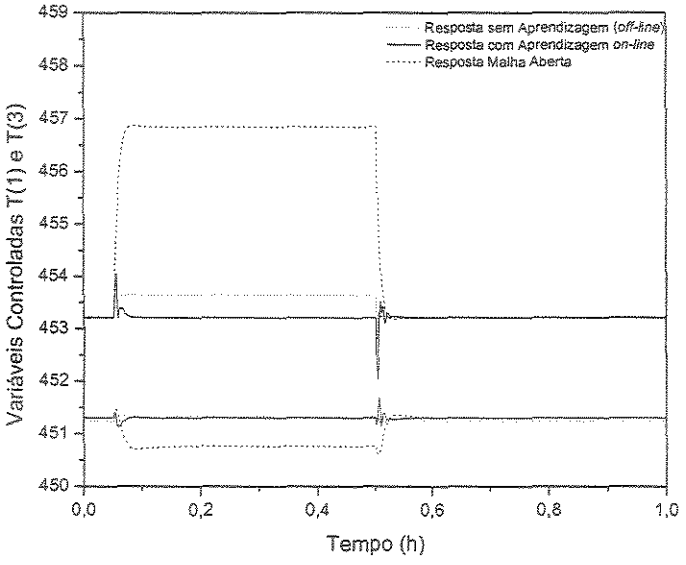


Figura 5.15 Resultado das Variáveis T(1) e T(3) para Problema Regulador em Tfo

Para o problema do tipo regulador na variável temperatura de alimentação do fluido refrigerante (Tro) foi acrescentado 4 Kelvin entre os tempos de 0,1 e 0,6 horas. Esta perturbação mostrou-se bastante forte no sistema e também pode-se observar a presença de resposta inversa intensa. Os resultados para este tipo de problema são apresentados nas Figuras 5.16 e 5.17 observa-se que mesmo na presença de resposta inversa a estratégia apresenta-se robusta.

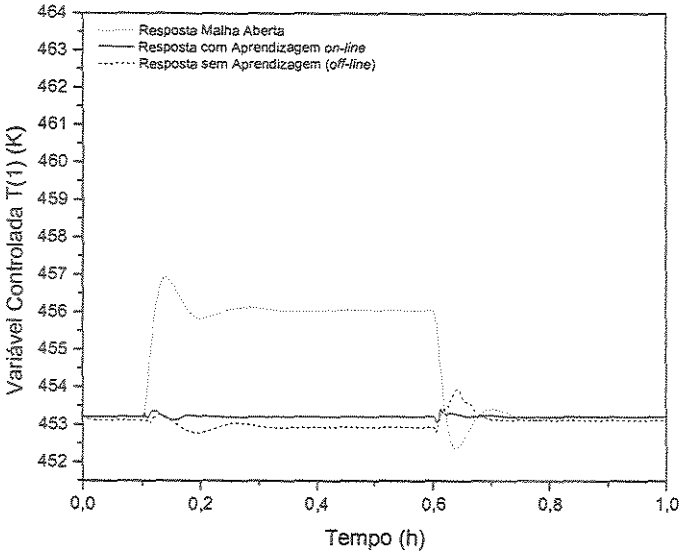


Figura 5.16 Resultado das Variáveis T(1) para Problema Regulador em Tro

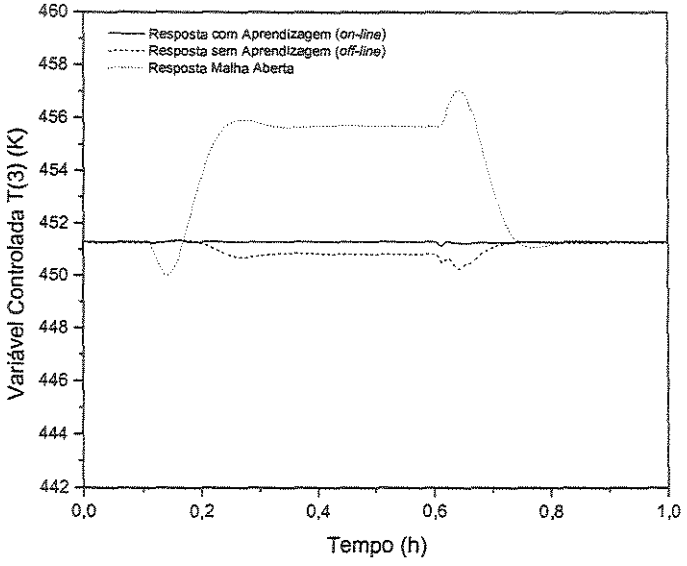


Figura 5.17 Resultado das Variáveis T(3) para Problema Regulador em Tro

5.9 CONCLUSÕES

O planejamento fatorial mostrou-se uma ferramenta interessante para análise de sensibilidade dos parâmetros de projeto determinando as variáveis importantes a serem

utilizadas na estratégia (manipuladas, controladas, perturbações), bem como a determinação das interações existentes entre as mesmas.

Com os resultados obtidos neste capítulo pode-se afirmar que a estratégia de controle proposta utilizando RNA's com aprendizagem *on-line* mostrou-se robusta e eficiente mesmo na presença de ruídos e resposta inversa acentuada.

CONCLUSÕES E SUGESTÕES

Este trabalho apresentou um estudo sobre controle não linear baseado em redes neurais artificiais. Teve como base o trabalho desenvolvido por ENDER (2002) onde buscou-se aprimorar a estratégia proposta pelo autor. A dissertação apresentou diferentes configurações para aprendizagem *on-line* que permitem uma melhor representação do processo e conseqüentemente um aumento na eficiência da estratégia.

A estratégia de controle proposta foi implementada em linguagem Fortran 90, para três configurações de aprendizagem *on-line* em dois casos de estudo. Os casos de estudo analisados são processos não lineares típicos na engenharia química. Estes são um processo extrativo de fermentação alcoólica e um reator catalítico de leito fixo.

Para o caso de estudo 2, reator de leito fixo catalítico foi realizado um planejamento fatorial completo para escolha das variáveis a serem utilizadas na estratégia. Este estudo foi realizado para aprendizagem da técnica que é uma poderosa ferramenta no planejamento de experimentos e também para um conhecimento do processo a ser pesquisado.

No estudo de redes neurais artificiais pode se observar a grande capacidade e seu potencial na representação e generalização de processos. As redes estudadas mostraram eficiência tanto no uso como modelos diretos e indiretos, por isso a grande aplicação na área de desenvolvimento de estratégias de controle.

No treinamento de redes observou-se que redes treinadas com parâmetros mensuráveis permitem o desenvolvimento de redes com melhor desempenho comparadas com redes treinadas com parâmetros não mensuráveis, principalmente quando trata-se de problema regulador acentuado.

Ainda no treinamento das redes observa-se que as redes usadas como modelo direto e indireto apresentam melhores resultados com diferentes funções de ativação de seus neurônios. As redes usadas para representar o processo apresentam melhores resultados quando treinadas com função sigmoidal e as redes usadas para representar o controlador tem um melhor resultado quando a função tangente hiperbólica é usada no seu treinamento.

A estratégia proposta por Ender (2002) mostrou-se bastante robusta. Na estratégia proposta nesta dissertação modificou-se a forma de ajuste do controlador permitindo o uso de redes treinadas com diversas funções de ativação.

A estratégia de aprendizagem *on-line* usando três redes neurais permitiu o acompanhamento das variações do processo, mesmo quando o número de iterações não for suficiente para convergência da aprendizagem *on*-line, pois utiliza-se a *RNA* padrão ou *RNA* base padrão. As três redes apresentam as seguintes características:

- Rede Padrão (RNA 1) – Apresenta uma "memória" geral do processo, ou seja, consegue representar bem o processo como um todo, mas muitas vezes não representa com exatidão uma região específica do processo;

- Rede Base-Padrão (RNA 2) – É uma rede intermediária entre as outras redes. Permite a retomada da aprendizagem *on-line* a partir dos pesos padrões sempre que estes apresentarem melhores resultados;

- Rede Nova (RNA 3) – Apresenta uma "memória" local, ou seja é a rede que melhor representa o processo nos últimos tempos de amostragem quando o número de iterações é suficiente. Como tem seus pesos modificados ao longo do tempo não consegue "lembrar" de modificações impostas para um tempo passado muito grande.

Para a aprendizagem *on-line* foram propostas três configurações diferenciadas nos vetores de aprendizagem e teste das redes. As três configurações propostas permitirão as redes acompanhar as modificações no processo. A configuração 3 pode ser considerada a forma mais adequada. As configurações 1 e 3 apresentaram melhor performance na estratégia de controle, porém pode-se afirmar que a forma de teste utilizada na configuração 3 permite selecionar a rede neural que apresenta uma "memória" de um horizonte passado maior do processo.

A estratégia proposta apresentou-se robusta para os dois casos de estudo analisados e permitiu que o controlador tivesse seus parâmetros atualizados ao longo do tempo, ou seja, permite um controle efetivo, pois consegue acompanhar a dinâmica do processo.

Finalmente para seqüência deste trabalho sugere-se como continuidade alguns trabalhos a serem explorados, tais como:

- Estudo avançado da arquitetura das redes neurais artificiais usadas, para determinar número de passados e neurônios da camada intermediária;
- Estudo e implementação de modelos híbridos na estratégia de controle;
- Estudo de técnicas de ajuste automático do controlador, associado a estratégia;
- Implementação experimental da estratégia proposta.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRIETA, S.R. Modelagem, Simulação e Controle de Fermentação Alcoólica contínua em Escala Industrial. São Paulo: Universidade Estadual de Campinas. Tese (Doutorado), 1994.

ÅSTROM, K. J., McAVOY, T., Intelligent Control, J. Proc. Cont. 2,3, 115-127, 1992.

ATALA, D. I. P. COSTA, A. C., MACIEL FILHO, R. MAUGERI, F. Kinetics of Ethanol Fermentation with High Biomass Concentration Considering the Effect of Temperature. Applied Biochemistry and Biotechnology, v. 91, p. 353-365, 2001.

BAKSHI, B. R. STEPHANOPOULOS, G. Wave-net: A Multiresolution, Hierarchical Neural Network with Localized Learning. AIChE Journal, v. 39-1, 1993.

BARROS NETO, B., SCARMINIO, I. S., BRUNS, R. E. Planejamento e Otimização de Experimentos. 2a ed. Campinas: Editora da UNICAMP, 1996. 299 p.

BAUGHMAN, D. R. LIU, Y. A. Neural Networks in Bioprocessing and Chemical Engieerin, London: Academic Press, 1995.

BISHOP, C. M. Neural Network and Their Applications. Rev. Science Istrum. V.65-6, p. 1803-1831, 1994.

BRAGA, A. de Pádua, LUDEMIR, T. B., CARVALHO A.C. P. L. Redes Neurais Artificiais – Teoria e Aplicações. Rio de Janeiro: LTC, 2000.

BRODERICK, G., BUDMAN, H. M., KNAPP, T. D. Adaptive Control of a CSTR with a Neural Network Model. Journal of Process Control, v.11, p.53-63, October, 1999.

BULSARI, A. B. Neural Networks for Chemical Engineers. Amsterdam: Elsevier, 1995.

CYBENDO, G. Approximation by Superposition of a Sigmoidal Function. Mathematics of Control, Signals and Systems. v. 2, p. 303-314, 1989.

CHEN, F. C. Backpropagation Neural Networks for Nonlinear Self- Tuning Adaptive Control. IEEE Control Systems Magazine, v.10-3, p.44-48, 1990.

CHEN, S., BILLINGS, S. A. Neural Networks for Nonlinear Dynamic System Modeling and Identification. International Journal Control, v.56-2, p. 319-346, 1992.

COSTA, A. C., ATALA, D. I. P., MACIEL FILHO, R. MAUGERI FILHO, F. Factorial Design and simulation for the Optimization and Determination of Control Structures for na Extractive Alcoholic Fermentation. Process Biochemistry, v37-2, p. 125-137, 2001.

COSTA, A. C. , MELEIRO, L. A. C., MACIEL FILHO, R. Non linear Predictive Control of na Extractive Alcoholic Fermentation Process. Process Biochemistry, v. 38-5, p. 743-750, 2002.

DOMINGUES, A., MACIEL FILHO, R. A Multitubular Reactor for Obtain of Acetaldehyde by Oxidation of Ethyl Alcohol. Chemical Engineering Science, v. 47, p. 2571-2576, 1992.

EBERHART, R., DOBBINS, R. Neural Networks PC Tools – A Practical Guide. Academic Press, San Diego, 1990.

ENDER, L., MACIEL FILHO, F. Use of Neural Networks in Control Strategies Applied to a Fixed Bed Catalytic Reactor, Proceedings of The Fourth Italian Conference on Chemical and Process Engineering, Florence, Italy, 1999.

ENDER, Laércio, MACIEL FILHO, Rubens. A Robust Procedure for On-line Learning of Neural Networks in Control Strategy Development. Proceedings of 3rd European Congress of Chemical Engineering. Germany, 2001a.

ENDER, Laércio, MACIEL FILHO, Rubens. A Robust Procedure for On-line Learning of Neural Networks in Control Strategy Development. Part II. Proceedings of 6th World Congress of Chemical Engineering. Melbourne, Austrália, 2001b.

ENDER, LAÉRCIO. Redes Neurais Aplicadas em Estratégias de Controle não Linear. São Paulo: Universidade Estadual de Campinas, 2002. Tese (Doutorado).

EPUB: http://www.epub.org.br/cm/n05/tecnologia/simulando.htm. Capturado em 15/03/2003.

GIARRATANO, J. C., CULBERT, C. and SAVELY, R. T., The State of the Art for Current and Future Expert Systems Tools, ISA Transactions, 29, 1, 17-25, 1990.

GOZDZIEJEWSKI JR. R., ENDER, L. Desenvolvimento de Software Didático em Redes Neurais Artificiais. Anais do Congresso Brasileiro de Engenharia Química, Porto Alegre, 1984.

HANG, C. C., GE, S. S., ZHANG, T. Design and Performance Analysis of a Direct Adapive Controller for Nonlinear Systems. Automatica, v.35, p. 1809-1817, March, 1999.

HEBB, D. O. The organization of the Behavior: A Neuropsychological Theory,. New York: Wiley, 1949.

HECHT-NIELSEN R., Theory of the Backpropagation Neural Network, IEEE int. Conf. on Neural Network, v. 1, p.593-605, 1989.

HECHT-NIELSEN R., Neurocomputing, Addisin-Wesley Publishing Company, Reading, MA, 1990.

HENRIQUE, H. M., LIMA, E. L. , SEBORG, D. E. Model Structure Determination in Neural Network Models, Chemical Engineering Science, v. 55-22 p. 5457-5469, 2000.

HORNIK, K., STINCHCOMBE, M., WHITE, H. Multi-layered Feedforward Neural Networks are Universal Approximations. Neural Networks, v.2, p. 359-366, 1990.

HOSKINS, J. C. and HIMMELBLAU, D. M., Artificial Neural Network Models of Knowledge Representation in Chemical Engineering, Comp. Chemical Engineering., 12, 9/10, 881-890, 1988.

HOPFIELD, J.J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities, Proceedings of the National Academy of Sciences of the USA, v.79, p. 2554-2558, 1982.

HUNT, K. J., SBARBARO, D., ZBIKOWSKI, R., GAWATHROP, P. J. Neural Networks in Control Systems: A survey, Automatica, v. 28, n.6, p. 1083-1112, 1992.

HUSSAIN M. A. Review of the applications of neural networks in chemical process control – simulation and online implementation. Artificial Intelligence in Engineering, v.13, p.55-68. June, 1999.

IANET – Inteligência Artificial. Versão 3.8. http://www.lokall.com/ianet/rdneuro.htm. Capturado em 14/08/20003.

KAUFMANN P., ANDERSSON G. G. Development of a generalized neural network. Chemo metrics and Intelligent Laboratory Systems, v. 50, p. 101-105, 2000.

KOLMOGOROV, A. N. On the Representation of Continuous Functions. URSS: Dokl. Acad. Nauk, 1957.

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

LIEW, A. C., WANG, F., CHANG, C. S. Simulated Hardware Design of Artificial Neural Networks for Adaptive Plant Control. Electric Power Systems Research, v.37, p. 231-240, March, 1996.

LOESCH, C., SARI, S. T. Redes Neurais: Fundamentos e Modelos. Blumenau: Universidade Regional de Blumenau, 1996.

MACIEL FILHO, R. Oxidação Catalítica de Etanol e Acetaldeído sobre Catalisador Óxido de Ferro-Molibdênio. São Paulo: Universidade Estadual de Campinas. Dissertação (Mestrado), 1985.

MACIEL FILHO, R. Modeling and Control of Multitubular Reactors, Leeds, 1989. Thesis (Ph.D, in Chemical Engineering).

McCULLOCH, W.S. e W. PITTS. A Lógica Calculus of the Ideas Immanent in Neurons Activity. Bulletin of Mathematical Biophysics, v.5, p. 115, 1943.

MELEIRO, L. A. C., MACIEL FILHO, R. Projeto e Aplicação de Controladores Baseados em Modelos Lineares, Neurais e Nebulosos. São Paulo: Universidade Estadual de Campinas, 2002. Tese (Doutorado).

MINSKY, M. L., PAPERT, S. A. Perceptrons. MIT Press, Cambridge, MA, 1969.

MORENO, C. A. F. Predição da Resistência à Compressão de um Cimento Industrial Utilizando Técnicas de Redes Neurais Artificiais, Universidade Estadual de Campinas, 2001. Dissertação (Mestrado).

NAHAS, E. P., HENSON, M. A., SEBORG, D. E. Nonlinear Internal Model Control Strategy for Neural Network Models. Compurters Chemical Engineering, v. 16-2, p. 1039-1057, 1992.

NARENDRA, S. K., CHEN, L. Nonlinear Adaptive Control Using Neural Networks and Multiple Models. Automatica, v. 37, p. 1245-1255, March, 2001.

NARENDRA, K., PARTHASARATHY, K. Identification and Control of Dynamical Systems Using Neural Networks. IEEE Trans. On Neural Networks, v. 1, p. 4-27, March, 1990.

NGUYEN, D. H., WIDROW, B. Neural Networks for Self-Learning Control Systems. IEEE Control Systems Magazine, p. 13-28, 1990.

PRETT, M. T., GARCIA, C.E. Fundamental Process Control, Butterworth Publishers, Stoneham, MA. USA, 1988.

PSICHOGIOS, D. C., UNGAR, L. H. Direct and Indirect Model – based Control Using Artificial Networks. Industrial Engineering & Chemical Research, v.30, n.12, p-2564-2573, 1991.

RAMIREZ-BELTRAN, N. D., JACKSON, H. Application of Neural Networks to Chemical Process Control. Computers & Industrial Engineering, v.37, p.387-390, 1999.

ROMARIZ, S. R. A. Representação e Aquisição de Regras em Sistemas Conexionistas. São Paulo: Universidade Estadual de Campinas, 1995. Dissertação (Mestrado).

ROSENBLATT, F. Principles of Neurodynamics, Washington, DC: Spartan Books.

RUMEHART, D. E., McCLELLAND, J. L. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1, Cambridge, MA: MIT Press, 1986.

SILVA, F. L. H. Modelagem, Simulação e Controle da Fermentação Alcoólica Contínua Extrativa. São Paulo: Universidade Estadual de Campinas. Tese (Doutorado), 1998.

SILVA, F. L. H., RODRIGUES M. I., MAUGERI F. Dynamic Modelling, Simulation and Optimization of an Extractive Continuous Alcoholic Fermentation, Journal Chem. Tech. Biotechnol., v-74-2, p. 176- 182, 1999.

SIMPSON, P. K., Artificial Neural Systems: Foundations, Paradigms, Applications and Implementations, Pergamon Press, Inc., New York, 1990.

DE SOUZA JÚNIOR, M. B. D. Redes Neurais Multicamadas Aplicadas a Modelagem e Controle de Processos Químicos. Rio de Janeiro: Universidade Estadual do Rio de Janeiro, 1993. Tese (Doutorado).

STEPHANOPOULOS, G. Chemical Process Control: An Introduction to Theory and Practice, Englewood Cliffs: Prentice-Hall, 1984.

STOCK, M. AI in Process Control, Intertext Pub.. McGraw-Hill Book Company, New York, 1989.

SU, H. T., McAVOY, T. J. , WERBOS, P. Long-Term Predictions of Chemical Process Using Recurrent Neural Networks: A Parallel Training Approach. Industrial Engineering Chemical Research, v.31-5, 1992.

TOLEDO, E. C. V. Modelagem, Simulação e Controle de Reatores Catalíticos de Leito Fixo. São Paulo: Universidade Estadual de Campinas, 1999. Tese (Doutorado).

VAN ZUBEN, V. J. F. Modelos Paramétricos e Não Paramétricos de Redes Neurais Artificiais e Aplicações. São Paulo: Universidade Estadual de Campinas, 1996. Tese (Doutorado).