

UNIVERSIDADE FEDERAL DE SÃO CARLOS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

INTELIGÊNCIA COMPUTACIONAL
APLICADA À AUTOMAÇÃO DE BIORREATOR
PARA PRODUÇÃO
DE PENICILINA G ACILASE (PGA)

Pedro Luiz Fernandes

SÃO CARLOS
2008

UNIVERSIDADE FEDERAL DE SÃO CARLOS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

INTELIGÊNCIA COMPUTACIONAL
APLICADA À AUTOMAÇÃO DE BIORREATOR
PARA PRODUÇÃO
DE PENICILINA G ACILASE (PGA)

Pedro Luiz Fernandes

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Biotecnologia da Universidade Federal de São Carlos como parte dos requisitos necessários à obtenção do título de Mestre em Biotecnologia.

Orientadores:

Profa. Dra. Heloísa de Arruda Camargo

Prof. Dr. Roberto de Campos Giordano

SÃO CARLOS
2008

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

F362ic

Fernandes, Pedro Luiz.

Inteligência computacional aplicada à automação de
biorreator para produção de Penicilina G Acilase (PGA) /
Pedro Luiz Fernandes. -- São Carlos : UFSCar, 2008.
112 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2008.

1. Sistemas baseados em conhecimento. 2. Automação.
3. Softsensor. 4. Inteligência computacional. 5. Fuzzy logic.
6. Redes neurais (Computação). I. Título.

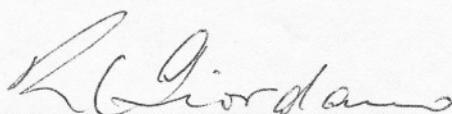
CDD: 006.33 (20^a)

PEDRO LUIZ FERNANDES

Dissertação de Mestrado submetida à Coordenação do Programa de Pós-Graduação em Biotecnologia, da Universidade Federal de São Carlos, como requisito parcial para a obtenção do título de Mestre em Biotecnologia

Aprovado em: 11/07/2008

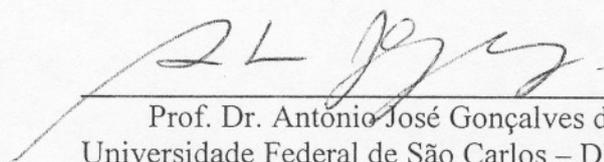
BANCA EXAMINADORA



Prof. Dr. Roberto de campos Giordano (Orientador)
Universidade Federal de São Carlos – DEQ/UFSCar



Profa. Dra. Heloisa de Arruda Camargo (Orientadora)
Universidade Federal de São Carlos - DC/UFSCar



Prof. Dr. Antônio José Gonçalves da Cruz
Universidade Federal de São Carlos – DEQ/UFSCar



Profa. Dra. Valéria Viana Murata
DEQ- Univ. Federal de Uberlândia

Ao nosso passado, onde muitos conhecimentos foram sedimentados, e as lembranças dos mortos nos trazem saudades (em especial aos meus pais), mas com o espírito leve, de quem sabe o conforto proporcionado pelos mortos para nossas vidas.

Ao nosso presente, onde o livre arbítrio na utilização dos conhecimentos passados permite a construção de novos conhecimentos, para utilização dos vivos, e nesse momento, gostaria de agradecer a qualidade de todos os que estão em nossa volta, na família, na universidade e na sociedade. Obrigado por existirem e participarem dessa construção coletiva.

Ao nosso futuro, onde a certeza da nossa morte é só mais um conhecimento. Aos nossos filhos, herdeiros desses conhecimentos.

Conta-se que Hegel, o mais elevado e poderoso dos filósofos românticos – ao mesmo tempo claro e crítico, abstrato e realista, enquanto jazia no seu leito de morte, em 1831, atingido prematuramente pela cólera, era consolado por um dos seus discípulos, amigo íntimo e destacado seguidor. Procurando tranquilizar o mestre, dizia-lhe que, caso viesse a falecer antes de completar sua notável obra enciclopédica, restariam seus fiéis discípulos para prosseguir-la. Hegel, sereno como o silêncio antártico, no exato momento de sua morte, apenas ergueu um pouco a cabeça e murmurou: “Tive um discípulo que me entendeu”. E enquanto todos presentes ficaram alertas para ouvir o nome que o venerado pronunciaria, sua cabeça afundou novamente no travesseiro: “Um discípulo que me entendeu – prosseguiu – e que me entendeu mal”.

Estas penetrantes histórias não precisam ser literalmente verídicas, mas como antigos hieróglifos, amiúde espelham algo verdadeiro.

Heinrich Zimmer, *Filosofias da Índia*, 1986.

Resumo

A Biotecnologia tem apresentado nos últimos anos um grande e rápido desenvolvimento. Constantemente novos processos biotecnológicos industriais são introduzidos, utilizando diferentes microrganismos e/ou enzimas. Neste contexto, a aplicação de técnicas de otimização e controle de processos tornou-se uma necessidade para a indústria de base biotecnológica.

Uma abordagem tecnológica foi desenvolvida nesta dissertação para monitoração e controle de um biorreator, operando em batelada alimentada (“fed-batch”), para produção da enzima penicilina G acilase (PGA) por cepa selvagem do microrganismo *Bacillus megaterium*. Essa enzima é de grande importância industrial, sendo empregada na manufatura de antibióticos β -lactâmicos semi-sintéticos. Este estudo de caso apresenta as principais dificuldades encontradas no controle de processos biológicos em geral: variabilidade dos parâmetros cinéticos e limitada disponibilidade de informação *on-line*. Para superar estas dificuldades foi proposta uma arquitetura não convencional para um controlador dinâmico e adaptativo utilizando filtros, alguns desenvolvidos neste trabalho, e aplicando metodologias da Inteligência Computacional (IC), tanto de forma direta como híbrida.

A inferência da variável de estado concentração microbiana (C_x), objetivo muito importante para a lógica do controlador, foi realizada por um *softsensor* que teve como entrada valores filtrados dos sinais dos sensores das frações molares de CO_2 (y_{CO_2}) e O_2 (y_{O_2}) nos gases efluentes, da vazão de alimentação de ar e da velocidade de agitação. O quociente respiratório (RQ), calculado a partir desses dados, foi também utilizado pelos algoritmos do *software* aqui desenvolvido. Para a inferência de C_x , o *softsensor* empregou um sistema híbrido inteligente (SHI) composto por um comitê de redes neurais artificiais (CRNAs) e por um sistema *fuzzy* baseado em regras (SFBR). Essas técnicas foram estruturadas de modo a se complementarem, pois o CRNAs infere a concentração microbiana (C_x) capturando dados do processo em tempo real (conhecimento empírico), e o SFBR corrige esse valor inferido utilizando conhecimento com base fenomenológica. Os resultados obtidos mostraram uma inferência mais robusta ao se utilizar esta arquitetura, suportando inclusive algum grau de extrapolação.

Outro parâmetro operacional importante é a definição dos momentos de início e fim da vazão de alimentação de meio suplementar. Para alcançar esse objetivo, foi empregada lógica que foi capaz de prever estes momentos com acuidade, utilizando o sinal de fração molar de CO_2 (y_{CO_2}), filtros e conjuntos *fuzzy* adaptativos.

Abstract

Biotechnology has presented, in the last years, a rapidly growing development. New biotechnological industrial processes are constantly introduced using different microorganisms and/or enzymes. In this context, the application of process control and optimization techniques has become a need for biotechnological based industry.

A technological approach was developed in this dissertation to bioreactor monitoring and control operating in fed-batch model to produce the penicillin enzyme G acilase (PGA) by means of wild cepa of the microorganism *Bacillus megaterium*. This enzyme is of great industrial importance, being used in the manufacture of semi-synthetic β -lactamic antibiotics. This case study presents the main difficulties faced in the control of biological processes in general: the variability of kinetic parameters and the limited availability of on-line information. In order overcome these difficulties, a unconventional architecture was proposed for a dynamic and adaptive controller using filters, some of them developed in this work, and applying Computational Intelligence (CI) methodologies both in direct and hybrid form.

The inference for the microbial concentration (C_x) state variable, a very relevant objective for the logic of the controller, was performed by a *softsensor* that had as input the filtered values of the sensors signals of the molar fractions of CO_2 (y_{CO_2}) and O_2 (y_{O_2}) in the effluent gases, of the air feeding flow and of agitation velocity. The respiratory quotient (RQ), calculated from these data, was also used by the algorithms of the *software* developed here. For the C_x inference, the *softsensor* employed a hybrid intelligent system (HIS) composed by a neural networks ensemble (RNE) and a fuzzy rule based system (FRBS). These techniques were structured to complement each other such that the RNE infers the microbial concentration (C_x) capturing real-time process data (empirical knowledge) and the FRBS corrects this inferred value using phenomenological based knowledge. The obtained results demonstrated a more robust inference by using this architecture, even supporting some degree of extrapolation.

Another important operational parameter is the definition of the initial and final instants of feeding flow of supplemental. In order to meet this goal, a logic was employed that is able to accurately predict this moments, using the CO_2 (y_{CO_2}) molar fraction signal, filters and adaptive fuzzy sets.

LISTA DE TABELAS

Tabela 1 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA) do ensaio 1.....	109
Tabela 2 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA) do ensaio 2.....	109
Tabela 3 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA) do ensaio 3.	110
Tabela 4 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA) do ensaio 4.....	110
Tabela 5 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais do ensaio 5.....	111
Tabela 6 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais do ensaio 6.....	112
Tabela 7 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais do ensaio 7.....	112
Tabela 8 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais do ensaio 8.....	113
Tabela 9 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais do do ensaio 9.....	113

LISTA DE FIGURAS

2. REVISÃO BIBLIOGRÁFICA	5
Figura 2.1 Estrutura química: (a) ácido 6-aminopenicilânico (6-APA) e (b) anel β -lactâmico.....	5
Figura 2.2 Estrutura química da: (a) penicilina G (PG) e (b) penicilina V (PV)	5
Figura 2.3 Estrutura química da: (a) amoxicilina e (b) ampicilina.....	6
Figura 2.4 Reação de hidrólise enzimática da penicilina G para a obtenção do ácido 6-aminopenicilânico (6-APA) e do ácido fenilacético (AFA) (RAJENDHRAN e GUNASEKARAN, 2004)	7
Figura 2.5 Esporo do <i>Bacillus megaterium</i> . Legenda: (1) cortex, (2) capa externa do cortex, (3) membrana plasmática, (4) parede celular, (5) Cobertura de esporo, (6) extremo germinativo e (7) nucleóide (Disponível em: < http://fai.unne.edu.ar/biologia/bacterias/micro7.htm >. Acesso em: 22/06/2008.)	8
Figura 2.6 Imagem feita em um microscópio eletrônico do <i>Bacillus megaterium</i> e <i>Escherichia coli</i> (VARY <i>et al.</i> , 2007)	9
Figura 2.7 Níveis de pesquisa e de modelagem de atividades em um bioprocesso (adaptado de REHM e REED, 1985)	10
Figura 2.8 (a) Esquema de uma célula e (b) Esquema do crescimento celular, onde as letras representam: (x) células, (s _i) substratos, (μ) velocidade específica de crescimento celular e (p _j) produtos (adaptado de NIELSEN E VILLADSEN, 1994).....	11
Figura 2.9 Esquema das interações entre os microrganismos e o seu ambiente (meio de cultivo) (adaptado de REHM e REED, 1985).	12
Figura 2.10 Esquema de um biorreator operando em batelada (SBTR).....	14
Figura 2.11 Esquema de um biorreator operando em batelada alimentada (“fed-batch”)	15
Figura 2.12 Ilustração da cinética de crescimento celular (x), de consumo de substrato (s) e de formação de produto (p) em um biorreator operando em batelada (SBTR)	16
Figura 2.13 Esquema das interações entre os microrganismos e o seu ambiente (meio de cultivo) (adaptado de anotações de aula de Alberto Colli Badino)	17
Figura 2.14 Gráficos de log C _x versus tempo e as fases de crescimento microbiano (1-6).....	18
Figura 2.15 Gráfico de μ versus C _s usando o modelo de MONOD.....	26

Figura 2.16 Gráficos dos diferentes momentos de formação do produto em relação ao crescimento microbiano: (a) se formam juntos; (b) produto é formado após a fase exponencial e (c) produto é formado no meio da fase exponencial e continua na fase estacionária.....	27
Figura 2.17 Gráfico de μ versus C_s usando o modelo de Andrews (MONOD com inibição).....	29
Figura 2.18 Exemplos de função de pertinência triangular.....	38
Figura 2.19 Exemplos de função de pertinência trapezoidal.....	38
Figura 2.20 Exemplos de função de pertinência gaussiana.....	39
Figura 2.21 (a) Conjuntos A e B, (b) união, (c) interseção e (d) complemento de A.....	42
Figura 2.22 Variável lingüística <i>velocidade</i> e sua divisão, em categorias, através de conjuntos <i>fuzzy</i>	45
Figura 2.23 Funcionamento de um neurônio computacional.....	50
Figura 2.24 Funções: (a) transferência de limite, (b) linear e (c) sigmoidal. Os eixos a e n correspondem a saída e a entrada da rede, respectivamente.....	51
Figura 2.25 Estrutura básica de treinamento supervisionado de uma rede neural.....	52
Figura 2.26 Estrutura esquemática do <i>perceptron</i> multicamadas (MLP).....	54
Figura 2.27 Estrutura esquemática do método Ensemble Average.....	55
Figura 2.28 Estrutura esquemática dos métodos (a) <i>mixture experts</i> e (b) <i>hierarchical mixture experts</i>	55
3. MATERIAIS E MÉTODOS.....	59
Figura 3.1 Esquema da preparação do experimento desde a etapa de preparação dos criotubos até o experimento em biorreator.....	59
Figura 3.2 Esquema do aparato experimental com sistema de aquisição de dados (CRUZ, 2000).....	60
Figura 3.3 Esquema do aparato experimental e variáveis, para cálculo do coeficiente respiratório, em um balanço material (OLIVEIRA, 2006).....	62
Figura 3.4 Esquema do treinamento de cada rede neural que pertence ao comitê (CRNAs).....	64
Figura 3.5 Estrutura dinâmica do comitê, com seus dois componentes (BD RNAs e MP), utilizado na inferência da concentração microbiana (C_x).....	65
Figura 3.6 Formato da 1ª partição de entrada <i>fuzzy</i> para a fração molar de CO_2 (y_{CO_2}).....	67
Figura 3.7 Formato da 2ª partição de entrada <i>fuzzy</i> para o quociente respiratório (RQ).....	68

Figura 3.8 Formato da partição de saída dos conjuntos <i>fuzzy</i> para correção do valor predito de crescimento microbiano (Cx), pelo comitê (CRNAs)	69
Figura 3.9 Resultado da simulação, feita no instante <i>i</i> , do ensaio2, mostrando o mecanismo de inferência do sistema <i>fuzzy</i> (SFBR).....	70
Figura 3.10 Esquema do método de validação cruzada utilizada nesta dissertação	71
Figura 3.11 Gráficos de y_{CO_2} versus tempo, para ensaio validado, cor azul, e para os ensaios treinados pelas RNAs especialistas do comitê (CRNAs), cor vermelha: (a) do ensaio um e (b) do ensaio dois. A seta de cor verde, nos gráficos, indica regiões de extrapolação.....	72
4. ARQUITETURA DO SISTEMA DE SUPERVISÃO E CONTROLE PROPOSTO	75
Figura 4.1 Esquema geral do funcionamento do controlador proposto	75
Figura 4.2 Esquema de funcionamento do módulo <i>Controle do Cultivo</i>	81
Figura 4.3 Gráficos de distribuição da derivada de y_{CO_2} nos ensaios um, dois, três e quatro	82
Figura 4.4 Distribuição central, da derivada de y_{CO_2} do ensaio um: (a) da 1 ^a hora, (b) da 2 ^a a 5 ^a hora, (c) da 6 ^a a 10 ^a hora e (d) da 11 ^a hora.....	83
Figura 4.5 Gráfico da y_{CO_2} , do tratamento dado pelo filtro “alisa pós”, a uma perda de sinal do sensor, ocorrido no ensaio um.....	85
Figura 4.6 Esquema de funcionamento do módulo <i>Controle da Vazão</i>	87
Figura 4.7 Gráfico dos filtros “alisa pós”, de cor vermelha, e “escada”, de cor preta	88
Figura 4.8 Esquema de funcionamento do módulo <i>Inferência de Cx</i>	89
Figura 4.9 Esquema de funcionamento do módulo <i>Núcleo Controlador</i>	91
5. RESULTADOS E DISCUSSÃO	92
Figura 5.1 Gráfico da y_{CO_2} versus tempo do sinal do sensor e do filtro “alisa pós”, para o tratamento da 1 ^a hora: (a) do ensaio cinco e (b) do ensaio seis	92
Figura 5.2 Gráfico da y_{CO_2} versus tempo do sinal do sensor e do filtro “alisa pós”, para o tratamento de vinte horas do cultivo: (a) do ensaio cinco e (b) do ensaio seis	93
Figura 5.3 Gráfico da y_{CO_2} versus tempo do filtro “alisa pós”, para a identificação das fases do crescimento celular (Cx): (a) do ensaio cinco e (b) do ensaio seis	93
Figura 5.4 Gráfico dos filtros “alisa pós” e “escada”, e os valores utilizados pela lógica <i>fuzzy</i> para inferir os momentos de início e fim da vazão de alimentação de meio suplementar.....	94

Figura 5.5 Domínio da magnitude do módulo da derivada de y_{CO_2} (Dy_{CO_2}) e seus conjuntos <i>fuzzy</i> : (a) do ensaio cinco e (b) do ensaio seis.....	95
Figura 5.6 Visualização da inferência <i>fuzzy</i> para os momentos de início e fim de vazão de alimentação do ensaio cinco: (a) com os filtros “alisa pós” e “escada” e (b) com os dados experimentais de crescimento celular (C_x) e consumo de substrato (C_s).....	95
Figura 5.7 Visualização da inferência <i>fuzzy</i> para os momentos de início e fim de vazão de alimentação do ensaio seis: (a) com os filtros “alisa pós” e “escada” e (b) com os dados experimentais de crescimento celular (C_x) e consumo de substrato (C_s).....	96
Figura 5.8 Visualização do resultado da inferência da concentração celular (C_x) do comitê (CRNAs) juntos dos dados experimentais: (a) do ensaios um e (b) do ensaio dois	97
Figura 5.9 Visualização do resultado da inferência da concentração celular (C_x) do comitê (CRNAs) juntos dos dados experimentais: (a) do ensaios três e (b) do ensaio quatro.....	97
Figura 5.10 Comparação dos erros, medidos pela diferença entre os dados experimentais da concentração celular (C_x) e os C_x inferidos pelo comitê (CRNAs) e pelo sistema híbrido (SHI): (a) para o ensaios um e (b) para o ensaio dois.....	99

SUMÁRIO

Resumo	i
Abstract.....	ii
LISTA DE TABELAS	iii
LISTA DE FIGURAS	iv
1. INTRODUÇÃO	1
2. REVISÃO BIBLIOGRÁFICA	4
2.1 Bioprocesso: cultivo microbiano	4
2.1.1 Penicilinas.....	4
2.1.2 Enzima Penicilina G Acilase (PGA)	7
2.1.3 <i>Bacillus Megaterium</i>	8
2.2 Engenharia das fermentações: cultivo microbiano.	9
2.2.1 Biorreatores.	13
2.2.1.1 Operação em batelada ou descontínua	14
2.2.1.2 Operação em batelada alimentada ou semicontínuo.....	15
2.2.2 Cinética de crescimento microbiano.....	16
2.2.2.1 Fases do crescimento microbiano.....	18
2.2.2.2 Coeficientes de Rendimento.....	20
2.2.2.3 Estequiometria para Balanços Elementais.....	21
2.2.2.4 Coeficiente respiratório	22
2.2.3 Modelos matemáticos para um cultivo microbiano.....	23
2.2.3.1 Classificação de Modelos	24
2.2.3.2 Modelos para velocidade específica de crescimento.....	25
2.2.3.3 Modelos para múltiplos substratos (s_i)	27
2.2.3.4 Modelos de formação de produtos (p_j)	27
2.2.4 Batelada versus batelada alimentada	30
2.3 Inteligência computacional para controle de bioprocessos.....	32
2.3.1 Introdução.....	32

2.3.2 <i>Softsensors</i> ou sensores virtuais	34
2.3.3 Inteligência computacional (IC)	35
2.3.4 Lógica <i>fuzzy</i> (nebulosa)	36
2.3.4.1 Conjuntos <i>fuzzy</i>	36
2.3.4.2 Tipos de funções de pertinência	38
2.3.4.3 Definições básicas	39
2.3.4.4 Operações em conjuntos <i>fuzzy</i>	40
2.3.4.5 Computação com regras <i>fuzzy</i>	43
2.3.4.6 Sistema <i>fuzzy</i> baseado em regras	49
2.3.5 Redes neurais artificiais	50
2.3.5.1 Comitê de redes neurais	54
2.3.6 Sistema Inteligente Híbrido (SIH)	56
3. MATERIAIS E MÉTODOS	57
3.1 Materiais	57
3.1.1 Microorganismo	57
3.1.2 Principais equipamentos	57
3.2 Métodos	59
3.2.1 Fermentação principal	59
3.2.2 Sistema de aquisição de dados	59
3.2.3 Cálculo do quociente respiratório (RQ)	61
3.2.4 Metodologia do Comitê de redes neurais artificiais (CRNAs)	64
3.2.4.1 BD RNAs especialistas	64
3.2.4.2 Mecanismo de ponderação	65
3.2.5 Metodologia do sistema <i>fuzzy</i> baseado em regras (SFBR)	67
3.2.5.1 Base de conhecimento <i>fuzzy</i>	67
3.2.5.2 Mecanismo de inferência	70
3.2.6 Método de validação cruzada	71
3.2.7 Metodologia do sistema híbrido inteligente (SHI)	72

4. ARQUITETURA DO SISTEMA DE SUPERVISÃO E CONTROLE PROPOSTO	74
4.1 Introdução	74
4.2 Descrição do controlador dinâmico e adaptativo	74
4.2.1 Funcionamento geral do controlador	76
4.2.2 Lógica da vazão de alimentação do meio suplementar	77
4.3 Base de Conhecimento	78
4.4 Módulo 1 – Controle do Cultivo	80
4.4.1 Filtro “alisa pós”	81
4.4.1.1 Etapas de construção do filtro	83
4.4.1.2 Identificação das fases de crescimento celular (Cx).....	86
4.5 Módulo 2 – Controle da Vazão de meio suplementar	86
4.5.1 Filtro “escada”	87
4.5.2 Inferência <i>fuzzy</i> para definição dos momentos de início e fim de vazão de alimentação	88
4.6 Módulo 3 – Inferência de Cx	89
4.7 Módulo 4 – Núcleo Controlador	90
5. RESULTADOS E DISCUSSÃO	92
5.1 Resultados do módulo Controle do Cultivo	92
5.1.1 Discussões dos resultados.....	93
5.2 Resultados do módulo Controle da Vazão de meio suplementar.....	94
5.2.1 Discussões dos resultados.....	96
5.3 Resultados do módulo Inferência de μ	97
5.3.1 Discussões dos resultados.....	99
6. CONCLUSÕES E SUGESTÕES	100
REFERÊNCIAS BIBLIOGRÁFICAS	103
APÊNDICE A – Descrição e dados dos ensaios	108

1. INTRODUÇÃO

A tecnologia, no sentido mais abrangente do seu significado, é o centro de uma revolução que tem como aliados a internet, os softwares e os computadores. As tecnologias da informação e da comunicação derrubaram as fronteiras entre países, criando novas oportunidades de mercado e trazendo, atreladas a si, uma maior concorrência. Conseqüentemente ocorreram mudanças na sociedade, para as quais o conhecimento é a principal mola propulsora (REZENDE, 2005). Em um ambiente mais competitivo, a opção de adquirir um determinado produto/serviço pode estar muito relacionada a uma maior qualidade e/ou menor custo. Novos conhecimentos sobre este produto/serviço podem gerar estes atributos.

Neste contexto, a Biotecnologia, ramo que integra as ciências naturais e de engenharia com vista à aplicação industrial de organismo, células, partes destas e análogos moleculares para a obtenção de produtos e serviços (definição dada pela EUROPEAN FEDERATION OF BIOTECHNOLOGY em 1994), tem apresentado um grande e rápido desenvolvimento. Novos processos biotecnológicos utilizando microrganismos estão sendo constantemente introduzidos pela indústria. Esses bioprocessos são usados para a produção de uma vasta gama de produtos, com diferentes valores comerciais, e têm atualmente uma grande aplicação em setores tais como o agro-alimentar, o de química fina e farmacêutico, de energia e ambiente (FERREIRA, 1995).

Produtos bioquímicos, especialmente penicilinas, têm um importante valor agregado que continua impulsionando as pesquisas em vários campos relacionados à sua produção. Mais de quatrocentas (400) cepas com a capacidade de produzir penicilinas foram classificadas e multiplicadas por produção de mutantes (PAUL, KENT, e THOMAS, 1993) e uso de engenharia genética (AGRAWAL, DEEPIKA, e JOSEPH, 1999). A melhora de produtividade chegou a cepas que fornecem 30 g/L de penicilina G. Apesar das novas cepas terem alta produtividade, suas fermentações devem obedecer a certas condições muito controladas.

A produção de enzimas é de grande importância para a indústria de base biotecnológica. Sua importância é o fato de serem catalisadores interessantes devido a sua alta especificidade, alta eficiência na conversão substrato-produto e sua elevada atividade em condições ambientais moderadas, e estas características explicam o grande interesse em sua utilização em processos industriais (ILLANES, 1994). O reconhecimento da importância atual das enzimas é citado no relatório final apresentado no Fórum de Competitividade em

Biotecnologia em 2006, realizado pelo Centro de Gestão e Estudos Estratégicos (CGEE). Este relatório prevê que diante das necessidades e tendências atuais, de utilização de matérias primas renováveis e de preservação ambiental, haverá um aumento significativo do consumo de enzimas a nível internacional. Neste cenário, o relatório alerta da importância do Brasil inserir-se de forma representativa como usuário de tecnologia enzimática.

Penicilina G acilase (PGA) é a segunda enzima mais utilizada industrialmente no mundo na forma imobilizada, e sua principal aplicação é na hidrólise da penicilina G e cefalosporina para a produção dos ácidos 6-aminopenicilânico (6-APA) e 7-aminocefalosporânico (7-ACA). O 6-APA é um composto-chave para a produção industrial dos antibióticos β -lactâmicos semi-sintéticos, tais como ampicilina, amoxicilina e cefalexina, com larga aplicação como agentes terapêuticos (BERAZAÍN, 1997). Dentre os microrganismos produtores da enzima PGA, *Bacillus megaterium* é um dos poucos que a produzem de modo extracelular, o que torna particularmente interessante o estudo do processo de produção de PGA utilizando este microrganismo, pois os procedimentos de purificação tornam-se sobremaneira simplificados e, conseqüentemente, menos caros. Há quatorze anos, o grupo de pesquisas da UFSCar, de Engenharia Bioquímica, possui uma linha de pesquisa voltada para o processo de produção de Penicilina G Acilase (PGA) por *Bacillus megaterium* em biorreatores.

O uso de biorreatores tipo tanque agitado e aerado (convencional), operando em batelada alimentada, é muito popular na indústria de fermentação, farmacêutica e de química fina. Este modo de operação é usualmente utilizado para superar principalmente os problemas relacionados com inibições e/ou por repressão catabólica (por exemplo, algum substrato ou algum catabólito reprime a rota metabólica responsável pela biossíntese de alguns produtos de interesse).

A implementação de estratégias (por vezes automáticas) de monitoração e controle da operação do biorreator, com o intuito de se assegurar de modo estável e reprodutível, com uma produtividade máxima com um mínimo de custos, tem sido uma imposição do ambiente altamente competitivo em que se inserem os processos produtivos baseados em biotecnologia. A importância econômica da obtenção de altos rendimentos justifica a introdução de estratégias de monitoração e controle, para acompanhar o processo de um modo regular e preciso, de forma a garantir certa reprodutibilidade. Estes requisitos ganham uma grande importância na situação de utilização de metabólitos secundários para fins terapêuticos (ROYCE, 1993). Neste caso, a aprovação por agências reguladoras (como a ANVISA no Brasil e o FDA nos EUA) implica a validação de todo processo, com seu controle

estabelecido de modo a garantir uma qualidade consistente do produto (GERSON *et al.*, 1988; WERNER e LANGLOUIS-GAU, 1989). Um segundo fator prende-se à necessidade de se proteger a propriedade intelectual já na fase de desenvolvimento de um produto biotecnológico: o melhoramento da monitoração do processo e a remoção de etapas limitantes poderão minimizar os custos de investigação e desenvolvimento normalmente associados a esta fase. A existência de um gargalo tecnológico, que conduzia muitas vezes a soluções pouco racionais ao se aumentar a escala de um processo patenteado e aprovado por agências regulatórias, foi reconhecida pela própria política de regulação industrial dos EUA nessa área. De fato, o FDA (*Food and Drug Administration*, EUA) lançou, em 2003, o *PAT Process Analytical Technology Initiative* (Disponível em <<http://www.fda.gov/cder/OPS/PAT.htm>>. Acesso em 10/01/2008). Com esse programa, procura-se conciliar a rapidez no patenteamento e aprovação pelo FDA de produtos/processos com a utilização de ferramentas avançadas para controle, programação de produção e otimização de processos industriais (ferramentas da chamada engenharia de sistemas e bioprocessos). Esse programa foi um grande incentivo à pesquisa nesse campo, no qual se inclui o tema desta dissertação.

O estudo cinético dos componentes de um bioprocessos, como o de produção de PGA, apresenta tipicamente um comportamento não-linear, com variabilidade dos parâmetros cinéticos e uma limitada disponibilidade de informação *on-line*. Modelos matemáticos desses sistemas, baseados em enfoques clássicos da engenharia de processos, resultam muitas vezes em problemas intratáveis em virtude do número excessivo de parâmetros, ou no extremo oposto, são simplificações da realidade que não conseguem captar aspectos essenciais do caso em estudo.

Para tratar desses problemas uma arquitetura não convencional de supervisão e controle foi proposta nesse trabalho utilizando uma abordagem tecnológica. Esta arquitetura de monitoração e controle do biorreator para o processo de produção de PGA, pelo microrganismo selvagem *Bacillus megaterium*, foi representada por um controlador dinâmico e adaptativo. Este controlador se utilizou de filtros desenvolvidos especialmente para este trabalho, de um *softsensor* que teve como entrada os valores filtrados dos sinais dos sensores das frações molares de CO₂ (y_{CO_2}) e O₂ (y_{O_2}) nos gases efluentes, velocidade de agitação e quociente respiratório (RQ) e das metodologias da Inteligência Computacional (IC), de forma simples, e de forma híbrida (duas ou mais metodologias utilizadas ao mesmo tempo), já que essas metodologias são complementares, o que torna freqüentemente vantajoso usá-las de forma combinada, levando aos chamados sistemas híbridos inteligentes (SHI).

2. REVISÃO BIBLIOGRÁFICA

2.1 Bioprocesso: cultivo microbiano

Um processo biotecnológico é definido como aquele em que há utilização de células, microrganismos e/ou enzimas, podendo ter como objetivo: a síntese de compostos químicos intracelulares ou extracelulares; a produção de biomassa (fermento de padeiro, por exemplo), de alimentos, de energia (biogás, bioetanol); ou ainda a despoluição biológica de efluentes, como na degradação da matéria poluente pelos microrganismos (FERREIRA, 1995). O uso industrial do termo fermentação se refere ao cultivo em geral de microrganismo em meio sólido ou líquido, quer em ambiente aeróbio ou anaeróbio. Do ponto de vista bioquímico, entretanto, essa nomenclatura carece de rigor e, nesta dissertação, utiliza-se a alternativa “cultivo microbiano”.

Nesta revisão bibliográfica será descrito o processo de cultivo microbiano para produção da enzima penicilina G acilase (PGA) pelo microorganismo *Bacillus megaterium*. Esta enzima é um produto chave para produção das penicilinas semi-sintéticas.

2.1.1 Penicilinas

Antibióticos são tipos específicos de substâncias químicas que combatem infecções usualmente bacterianas em seres humanos e animais. Penicilinas são antibióticos que interferem no metabolismo celular da bactéria, inibindo as enzimas transpeptidase e carboxipeptidase, responsáveis pela síntese dos peptidoglicanos. O peptidoglicano é a estrutura que confere rigidez à parede celular de bactérias, determina sua forma e a protege da lise osmótica, quando em meio hipotônico. Sem a proteção da parede celular, a célula provavelmente morrerá por absorção de água, aumento de volume e conseqüentemente rompimento. Bactérias gram-negativas têm o peptidoglicano situado entre a membrana interior e a membrana exterior. Bactérias gram-positivas não possuem membrana exterior, cumprindo o peptidoglicano esta função e tendo em geral uma espessura superior à encontrada em bactérias gram-negativas (LEHNINGER *et al.*, 2006).

A penicilina é o antibiótico mais amplamente utilizado, sendo em geral produzida pelo fungo *Penicillium chrysogenum*. A sua descoberta tornou possível o tratamento de várias infecções provocadas por bactérias, representando um grande avanço no tratamento de várias doenças. As penicilinas possuem um núcleo comum denominado ácido 6-aminopenicilânico (6-APA), (Figura 2.1a), que embora seja biologicamente inativo, é o

responsável pela atividade de todas as penicilinas. O 6-APA contém dois anéis: um de quatro átomos, chamado anel β -lactâmico (Figura 2.1b), e outro de cinco, chamado anel tiazolidínico. A existência do anel β -lactâmico em todas as penicilinas naturais ou semi-sintéticas é responsável pela designação geral de antibióticos β -lactâmicos. Atualmente, os antibióticos β -lactâmicos são os mais utilizados devido à sua eficiência contra várias infecções bacterianas e aos poucos efeitos colaterais por eles causados (BAGGALEY *et al.*, 1997).

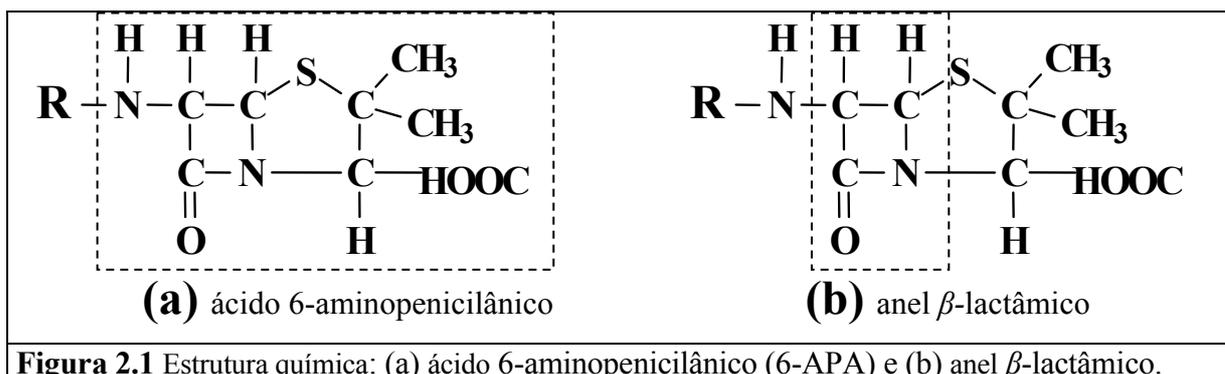


Figura 2.1 Estrutura química: (a) ácido 6-aminopenicilânico (6-APA) e (b) anel β -lactâmico.

As várias penicilinas existentes apresentam estruturas químicas semelhantes, com propriedades e atividades ligeiramente diferentes. Diferenciam-se entre si pela cadeia lateral ligada ao 6-APA (o radical R, na Figura 2.1a). Estas diversas cadeias laterais conferem propriedades especiais a cada derivado. Dependendo da forma de obtenção, as penicilinas podem ser classificadas em biossintéticas e semi-sintéticas (SILVA, 2003).

As penicilinas biossintéticas, ou naturais, são obtidas a partir do cultivo de microrganismos. As de maior interesse industrial e econômico, devido a sua importância terapêutica, são as penicilina G (Figura 2.2a) e penicilina V (Figura 2.2b). Ambas apresentam um espectro de atividade contra bactérias gram-positivas. Apenas a penicilina V é administrada via oral uma vez que a penicilina G é degradada pelo meio ácido do estômago, sendo portanto administrada por via intramuscular. Para a produção exclusiva de penicilina G, durante um cultivo microbiano, devemos fornecer o ácido fenilacético (AFA). Já para a produção exclusiva de penicilina V, devemos fornecer o ácido fenoxiacético (AFNA).

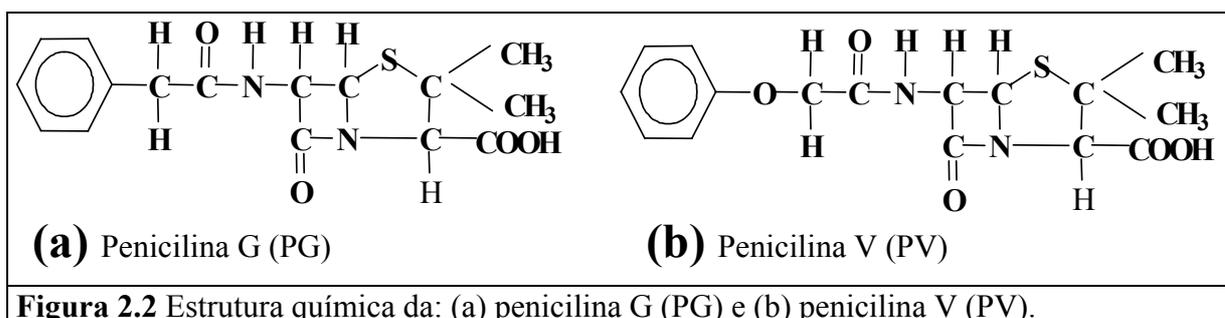
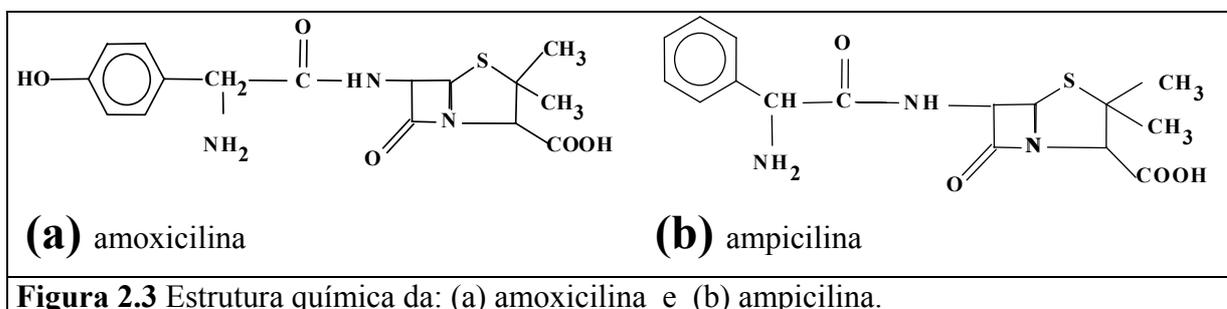


Figura 2.2 Estrutura química da: (a) penicilina G (PG) e (b) penicilina V (PV).

As penicilinas naturais podem ser destruídas por enzimas denominadas penicilases. Estas enzimas hidrolisam o anel β -lactâmico do núcleo básico da molécula e por esta razão são também denominadas β -lactamases. A abertura deste anel impede a penicilina de atuar no bloqueio da formação da parede celular de bactérias, não combatendo assim a infecção. Bactérias que produzem a enzima penicilase são assim resistentes a antibióticos β -lactâmicos.

Para combater microrganismos patogênicos, tem-se buscado novos antibióticos. As penicilinas semi-sintéticas têm se destacado por apresentarem estabilidade em meio ácido, podendo assim ser administradas oralmente, melhor absorção, efetividade a baixa concentração para uma ampla variedade de organismos gram-negativos e gram-positivos (HOU e POOLE, 1969), além de apresentarem menor número de efeitos colaterais que as penicilinas naturais. O desenvolvimento de penicilinas semi-sintéticas passa pela produção do precursor 6-APA, que pode ser obtido a partir da desacilação de penicilinas naturais, principalmente as penicilinas G ou V. Para obter o 6-APA foi desenvolvido um método químico de alto rendimento que recorria ao uso de compostos muito tóxicos e corrosivos e utilizavam temperaturas extremamente baixas (de -30°C a -40°C). Posteriormente foi desenvolvido o método enzimático, com enzimas imobilizadas para produção de 6-APA. O método enzimático representou um avanço em relação ao método químico devido à estereoespecificidade das enzimas, e comparações econômicas realizadas da manufatura do 6-APA pelos dois métodos mostraram que o processo enzimático é pelo menos 9% mais barato (ANON, 1992 apud PARMAR *et al.*, 2000). As vantagens econômicas, ambientais e operacionais do processo enzimático sobre a rota química fizeram com que praticamente todos os produtores de 6-APA utilizassem o método enzimático.

Entre alguns dos antibióticos semi-sintéticos produzidos a partir do 6-APA destacam-se a amoxicilina (Figura 2.3a) e a ampilina (Figura 2.3b).



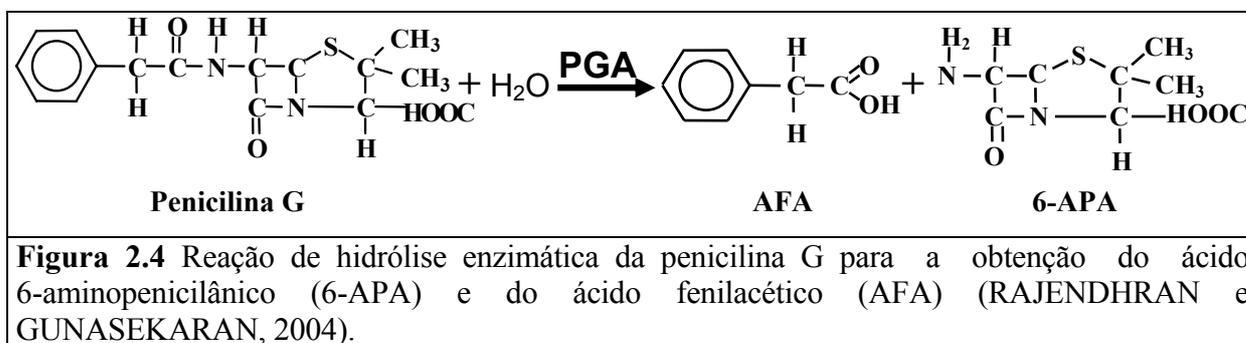
O grande aumento do uso de enzimas em processos industriais é devido à sua alta especificidade, alta eficiência na conversão substrato-produto e elevada atividade em condições ambientais moderadas (fisiológicas). Uma grande quantidade de microorganismo

pode produzir enzimas para hidrolisar a ligação amida da penicilina, para se obter o 6-APA e sua respectiva cadeia lateral.

2.1.2 Enzima Penicilina G Acilase (PGA)

Enzimas são um grupo de substâncias orgânicas de natureza normalmente protéica (existem também enzimas constituídas de RNA, as ribozimas). Possuem atividade intra ou extracelular com funções catalisadoras, promovendo reações químicas que dificilmente aconteceriam na sua ausência. Elas atuam de forma a diminuir a energia de ativação necessária para que se dê uma reação química, resultando no aumento da velocidade da reação e possibilitando o metabolismo dos seres vivos. A capacidade catalítica das enzimas torna-as adequadas para aplicações industriais. Em sistemas vivos, a maioria das reações bioquímicas ocorre em rotas metabólicas, que são seqüências de reações em que o produto de uma reação é utilizado como reagente na reação seguinte. Diferentes enzimas catalisam diferentes passos de rotas metabólicas, agindo de forma concertada de modo a não interromper o fluxo nessas vias. Cada enzima pode sofrer regulação da sua atividade, aumentando-a, diminuindo-a ou mesmo interrompendo-a, de modo a modular o fluxo da via metabólica em que se insere.

A enzima penicilina acilase foi isolada inicialmente a partir de *Penicillium chrysogenum* Q176 (PARMAR *et al.*, 2000). Sua classificação é feita de acordo com a fonte a partir da qual a enzima penicilina acilase foi isolada. As provenientes de bactérias têm maior afinidade para hidrolisar a penicilina G e as provenientes de actinomicetos e fungos têm maior afinidade para hidrolisar a penicilina V. A penicilina G acilase (PGA) é uma enzima de grande importância na produção industrial de antibióticos β -lactâmicos (BRUGGINK *et al.*, 1998) e seu nome oficial é penicilina amido hidrolase (E.C.3.5.1.11). Sua grande importância reside na sua utilização como catalisador na hidrólise da penicilina G para a obtenção de 6-APA e ácido fenilacético (AFA) (Figura 2.4), normalmente em processo com enzimas imobilizadas (SAVIDGE, 1984).



A PGA apresenta diferentes características e propriedades, como diferentes pesos moleculares e afinidades pelos substratos, dependendo do microorganismo que a produz. As

bactérias gram-negativas, como a *Escherichia coli*, produzem a enzima PGA acumulada no espaço periplasmático (intracelular). A purificação da enzima intracelular passa pela ruptura das células, encarecendo os custos para sua recuperação. As bactérias gram-positivas, como o *Bacillus megaterium*, excretam a enzima PGA para o meio de cultura (extracelular), simplificando os procedimentos de purificação e barateando os custos de sua recuperação. Segundo SAVIDGE e COLE, 1975, a massa molecular da PGA encontra-se em torno de 120.000 Da.

2.1.3 *Bacillus megaterium*

VARY, 1994, descreve que *Bacillus megaterium* tem fascinado microbiologistas por causa da sua fisiologia e da ampla extensão de seu habitat ecológico. Apesar de ser considerado proveniente do solo, pode ser encontrada em diversos ambientes, tais como alimentos desidratados, água marinha, sedimentos, peixes e até mesmo em mel de abelha (SILVA, 2003). Aproximadamente 1000 mutantes geneticamente caracterizados estão disponíveis do *Bacillus megaterium* no Genetic Stock Center e no Laboratório de Vary (Disponível em <http://www.bios.niu.edu/b_megaterium/>. Acesso em 10/01/2008). É um microrganismo aeróbio e segundo BERGEY, 1974, pode ser encontrado nas formas oval ou cilíndrica, em pequenas cadeias, podendo apresentar diferentes dimensões dependendo do meio de cultura. Apresenta capacidade de esporular e seus esporos, como mostra a Figura 2.5, podem assumir formas ovaladas ou alongadas (VARY, 1994).

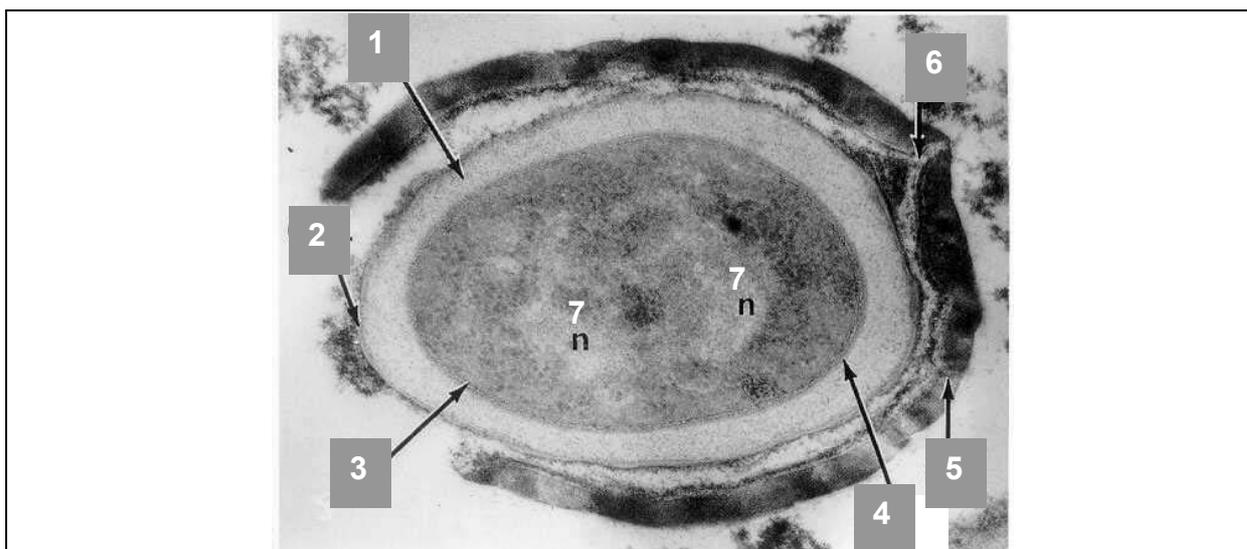


Figura 2.5 Esporo do *Bacillus megaterium*. Legenda: (1) cortex, (2) capa externa do cortex, (3) membrana plasmática, (4) parede celular, (5) Cobertura de esporo, (6) extremo germinativo e (7) nucleóide (Disponível em:<<http://fai.unne.edu.ar/biologia/bacterias/micro7.htm>>. Acesso em: 22/06/2008.).

A Figura 2.6 mostra uma imagem do *Bacillus megaterium* junto de uma *Escherichia coli*. Enquanto um *Bacillus megaterium* pode atingir um volume maior do que $60 \mu\text{m}^3$ ($2,5 \times 2,5 \times 10$) uma *Escherichia coli* atinge um volume de $0,5 \mu\text{m}^3$ ($0,5 \times 0,5 \times 2$) (VARY *et al.*, 2007).

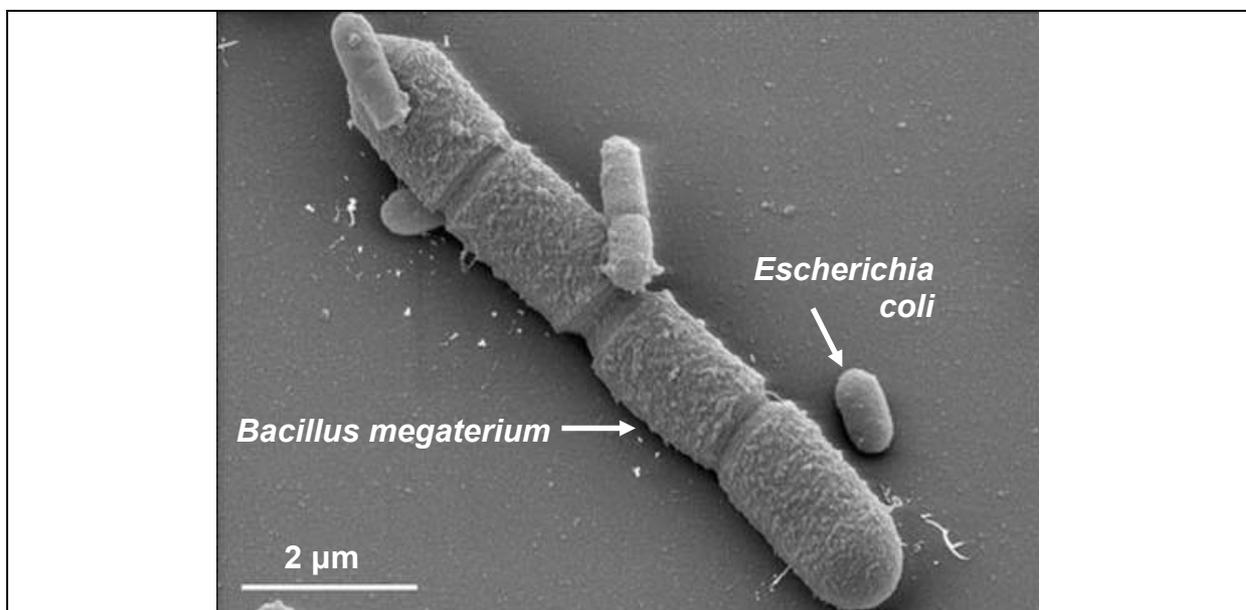


Figura 2.6 Imagem feita em um microscópio eletrônico do *Bacillus megaterium* e *Escherichia coli* (VARY *et al.*, 2007).

Sua importância econômica se deve ao fato de produzir penicilina acilase, amilases, glicose desidrogenase, antibióticos (emicina, oxitanocina), esteróides, hidrolases e de ser o microorganismo aeróbio maior produtor de vitamina B₁₂ (HOJO, 1997). Além disso, é um microorganismo promissor para manipulação genética, seja como hospedeiro ou como fornecedor de genes (KANG *et al.*, 1991).

Quanto às suas propriedades fisiológicas, sabe-se que sua temperatura ótima de crescimento está entre de 28° a 35 °C, não ocorrendo nenhum crescimento acima de 45 °C. O pH ótimo está entre 7,0 e 8,5 e nenhum crescimento ocorre a pH 5,0 ou menor (MURAO *et al.*, 1964).

2.2 Engenharia das fermentações: cultivo microbiano

Existem quatro padrões gerais de crescimento microbiano, associados às bactérias, leveduras, fungos filamentosos e vírus. Vírus não seguem crescimento normal porque requerem um organismo hospedeiro, e seu crescimento exponencial tem expoente muito maior do que dois. Bactérias, leveduras e fungos filamentosos crescem exponencialmente em um ambiente não limitante. Bactéria se divide por fissão, leveduras por germinação e fungos filamentosos por alongamento encadeado e ramificação. Sob o ponto de vista cinético, a

diferenciação morfológica é um item adicional, e deve ser levado em consideração. Por isto, os conceitos relacionados à cinética abordados nesta dissertação, têm ênfase no crescimento de bactérias.

A cinética do cultivo microbiano pode ser considerada em cinco níveis diferentes, como mostra a Figura 2.7.

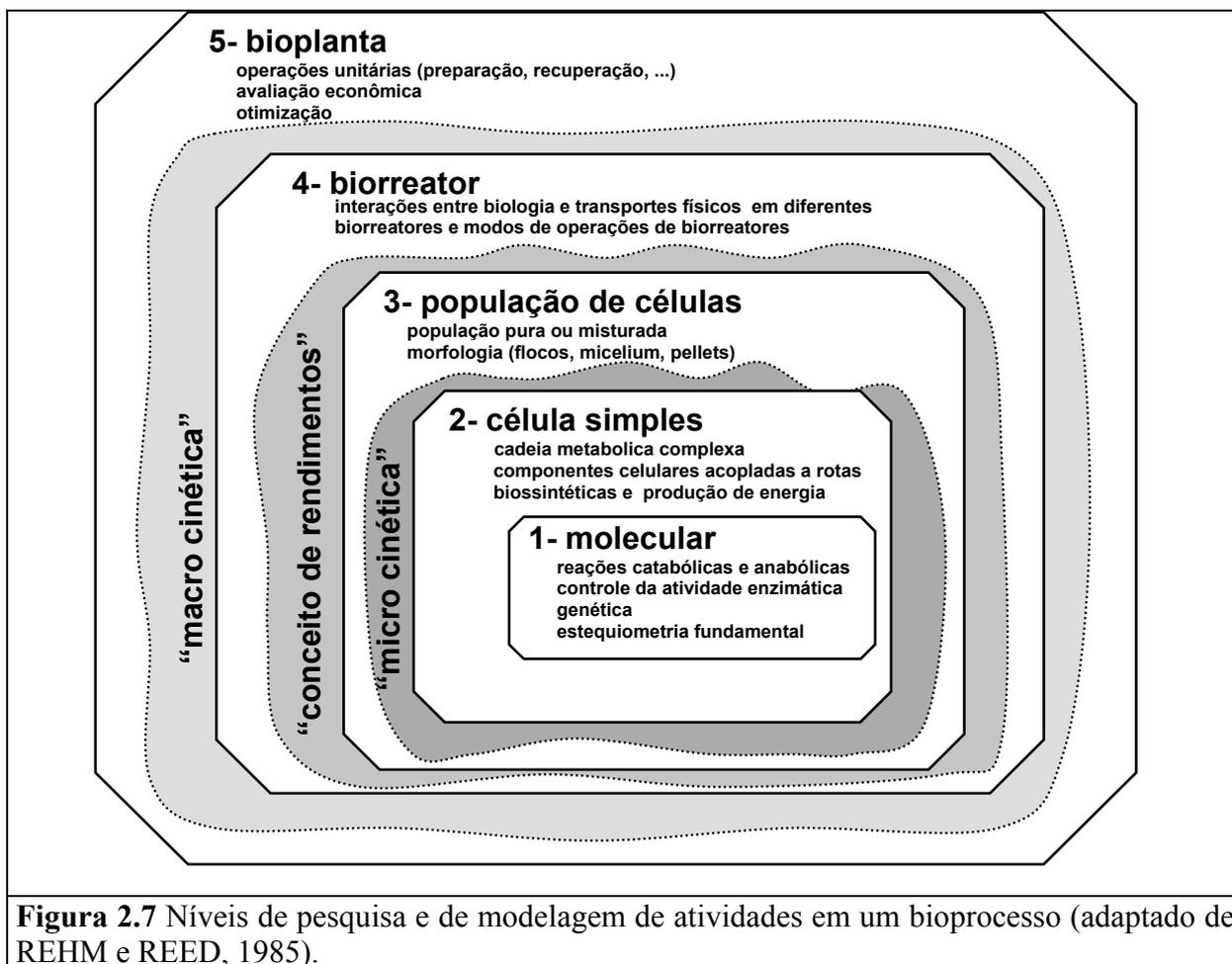


Figura 2.7 Níveis de pesquisa e de modelagem de atividades em um bioprocessamento (adaptado de REHM e REED, 1985).

Cada nível tem uma característica única que conduz para um tratamento cinético específico. Nos níveis um (molecular) e dois (célula simples), onde é definida a visão micro cinética do bioprocessamento, é observado que um microorganismo possui um sistema complexo para conseguir energia para sua manutenção, crescimento e reprodução. Os vários eventos que ocorrem durante o crescimento de uma simples célula, desde seu início até o começo de sua divisão, são referidos como ciclo celular. Durante o ciclo celular, diversos trabalhos são executados pelo maquinário celular, como a síntese de proteínas, a produção de membrana, a duplicação do DNA, a movimentação dos cílios ou flagelos ou ainda o transporte ativo de substâncias através da membrana. Para produzir energia e realizar trabalho, microorganismos consomem nutrientes e geram produtos, através de um número muito grande de reações químicas que ocorrem no seu interior. Este conjunto de todas as reações é referido como

metabolismo. Estas reações são responsáveis pelos processos de síntese e degradação dos nutrientes e constituem a base da vida. O metabolismo está organizado em rotas, que são seqüências de reações enzimáticas em que o produto de uma reação é utilizado como reagente na reação seguinte. Diferentes enzimas catalisam diferentes passos de rotas metabólicas, agindo de forma concertada de modo a não interromper o fluxo nessas rotas. As enzimas são vitais para o metabolismo porque permitem a realização de reações desejáveis, mas termodinamicamente desfavoráveis, ao acoplá-las a reações mais favoráveis. As enzimas regulam as rotas metabólicas em resposta a mudanças no ambiente celular ou a sinais de outras células.

O conteúdo bioquímico da célula está constantemente mudando no ciclo celular. Em consequência deste metabolismo, o microrganismo é capaz de converter moléculas simples em produtos químicos de interesse comercial. Microrganismos aeróbios, como *Bacillus megaterium*, geram a energia necessária através de rotas catabólicas nas quais elétrons são transferidos durante o processo de respiração aeróbica (transformação de O_2 e das fontes de C em CO_2 e água).

O processo de cultivo microbiano, esquematizado na Figura 2.8b, envolve dois sistemas que interagem continuamente: a fase biológica (fase biótica), composta pela população microbiana (x) e o meio de cultura (fase abiótica), composto pelos substratos (s_i) e produtos (p_j) formados no processo.

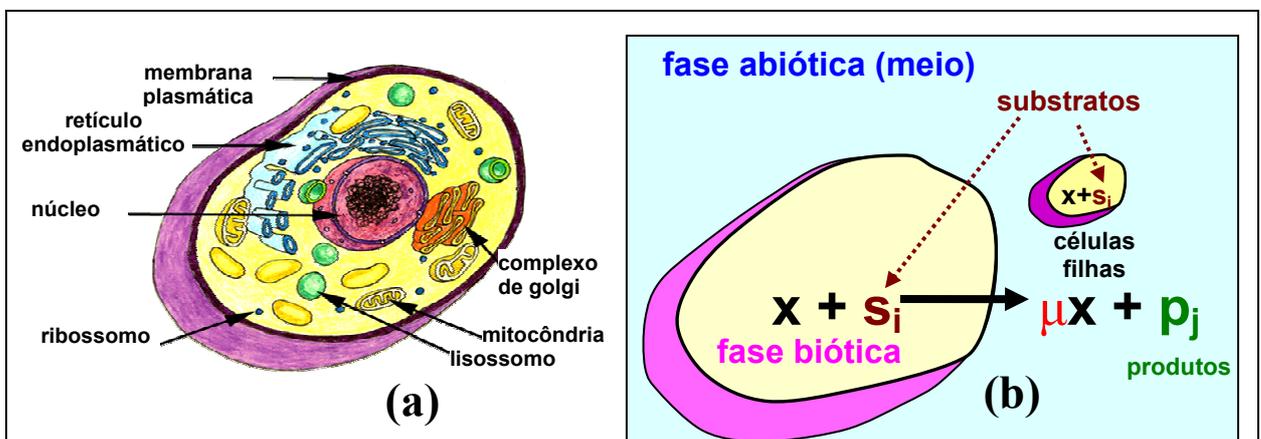


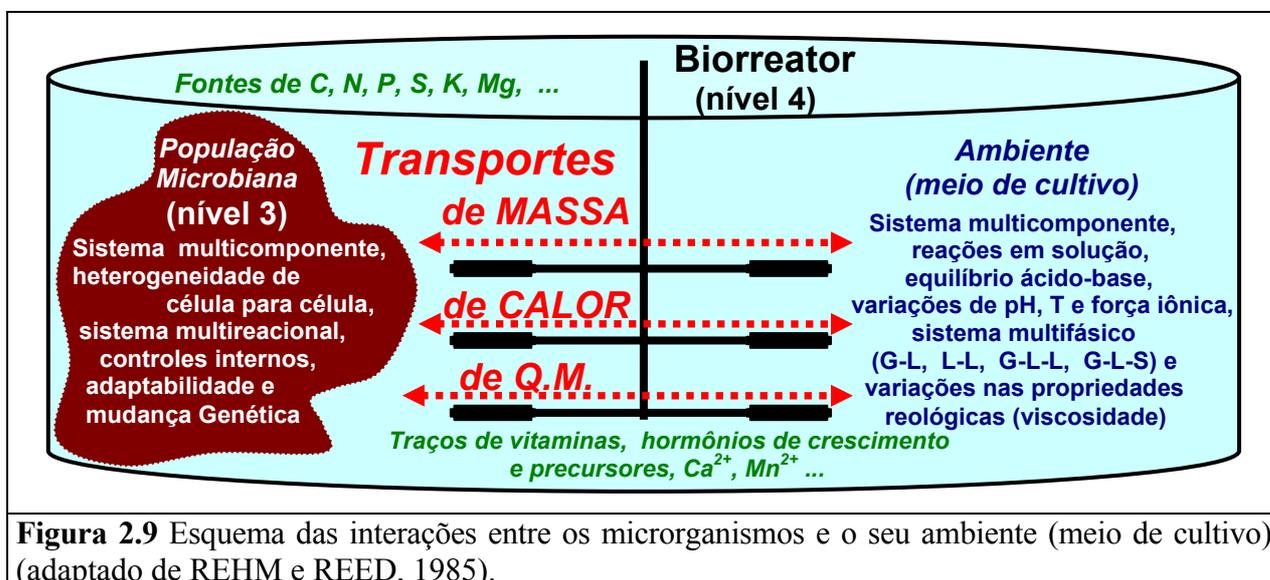
Figura 2.8 (a) Esquema de uma célula e (b) Esquema do crescimento celular, onde as letras representam: (x) células, (s_i) substratos, (μ) velocidade específica de crescimento celular e (p_j) produtos (adaptado de NIELSEN E VILLADSEN, 1994).

O crescimento celular envolve o transporte do substrato para dentro da célula, seguido pela conversão do substrato intracelular em biomassa e produtos metabólicos. Finalmente, o produto metabólico é excretado para o meio extracelular. A formação de células filhas pela reprodução é considerada como produção de espécies químicas adicionais. O crescimento microbiano é um exemplo de reação autocatalítica. A velocidade de crescimento

está diretamente relacionada com a concentração celular e a reprodução celular é uma consequência normal dessa reação.

O metabolismo de uma população microbiana (nível três) pode se adaptar a um grande número de condições ambientais (físicas, químicas e nutricionais), e conseqüentemente mostrar vários comportamentos, com o uso de diversas rotas metabólicas. Os nutrientes suprem necessidades específicas de cada microrganismo. Em todas as culturas microbianas é preciso que o meio contenha água, fontes de energia, carbono, nitrogênio, sais minerais, vitaminas e oxigênio (no caso de cultivo aeróbio). Esses requisitos são essenciais para que as células sintetizem todos os componentes necessários à manutenção e divisão celular. Alguns desses substratos serão utilizados ativamente para a formação do produto final, como as fontes de carbono e nitrogênio, enquanto outros, sais minerais e vitaminas, permitem que reações importantes ocorram.

Para controlar uma conversão biológica não é suficiente quantificar apenas a micro cinética implícita (nível um e dois), mas também o fenômeno físico de transporte, como mostra a Figura 2.9, que ocorre entre uma população de células (nível três) com o seu ambiente, em um biorreator (nível quatro), onde está definido o princípio macroscópico da cinética do bioprocessos.



Devido à complexidade do fenômeno biológico e físico, simplificações devem ser feitas, mas com a menor perda de informação possível, usando analogias formais para acompanhar o bioprocessos com base em variáveis macroscópicas. Esta aproximação é uma consequência do princípio macroscópico.

A aproximação macroscópica usando a análise de sistemas como uma analogia da engenharia das reações químicas tem sido aplicada e aceita crescentemente para problemas de

engenharia de bioprocessos (KAFAROW *et al.*, 1979, ROELS, 1980, MOSER, 1981 e VOTRUBA, 1982). Sob a óptica macroscópica, o cultivo microbiano pode ser descrito observando o comportamento das variáveis de estado do processo, como por exemplo, mudanças dos componentes: celular ou de biomassa (x), substratos (s_i), oxigênio (o_2), produtos (p_j), dióxido de carbono (co_2), em função do tempo deste cultivo, no biorreator. Segundo REHM e REED, 1985, o quarto nível, o do biorreator, é o mais essencial de acordo com o princípio macroscópico (“a aproximação macro formal do bioprocesso”), mas todos os outros níveis contribuem para a modelagem cinética do fenômeno (cultivo microbiano).

O primeiro item descreve o biorreator e suas formas de operação. A cinética de crescimento microbiano, segundo item, definirá as variáveis do cultivo, assim como a forma de sua medida. O terceiro item fala sobre a modelagem matemática do cultivo microbiano, dando ênfase aos modelos que são de interesse desta dissertação. No quarto e último item relacionado ao cultivo microbiano, um balanço será feito para o componente celular (x), nas duas formas de operação, em batelada e batelada alimentada, seguido de uma análise comparativa entre as formas de operar o biorreator e suas conseqüências para o crescimento celular.

2.2.1 Biorreatores

Biorreator é o equipamento central do processo de cultivo microbiano industrial (também denominado processo fermentativo, no jargão da área), onde ocorrem as reações de transformação dos substratos nos produtos de interesse através da ação de microrganismos. Um bioprocessos industrial se divide em três etapas básicas: upstream, etapa de preparo do meio e do material e de esterilização; etapa de cultivo (nos biorreatores), onde ocorre a transformação do substrato em produto; e downstream, etapa da separação de células, isolamento, purificação e concentração do produto. Um biorreator deve propiciar condições adequadas de mistura, proporcionando uniformidade na distribuição de nutrientes e aeração e baixo cisalhamento, condições adequadas de pH e temperatura e projeto adequado para facilitar operações das etapas de upstream e downstream.

Um cultivo microbiano pode ser considerado um sistema de três fases: a fase líquida, que contém sais, substratos e metabólitos dissolvidos. A fase sólida que consiste das células, dos substratos ou produtos insolúveis, e fase gasosa que inclui oxigênio e CO_2 . Assim uma agitação adequada é essencial em um cultivo, pois proporciona os seguintes efeitos nas três fases: dispersão do ar no meio de cultivo; homogeneização para igualar temperatura, pH, e concentração de nutrientes; suspensão dos microrganismos e dos nutrientes sólidos e

dispersão de líquidos imiscíveis. Pela dificuldade em obter informações cinéticas úteis em biorreator com condições espaciais não uniformes, o estudo cinético é geralmente feito em biorreator descrito como ideal que é considerado como de “mistura perfeita”. Entende-se como “mistura perfeita” um caldo de fermentação homogêneo, sem zonas mortas no interior do biorreator. O biorreator mais comumente utilizado é o de cultivo submerso, em que o microrganismo está disperso em meio líquido (outro tipo importante de reator é o semi-sólido, que não será tratado nesta revisão). O responsável pela “mistura perfeita” em um biorreator de cultivo submerso, é o agitador ou impelidor, cuja velocidade de rotação (geralmente em rpm) é uma importante variável de estado desse sistema.

No desenvolvimento de um cultivo microbiano em geral, uma das etapas principais é a escolha do tipo do biorreator a ser utilizado no processo, bem como do modo com que este deva ser operado. Os tipos de biorreatores mais utilizados em processos de cultivo submerso (onde o microrganismo está disperso em meio de cultura líquido) são: o convencional, tipo tanque agitado e aerado, e os não convencionais, pneumáticos tipo torre ou coluna de bolhas e “air-lift. A análise de biorreatores trata em parte das formas de operação ou dos procedimentos de entrada e saída de materiais. Um biorreator pode ser operado nas formas descontínua (batelada), contínua (com ou sem reciclo de células) e semi-contínua (batelada alimentada). Neste trabalho foi abordado com mais detalhes o biorreator convencional e os modos de operações descontínuo (batelada) e o semi-contínuo (batelada alimentada).

2.2.1.1 Operação em batelada ou descontínua

A Figura 2.10 mostra um esquema de um biorreator, operando em batelada, e de algumas variáveis importantes para a cinética do processo.

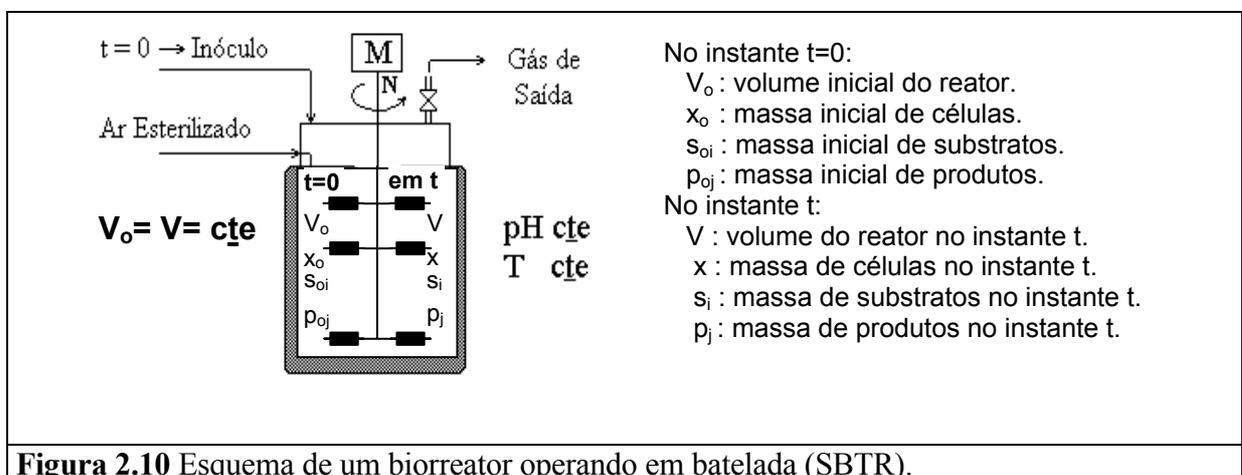


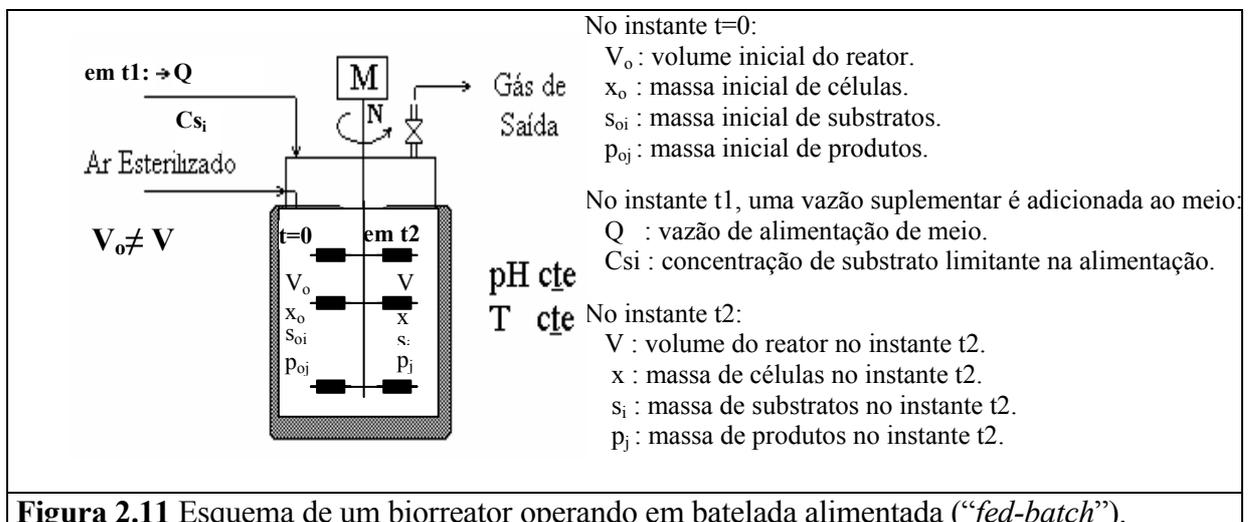
Figura 2.10 Esquema de um biorreator operando em batelada (SBTR).

Operação em batelada (SBTR – “stirred batch tank reactor”) é amplamente usada em cultivos comerciais. Em um processo em batelada convencional, o inóculo e todos os nutrientes, exceto o oxigênio molecular, são adicionados ao meio antes do início do processo, em $t=0$, e os produtos são removidos somente no fim do cultivo. Durante o processo não há entrada ou saída de material do biorreator (exceto pela aeração), sendo o término do cultivo, na grande maioria, associado ao esgotamento de um substrato limitante (s_i). Como não há variação de volume ($V_o=V$) na operação em batelada, a densidade (ρ) dentro do biorreator é constante. Esta forma de operação, do biorreator, é interessante para os estudos cinéticos.

Nesta dissertação, a letra minúscula de um componente representa a sua massa e quando precedido pela letra C, maiúscula, representa a sua concentração.

2.2.1.2 Operação em batelada alimentada ou semicontínuo

Batelada alimentada (“*fed-batch*”) é uma forma de operação muito utilizada pela indústria, onde um ou mais nutrientes são supridos, durante o cultivo, enquanto os produtos permanecem no interior do reator até o final do processo. O modo de operação, como mostra Figura 2.11, é o seguinte: o reator é iniciado ($t=0$) como uma batelada, e uma grande concentração de biomassa é obtida pelo consumo inicial de substrato. A velocidade de formação do produto é tipicamente baixa durante o cultivo em batelada. Em certo tempo (t_1), quando o nível de substrato está baixo, uma vazão de alimentação, usualmente muito concentrada, de substratos é iniciada. No mesmo tempo um indutor pode ser adicionado para ativar a rota metabólica que estimule a formação do produto desejado. Durante todo o período da batelada alimentada nenhum produto é retirado do reator e o volume do meio vai aumentando.



2.2.2 Cinética de crescimento microbiano

Quando um meio de cultura é inoculado (entende-se por inoculação a introdução de uma linhagem pura no meio reacional) os microrganismos consomem seletivamente os nutrientes dissolvidos e os convertem, em parte, em biomassa, aumentando a concentração celular. Em um cultivo em batelada (SBTR), sendo o volume constante ($V_0=V$), utiliza-se a molaridade para medir as concentrações dos componentes. Neste cultivo é possível, a partir de amostras retiradas ao longo do tempo (pontos experimentais), visualizar as cinéticas de crescimento celular (C_x), de consumo de substratos limitantes (C_s) e de formação de produtos (C_p), como mostra a Figura 2.12.

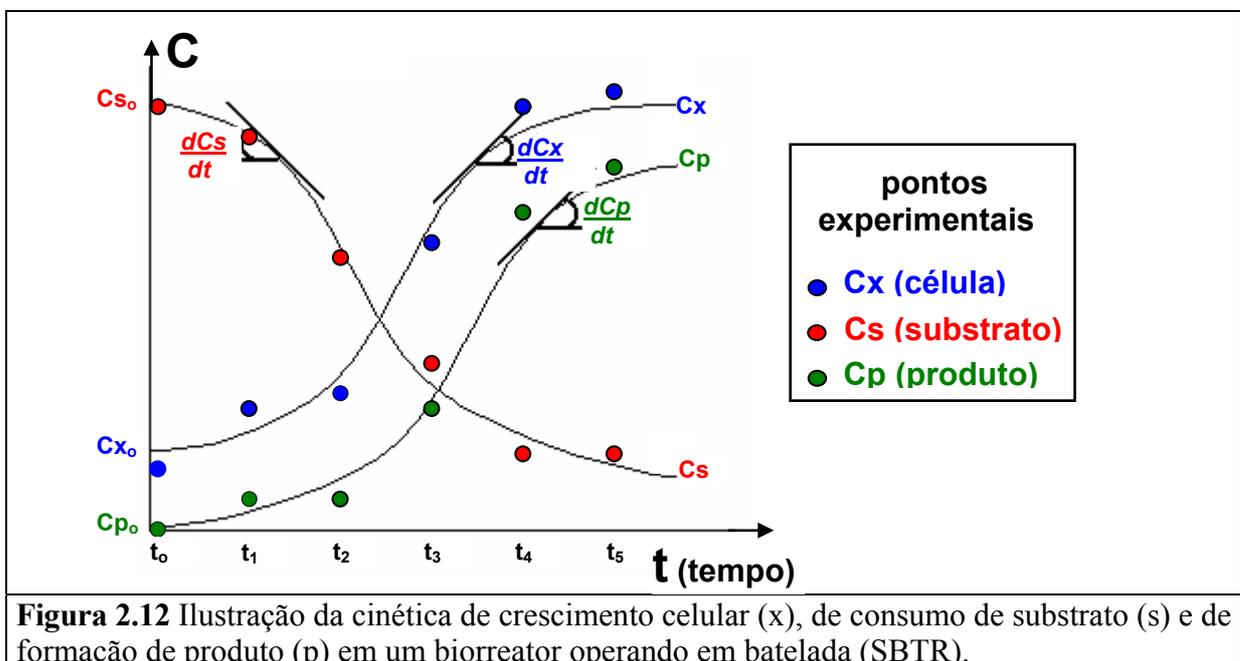


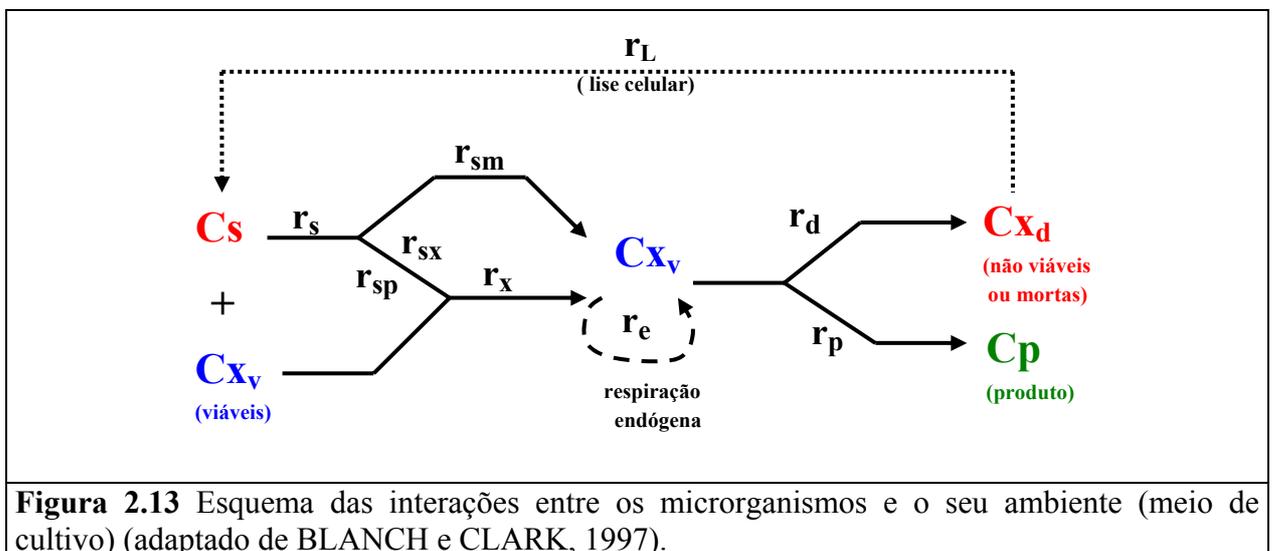
Figura 2.12 Ilustração da cinética de crescimento celular (x), de consumo de substrato (s) e de formação de produto (p) em um biorreator operando em batelada (SBTR).

A tentativa de escrever uma equação de balanço para cada componente do cultivo, onde um conjunto de reações em série e em paralelo acontece no meio de crescimento, é impraticável. Desta forma, os balanços são escritos em termos do substrato limitante, pois ele é o primeiro, dentre os substratos, a ser consumido quando o microorganismo está em crescimento. Além disso, para muitos organismos, a fonte de energia e carbono para crescimento e formação de produto pode ser encontrada no mesmo composto orgânico (BLANCH e CLARK, 1997). Ao ser consumido ao longo do tempo, o substrato limita a velocidade de crescimento. Do ponto de vista cinético, o fenômeno é equacionado através de um conjunto, de equações diferenciais, correspondentes aos balanços materiais. A partir de balanços materiais, é possível calcular as concentrações de células (C_x), de substratos (C_{s_i}), e de produtos (C_{p_j}), e em alguns casos preverem variações de compostos intermediários que não são quantificados durante o processo. Com o conhecimento das condições iniciais (C_{x_0} ,

$C_{s_{oi}}$ e $C_{p_{oj}}$) e de parâmetros estimados, o sistema é resolvido numericamente. A estimação de parâmetros ocorre minimizando-se os desvios entre os valores calculados por integração numérica das equações de balanço e os pontos experimentais.

A partir da estequiometria da atividade celular, associada às possíveis reações globais envolvidas, é possível modelar a cinética do bioprocessamento e prever variações nas concentrações de seus componentes ao longo de cultivos.

Embora o crescimento celular e a formação de produto sejam processos complexos, a estequiometria de conversão de substratos em produtos e materiais celulares é representada por simples equações “pseudo-estequiométricas”, como mostradas Figura 2.13:



r_s : velocidade global de consumo de substrato. onde $r_s = \left(\frac{dC_s}{dt} \right)_{\text{global}}$.

r_{sm} : velocidade de consumo de substrato para produção de energia de manutenção.

$$r_{sm} = m_s \cdot C_{xv}, \text{ onde } m_s \text{ é o coeficiente de manutenção.}$$

r_{sx} : velocidade de consumo de substrato para a produção de biomassa ou para o crescimento celular.

r_{sp} : velocidade de consumo de substrato para a formação de produto.

r_x : velocidade global de crescimento celular .

r_e : velocidade de consumo de material celular pela respiração endógena.

$$r_e = k_e \cdot C_{xv}, \text{ onde } k_e \text{ é o coeficiente de respiração endógena.}$$

r_d : velocidade de morte celular.

r_p : velocidade global de formação de produto.

Energia de manutenção, relacionada a r_{sm} , representa a energia gasta para a troca osmótica, mobilidade, transferência de nutrientes e produtos para dentro e para fora da célula.

A respiração endógena, relacionada a Γ_e , pode acontecer no início do cultivo ou após o término do substrato, é o consumo de reservas celulares para a manutenção da célula.

As velocidades na Figura 2.13 variam ao longo do tempo no cultivo microbiano, como indicado no próximo item, fases do crescimento microbiano.

2.2.2.1 Fases do crescimento microbiano

Em um cultivo em batelada, as seguintes fases do crescimento celular ocorrem no tempo, Figura 2.14.

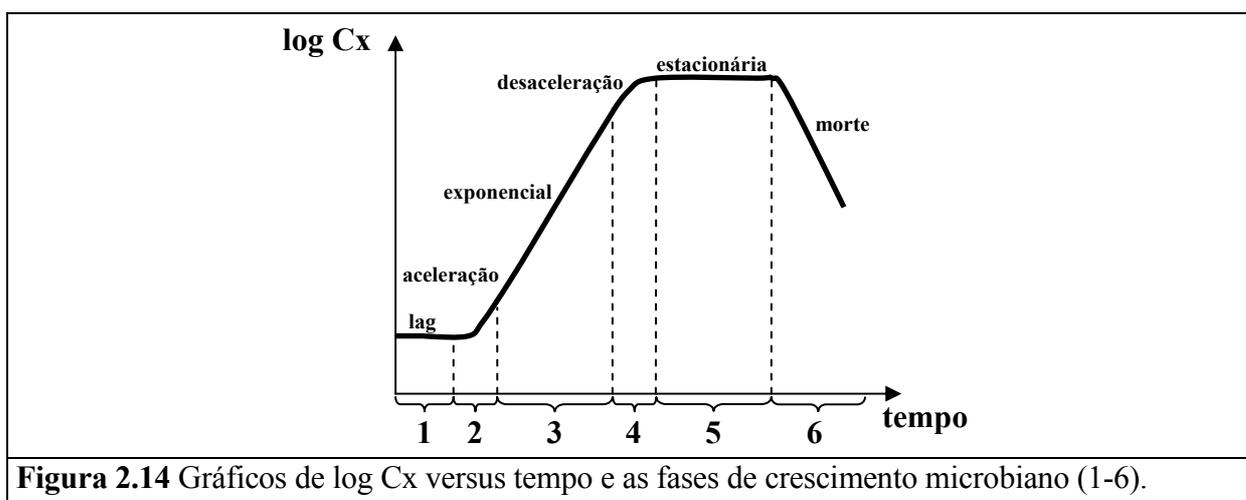


Figura 2.14 Gráficos de $\log C_x$ versus tempo e as fases de crescimento microbiano (1-6).

A primeira fase, conhecida como fase lag (1), ocorre logo após a inoculação. Os microrganismos reorganizam seus constituintes quando são transferidos para um novo meio. Dependendo da composição do meio, novas enzimas são sintetizadas, as sínteses de algumas enzimas são reprimidas e o “maquinário” celular se adapta às novas condições ambientais. Durante a fase lag usualmente pode-se considerar a concentração de células constante:

$$C_x = C_{x0} = \text{constante e } \frac{dC_x}{dt} = 0.$$

As causas e o tempo de duração da fase de adaptação podem resultar de diversos fatores. Se o inóculo cresce em um meio contendo uma fonte de carbono diferente da anterior, novas enzimas podem ter que ser induzidas para o catabolismo do novo substrato, contribuindo assim para uma maior duração da fase lag. Baixas concentrações de alguns nutrientes e de fatores de crescimento também contribuem para a ocorrência desta fase. A idade do inóculo tem um forte efeito na duração da fase lag, que tende a aumentar com o aumento dessa idade. Em alguns casos há uma idade ótima, que resulta numa fase lag mínima. Múltiplas fases lag podem ser observadas quando as células são inoculadas em um meio com várias fontes de carbono. Este fenômeno é conhecido como efeito diáuxico, onde as células preferencialmente utilizam a fonte

de carbono mais facilmente assimilável para então adaptar o seu metabolismo para utilizar a segunda fonte. Isto ocorre devido à repressão catabólica das enzimas necessárias para o metabolismo do segundo substrato. No fim da fase lag, a população de microorganismos está bem adaptada ao novo ambiente, e depois deste período, as células começam a se multiplicar, com o aumento de massa celular, determinando o início da fase de aceleração (2) do crescimento.

A fase exponencial (3) é um período de “crescimento balanceado”, onde a composição média de cada célula permanece aproximadamente constante. Com abundância de concentrações de substratos, a velocidade de crescimento celular independe dessas concentrações. Assim, nesta fase, a massa celular aumenta exponencialmente com o tempo sem nenhum indício de inibição. A velocidade de crescimento exponencial é de 1ª ordem em relação a C_X :

$$r_x = \left(\frac{dC_X}{dt} \right)_{\log} = \mu \cdot C_X \quad (2.1)$$

, onde r_x é a velocidade global de crescimento celular, e μ é a velocidade específica de crescimento, e nesta fase, é constante e máxima ($\mu = \mu_{\max}$). Na equação 2.1, separando os termos comuns, célula e tempo e integrando, em $t=0$, onde $C_X = C_{X_0}$ e $t > t_{\text{lag}}$, tem-se:

$$\ln \frac{C_X}{C_{X_0}} = \mu_{\max} \cdot (t - t_{\text{lag}}), \text{ e assim, } C_X = C_{X_0} \cdot e^{\mu_{\max} \cdot (t - t_{\text{lag}})} \quad (2.2)$$

A fase exponencial também é caracterizada pelo tempo de geração τ_d , que é o tempo necessário para a massa celular dobrar ($C_X = 2 \cdot C_{X_0}$):

$$\tau_d = \frac{\ln 2}{\mu} \quad (2.3)$$

Seguindo a fase de crescimento exponencial tem-se a fase de desaceleração (4) ou declínio, onde há a diminuição de um ou mais nutrientes essenciais (fonte de carbono, nitrogênio ou oxigênio), ou o acúmulo de algum produto inibidor ou tóxico.

Na fase estacionária (5), a concentração celular atinge o valor máximo e constante, e a velocidade de crescimento é zero (sem divisão celular) ou a velocidade de crescimento é igual à velocidade de morte. Mesmo com velocidade de crescimento igual a zero durante esta fase, as células ainda estão metabolicamente ativas e metabólitos secundários (não associados ao crescimento) são produzidos.

Finalmente a fase de morte (6), não há mais crescimento e ocorre a lise celular com uma redução significativa das células viáveis. Células morrem também durante a fase

estacionária e uma clara demarcação entre essas duas fases nem sempre é possível. A velocidade de morte (r_d) pode seguir uma cinética de 1ª ordem em relação a C_x e k_d é a constante de morte.

$$r_d = \left(\frac{dC_x}{dt} \right)_{\text{morte}} = -k_d \cdot C_x \quad (2.4)$$

2.2.2.2 Coeficientes de Rendimento

Equações de balanço são escritas em termos do substrato limitante. Para vários cálculos de processo, deseja-se comparar o substrato potencial (disponível) em termos do rendimento celular, do rendimento em produto e de geração de calor. Para melhor descrever a cinética de crescimento, definem-se alguns parâmetros relacionados com a estequiometria. Em um determinado tempo t de cultivo, os correspondentes valores de C_x , C_s e C_p podem ser relacionados pelos fatores de conversão (ou pseudo-estequiométricos) globais, expressos matematicamente. Por exemplo, o rendimento celular em um cultivo microbiano é dado por:

$$Y_{X/S} = -\frac{\Delta C_x}{\Delta C_s} = \frac{C_x - C_{x_0}}{C_{s_0} - C_s} \quad (2.5)$$

Onde $Y_{X/S}$ é o coeficiente global de conversão de substrato a células, também conhecido como coeficiente aparente ou observado. Seu valor varia ao longo de um cultivo

O valor de $Y_{X/S}$ pontual é definido por:

$$Y_{X/S} = -\frac{dC_x}{dC_s} \quad (2.6)$$

Bacillus megaterium, um organismo quimiotrófico, que possui uma fonte química de energia (quebra de ligações), utiliza um substrato simples, uma solução de aminoácidos, por exemplo, como fonte para consumir carbono e produzir energia, representado matematicamente pela Equação 2.7.

$$\Delta C_{s_{\text{total}}} = \Delta C_{s_1} + \Delta C_{s_2} + \Delta C_{s_3} + \Delta C_{s_4} \quad (2.7)$$

Onde o substrato é utilizado para: (1) assimilação como biomassa, (2) energia de crescimento, (3) energia de manutenção e (4) assimilação como produto. Enquanto o verdadeiro coeficiente de rendimento celular $Y_{\text{verdadeiro}} = -\frac{\Delta C_x}{\Delta C_s}$ é praticamente constante, ou seja, uma quantidade estequiométrica bem definida, o coeficiente global ($Y_{X/S}$) apresenta variações na alocação do substrato, ao utilizar-se dos quatro componentes indicados na Equação 2.7. Uma população crescendo rapidamente utiliza o substrato para assimilação

como biomassa (1) e para energia de crescimento (2), ou seja, $Y_{X/S} \cong Y_{\text{verdadeiro}}$, enquanto que uma população em repouso ou em fase estacionária freqüentemente consome substrato para manutenção (3) sem nenhum crescimento e $Y_{X/S} = 0$.

Da mesma forma, $Y_{X/P}$ é o coeficiente global de conversão de células a produto, Equação 2.8, e $Y_{P/S}$ é o coeficiente global de conversão de substrato à produto, Equação 2.9.

$$Y_{X/P} = \frac{C_X - C_{X_0}}{C_P - C_{P_0}} \quad (2.8)$$

$$Y_{P/S} = \frac{P - P_0}{S_0 - S} \quad (2.9)$$

As relações abaixo mostram outros coeficientes de rendimentos importantes, resultantes da bioconversão, no cultivo microbiano. Esses coeficientes globais são respectivamente: de consumo de O_2 , Equação 2.10, de geração de ATP no catabolismo, Equação 2.11, e de liberação de calor, Equação 2.12.

$$Y_{X/O_2} = \frac{\text{quantidade de células produzidas}}{\text{quantidade de } O_2 \text{ consumido no processo aeróbio}} \quad (2.10)$$

$$Y_{X/ATP} = \frac{\text{quantidade de células produzidas}}{\text{quantidade de ATP gerados no catabolismo}} \quad (2.11)$$

$$Y_{X/H} = \frac{\text{quantidade de células produzidas}}{\text{quantidade de calor liberado}} \quad (2.12)$$

2.2.2.3 Estequiometria para Balanços Elementais

Para escrever balanço de espécies atômicas para crescimento celular foi estabelecida em convenção uma composição elementar para uma célula, tendo como base um grama-átomo de carbono e contendo somente os principais elementos. A seguinte fórmula-célula foi definida: $CH_aO_bN_c$.

Quando *Bacillus megaterium*, um microorganismo aeróbico, cresce tendo aminoácidos como fonte única de carbono, nitrogênio e de energia, uma forma geral da reação de crescimento celular e formação de produtos pode ser escrita:



onde: $CH_aO_bN_c$ é a composição elementar da célula

$CH_lO_mN_n$ é a composição elementar da fonte de carbono (aminoácidos)

$CH_pO_qN_r$ é a composição elementar do produto

A Equação global 2.13 é a equação líquida resultante de centenas de reações metabólicas na vida celular. Os vários ciclos e cadeias do metabolismo, incluindo o sistema ATP e outros sistemas de manipulação de energia, que não resultam em novas células ou produtos, não contribuem para a equação líquida. Por isso conhecimento detalhado destes ciclos não é necessário em um tratamento usando a equação líquida estequiométrica, macroscópica. Assim, a complexidade da análise do bioprocessamento é reduzida consideravelmente. Todas as células microbianas suspensas na fase líquida do biorreator são tratadas como “caixas pretas”, e apenas seu comportamento macroscópico, manifestado nas mudanças das concentrações, é levado em consideração.

Também pela Equação 2.13, fica claro que a produção de CO₂ reflete o consumo de substratos para crescimento celular. Portanto, o teor de CO₂ nos gases efluentes é um indicador importante da evolução do cultivo no biorreator.

2.2.2.4 Coeficiente respiratório

Outro parâmetro importante em cultivos aeróbios é o coeficiente respiratório (RQ, da sigla em inglês), definido como a razão da velocidade de formação de gás carbônico pela velocidade de consumo de oxigênio. O coeficiente respiratório é dado pela Equação 2.14:

$$RQ = \frac{CO_2 \text{ produzido}}{O_2 \text{ consumido}} \quad (2.14)$$

A relação entre os coeficientes estequiométricos da produção de gás carbônico e do consumo de oxigênio permite que o coeficiente respiratório, determinado experimentalmente, seja relacionado com a estequiometria do crescimento microbiano; a partir da Equação 2.13 tem-se:

$$RQ = \frac{\kappa}{\gamma} \quad (2.15)$$

Como se pode concluir com base nos itens anteriores, a evolução do teor de CO₂ nos gases efluentes do biorreator está intimamente ligada à dinâmica de crescimento celular. Seu emprego para inferir crescimento celular (Cx), como será visto nesta dissertação, é uma escolha lógica, assim como o quociente respiratório (RQ), é uma boa base fenomenológica, para usar uma lógica que mensure a evolução do cultivo.

2.2.3 Modelos matemáticos para um cultivo microbiano

Os modelos matemáticos desempenham um papel importante na síntese e projeto de um sistema de controle. Eles fornecem uma relação funcional entre as variáveis do processo. Podem ser usados para a simulação do processo, com ou sem sistema de controle, para um planejamento experimental com objetivo de obter máxima informação com um mínimo de tempo e gastos, para aperfeiçoamento da estrutura da planta e de parâmetros de operação, para cálculos de perfis de tempo ótimos para adição de substratos e outros nutrientes, para estimativa de variáveis ou parâmetros que não podem ser medidos (métodos indiretos de medida). O modelo de um processo é peça fundamental no cálculo das regras de controle de um controlador adaptativo baseado em modelos.

Ao se propor um modelo para o cultivo, pretende-se descrever o processo de crescimento do microorganismo, a síntese de determinado produto e, se possível, prever os efeitos observados ao se variar composição do meio e condições operacionais do cultivo. O modelo estabelecerá as relações de causa e efeito entre as variáveis. A escolha do modelo cinético para o processo depende do nível de complexidade com que se deseja caracterizar a população microbiana do sistema e deve atender a pelo menos um dos seguintes requisitos: discriminar possíveis mecanismos de controle do metabolismo celular, auxiliar no projeto do biorreator e na otimização das condições de operação ou permitir um posterior desenvolvimento de técnicas de controle do processo (ARAÚJO, 1996).

Os modelos matemáticos de bioprocessos são muitas vezes formados por um conjunto de equações diferenciais não lineares de primeira ordem com parâmetros desconhecidos. Os parâmetros precisam ser estimados utilizando os resultados experimentais obtidos. Um ajuste satisfatório da resposta do modelo aos dados experimentais é desejável.

O desenvolvimento de modelos para representar um determinado processo passa pelas seguintes etapas: postulação do modelo, experimentos para a coleta dos dados empíricos, “ajuste” do modelo (estimativa dos parâmetros do modelo) e finalmente a avaliação dos resultados (HIMMELBLAU, 1970).

A preocupação atual na investigação na área de modelagem (BASTIN e MORRIS, 1992) está focada na análise estrutural, nos problemas de identificação de sistemas, no planejamento de experimentos e na aplicação de métodos alternativos (sistemas híbridos).

2.2.3.1 Classificação de Modelos

ROELS e KOSSEN (1978) apresentam uma perspectiva sobre os vários tipos de abordagens para modelagem dos processos microbianos em biotecnologia. Os vários tipos de modelos são normalmente distinguidos em termos de classes.

Dois grandes grupos de modelos matemáticos de cultivo microbiano são os modelos fenomenológicos e empíricos. O modelo fenomenológico utiliza equações de conservação (de massa, energia, quantidade de movimento) e tem sua base na formulação de hipóteses simplificadoras e na utilização de correlações empíricas para predição de propriedades e parâmetros necessários para quantificar as variáveis significativas do processo. Já o modelo empírico estabelece relações empíricas para correlacionar o efeito de variações nas variáveis de entrada com os valores das variáveis de saída ou medidas do processo, sem preocupação em explicar o comportamento do sistema às leis físicas e físico-químicas. Neste tipo de modelo a entrada é correlacionada com a saída e o fenômeno é tratado como uma “caixa preta” (em oposição à “caixa branca” de um modelo fenomenológico *stricto sensu*) (NELLES, 2001).

O modelo fenomenológico pode ser dividido quanto a sua representação sobre a célula e sobre a população das células. Nos modelos não estruturados as células são vistas como caixas pretas no que se refere aos processos fisiológico e metabólico. O material celular da população microbiana é considerado como tendo a mesma composição e qualidade. O crescimento é expresso unicamente como sendo o aumento de biomassa. Este tipo de modelo é usado, normalmente, na descrição de fenômenos de crescimento, em que num dado intervalo de tempo, todas as propriedades extensivas aumentem do mesmo fator - crescimento balanceado. Se o crescimento ocorre em um ambiente não balanceado, então o modelo a usar deve descrever a alteração na estrutura do microrganismo em resposta a alterações nas condições ambientais - modelo estruturado. Este tipo de modelo permite descrever o estado das células e sua adaptação às mudanças do meio ambiente.

Quanto à heterogeneidade da população microbiana existe o modelo não segregado, onde a população celular é considerada homogênea, e o modelo segregado, onde as células são consideradas discretas, como indivíduos de uma população heterogênea, com distribuição de idade, tamanho e de propriedades celulares.

Um modelo verbal faz uma descrição qualitativa do processo biológico e é normalmente usado na comunicação entre cientistas com especialidades diferentes (multidisciplinares).

Em um modelo descritivo utiliza-se um ajuste numérico de dados experimentais para fins interpolativos. Deve ser usado somente dentro da região onde o modelo foi testado experimentalmente. Este tipo de modelo é muitas vezes também referido como modelo caixa preta.

Um modelo preditivo (ou explicativo) tem como finalidade a extrapolação de dados. Usualmente combina equações fenomenológicas (do modelo caixa branca) com correlações empíricas. Por isso, esse tipo de modelo também é designado por modelo caixa cinza.

Um modelo determinístico tem um significado de ausência de incerteza, possibilitando reconstituir o comportamento passado e prever o comportamento futuro de uma propriedade, conhecidos que sejam os parâmetros do modelo e o estado atual. Em um modelo estocástico considera-se a incerteza, atribuindo a um acontecimento uma dada probabilidade de ocorrência (FERREIRA, 1995).

A modelagem de um processo com gradientes espaciais de concentração (ISAACS *et al.*, 1987; DOCHAIN, 1994), ou outra propriedade, é efetuada com recurso aos chamados modelos de parâmetros distribuídos. Na situação de inexistência de variações, espaciais usam-se os modelos de parâmetros concentrados.

Desenvolvimentos recentes na área de modelagem de processos biotecnológicos apontam caminhos em direção à utilização de modelos baseados em sistemas inteligentes. Estes modelos incorporam equações matemáticas para as relações quantitativas e relações qualitativas (regras) baseadas em conhecimento (MAVROVOUNIOTIS *et al.*, 1989; ZHANG *et al.*, 1994).

2.2.3.2 Modelos para velocidade específica de crescimento

Vários modelos para representar a velocidade específica de crescimento (μ) microbiano foram construídos. O modelo de MONOD, de 1942, é talvez a formulação mais conhecida e utilizada pela comunidade científica das áreas de biotecnologia e microbiologia. Observando o crescimento de *E. coli* em várias concentrações de glicose, MONOD desenvolveu um dos mais simples modelos que inclui o efeito da concentração de nutrientes, com base nos modelos de cinética enzimática. Nesse modelo é assumido que somente um substrato (s), o que limita o crescimento, é importante na determinação da velocidade de crescimento celular. Ele despreza o efeito das diferenças entre as células e das alterações na composição celular. Não é adequado para a fase lag e nem para a fase de morte celular. É um modelo com uma abordagem determinística, distribuída e não estruturada.

A Equação 2.16, modelo de MONOD, tem a forma de uma equação hiperbólica de saturação semelhante a outras utilizadas em cinética enzimática e em fenômenos de adsorção (FERREIRA, 1995).

$$\mu = \mu_{\text{máx}} \frac{C_s}{K_s + C_s} \quad (2.16)$$

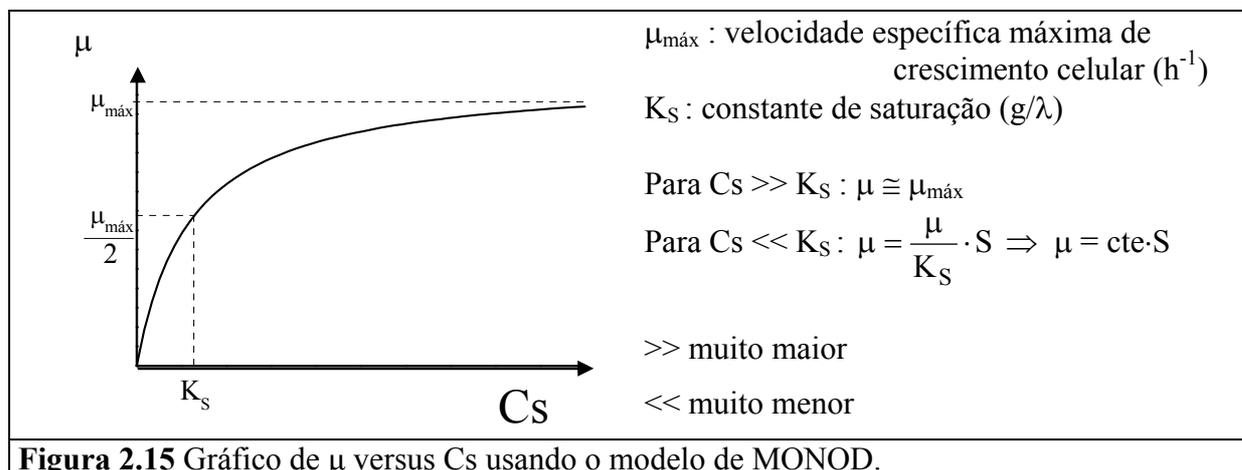


Figura 2.15 Gráfico de μ versus C_s usando o modelo de MONOD.

Neste modelo, $\mu_{\text{máx}}$ é a velocidade máxima específica de crescimento microbiano, C_s é a concentração do substrato limitante e K_s é a constante de saturação ou de MONOD. K_s representa o valor da concentração do nutriente limitante para a qual a velocidade específica de crescimento é a metade de seu valor máximo, como mostra a Figura 2.15. Um baixo K_s significa que o microorganismo tem muita afinidade pelo substrato limitante, crescendo rápido e com uma maior duração da fase exponencial.

O modelo de MONOD é uma grande simplificação do complexo sistema bioquímico das células, mas ainda assim descreve bem o comportamento da fase exponencial do crescimento de um grande número de microrganismo.

A relação entre o crescimento e a concentração de substrato é muito complexa e, portanto, outros modelos foram construídos, com base em observações, para representar a resposta das células ao seu ambiente. Entre eles estão os seguintes modelos: de Blackman, Equação 2.17; de Tessier, Equação 2.18; de Moser, Equação 2.19 e de Contois, Equação 2.20 (BAILEY e OLLIS, 1986).

Modelo de Blackman (2.17)

$$\mu = \frac{\mu_{\text{máx}}}{2 K_s} \cdot S, \text{ se } S < 2 K_s$$

$$\mu = \mu_{\text{máx}}, \text{ se } S \geq 2 K_s$$

$$\begin{aligned} &\text{Modelo de Tessier} && (2.18) \\ &\mu = \mu_{\text{máx}} (1 - e^{-S/K_S}) \end{aligned}$$

$$\text{Modelo de Moser} \quad (2.19)$$

$$\mu = \mu_{\text{máx}} \frac{S^n}{K_S + S^n} = \mu_{\text{máx}} (1 + K_S \cdot S^{-n})^{-1}$$

$$\text{Modelo de Contois} \quad (2.20)$$

$$\mu = \mu_{\text{máx}} \frac{S}{K_{SX} \cdot X + S}$$

2.2.3.3 Modelos para múltiplos substratos (s_i)

Existem situações onde múltiplos substratos podem limitar simultaneamente a velocidade de crescimento. Isto pode acontecer quando diferentes metabólitos entram em uma rota metabólica e cada metabólito pode ser limitado por uma diferente enzima.

Quando o crescimento celular continua mesmo quando um dos substratos limitantes é exaurido a cinética aditiva é escolhida, como mostra a Equação 2.21.

$$\mu = \sum_{i=1}^n \left(\frac{\mu_{\text{máx}i} \cdot Cs_i}{Ks_i + Cs_i} \right) \quad (2.21)$$

Quando o crescimento celular pára quando um dos substratos limitantes é exaurido, a cinética é multiplicativa, como mostra a Equação 2.22 (Blanch e Clark, 1997).

$$\mu = \mu_{\text{máx}i} \cdot \prod_{i=1}^{n^{\circ}S} \left(\frac{Cs_i}{Ks_i + Cs_i} \right) \quad (2.22)$$

2.3.3.4 Modelos de formação de produto (p_j)

Os modelos cinéticos para velocidade específica de formação de produto (μ_p) dependem do momento da formação do produto (p) em relação à fase de crescimento celular (x).

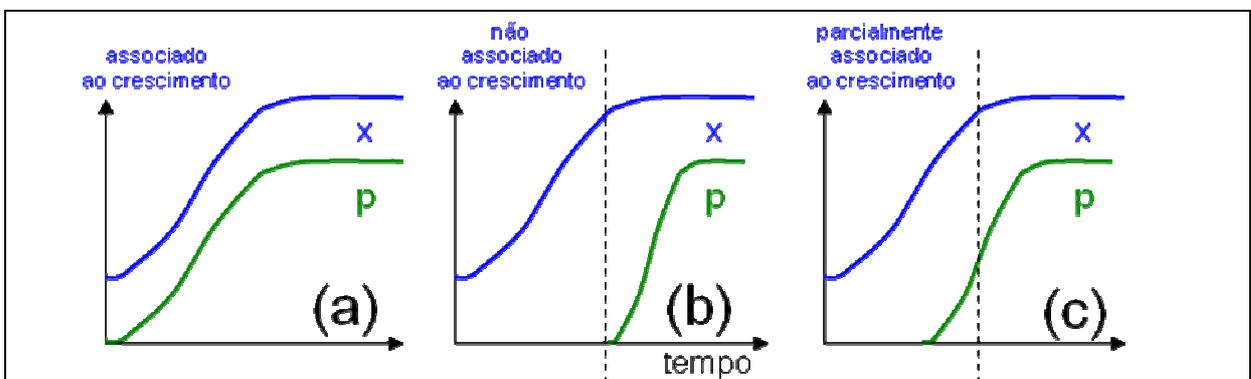


Figura 2.16 Gráficos dos diferentes momentos de formação do produto em relação ao crescimento microbiano: (a) se formam juntos; (b) produto é formado após a fase exponencial e (c) produto é formado no meio da fase exponencial e continua na fase estacionária.

Produtos que são formados no mesmo momento da fase exponencial, mostrado no gráfico da Figura 2.16a, ou seja, são associados ao crescimento da massa celular, são referidos como metabólitos primários, e o modelo que representa a velocidade específica de formação de produto (μ_p) é dado pela Equação 2.23.

$$\mu_p = \frac{1}{C_X} \cdot \frac{dC_p}{dt} = \alpha \cdot \mu \quad (2.23)$$

$$\mu_p = \frac{Y_{p/s}}{Y_{x/s}} \cdot \mu = Y_{p/x} \cdot \mu \quad (2.24)$$

São exemplos deste modelo: a produção de uma enzima constitutiva, produção de etanol, aminoácidos, vitaminas. Esses produtos estão, normalmente, associados ao catabolismo celular.

Produtos que são formados após a fase exponencial, mostrado no gráfico da Figura 2.16b, surgem na fase de desaceleração ou na fase estacionária, e são descritos como metabólitos secundários. A produção ocorre na fase estacionária de crescimento celular, quando $\frac{dC_X}{dt} = 0$. O modelo mais simples, representado pela Equação 2.25, supõe que a velocidade específica de formação de produto (μ_p) é constante.

$$\mu_p = \beta \rightarrow \frac{1}{C_X} \frac{dC_p}{dt} = \beta \rightarrow \frac{dC_p}{dt} = r_p = \beta \cdot C_X \quad (2.25)$$

A velocidade de formação de produto (r_p) é proporcional à concentração celular (C_X). Os antibióticos e as vitaminas são exemplos deste modelo (BLANCH e CLARK, 1997).

Uma classe intermediária pode também ser identificada, onde a formação de produtos fica entre as duas classes. A formação de produto ocorre quando o crescimento é menor e na fase estacionária. São os produtos parcialmente associados ao crescimento, como mostra no gráfico da Figura 2.16c. Equação 2.26, de Luedeking-Piret (1959), que descreve bem a produção de ácido láctico em cultivo em batelada a pH constante, expressa este modelo.

$$\mu_p = \alpha \cdot \mu + \beta \quad (2.26)$$

A produção de ácido láctico e de aminoácidos são exemplos deste modelo.

2.2.3.5 Modelos com inibição

A velocidade específica de crescimento (μ) pode ser inibida pelo substrato e/ou pelo produto. Um inibidor pode afetar o crescimento de várias maneiras: modificando o potencial químico de substratos, intermediários ou produtos; alterando a permeabilidade da

célula ou da atividade enzimática; afetando a síntese de enzimas, dissociando enzimas agregadas e influenciando a atividade funcional da célula (BLANCH e CLARK, 1997).

Muitos substratos presentes em altas concentrações podem inibir o crescimento microbiano por um ou mais mecanismos. Qualquer mudança nas propriedades físico-químicas do meio por alta concentração do substrato pode ter algum efeito no crescimento. Estas propriedades incluem a pressão osmótica, força iônica, constante dielétrica e atividades de solutos presentes. Qualquer uma destas mudanças pode resultar em alteração da fluidez da membrana e afetar o fenômeno de transporte de substrato do meio para dentro da célula. O modelo de Andrews, também conhecido como modelo de MONOD com inibição, é comparado com o de MONOD no gráfico da Figura 2.17. O modelo de Andrews é representado pela equação 2.27.

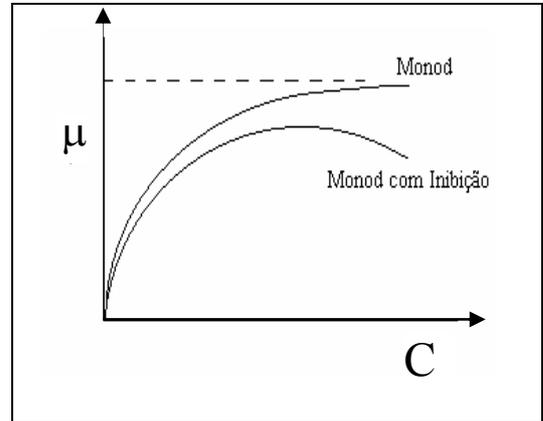


Figura 2.17 Gráfico de μ versus C_s usando o modelo de Andrews (MONOD com inibição).

$$\mu = \mu_{\text{máx}} \frac{C_s}{K_s + C_s + \frac{C_s^2}{K_{IS}}} \quad (2.27)$$

Outro modelo para inibição pelo substrato é o modelo de Aiba, representado pela Equação 2.28.

$$\mu = \mu_{\text{máx}} \frac{C_s}{K_s + C_s} \cdot e^{-C_s/K_{IS}} \quad (2.28)$$

Produtos podem também causar inibição. Produção alta de um metabólito por uma rota pode resultar em inibição em outra rota do metabolismo. Os modelos de Aiba-Shoda-Nagatani, Equação 2.29, e de Levenspiel, Equação 2.30 são exemplos de modelos de inibição pelo produto.

$$\mu = \mu_{\text{máx}} \frac{C_s}{K_s + C_s} \cdot \frac{K_p}{K_p + C_p} \quad (2.29)$$

$$\mu = \mu_{\text{máx}} \frac{C_s}{K_s + C_s} \cdot \left(1 - \frac{C_p}{C_p^*}\right)^n \quad (2.30)$$

Onde C_p^* é igual ao valor de C_p a partir do qual $\frac{dC_x}{dt} = 0$.

Na fermentação alcoólica, por exemplo, quando $80 < C_p^* < 100 \text{ g}/\lambda \rightarrow \frac{dC_x}{dt} = 0$,

mesmo com substrato (C_s) ainda disponível.

2.2.4 Batelada versus batelada alimentada

Neste ponto da dissertação têm-se os elementos necessários para fazer um balanço de massa, para as células, no biorreator, com operação tanto em batelada (SBTR) como batelada alimentada (“fed-batch”), para comparação dos seus resultados.

Na operação batelada (SBTR), o inóculo (massa de células - x) é adicionado ao meio de volume (V) no biorreator, no início do processo e retirado apenas no fim da operação. Assim, como $\text{entra}=\text{sai}=0$, a equação que representa o balanço ($\text{entra} - \text{sai} + \text{gerado} - \text{consumido} - \text{acúmulo} = 0$), fica reduzida para:

$$\begin{array}{l} \text{acúmulo} = \text{gerado} - \text{consumido} \\ \boxed{\frac{d(x \cdot V)}{dt}} = \boxed{r_x \cdot V} - \boxed{(r_d \cdot V + r_e \cdot V)} \\ \text{acúmulo} = \text{crescimento} - (\text{morte} + \text{respiração endógena}) \end{array}$$

$$\left. \begin{array}{l} r_d = k_d \cdot x \\ r_e = k_e \cdot x \end{array} \right\} \text{assume-se } r_d \text{ e } r_e \text{ desprezíveis (igual a zero) na fase exponencial.}$$

O resultado da derivada do lado esquerdo da equação é igual a:

$$x \frac{dV}{dt} + V \frac{dx}{dt} = r_x \cdot V. \text{ Em operação batelada o volume é constante } (V=V_0), \text{ podendo}$$

assim usar a concentração (C_x) de células no lugar de massa (x), e balanço fica:

$$\frac{dC_x}{dt} = r_x. \text{ Como foi visto na Equação 2.1, } r_x = \mu \cdot C_x, \text{ logo:}$$

$$\frac{dC_x}{dt} = \mu \cdot C_x \tag{2.31}$$

Na operação batelada alimentada, ocorre o mesmo a massa de células (x): $\text{entra}=\text{sai}=0$ e foi desprezado a morte (r_d) e a respiração endógena (r_e) na fase exponencial.

Assim o balanço de massa é o seguinte:

$$\begin{array}{l} \text{acúmulo} = \text{gerado} \\ \boxed{\frac{dx}{dt}} = \boxed{r_x} \\ \text{acúmulo} \qquad \qquad \text{crescimento} \end{array}$$

$r_x = \mu \cdot x$ como visto na Equação 2.1, e o balanço fica:

$$\frac{dx}{dt} = \mu \cdot x \quad (2.32)$$

sabe-se que concentração de células é igual a massa sobre o volume ($C_x = \frac{x}{V}$).

Isolando x ($x = C_x \cdot V$) e substituindo o nos dois lados da Equação 2.32, têm-se:

$$\frac{d(C_x \cdot V)}{dt} = \mu \cdot (C_x \cdot V)$$

O resultado da derivada do lado esquerdo da equação é igual a:

$$\frac{V \cdot dC_x}{dt} + \frac{C_x \cdot dV}{dt} = \mu \cdot C_x \cdot V$$

$$\frac{dC_x}{dt} = \mu \cdot C_x - \frac{C_x}{V} \cdot \frac{dV}{dt} \quad (2.33)$$

, mas $\frac{dV}{dt}$ é a vazão de alimentação (Q) no biorreator,

e a diluição (D) é igual a vazão dividido pelo volume ($D = \frac{Q}{V}$).

Substituindo D na Equação 2.33 e colocando C_x em evidência:

$$\frac{dC_x}{dt} = (\mu - D) \cdot C_x \quad (2.34)$$

Existe uma similaridade entre os balanços de massa de uma batelada alimentada, equação 2.34, e de uma batelada, equação 2.31. Na equação 2.34, a velocidade de diluição D pode ser uma função do tempo e o modelo de batelada alimentada em princípio é puramente transiente. Se no processo em batelada a concentração de substrato diminui com o tempo, na batelada alimentada ela é uma função do tempo não necessariamente decrescente. Cultivos em batelada alimentada são usualmente utilizados para superar problemas relacionados principalmente com a inibição pelo substrato e por repressão catabólica. Processos em batelada alimentada têm sido utilizados na produção de uma ampla variedade de produtos com a perspectiva de se aumentar produtividades; entre eles estão os antibióticos, vitaminas, aminoácidos, enzimas e outros.

Uma batelada alimentada pode ser vista como um problema de controle. Com certas restrições é possível escolher uma função que controle a vazão de alimentação (Q) de tal forma que um objetivo possa ser alcançado. Este objetivo pode ser definido no fim do processo de fermentação, quando o volume total alcançou um determinado valor. Este é um problema clássico na teoria de controle e tem como primeira tarefa a escolha do tempo de início da vazão de alimentação (Q), correspondente a determinadas concentrações de

biomassa (C_x) e substratos (C_{s_i}) no biorreator. As concentrações que caracterizam o estado de início da vazão são descritas como C_{x_0} e C_{s_0} respectivamente. A segunda tarefa, com certeza a mais árdua, é determinar os valores da concentração de substratos em função do tempo que otimizem a formação de produtos (C_{p_j}) – entendendo otimizar como maximizar C_{p_j} minimizando o tempo necessário para atingir esse valor máximo (ou, em outras palavras, maximizar a produtividade do biorreator). Uma dificuldade encontra-se no fato de que o microrganismo tem seus próprios mecanismos regulatórios internos. O controle do processo é feito manipulando o ambiente extracelular, esperando-se que este afete de maneira satisfatória os mecanismos intracelulares. O próximo passo da operação otimizada da batelada alimentada é a escolha do momento de finalizar a vazão de alimentação. A terceira tarefa é o momento de finalizar o cultivo microbiano.

2.3 Inteligência computacional para controle de bioprocessos

2.3.1 Introdução

Controle pressupõe utilizar os dados adquiridos na etapa de monitoração. O objetivo do controle é manter algumas variáveis de estado do processo (ou uma função dessas variáveis) próximas de valores de referência, compensando perturbações. Um monitoramento avançado compreende a aquisição e análise de informação, dados do processo, obtida de várias formas, com vistas a gerar novo conhecimento, e com isto melhorar a capacidade de acompanhamento de variáveis críticas que afetam o processo biotecnológico. Várias metodologias de controle avançado têm sido extensivamente aplicadas em sistemas de controle, incluindo processos mecânicos, aeroespaciais, químicos e elétricos. Entre as aplicações de metodologias de controle avançado, podem-se citar os algoritmos de controle preditivo baseado em modelo (OLIVEIRA e CARRAPICO, 1998; BORDENEUVE-GUIBE e VAUCORET, 2001; PONCIA e BITTANTI, 2001), controle de estrutura variável (JAFAROV, 2000; SAAJ & BANDYOPADHYAY, 2002), controle robusto (SANA & RAO, 2001; TAN *et al.*, 2002), controle adaptativo (HADDAD *et al.*, 2002; YU e ORLOV, 2002) e controle baseado em metodologias da inteligência computacional (KING, 1999; JAMSHIDI *et al.*, 2002).

Monitorar e controlar um cultivo de células, em um biorreator, não é tarefa simples. Muitos problemas são encontrados, nas aplicações da engenharia de controle a bioprocessos. Isto se deve à complexidade destes sistemas, que exibem as seguintes características: não-linearidade; incertezas no modelo e/ou seu fraco desempenho em

generalizações; medidas apresentando incertezas e inadequação; e necessidade de tratamento de aspectos de segurança do processo, diagnóstico de falhas e restrições ambientais.

Outra dificuldade encontra-se na existência de mecanismos regulatórios internos do microorganismo. Do ponto de vista da engenharia de controle, a limitação reside no fato de que somente é possível manipular o ambiente extracelular, esperando que este afete de maneira satisfatória os mecanismos intracelulares.

Nesta dissertação, um projeto de controlador será proposto, para a supervisão e controle do processo de cultivo de *Bacillus megaterium* para a produção da enzima PGA. Este controlador possui uma arquitetura baseada em sistemas híbridos das metodologias da inteligência computacional. A literatura tem apresentado muitos estudos referentes ao projeto de controladores usando metodologias da inteligência computacional, a citar as aplicações de redes neurais (GE *et al.*, 2001; YU *et al.*, 2002), algoritmos evolutivos (CHIPPERFIELD e FLEMING, 1996; ZHANG *et al.*, 2002), sistemas *fuzzy* (TANAKA & IKEDA, 1998; GOLEA e GOLEA, 2002) e sistemas híbridos inteligentes (CRAVEN *et al.*, 1999; KAWANAKA *et al.*, 2000).

Pelo princípio macroscópico, o cultivo microbiano pode ser descrito observando o comportamento das variáveis de estado do processo. A concentração de células viáveis, ou de biomassa (Cx), é uma variável de estado muito importante, da qual dependem velocidades de crescimento, consumo de substrato e formação de produto. Muitas estratégias de controle visam a maximizar a produção de biomassa. Este é o caso do processo analisado nesta dissertação, pois Silva (2003) constatou que a produção de PGA, principal objetivo do processo, é associada ao crescimento, com atraso de duas horas.

O monitoramento avançado do bioprocessamento se baseia em dados medidos em tempo real, que permitam inferir as variáveis de estado críticas. Dados obtidos através de sensores em biorreatores geralmente correspondem a grandezas físicas e químicas, como temperatura, pressão, pH, vazão de ar e de meio suplementar, concentração de oxigênio dissolvido, composição dos gases efluentes. Variáveis de relevância biológica, como a biomassa (Cx), são mais difíceis e complexas de se medir. Seus sensores, além de outras exigências, devem resistir a sucessivos procedimentos de esterilização, e devem ser estáveis durante períodos extensos. Estas dificuldades têm levado a indústria a ficar reticente na introdução de novas estratégias de monitoração e controle, além da preocupação em relação à robustez e viabilidade dos sensores biológicos. A consequência do não uso destas estratégias avançadas é a utilização de modelos parciais, empíricos, e de análises a posteriori, feitas em laboratório. Estas análises são caras e envolvem atrasos consideráveis, inviabilizando a operação de

controle e supervisão *on-line*. Um aspecto avançado do monitoramento é a incorporação das medidas disponíveis em algoritmos de inferência de estado (observadores ou *softsensors*) que forneçam informação não diretamente mensurável (FERREIRA, 1995). Estes algoritmos são também denominados por filtros, para os quais a estimação pode ser vista como a extração de um sinal puro a partir de sinais com ruído.

2.3.2 *Softsensors* ou sensores virtuais

Softsensor ou sensor virtual é outro nome usual para o *software* que, processando várias medidas em tempo real, utiliza-as (e a suas interações) para calcular valores de variáveis que não podem ser mensuradas diretamente.

Softsensors são usados especialmente em fusão de dados, onde medidas de características e dinâmicas diferentes são combinadas. *Softsensors* ajudam a supervisão humana dos processos e o desenvolvimento de controladores automáticos que podem eventualmente aliviar a necessidade de supervisão humana no processo. O desenvolvimento na área de controle avançado de bioprocessos tem demonstrado a aplicabilidade da inferência das variáveis de estado do bioprocessos (concentração de células, substrato e produtos) através de variáveis secundárias que podem facilmente ser monitoradas *on-line*.

A aplicação de *softsensors* tem como motivações: inexistência, custo proibitivo ou inadequação de sensores para a medição em linha de algumas das principais variáveis de estado (CARLEYSMITH, 1986; SONNLEITNER, 1991); aumentar a competitividade do processo biotecnológico, aplicando metodologias operacionais capazes de assegurar de modo estável e reprodutível uma produtividade máxima com um mínimo de custos; e as restrições impostas por autoridades regulatórias, tais como o FDA (*Food and Drug Administration*, EUA) e a ANVISA, no Brasil, que para permitir a comercialização dos produtos bioquímicos, produzidos por estes processos, exigem prova de que as operações seguiram procedimentos que garantem a consistência química dos produtos.

O desenvolvimento de um *softsensor* para o processo de produção de PGA é uma tarefa muito difícil devido ao caráter dinâmico, não linear e com parâmetros variando com o tempo. Uma tradicional técnica no campo de sistemas de controle é procurar por um sistema linear equivalente, ou seja, ajustar parâmetros de controle de acordo com um sistema linearizado ao redor de seu estado de equilíbrio, através de uma função transferência. Esta solução simplifica o problema, mas a perda de informação pode ser crítica no processo de produção de PGA. A grande complexidade e incerteza dos processos biológicos requerem uma sofisticada lógica que não ajusta facilmente na estrutura matemática dos modelos de

controle tradicionais (SHIOYA, SHIMIZU, e YOSHIDA, 1999). Como já foi dito, modelos matemáticos baseados em enfoques clássicos da engenharia de processos resultam muitas vezes em problemas intratáveis em virtude do número excessivo de parâmetros, ou no extremo oposto, são simplificações da realidade que não conseguem captar aspectos essenciais do caso em estudo. Para solucionar estes problemas, o *softsensor*, pode utilizar-se de técnicas adaptativas. Estas técnicas incorporam normalmente mecanismos de reajuste de parâmetros *on-line*, que caracterizam o estimador.

A aplicação de técnicas de inteligência computacional (IC) constitui-se em outra alternativa aos esquemas convencionais de controle. Essas técnicas procuram incorporar características inspiradas na natureza e na inteligência humana a sistemas de controle automático.

É esperado que um sistema de controle dito “inteligente” seja capaz de replicar (mesmo que rudimentarmente) algumas habilidades humanas, entre as quais aprendizado, tomada de decisão, percepção, adaptação, inferência. Metodologias da inteligência computacional são bem compatíveis para modelar e controlar processos biológicos, pois permitem trabalhar com conhecimento fragmentado, qualitativo, incerto, aproximado, inexato e até mesmo contraditório; com estruturas de dados complexas, típicas de tais processos. As vantagens oferecidas por estas metodologias são: algoritmos estruturados de maneira simples para a tomada de decisões; capacidade de aprendizagem, permitindo a sistematização do aprendizado *on-line* e *off-line*, levando conseqüentemente à melhoria gradual da capacidade preditiva dos modelos; e maior grau de autonomia.

2.3.3 Inteligência computacional (IC)

A idéia da construção de máquinas inteligentes é fascinante por duas razões: imitar o comportamento inteligente e produzir soluções funcionais para tarefas complexas. Inteligência computacional (IC) é um dos produtos desta busca. O termo inteligência computacional foi proposto por ZADEH (1994) e BEZDEK (1994), que atribuem esta denominação a uma coleção de metodologias que visam a tolerância a falhas, imprecisão e incerteza, proporcionando robustez e soluções de baixo custo. Ela não é ainda considerada um ramo sólido da ciência, por não ter base de compreensão totalmente estabelecida (DUCH, 2007) e ocasionalmente é classificada como sinônimo de Inteligência Artificial (IA), mas seus limites ainda não foram estabelecidos, estão em formação. No presente, IC é um imenso guarda-chuva para três metodologias, redes neurais artificiais (RNAs), lógica *fuzzy* (ou nebulosa) e computação evolucionária. RNAs se inspiram na fisiologia do sistema nervoso,

lógica *fuzzy* em modelos lingüísticos e computação evolucionária em diversos paradigmas relacionados ao princípio darwiniano da evolução das espécies. O uso simultâneo de duas ou mais destas metodologias resulta nos chamados sistemas híbridos inteligentes (SHI).

Faz-se a seguir a descrição das metodologias computacionais que foram utilizadas nesta dissertação.

2.3.4 Lógica *fuzzy* (nebulosa)

Seres humanos são capazes de lidar com processos bastante complexos fundamentado em informações imprecisas ou aproximados. A estratégia adotada pelos operadores humanos é também de natureza imprecisa e geralmente possível de ser expressa em termos lingüísticos. A teoria dos conjuntos clássicos, que nesta dissertação será referenciado pelo termo do inglês *crisp*, não é capaz de tratar o aspecto vago da informação e a teoria de probabilidades, na qual a probabilidade de um evento determina completamente a probabilidade do evento contrário, é mais adaptada para tratar de informação frequentistas do que aquelas fornecidas por seres humanos. Os conceitos de conjuntos *fuzzy* e lógica *fuzzy* (ZADEH, 1965), foram implementados para representar e processar informações lingüísticas, e fornecem mecanismos para tratar a imprecisão e a incerteza inerentes ao conhecimento do mundo real. Por isso têm sido aplicados com sucesso em diversas áreas, como: controle automático, classificação e reconhecimento de padrões, tomada de decisão, sistemas inteligentes, previsão de séries temporais e robótica (PEDRYCZ, 1996; PEDRYCZ & GOMIDE, 1999; DUMITRESCU *et al.*, 2000).

2.3.4.1 Conjuntos *fuzzy*

A noção básica de conjunto como uma coleção de objetos está presente em todos os aspectos da nossa vida diária. Utilizamos a noção de conjunto sempre que precisamos organizar, resumir e generalizar o conhecimento sobre objetos.

Nos conjuntos *crisp*, existe uma dicotomia: um objeto pertence ou não pertence ao conjunto. Um objeto pertencente a um conjunto é dito um elemento do conjunto. Escrevemos $x \in A$ para indicar que o objeto x pertence ao conjunto A e escrevemos $x \notin A$ para indicar que o objeto x não pertence ao conjunto A . Uma das maneiras de se definir um conjunto é por meio de uma *função característica*, que associa o valor 1 (um) ao elemento se ele pertence ao conjunto ou 0 (zero) se ele não pertence ao conjunto. Por exemplo, um conjunto A sobre um universo de discurso X pode ser definido por uma função $A: X \rightarrow \{0,1\}$ como:

$$A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases}$$

Assim, existe uma distinção clara e bem determinada entre os objetos que estão e os que não estão no conjunto.

Entretanto, o mundo real apresenta situações em que os conjuntos clássicos não são flexíveis o suficiente para uma descrição apurada de significado, tendo em vista que eles forçam de forma abrupta a transição de pertinência absoluta para a não pertinência absoluta.

Os conjuntos *fuzzy*, introduzidos por ZADEH (1965), são conjuntos cujos limites não são precisos, permitindo a pertinência parcial de cada elemento ao conjunto. Sendo A um conjunto *fuzzy*, um objeto x pode pertencer ao conjunto A com um grau, que é chamado de *grau de pertinência* de x em A e usualmente expresso por um número no intervalo $[0,1]$. Os valores extremos desse intervalo, 0 (zero) e 1 (um), representam, a não pertinência total ao conjunto e a pertinência total ao conjunto respectivamente. Assim, a teoria de conjuntos *crisp* pode ser considerada um caso restrito da teoria dos conjuntos *fuzzy*, em que somente os dois graus extremos de pertinência são permitidos. Uma das principais utilidades da transição gradual entre pertinência e não pertinência é permitir a representação de conceitos imprecisos expressos em linguagem natural, que não têm uma distinção rigorosa entre membros e não membros. Como exemplo, podemos citar os conceitos de pessoa alta, imóvel caro, temperatura baixa, tráfego intenso e outros.

Os conjuntos *fuzzy* são caracterizados formalmente por uma *função de pertinência*, que é uma generalização da função característica dos conjuntos clássicos. A função de pertinência associa, a cada elemento do universo de discurso, um valor dentro de um intervalo específico, que indica o grau de pertinência do elemento ao conjunto. O intervalo de valores mais utilizado é o intervalo $[0,1]$. Indicamos que a função de pertinência de um conjunto A mapeia elementos do universo \mathbf{X} em graus de pertinência no intervalo $[0,1]$ escrevendo (Klir & Yuan, 1995)

$$A: \mathbf{X} \rightarrow [0,1].$$

Um conjunto *fuzzy* A sobre o universo \mathbf{X} , pode ser representado por um conjunto de pares ordenados tais que o segundo elemento do par é um objeto $x \in \mathbf{X}$, e o primeiro elemento do par é seu respectivo grau de pertinência no conjunto *fuzzy* A , ou seja:

$$A = \{A(x)/x \mid x \in \mathbf{X}\}$$

2.3.4.2 Tipos de funções de pertinência

Em geral, conjuntos *fuzzy* são definidos para representar conceitos imprecisos, o que depende do entendimento que se tem desse conceito e também do contexto em que será utilizado. Apesar de teoricamente qualquer função da forma $A: X \rightarrow [0,1]$ descrever uma função de pertinência associada a um conjunto *fuzzy* A , a representação de conceitos como conjuntos *fuzzy* no contexto de uma aplicação pode exigir que essas funções possuam certas formas específicas ou até mesmo certas propriedades. A rigor, a forma mais adequada de uma função de pertinência só pode ser definida no contexto de uma aplicação, mas algumas formas de funções de pertinência tornaram-se mais conhecidas e são utilizadas freqüentemente, pela sua simplicidade. Entre as formas de funções mais conhecidas estão as famílias de funções parametrizadas, sendo que algumas delas são apresentadas a seguir:

Função triangular: parâmetros (a, m, b) , onde $a \leq m \leq b$.

$$A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } x \in (a, m) \\ 1 & \text{se } x = m \\ \frac{b-x}{b-m} & \text{se } x \in (m, b) \\ 0 & \text{se } x \geq b \end{cases}$$

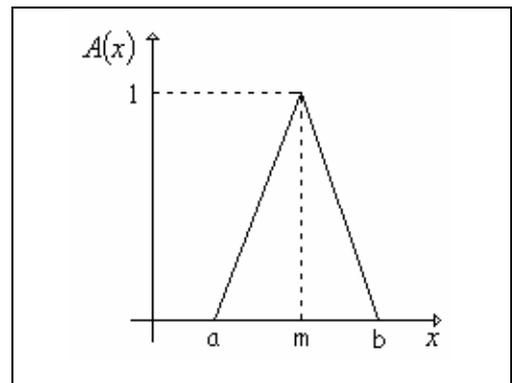


Figura 2.18 Exemplos de função de pertinência triangular.

Função trapezoidal: parâmetros (a, m, n, b) , onde $a \leq m < n \leq b$.

$$A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } x \in (a, m) \\ 1 & \text{se } x \in [m, n] \\ \frac{b-x}{b-n} & \text{se } x \in (n, b) \\ 0 & \text{se } x \geq b \end{cases}$$

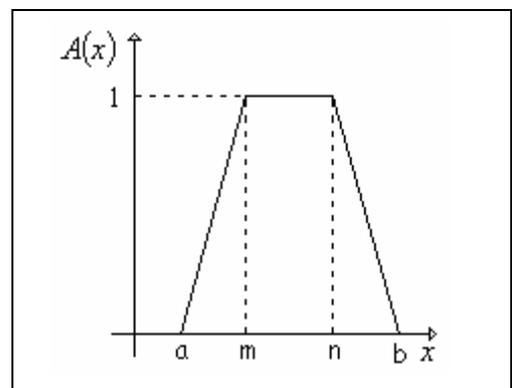
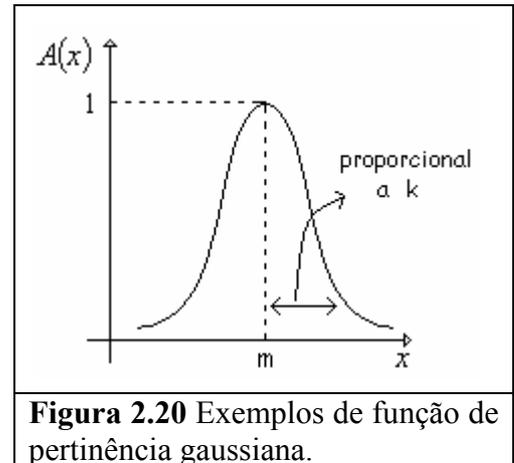


Figura 2.19 Exemplos de função de pertinência trapezoidal.

Função gaussiana: parâmetros (m, k) , onde $k > 0$.

$$A(x) = e^{-k(x-m)^2}$$



Um tipo bastante específico de conjunto *fuzzy* que pode ter utilidade em aplicações é o chamado conjunto *fuzzy* unitário, ou *fuzzy singleton*. Nesse conjunto, existe apenas um ponto x no universo \mathbf{X} para o qual o valor da função de pertinência é não-nulo.

2.3.4.3 Definições básicas.

Um conjunto *fuzzy* pode ser caracterizado por alguns atributos adicionais. O suporte de um conjunto *fuzzy* A (S_A): conjunto *crisp* que contém todos os elementos do conjunto universo \mathbf{X} que têm grau de pertinência diferente de 0 (zero) em A .

$$S_A = \{x \mid A(x) > 0\}$$

O núcleo de um conjunto *fuzzy* A (N_A) é um conjunto *crisp* que contém todos os elementos do conjunto universo \mathbf{X} que têm grau de pertinência igual a 1 (um) em A .

$$N_A = \{x \mid A(x) = 1\}$$

A altura de um conjunto *fuzzy* A (H_A) é o maior grau de pertinência obtido por qualquer elemento do conjunto A (sup denota supremo).

$$H_A = \sup_x \{A(x)\}$$

O conjunto normal de um conjunto *fuzzy* é dito normal se sua altura for igual a 1 (um). Caso contrário é chamado de subnormal. O núcleo de um conjunto subnormal é um conjunto vazio.

Um dos conceitos mais importantes relacionados a conjuntos *fuzzy* é o conceito de corte- α e sua variante corte- α forte. O conjunto corte- α contém todos os elementos do conjunto universo \mathbf{X} que têm grau de pertinência maior ou igual que α , para algum $\alpha \in [0,1]$.

$$A_\alpha = \{x \mid A(x) \geq \alpha\}$$

O conjunto corte- α forte é o conjunto *crisp* que contém os elementos do conjunto universo X que têm grau de pertinência maior que α , para algum $\alpha \in [0,1]$.

$$A_{\alpha+} = \{x \mid A(x) > \alpha\}$$

Associado ao conceito de corte- α é definido o conjunto de níveis de A , como o conjunto de todos os valores $\alpha \in [0,1]$ que representam cortes- α distintos de um dado conjunto *fuzzy* A , ou seja:

$$\eta_A = \{\alpha \mid A(x) = \alpha, \forall x \in X\}$$

2.3.4.4 Operações em conjuntos *fuzzy*

As operações sobre conjuntos *fuzzy* são essenciais no processo de representação e processamento de conhecimento na forma de conjuntos *fuzzy*. As operações sobre conjuntos *fuzzy* são distinguidas entre os casos das operações de argumento único, que modificam o formato da função de pertinência que define o conjunto *fuzzy*, e das operações de múltiplos argumentos, que relacionam dois ou mais conjuntos *fuzzy*.

As operações simples possuem um argumento único que modificam o formato da função de pertinência. A normalização é a transformação de um conjunto *fuzzy* subnormal A em um conjunto *fuzzy* normal A^N por:

$$A^N(x) = A(x) / H_A, \text{ sendo } H_A \text{ a altura de } A.$$

A concentração define um conjunto *fuzzy* A^C tornando os valores de pertinência de A relativamente menores.

$$A^C(x) = A^2(x)$$

A generalização da concentração.

$$A^C(x) = A^p(x), p > 1$$

A dilatação define um conjunto *fuzzy* A^D tornando os valores de pertinência de A relativamente maiores.

$$A^D(x) = A^{0,5}(x)$$

A generalização da dilatação.

$$A^D(x) = A^p(x), p \in (0;1)$$

A intensificação de contraste resulta no conjunto A^1 combinando as operações de dilatação e concentração. Os valores de pertinência menores que $\frac{1}{2}$ diminuem e os maiores aumentam.

$$A^1(x) = \begin{cases} 2 * A^2(x) & \text{se } A(x) \in [0;0,5] \\ 1 - 2 * (1 - A(x))^2 & \text{se } A(x) \in (0,5;1] \end{cases}$$

As operações compostas, conhecidas da teoria de conjuntos *crisp* – união, interseção e complemento – podem ser generalizadas para o contexto dos conjuntos *fuzzy*. Essa generalização pode ser feita de formas diferentes, mas uma delas em particular tem um significado especial na teoria de conjuntos *fuzzy*, por isso as operações assim obtidas são chamadas de operações padrão. Outro tipo são as operações generalizadas.

As operações sobre conjuntos *fuzzy* referidas como operações padrão têm uma importância destacada na teoria de conjuntos *fuzzy* por possuírem certas propriedades que dá a elas um significado especial (KLIR e YUAN, 1995). Na seqüência são apresentadas as definições de tais operações.

Sejam A e B dois conjuntos *fuzzy* no universo X com suas funções de pertinência $A(x)$ e $B(x)$, respectivamente.

A interseção nebulosa padrão ($A \cap B$) dos conjuntos A e B é definida por:

$$(A \cap B)(x) = \min[A(x), B(x)] = A(x) \wedge B(x)$$

onde \min e \wedge denotam o operador de mínimo.

A união nebulosa padrão ($A \cup B$) dos conjuntos A e B é definida por:

$$(A \cup B)(x) = \max[A(x), B(x)] = A(x) \vee B(x)$$

onde \max e \vee denotam o operador de máximo.

O complemento *fuzzy* de A indica o grau com que x não pertence ao conjunto A, e pode ser definido por:

$$\overline{A}(x) = 1 - A(x)$$

A figura 2.21 mostra exemplos de união, interseção e complemento padrão aplicados a dois conjuntos *fuzzy* A e B.

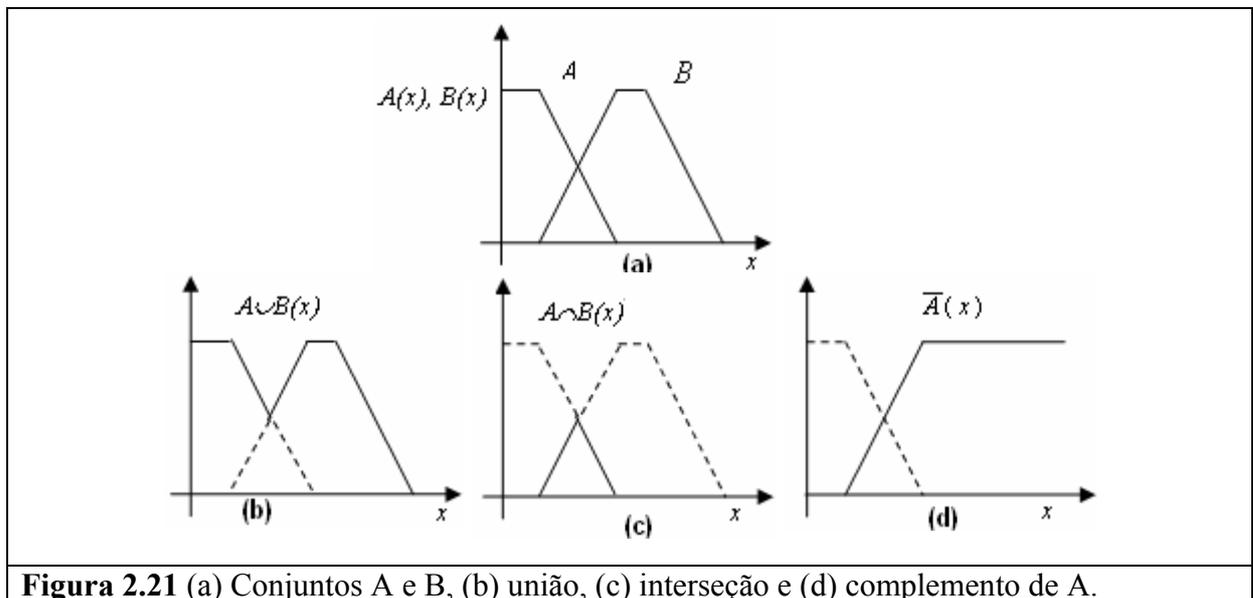


Figura 2.21 (a) Conjuntos A e B, (b) união, (c) interseção e (d) complemento de A.

As operações *fuzzy* padrão não são as únicas generalizações possíveis para as operações sobre conjuntos *fuzzy*. Para cada uma das três operações, existe uma classe de funções que podem ser utilizadas na generalização dessas operações. As funções que podem ser utilizadas na generalização das operações de interseção e união são referidas na literatura como t-normas e t-conormas, respectivamente.

A interseção de conjuntos *fuzzy* pode ser realizada usando operadores que pertencem a uma classe de funções conhecidas como t-normas (KLIR & YUAN, 1995). As t-normas são operações binárias da forma $i: [0,1]^2 \rightarrow [0,1]$ que satisfazem as seguintes propriedades:

Comutatividade: $i(x,y) = i(y,x)$;

Associatividade: $i(x,i(y,z)) = i(i(x,y),z)$;

Monotonicidade: se $x \leq y$ e $z \leq w$, então $i(x,z) \leq i(y,w)$;

Condição Limite: $i(0,x) = 0$ e $i(1,x) = x$.

Alguns exemplos de t-normas são:

Intersecção Padrão: $a \ t_1 \ b = \min\{a, b\}$

Produto Algébrico: $a \ t_2 \ b = ab$

Diferença Limitada: $a \ t_3 \ b = \max\{0, a + b - 1\}$

Logo, interseção nebulosa generalizada ($A \cap B$) dos conjuntos A e B é definida por:

$(A \cap B)(x) = A(x) \ t \ B(x)$

onde t denota uma t -norma.

A união de conjuntos *fuzzy* pode ser realizada usando operadores que pertencem a uma classe de funções conhecidas como s -normas ou t -conormas (KLIR & YUAN, 1995). As s -normas são operações binárias da forma $u : [0,1]^2 \rightarrow [0,1]$ que satisfazem as seguintes propriedades:

Comutatividade: $u(x,y) = u(y,x)$;

Associatividade: $u(x,u(y,z)) = u(u(x,y),z)$;

Monotonicidade: se $x \leq y$ e $z \leq w$, então $u(x,z) \leq u(y,w)$;

Condição Limite: $u(0,x) = x$ e $u(1,x) = 1$.

Alguns exemplos de s -normas são:

União Padrão: $a \text{ s } b = \max\{a,b\}$

Soma Algébrica: $a \text{ s } b = a + b - ab$

Soma Limitada: $a \text{ s } b = \min\{1, a + b\}$

Logo, a união nebulosa generalizada ($A \cup B$) dos conjuntos A e B é definida por:

$(A \cup B)(x) = A(x) \text{ s } B(x)$

onde s denota uma s -norma.

Devido à propriedade de associatividade das t -normas e s -normas, as operações de união e interseção nebulosas podem ser aplicadas a n conjuntos, com $n \geq 2$.

De forma análoga, a operação generalizada de complemento fuzzy pode ser realizada usando-se um operador da classe de operadores definido como $c : [0,1] \rightarrow [0,1]$ que satisfazem as seguintes propriedades:

Monotonicidade: para todo $x,y \in [0,1]$, se $x \leq y$, então $c(x) \geq c(y)$;

Condição Limite: $c(0) = 1$ e $c(1) = 0$;

Involução: $c(c(x)) = x$ para $x \in [0,1]$.

2.3.4.5 Computação com regras fuzzy

As regras são utilizadas como um mecanismo formal para representar conhecimento de um domínio. As regras *fuzzy*, mais especificamente, permitem a representação de conhecimento impreciso, o que muitas vezes torna a modelagem do problema mais fiel e adequada ao mundo real. Todo processo computacional envolvendo regras consiste em aplicar os chamados mecanismos de raciocínio que usam as regras *fuzzy* juntamente com os fatos, para fazer inferências.

Variáveis Lingüísticas

Grande parte da experiência e do conhecimento dos seres humanos pode ser mais bem representada na forma lingüística, mais geral e imprecisa. Por outro lado, as representações numéricas são mais precisas e mais facilmente tratáveis por computador. As variáveis lingüísticas permitem que problemas complexos ou mal definidos sejam adequadamente modelados com a representação apropriada das imprecisões naturalmente presentes nesses domínios pelo uso de termos lingüísticos, ao mesmo tempo em que possibilita seu tratamento computacional. Essas variáveis provêm mecanismos de tradução entre descrições lingüísticas imprecisas e descrições numéricas, passíveis de tratamento computacional. Essa tradução se baseia no mapeamento entre termos lingüísticos e conjuntos *fuzzy*.

Uma variável lingüística pode ser definida, de maneira informal, como uma variável cujos valores são palavras ou sentenças em linguagem natural ao invés de números (ZIMMERMANN, 1991). Elas são definidas sobre um determinado domínio, o qual é granularizado em termos lingüísticos definidos por conjuntos *fuzzy*. O processo de granularização de um domínio de uma variável em conjuntos *fuzzy* define a chamada partição nebulosa.

O conjunto de termos lingüísticos que a variável lingüística assume pode conter termos primários e não-primários. Os termos primários são aqueles que representam as categorias mais elementares (por exemplo, baixa, média, alta) e têm uma associação direta com cada conjunto *fuzzy* da partição definida sobre o domínio da variável lingüística. Os termos não-primários são formados a partir dos termos primários com o uso de modificadores (muito, mais ou menos, etc.) e conectivos (e, ou não), como por exemplo, muito baixa, não muito baixa e não muito alta, etc. Usualmente, as funções de pertinência associada aos termos primários têm formatos conhecidos (triangular, trapezoidal, gaussiana). As funções de pertinência associada aos termos não primários são obtidas pela aplicação de operadores sobre as funções dos termos primários. Usualmente, os modificadores são implementados pelas operações de um único argumento e os conectivos, pelas operações de interseção, união e complemento de conjuntos *fuzzy* (detalhes em PEDRYCZ e GOMIDE, 1999).

Na Figura 2.22, é mostrado um exemplo de partição *fuzzy* de uma variável lingüística *velocidade*. O universo de discurso utilizado é um segmento de escala km/h, entre 0 e 100 Km.

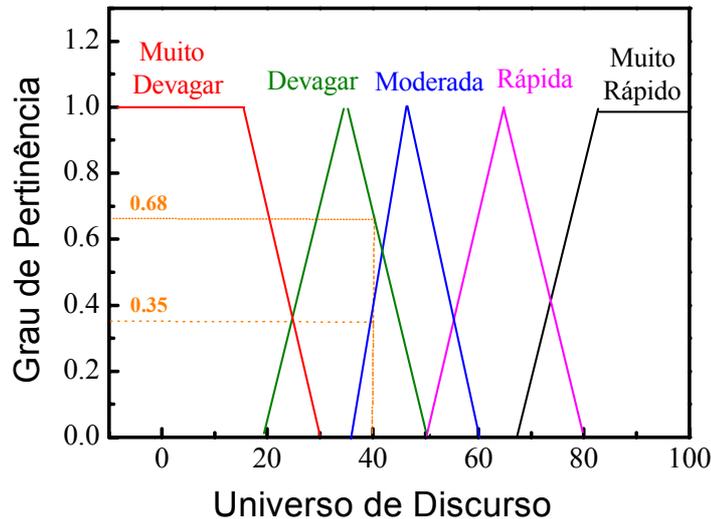


Figura 2.22 Variável lingüística *velocidade* e sua divisão em categorias através de conjuntos *fuzzy*.

Sintaxe das regras *fuzzy*

O uso de regras como mecanismo formal para representação do conhecimento remonta às primeiras décadas da criação da área de Inteligência Artificial quando as regras de produção foram propostas como um par condição-ação que define uma porção de conhecimento para solução de um problema (LUGER, 2002). As regras são uma das técnicas mais antigas para representação de conhecimento e permanecem sendo largamente utilizadas. Sua popularidade vem do fato de possibilitar a expressão clara de diretivas e estratégias, captar o conhecimento experimental de especialistas humanos e possuir um formato lingüístico de fácil compreensão.

No contexto de sistemas *fuzzy* as regras adquirem sintaxe e semântica particulares e são chamadas de regras *fuzzy*. Sua principal propriedade é a capacidade de representar conhecimento impreciso, pelo uso de termos lingüísticos.

Em geral, as regras nebulosas têm a seguinte forma:

SE <antecedente> ENTÃO <conseqüente>

e têm a finalidade de estabelecer relações entre as variáveis que aparecem no antecedente – também chamado de condição ou premissa – e as que aparecem no conseqüente – também chamado de conclusão ou ação. Por exemplo, a regra:

SE a distância é grande ENTÃO o tempo de viagem é longo estabelece uma relação entre as variáveis distância e tempo_de_viagem usando os termos grande e longo, que são inerentemente imprecisos.

Com a finalidade de detalhar melhor a sintaxe de uma regra nebulosa, discutimos inicialmente a proposição atômica *fuzzy*. A proposição atômica *fuzzy* é a parcela de informação básica que pode aparecer em uma regra e seu formato é:

$$X \text{ é } A$$

onde X é uma variável linguística sobre um domínio X e A é um termo linguístico que representa um conjunto *fuzzy*.

Por exemplo, se X for a variável distância e A o termo grande, a proposição $X \text{ é } A$ representa formalmente a parcela de informação do exemplo anterior que diz a distância é grande.

Proposições *fuzzy* podem ser compostas, isto é, construídas a partir de proposições atômicas pelos operadores de conjunção ou disjunção. Por exemplo, as proposições:

$$X_1 \text{ é } A \text{ E } X_2 \text{ é } B \text{ e}$$

$$X_1 \text{ é } C \text{ OU } X_2 \text{ é } D$$

podem representar conhecimento do tipo:

“A temperatura está alta E a umidade está baixa” e “O tempo está nublado OU o volume de trabalho é muito grande”, respectivamente.

Uma regra *fuzzy* assume o formato de uma proposição condicional, que no caso mais simples fica:

$$\text{SE } X \text{ é } A \text{ ENTÃO } Y \text{ é } B$$

onde X e Y são variáveis linguísticas sobre os domínios X e Y , respectivamente, e A e B são termos linguísticos, que rotulam conjuntos fuzzy definidos sobre os domínios X e Y , respectivamente.

Em uma regra *fuzzy*, tanto o antecedente quanto o conseqüente podem ser formados por proposições compostas, o que dá origem, por exemplo, a regras da forma:

$$\text{SE } X_1 \text{ é } A_1 \text{ E } X_2 \text{ é } A_2 \text{ E } X_3 \text{ é } A_3 \text{ ENTÃO } Y_1 \text{ é } B_1 \text{ E } Y_2 \text{ é } B_2$$

No caso mais geral, tanto o antecedente como o conseqüente da regra podem ser formados por um número finito de proposições atômicas relacionadas por conjunção ou disjunção. Assim, assumindo que p e q denotam proposições *fuzzy*, uma regra *fuzzy* pode ser da forma SE p ENTÃO q sendo que a proposição p pode ter a forma:

$$X_1 \text{ é } A_1 \text{ E } X_2 \text{ é } A_2 \text{ E } \dots \text{ E } X_n \text{ é } A_n, n \geq 1 \text{ ou}$$

$$X_1 \text{ é } A_1 \text{ OU } X_2 \text{ é } A_2 \text{ OU } \dots \text{ OU } X_n \text{ é } A_n, n \geq 1$$

e a proposição q pode ter a forma:

$$Y_1 \text{ é } B_1 \text{ E } Y_2 \text{ é } B_2 \text{ E } \dots \text{ E } Y_m \text{ é } B_m, m \geq 1 \text{ ou}$$

$$Y_1 \text{ é } B_1 \text{ OU } Y_2 \text{ é } B_2 \text{ OU } \dots \text{ OU } Y_m \text{ é } B_m, m \geq 1$$

onde A_1, A_2, \dots, A_n e B_1, B_2, \dots, B_m são conjuntos fuzzy nos universos $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m$, respectivamente, e X_1, X_2, \dots, X_n e Y_1, Y_2, Y_m são variáveis sobre os domínios $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m$, respectivamente.

Semântica das regras fuzzy

Considerando uma regra *fuzzy* no seu formato mais simples:

$$\text{SE } X \text{ é } A \text{ ENTÃO } Y \text{ é } B$$

Entende-se que esta regra descreve uma relação entre as variáveis X e Y , que pode ser descrita por uma relação *fuzzy* R cuja função de pertinência $R(x,y)$ descreve o grau com que o par $(x,y) \in X \times Y$ é compatível com a relação entre as variáveis X e Y . Diz-se que uma regra *fuzzy* nessas condições induz uma relação *fuzzy* R .

Sendo A e B conjuntos *fuzzy* sobre X e Y , respectivamente a relação R pode ser determinada por

$$R(x,y) = f(A(x), B(y)), \forall (x,y) \in X \times Y$$

onde f é uma função na forma $f : [0,1]^2 \rightarrow [0,1]$.

A escolha da função f é fundamental para o sistema *fuzzy*, pois determina a semântica da regra e diversos aspectos de processamento do sistema.

Em geral as relações induzidas por regras *fuzzy* derivam de três classes principais de funções: conjunção *fuzzy*, disjunção *fuzzy* ou implicação *fuzzy* (PEDRYCZ e GOMIDE, 1999).

Conjunção *fuzzy* é uma função $f_t : [0,1]^2 \rightarrow [0,1]$ definida por:

$$f_t(A(x), B(y)) = A(x) \text{ t } B(y), \text{ onde } t \text{ denota uma t-norma.}$$

As definições mais comuns são as de Mamdani (MAMDANI & ASSILIAN, 1975; MAMDANI, 1977) e Larsen (LARSEN, 1980), que utilizam o mínimo e o produto algébrico.

$$\text{Mamdani: } f_m(A(x), B(y)) = A(x) \wedge B(y).$$

$$\text{Larsen: } f_p(A(x), B(y)) = A(x) \cdot B(y).$$

Disjunção *fuzzy* é uma função $f_s : [0,1]^2 \rightarrow [0,1]$ definida por:

$$f_s(A(x), B(y)) = A(x) \text{ s } B(y), \text{ onde } s \text{ denota uma s-norma.}$$

Implicação *fuzzy* é uma função $f_l : [0,1]^2 \rightarrow [0,1]$ definida por um conjunto de propriedades e classificada em categorias caracterizadas pelo formalismo da lógica clássica e

da lógica intuicionista. Maiores detalhes podem ser encontrados em PEDRYCZ e GOMIDE (1999). Um dos exemplos mais conhecidos de implicação *fuzzy* é a implicação de LUKASIEWICZ (1970):

$$f_l(A(x), B(y)) = \min[1, 1 - A(x) + B(y)]$$

Caso a regra possua múltiplas variáveis no antecedente ou no conseqüente, a relação R é definida de forma semelhante. Para a regra:

$$\text{SE } X_1 \text{ é } A_1 \text{ E } X_2 \text{ é } A_2 \text{ E } \dots \text{ E } X_n \text{ é } A_n \text{ ENTÃO } Y_1 \text{ é } B_1 \text{ E } Y_2 \text{ é } B_2 \text{ E } \dots \text{ E } Y_m \text{ é } B_m$$

definimos a relação R assumindo que a proposição composta $X_1 \text{ é } A_1 \text{ E } X_2 \text{ é } A_2 \text{ E } \dots \text{ E } X_n \text{ é } A_n$ induz uma relação n-ária definida por:

$$P_a(x_1, x_2, \dots, x_n) = A_1(x_1) \text{ t } A_2(x_2) \text{ t } \dots \text{ t } A_n(x_n)$$

e a proposição composta $Y_1 \text{ é } B_1 \text{ E } Y_2 \text{ é } B_2 \text{ E } \dots \text{ E } Y_m \text{ é } B_m$ induz uma relação m-ária definida por:

$$P_c(y_1, y_2, \dots, y_m) = B_1(y_1) \text{ t } B_2(y_2) \text{ t } \dots \text{ t } B_m(y_m)$$

A regra induz então uma relação de ordem n+m definida por:

$$R(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = f(P_a(x_1, x_2, \dots, x_n), P_c(y_1, y_2, \dots, y_m))$$

Caso o antecedente ou conseqüente da regra seja formado pelo conectivo OU em vez de E, substituímos a t-norma por uma s-norma no cálculo das relações induzidas P_a ou P_c .

Inferência com regras *fuzzy*

A computação com regras *fuzzy* envolve a aplicação de um método de raciocínio, o qual está baseado em regras de inferência. A regra da inferência composicional é a regra de inferência básica para os processos de raciocínio nos sistemas *fuzzy*. A regra da inferência composicional, também chamada de regra da composição, será apresentada para o caso simples, isto é, quando a regra tem apenas uma variável no antecedente e uma no conseqüente.

O esquema de inferência de interesse nesse contexto é aplicado a partir de uma regra, que estabelece um conhecimento genérico sobre algum domínio, e de um fato, que estipula uma informação específica sobre a situação da qual se deseja inferir alguma conclusão.

Seja A um conjunto fuzzy sobre X, e B um conjunto fuzzy sobre Y. Conforme o que foi descrito anteriormente sobre a semântica das regras, a regra

$$\text{SE } X \text{ é } A \text{ ENTÃO } Y \text{ é } B$$

induz uma relação nebulosa $R(x,y)$ sobre $X \times Y$. Para usar essa regra em uma inferência, é necessário conhecer algum fato que estabeleça um valor específico para a variável do antecedente X .

Dado o fato X é A' , pode-se inferir a conclusão Y é B' , sendo que B' é definido por (PEDRYCZ e GOMIDE, 1999):

$$B'(y) = \sup_{x \in X} [A'(x) \wedge R(x,y)] \text{ para todo } y \in Y$$

Onde \sup denota o operador de supremo.

A aplicação da regra da inferência composicional pode ser facilmente estendida para regras com múltiplas variáveis, e também para situações em que o conhecimento do problema é descrito por um conjunto de regras. O processo geral de computação usando regras nebulosas é baseado na regra de inferência composicional.

Alternativamente a este, pode ser aplicado o chamado método de inferência escalonada, que envolve um número pequeno de cálculos numéricos com relação ao método geral. Esses dois processos gerais de computação com regras estão descritos em PEDRYCZ e GOMIDE (1999).

2.3.4.6 Sistema *fuzzy* baseado em regras

Um sistema *fuzzy* é qualquer sistema que incorpora algum mecanismo derivado da lógica nebulosa e que pelo menos uma das variáveis do problema pode assumir valores lingüísticos definidos por conjuntos *fuzzy* (ZADEH, 1973; KLIR & YUAN, 1995).

O tipo particular de sistemas *fuzzy* de interesse para este trabalho são os sistemas *fuzzy* baseados em regras (SFBR). Os SFBR têm sido usados com sucesso em diversas áreas de aplicação, principalmente em problemas de controle, classificação e modelagem (ver, por exemplo: PEDRYCZ, 1996; PEDRYCZ e GOMIDE 1999; DUMITRESCU *et al.*, 2000). Um SFBR é composto por dois componentes principais: a base de conhecimento (BC) e o mecanismo de inferência (MI).

A base de conhecimento é composta pela base de dados *fuzzy* (BDF), a qual contém as definições dos conjuntos *fuzzy* relacionados aos termos lingüísticos usados na regras *fuzzy* e pela base de regras *fuzzy* (BRF), que armazena o conjunto de regras *fuzzy* que modelam um determinado problema.

O mecanismo de inferência é responsável pelo processamento das regras, o qual é realizado por um método de raciocínio. Este consiste da aplicação de um procedimento de inferência para derivar conclusões a partir das regras e de fatos conhecidos.

Com a definição dos diversos parâmetros envolvidos em um SNBR, obtêm-se o que é chamado de modelo de sistema *fuzzy*, sendo que alguns desses modelos recebem destaque na literatura, devido a seu sucesso em aplicações práticas. Nesse sentido, podemos citar dois modelos principais de sistemas *fuzzy*: os de Mamdani (MAMDANI e ASSILIAN, 1975; MAMDANI, 1977) que possuem na parte conseqüente das regras nebulosas uma proposição nebulosa e os de TAKAGI-SUGENO (TAKAGI e SUGENO, 1983; 1985) que possuem na parte conseqüente uma função aplicada sobre os valores de entrada.

2.3.5 Redes neurais artificiais

Redes neurais artificiais (RNA) são modelos matemáticos que se assemelham às estruturas neurais biológicas e que tem capacidade computacional adquirida por meio de aprendizado e generalização (HAYKIN, 2001). A habilidade de um ser humano em realizar funções complexas e principalmente a sua capacidade de aprender advém do processamento paralelo e distribuído da rede de neurônios do cérebro. Um novo conhecimento ou uma nova experiência pode levar as alterações estruturais no cérebro. Tais alterações são efetivadas por meio de rearranjo das redes de neurônios, reforçando ou inibindo algumas sinapses (HAYKIN, 2001). A busca por um modelo computacional que simule o funcionamento do cérebro data dos anos 40, com o trabalho de MCCULLOCH e PITTS (1973). O modelo de neurônio artificial da Figura 2.23 é uma simplificação do modelo apresentado por HAYKIN (2001).

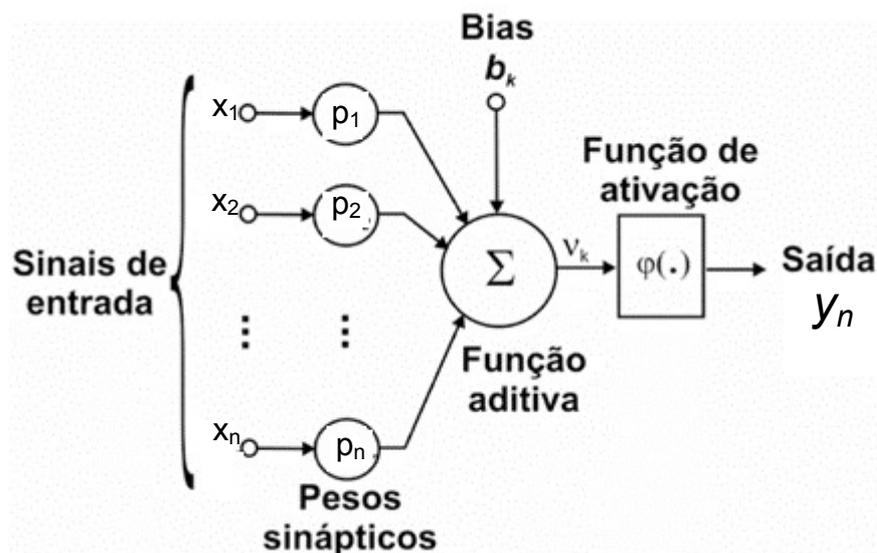


Figura 2.23 Funcionamento de um neurônio computacional.

Este modelo é composto por três elementos básicos: um conjunto de n conexões de entrada (x_1, x_2, \dots, x_n), caracterizadas por pesos (p_1, p_2, \dots, p_n); um somador (Σ) para acumular os sinais de entrada e uma função de ativação (φ).

As diferentes camadas podem possuir uma conexão adicional, chamada de *bias*. *Bias* de aprendizado é qualquer preferência de uma hipótese sobre a outra, além da simples consistência com os exemplos. Devido ao fato de quase sempre existir um número grande de hipóteses consistentes, todos os indutores possuem alguma forma de *bias*. MITCHELL, 1982, disse que aprendizado sem *bias* é impossível.

A função de ativação (φ) limita o intervalo permissível de amplitude do sinal de saída (y_n) a um valor fixo. Existem vários tipos de funções de ativação, as mais comuns são a função de transferência de limite, linear ou sigmoideal, como mostra a Figura 2.24.

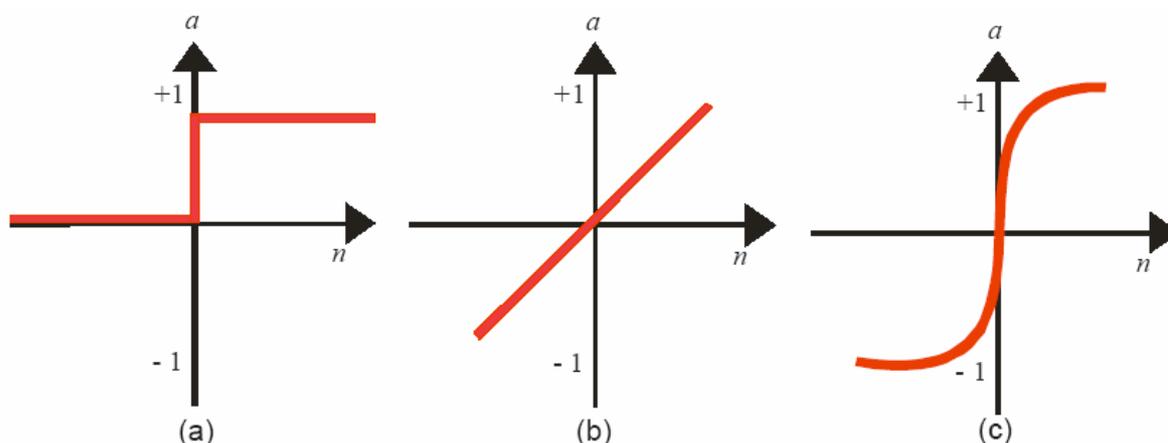


Figura 2.24 Funções: (a) transferência de limite, (b) linear e (c) sigmoideal. Os eixos a e n correspondem a saída e a entrada da rede, respectivamente.

A função de transferência de limite limita a saída do neurônio para zero (se a entrada n for menor do que zero) ou um (se a entrada n for maior ou igual a um). Ela é utilizada para fazer decisões em classificação. A função linear é utilizada por neurônios como aproximadores lineares em filtros lineares. A função sigmoideal recebe a entrada de qualquer valor entre mais ou menos infinito, e retorna a saída entre zero e um. Ela é mais usada em redes *backpropagation*, como será visto adiante, porque é diferenciável.

O comportamento das conexões entre os neurônios é simulado por meio dos seus pesos. O efeito de um sinal proveniente de um outro neurônio é determinado pela multiplicação do valor (intensidade) do sinal recebido pelo peso correspondente ($x_i \times p_i$). É efetuada a soma dos valores $x_i \times p_i$ de todas as conexões, e o valor resultante é enviado para a função de ativação, que define a saída (y) do neurônio. Combinando diversos neurônios, forma-se uma rede neural artificial (RNA).

As redes neurais artificiais se diferenciam pela sua arquitetura e pela forma como os pesos associados às conexões são ajustados durante o processo de aprendizado. A arquitetura de uma rede neural restringe o tipo de problema no qual a rede poderá ser utilizada, e é definida pelo número de camadas (camada única ou múltiplas camadas), pelo número de nós em cada camada, pelo tipo de conexão entre os nós (*feedforward* ou *feedback*) e por sua topologia (HAYKIN, 2001).

Uma das propriedades mais importante de uma RNA é a capacidade de aprender por intermédio de exemplos e fazer inferências sobre o que aprendeu, melhorando gradativamente o seu desempenho. As redes neurais utilizam um algoritmo de aprendizagem cuja tarefa é ajustar os pesos de suas conexões. O aprendizado em RNA é feito de forma empírica, ou seja, está associado a capacidade das mesmas adaptarem os seus parâmetros como consequência da sua interação com o meio externo. O processo de aprendizado é interativo. Uma apresentação completa do conjunto de dados é chamada de época de aprendizado. O critério de desempenho que determina a qualidade do modelo neural e o ponto de parada de treinamento são preestabelecidos pelos parâmetros de treinamento e podem estar associados, por exemplo, ao valor do erro quadrático médio das respostas da RNA para um determinado conjunto de dados. Por sua vez, a generalização de uma RNA está associada à sua capacidade de dar respostas coerentes para dados não apresentados a ela previamente durante o treinamento.

Os métodos de aprendizado neural podem ser divididos em três grandes classes, segundo o grau de controle dado ao usuário que são o aprendizado supervisionado, não supervisionado e semi-supervisionado.

No aprendizado supervisionado, Figura 2.25, o usuário dispõe de um comportamento de referência preciso que ele deseja ensinar a rede. Sendo assim, a rede deve ser capaz de medir a diferença entre seu comportamento atual e o comportamento de referência, e então corrigir os pesos de maneira a reduzir este erro (desvio de comportamento em relação aos exemplos de referência).

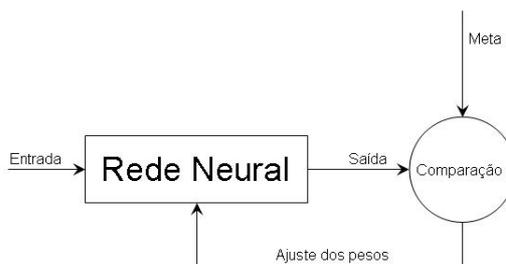


Figura 2.25 Estrutura básica de treinamento supervisionado de uma rede neural.

O aprendizado supervisionado utiliza conhecimentos empíricos, habitualmente representados por um conjunto de exemplos etiquetados, ou seja, exemplos com pares de dados de entrada com a respectiva saída associada. Nos casos de problemas de classificação, a saída é a classe à qual cada exemplo está associado. Nesta dissertação, as RNAs foram treinadas de modo supervisionado, e tiveram como entrada, os dados das variáveis de tempo do ensaio, fração molar de CO₂ (y_{CO_2}) e velocidade de agitação, associado ao valor de biomassa do microrganismo (Cx), dado de saída.

No aprendizado não-supervisionado os pesos da rede são modificados em função de critérios internos, tais como, por exemplo, a repetição de padrões de ativação em paralelo de vários neurônios. O comportamento resultante deste tipo de aprendizado é usualmente comparado com técnicas de análise de dados empregadas na estatística (*clustering*).

No aprendizado semi-supervisionado o usuário possui apenas indicações imprecisas (por exemplo: sucesso/insucesso da rede) sobre o comportamento final desejado. As técnicas de aprendizado semi-supervisionado são chamadas também de aprendizado por reforço. Neste tipo de aprendizado se tem apenas de uma avaliação qualitativa do comportamento do sistema, sem, no entanto poder medir quantitativamente o erro (desvio do comportamento em relação ao comportamento de referência desejado).

Os diferentes tipos de modelos de redes neurais são determinados pela maneira que os neurônios da rede estão estruturados e são treinados.

O modelo *perceptron* é um modelo utilizado para a classificação de padrões linearmente separáveis (HAYKIN, 2001), ou seja, ele aprende uma função de decisão para classificar dois conjuntos linearmente separáveis. Este modelo consiste apenas de um neurônio e utiliza o método de aprendizado supervisionado.

O *perceptron* multicamadas (*Multilayer Perceptron* – MLP) representa uma generalização do *perceptron*, sendo também conhecido como rede multicamada progressiva (*feedforward*). O MLP tem esta denominação pelo fato dos neurônios não possuírem suas saídas conectadas aos neurônios localizados nas camadas anteriores. O MLP é uma importante arquitetura de rede neural, pelo fato de ser aplicável a problemas de reconhecimento de padrões, independente das classes serem separáveis linearmente (HAYKIN, 2001). Suas três características básicas são: cada neurônio na rede inclui uma função de decisão linear, como por exemplo, uma função de ativação sigmoideal, onde os valores de saída poderão oscilar entre menos um e um, prevalecendo como resposta a classe relacionada ao neurônio da camada de saída que apresentar o maior valor; a rede contém uma

ou mais camadas ocultas e a rede apresenta um alto grau de conectividade entre as camadas, como mostra a Figura 2.26.

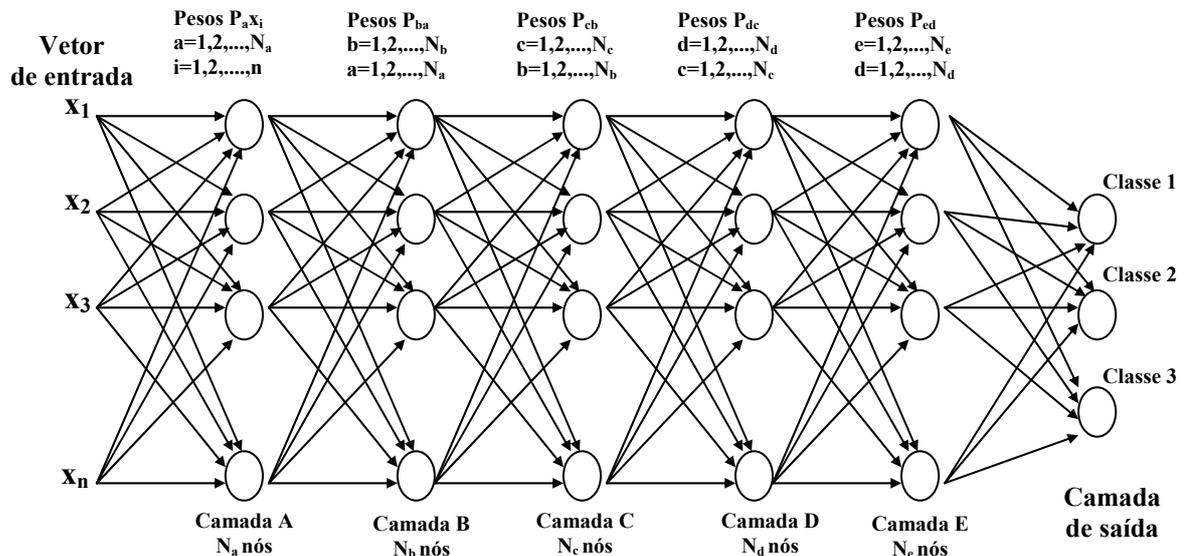


Figura 2.26 Estrutura esquemática do *perceptron* multicamadas (MLP).

O MLP é uma rede supervisionada, tendo o seu algoritmo de treinamento composto por dois passos. O primeiro passo é o *feedforward*, onde um vetor de padrões é apresentado para a camada de entrada, sendo propagado pela rede, camada por camada. Assim, cada neurônio ao receber um sinal, calcula sua função de ativação e envia o valor para os neurônios da próxima camada, até alcançar os neurônios de saída. Os neurônios de saída calculam suas ativações resultando no valor da saída neural. O segundo passo é o *backpropagation*. Como se trata de uma rede supervisionada que requer uma resposta desejada para ser treinada, o erro para cada neurônio de saída é calculado comparando o valor obtido para a saída com o valor alvo. Esse erro é então propagado para as camadas inferiores da rede, obtendo-se a ponderação dos erros da camada superior. Esses passos são repetidos até os erros dos neurônios de saída serem considerados satisfatórios.

Nesta dissertação, fez-se uma alternativa a esse enfoque: a aplicação de um consórcio, ou comitê de redes neurais. Será utilizado um banco de redes MLP (*Multilayer Perceptron*), com uma camada oculta, do tipo *feedforward*, já treinadas (de forma supervisionada) em ensaios-padrão, para se obter uma resposta ponderada, mais adequada ao cultivo em andamento.

2.3.5.1 Comitê de redes neurais

De acordo com o princípio de “dividir para conquistar”, uma tarefa computacional complexa pode ser resolvida mais eficientemente se for dividida em tarefas mais simples,

tendo suas respostas combinadas posteriormente (HAYKIN, 2001). Como o objetivo é ter tarefas simplificadas, no treinamento, as tarefas são distribuídas entre um número de redes, denominadas “especialistas”. A combinação das especialistas constitui o comitê de redes neurais (CRNAs). Basicamente, CRNAs é a fusão do conhecimento adquirido pelas redes especialistas para chegar a uma decisão conjunta que é supostamente superior à ação de uma única rede.

Dois tipos principais de estrutura são encontrados na literatura (HAYKIN, 2001). A primeira estrutura é a estática onde as respostas das redes individuais são combinadas sem o envolvimento dos dados de entrada. Esta categoria inclui dois métodos. O método *ensemble averaging* (Figura 2.27) onde as saídas são combinadas linearmente para produzir uma saída global e o método *boosting* onde o algoritmo de treinamento elimina as redes fracas, filtrando as entradas, para atingir uma maior precisão.

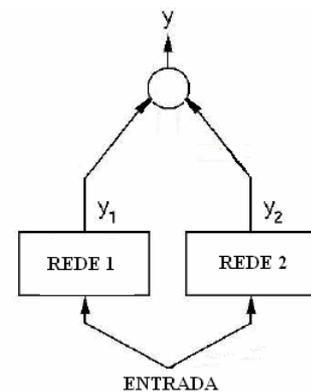


Figura 2.27 Estrutura esquemática do método Ensemble Average.

A segunda estrutura é a dinâmica onde são utilizados “pesos”, baseados nos dados de entrada, para a combinação dos dados das redes individuais. Esta estrutura também engloba dois métodos. O método *mixture experts* (Figura 2.28 a) onde cada rede individual é combinada por um simples peso e o método *hierarchical mixture experts* (Figura 2.28 b) onde cada rede individual é responsável por uma combinação não linear, ponderada por pesos arranjados em uma estrutura hierárquica.

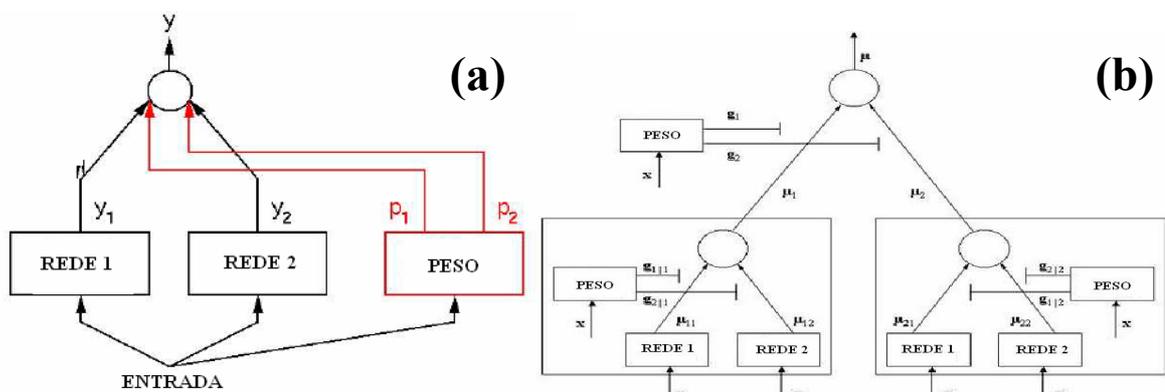


Figura 2.28 Estrutura esquemática dos métodos (a) *mixture experts* e (b) *hierarchical mixture experts*.

A utilização de um CRNA, para a inferência de estado em bioprocessos, nesta dissertação, tem como objetivo, testar esta ferramenta em inferência, tendo em vista que

normalmente é utilizada como critérios de classificação ou identificação. Outra vantagem é que a cada novo ensaio (cultivo), somente o novo ensaio passa por treinamento e depois seu resultado é adicionado ao banco de dados do comitê. Não existe a necessidade de treinar todas as redes novamente.

2.3.6 Sistema Híbrido Inteligente (SHI)

Os sistemas híbridos inteligentes (SHI), que resultam da combinação de duas ou mais metodologias distintas, são uma tendência presente no contexto da inteligência computacional (IC). Esta tendência, na área de controle de bioprocessos, tem como explicação o fato de que a abordagem tradicional da produção de *software*, para a resolução de problemas particulares, não pode ser aplicada facilmente, em geral, nas áreas não-lineares complexos. Além disto, as principais limitações de um sistema, nesta área, é o seu poder de representação e manipulação de conhecimentos. Os sistemas híbridos inteligentes são relevantes quando domínios complexos e que apresentam problemas com componentes diferentes são considerados, os quais requerem diversos tipos de processamento (ou finalidades). Eles permitem o uso de todo conhecimento disponível, sejam eles conhecimentos básicos e empíricos (exemplo: cultivos já realizados), sejam conhecimentos teóricos, já tratados e estruturados (regras e outros conhecimentos simbólicos). Assim, a integração de duas técnicas complementares, permite que uma complete as deficiências da outra, de forma a obter um melhor desempenho. O uso de diferentes técnicas de aquisição e de representação de conhecimentos amplia a capacidade do sistema de adquirir novas informações, evitando que o sistema tenha uma visão parcial, limitada pelas imposições e restrições de um único método. A concepção de que as metodologias da inteligência computacional são complementares, mais que competitivas, gerou uma ampla diversidade de configurações de sistemas híbridos inteligentes. Vale citar, por exemplo, os sistemas neuro *fuzzy*, *fuzzy*-evolutivos, neuro-evolutivos e neuro *fuzzy*-evolutivos, com diversas aplicações em identificação e controle de processos, e outras áreas afins (SENG *et al.*, 1999; BONISSONE *et al.*, 1999; COELHO, 2001; NUCCI *et al.*, 2007).

Nesta dissertação foi utilizado um sistema híbrido inteligente (SHI), composto por um comitê de redes neurais artificiais (CRNAs) e um sistema *fuzzy* baseado em regras (SFBR). O motivo desta arquitetura proposta é capturar conhecimento do processo, tanto empírico, pelo CRNAs, quanto pelo raciocínio lógico com base fenomenológica, pelas regras *fuzzy*.

3. MATERIAIS E MÉTODOS

3.1 Materiais

Os principais reagentes utilizados foram: aminoácidos (Synth e Sigma), soro de queijo (Sigma), ácido fenil acético -AFA (Carlo Erba), sais (Synth-Mallinckrodt) e penicilina G acilase (Paraquímica do Brasil). Demais reagentes utilizados eram de grau analítico de diferentes marcas comerciais. As quantidades de reagentes utilizadas, e outros dados e informações de cada ensaio, estão especificadas no APÊNDICE A, desta dissertação.

3.1.1 Microorganismo

A linhagem de *Bacillus megaterium* ATCC 14945 foi cedida pela Fundação Tropical de Campina, SP e mantida em criotubos a -50°C . O inóculo foi preparado em erlenmeyers com meio cuja composição é mostrada na Tabela 1. Após a adição do microrganismo, o meio líquido de germinação foi incubado em câmara rotativa (New Brunswick Scientific) por 24 horas a 300 rpm e 30°C . O meio germinado representa 10% do volume total do meio de fermentação para a produção da enzima.

3.1.2 Principais equipamentos

Os principais equipamentos utilizados foram:

Autoclaves

Marca FABBE modelo 103, 2000W e 4000W.

As autoclaves foram utilizadas para a esterilização dos meios de cultura, soluções em geral e todo material a ser descartado utilizando vidrarias tipo erlenmeyers, biorreator, tubos de ensaio, etc.

Câmara rotativa

Com controle de temperatura e agitação (“*shaker*”).

As etapas de crescimento do microrganismo (pré-inóculo e inóculo) foram realizadas em frascos agitados (*erlenmeyers*) em câmara rotativa climatizada marca NEW BRUNSWICK Co., Inc., Edison, Nj, EUA.

Estufa de secagem e esterilização

Marca FANEM 320 – SE e 315/9.

Foi utilizada a estufa, com temperatura controlada em 60°C , para secagem de vidrarias, mangueiras e conexões utilizadas durante a fermentação. Para a secagem da sílica

utilizada no analisador de CO₂ utilizou-se uma estufa a 100°C.

Câmara de fluxo laminar

Marca VECO.

Nas etapas de manipulação do microrganismo, visando-se evitar contaminações no meio de germinação e produção, utilizou-se a câmara asséptica de fluxo laminar contendo bico de Bunsen e lâmpada germicida Ultra-Violeta (UV).

Biorreator de bancada

Os ensaios (APÊNDICE A), utilizados nesta dissertação, foram realizados em biorreatores de bancada distintos. Os ensaios identificados por 5, 6, 7, 8 e 9 foram realizados no biorreator de bancada marca *New Brunswick BIOFLO II-C* e validaram o módulo 2, controle da vazão de alimentação do meio suplementar. Os ensaios identificados por 1, 2, 3 e 4 foram realizados no biorreator de bancada marca Applikon, e validaram o módulo 3, da arquitetura que será proposta na seção 4, inferência de μ .

New Brunswick BIOFLO II-C

Biorreator de bancada tipo tanque agitado e aerado (STR) de 5 litros de volume útil, acoplado a um sistema de aquisição de dados e analisador dos gases de saída, dióxido de carbono e oxigênio.

Applikon

Biorreator de bancada tipo tanque agitado e aerado (STR) de 2 litros de volume útil, integrado com acessórios de medidas de temperatura, pH, oxigênio dissolvido, velocidade de agitação, vazão de ar e sistema de análise de gases de exaustão da SICK-MAIHAK da concentração de oxigênio e dióxido de carbono. Sistema de aquisição e controle implementado utilizando como *hardware* de aquisição de dados o equipamento *compact Field Point* (cFP-2020, National Instruments) que contém módulos para entrada e saída de sinais analógicos e digitais. A interface de comunicação com o usuário foi elaborada com a utilização do software LabView 7.1 (National Instruments).

Analisador de gases

Analisador da SICK-MAIHAK, Modelo S710, responsável pelas medidas das concentrações de CO₂ (sensor infravermelho) e O₂ (sensor paramagnético) nos gases efluentes. O equipamento envia os valores medidos, das concentrações dos gases, ao computador via comunicação serial RS232. Os dados são recebidos tratados e armazenados, em um banco de dados, pela interface programada em LabVIEW 7.1 (National Instruments).

Balanças analíticas

Marcas METTLER PB – 3002 e KERN 410.

PHmetro: Marca ORION modelo 420 A.

PHstato: Marca METROHM modelo 718 STAT Titrino.

Espectrofotômetro

Marca PHARMACIA BIOTECH modelo Ultropec 2000.

Centrífuga

Marca EPPENDORF modelo 5403.

Reator de vidro

Com camisa de água volume 50 ml.

Demais equipamentos e vidrarias comuns de laboratório (dessecadores, tubos de ensaio, provetas, balões volumétricos, erlenmeyers, béqueres, pipetas, cubetas, tampões de algodão).

Desenvolvimento de algoritmos com o *software* **MATLAB® 6.5 (Mathworks)**.

3.2 Métodos

3.2.1 Fermentação principal

As fermentações foram realizadas em biorreator do tipo agitado e aerado acoplado a um sistema de aquisição de dados. Após a obtenção do inóculo e do meio de produção, estes foram transferidos para o biorreator por gravidade. Na Figura 3.1 está apresentado um esquema do preparo do inóculo e pré-inóculo partindo de criotubos até a fermentação final em biorreator.

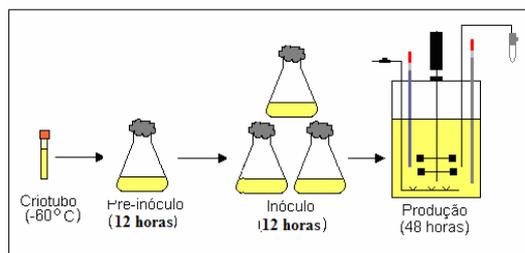


Figura 3.1 Esquema da preparação do experimento desde a etapa de preparação dos criotubos até o experimento em biorreator.

3.2.2 Sistema de aquisição de dados

Um esquema do aparato experimental com o sistema de aquisição de dados, utilizado nos ensaios cinco, seis, sete, oito e nove, e responsável pela validação do módulo 2,

Controle da Vazão de alimentação de meio suplementar, é ilustrado na Figura 3.2.

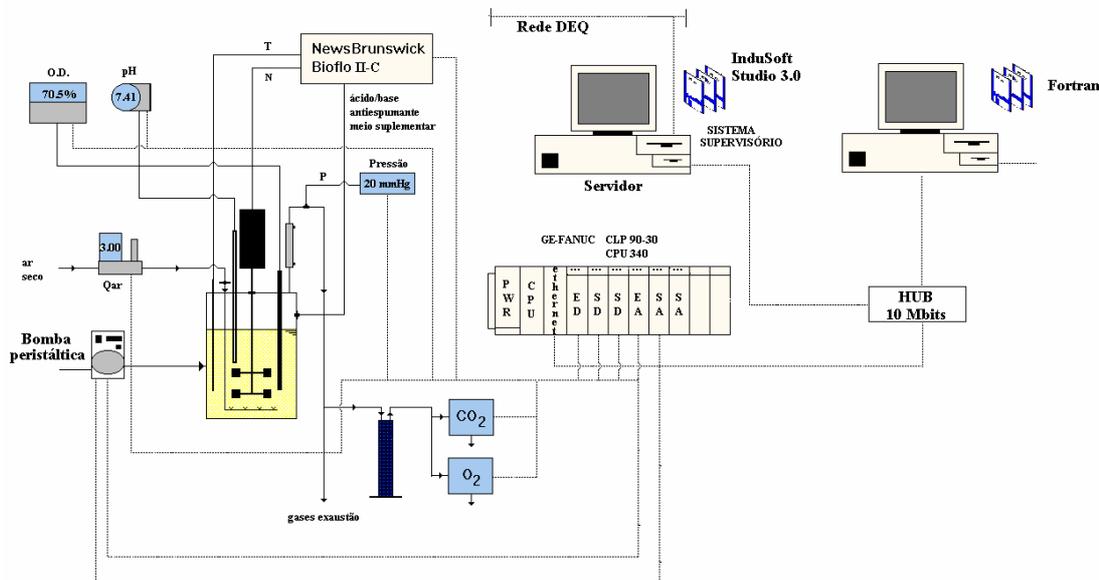


Figura 3.2 Esquema do aparato experimental com sistema de aquisição de dados (CRUZ, 2000).

O sistema de aquisição de dados é formado pelos seguintes equipamentos (CRUZ, 2000):

- Eletrodo de pH (tipo 465-90, Mettler Toledo) conectado a Medidor e Transmissor de pH (modelo TH-41, Digimed);
- Eletrodo de oxigênio dissolvido (tipo 12/320T,, Metler Toledo) conectado a Medidor e Transmissor de Oxigênio Dissolvido (O₂ 4500, Metler Toledo);
- Bombas peristálticas (BVP, Ismatec) para adição de meio suplementar, ácido e/ou base;
- Fluxômetro de massa (modelo 33116-20, 0 a 5,0 NLPM, Cole Parmer®) para monitorização da vazão de ar no sistema;
- Transdutor de pressão (modelo 560, T&S Equipamentos) para monitorizar a sobre pressão no interior do biorreator;
- Sensor de temperatura (Pt – 100), inserido em um poço dentro do reator;
- Analisador de oxigênio e de dióxido de carbono (Rosemount Analytical) acoplados para análise dos gases de saída do fermentador. Os analisadores são calibrados com gases padrões (N₂ 99,99% e CO₂ 5,0% / 95%).

Descrição do sistema de controle de pH

O pH é controlado na faixa de 2,00 a 12,00 ($\pm 0,01$). Sensor amperométrico (pHE - 200) é imerso no meio da reação para determinar o valor do pH, sendo o sinal deste sensor

enviado ao transmissor. Para controle de pH, emprega-se um controlador por faixas (com limites inferior e superior) (pHC - 200) o qual aciona duas bombas peristálticas para adição de ácido ou base, dependendo do pH do meio e do “set-point” determinado pela função lógica programável (CRUZ, 2000).

Descrição do sistema de controle de oxigênio dissolvido

O oxigênio dissolvido é controlado na faixa de 5 a 95% (± 1) de saturação. É medido através de um eletrodo imerso no interior do biorreator e o controle é realizado através de controlador PID, que atua sobre a velocidade de agitação, mantendo-se constante a vazão de alimentação de ar.

Descrição do sistema de controle de temperatura

A temperatura do meio de cultura pode ser selecionada na faixa de 20 a 60°C ($\pm 0,1^\circ\text{C}$) e controlada por um microprocessador baseado em controlador PID. A temperatura média é medida por um sensor RTD (*Resistance Temperature Sensor*), imerso no biorreator. Quando a temperatura do sistema está abaixo do “set-point”, o controlador atua abrindo uma válvula solenóide e ligando o aquecedor. O sistema dispõe ainda de um dispositivo de proteção que impede que o aquecedor entre em funcionamento na ausência do fluido de refrigeração, ou seja, a falta de água.

3.2.3 Cálculo do quociente respiratório (RQ)

As fermentações aeróbicas são processos que ocorrem em presença de oxigênio pelo fato dos microrganismos necessitarem de oxigênio para se desenvolver. O oxigênio molecular é utilizado pelos microrganismos aeróbios como um oxidante para as suas necessidades energéticas. Este processo denominado respiração pode ocorrer através de diferentes vias metabólicas, dependendo do tipo de substrato e do tipo de microrganismo.

A medição da concentração de oxigênio dissolvido é normalmente efetuada por eletrodos químicos cuja calibração decorre por imersão de uma sonda em meio líquido supostamente em equilíbrio com a fase gasosa de concentração conhecida, por exemplo, o ar. A velocidade de consumo de oxigênio (do inglês *Oxygen Uptake Rate*, OUR) depende da atividade metabólica do microrganismo (FERREIRA, 1995).

Apesar da concentração de oxigênio na fase líquida ser considerada a principal variável que controla o ambiente de uma fermentação, há situações evidentes de que o dióxido de carbono desempenha um importante papel em várias fermentações, sejam em processos

anaeróbio ou aeróbio, devido ao seu papel no metabolismo celular.

O dióxido de carbono é bem mais solúvel em água que o oxigênio (aproximadamente 25 vezes mais solúvel). O dióxido de carbono produzido no interior da célula pelo seu metabolismo passa para o meio na forma de dióxido de carbono dissolvido uma vez que a membrana celular é relativamente impermeável às substâncias iônicas.

A velocidade de produção de dióxido de carbono (do inglês *Carbon Evolution Rate*, CER) depende da velocidade de crescimento do microrganismo. O método do balanço gasoso amplamente divulgado na literatura permite a determinação de um parâmetro muito importante que é a velocidade global de consumo de oxigênio ($Q_{O_2} \cdot X$) durante o cultivo envolvendo microrganismos aeróbios e aeróbios facultativos. A técnica baseia-se em um balanço molar de oxigênio realizado no sistema, utilizando as correntes gasosas de entrada e saída do biorreator (FERREIRA, 1995).

O biorreator foi tomado como volume de controle para a realização do balanço gasoso, como mostra a Figura 3.3. As frações molares dos gases de saída do biorreator ($Y_{O_2}^S$ e $Y_{CO_2}^S$) são quantificadas através de analisadores de gás (método paramagnético para o oxigênio e infravermelho para o dióxido de carbono). A partir da vazão molar de entrada de ar (medida em fluxômetro de massa) e com frações molares dos gases de saída do biorreator é possível determinar a vazão molar do gás de saída.

Considera-se que o ar de entrada é seco, sendo constituído de 79% de N_2 e 21% de O_2 e que o gás de saída do reator é constituído de N_2 , O_2 e CO_2 , este último produto do metabolismo celular. Através deste balanço gasoso é possível determinar a fração molar de nitrogênio na corrente de saída a partir das frações de oxigênio e gás carbônico, medidas experimentalmente. A fração é dada pela Equação 3.1:

$$Y_{N_2}^S = 1 - Y_{O_2}^S - Y_{CO_2}^S \quad (3.1)$$

É feito um balanço molar para o nitrogênio que é um gás inerte, indicando que a taxa mássica de nitrogênio na corrente de entrada do biorreator é a mesma da corrente de

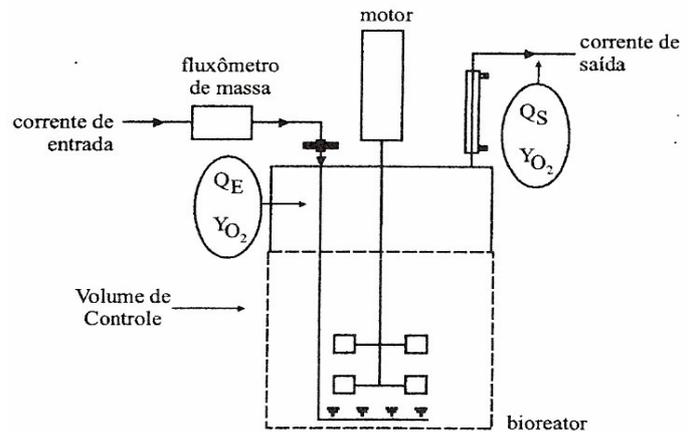


Figura 3.3 Esquema do aparato experimental e variáveis, para cálculo do coeficiente respiratório, em um balanço material (OLIVEIRA, 2006).

saída. A fração molar de N₂ na corrente é considerada como sendo de 0,79.

$$Q_s = \frac{Y_{N_2}^e Q_e}{1 - Y_{O_2}^s - Y_{CO_2}^s} \quad (3.2)$$

A quantidade total de CO₂ gerada ao longo do cultivo pode ser calculada pela equação 3.3.

$$nCO_2 = \int_{t=0}^{t=t_f} Q_s Y_{CO_2}^s dt \quad (3.3)$$

A vazão molar de saída (Q_s) pode ser calculada pela Equação 3.2 e a Equação 3.3 torna-se:

$$nCO_2 = \int_{t=0}^{t=t_f} \left(\frac{Q_e Y_{N_2}^e}{1 - Y_{O_2}^s - Y_{CO_2}^s} \right) Y_{CO_2}^s dt \quad (3.4)$$

A quantidade total de O₂ consumida no cultivo pode ser determinada também de forma análoga ao que foi feito para o CO₂.

$$nO_2 = \int_{t=0}^{t=t_f} \left(Q_e Y_{O_2}^e - \left(\frac{Q_e Y_{N_2}^e}{1 - Y_{O_2}^s - Y_{CO_2}^s} \right) Y_{O_2}^s \right) dt \quad (3.5)$$

Se forem conhecidos os valores de nCO₂ e nO₂ é possível calcular o quociente respiratório (RQ) definido por:

$$RQ_{\text{global}} = \frac{nCO_2}{nO_2} \quad (3.6)$$

Este será o RQ_{global} no cultivo como um todo, desde o tempo inicial t₀ até o final do cultivo t_f. No entanto para efeito de controle, é melhor que seja utilizado o RQ_{pontual}, isto é, em pontos discretos do processo.

Deste modo o RQ_{pontual} pode ser calculado pela Equação 3.7, no entanto, $\dot{n}CO_2$ será a vazão molar de dióxido de carbono medida naquele instante, e $\dot{n}O_2$ será a vazão molar de oxigênio no mesmo instante.

$$RQ_{\text{pontual}} = \frac{\dot{n}CO_2}{\dot{n}O_2} \quad (3.7)$$

Nesta dissertação, cada RQ_{pontual} foi calculado a cada 1 minuto, e os dados, de RQ dos ensaios selecionados, foram armazenados na base de dados de Ensaio Selecionados (BD SE), descrito no APÊNDICE A desta dissertação.

3.2.4 Metodologia do comitê de redes neurais artificiais (CRNAs)

O comitê de redes neurais artificiais (RNAs), denominado nesta dissertação de comitê ou CRNAs, foi composto por dois componentes principais: uma base de dados de redes neurais artificiais, ou BD RNAs e um mecanismo de ponderação (MP).

3.2.4.1 BD RNAs especialistas

No BD RNAs foi armazenados as redes neurais artificiais “especialistas”. Cada rede neural foi treinada com dados de um ensaio (cultivo), realizado com *Bacillus megaterium* para produção de PGA. Nestas redes neurais utilizou-se a arquitetura MLP (*multi layer perceptron*), do tipo *feedforward*, com topologia de oito neurônios na camada oculta e um neurônio na camada de saída. Esta topologia foi definida após vários testes, com diferentes números de neurônios na camada oculta, e diferentes funções de transferência. Menor erro na saída, menor número de neurônios e menor tempo foram as variáveis utilizadas para definição desta topologia. A Figura 3.4 mostra o esquema de treinamento de cada rede “especialista” do comitê. O vetor de entrada (x_i), onde i é um instante do ensaio (com início no sexto minuto), de cada rede é composto das seguintes variáveis: o tempo (valor atual, com intervalos de amostragem de um minuto), Y_{CO_2} e velocidade de agitação (valor atual e cinco valores atrasados, e por isto, o início no sexto minuto do cultivo). Os dados medidos de y_{CO_2} e velocidade de agitação foram suavizados, antes de serem carregados no vetor de entrada (x_i), pelo filtro “alisa pós”, desenvolvido para este trabalho, e descrito no item 4.4.1, da seção 4.

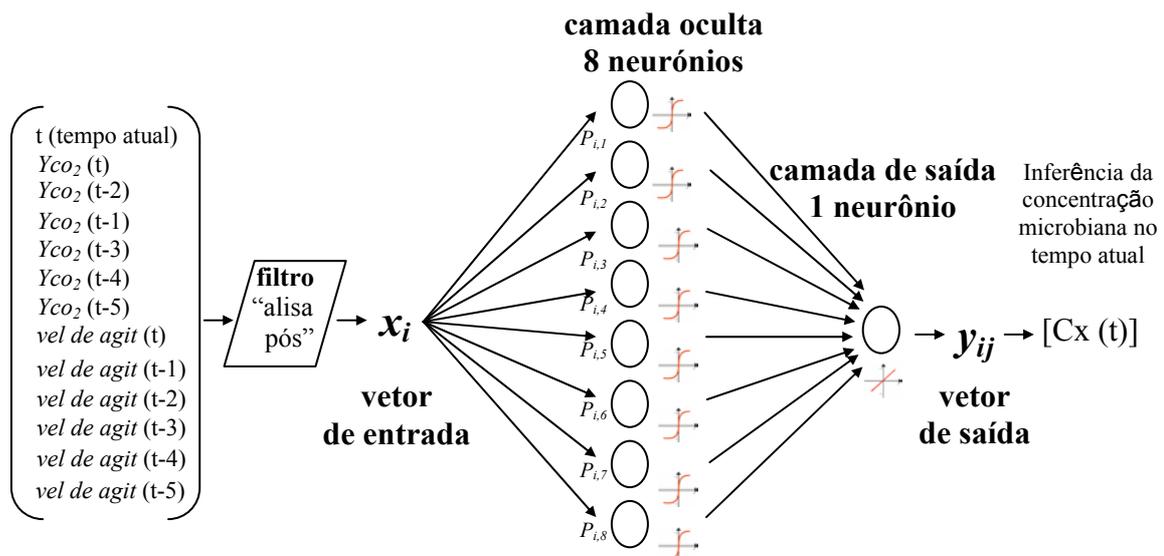


Figura 3.4 Esquema do treinamento de cada rede neural que pertence ao comitê (CRNAs).

As funções de transferência utilizadas foram a tangente sigmoideal (na camada oculta) e a linear (nos neurônios de saída). O vetor de saída da rede (y_{ij}), onde j identifica o

ensaio, foi a variável concentração microbiana (C_x) no tempo presente. Para treinamento das redes neurais foi utilizado o algoritmo de Levenberg-Marquardt no MATLAB® 6.5 (*Mathworks*).

A parada do treinamento de cada rede especialista foi estabelecida como quinhentas (500) épocas, sem teste de *overfitting*, uma vez que se trata de uma rede *especialista*. Na concepção utilizada neste trabalho, o mecanismo de ponderação do comitê será o responsável pela capacidade de generalização do *softsensor*. Em outras palavras, a RNA especializada em representar um ensaio específico deve fazê-lo com precisão (e para isso 500 épocas foram sempre suficientes, levando em conta os erros experimentais). Mas não se exige que tenha a capacidade de representar outro ensaio da base de dados.

3.2.4.2 Mecanismo de ponderação

A Figura 3.5 mostra os dois componentes do comitê (BD RNAs e MP), e sua estrutura dinâmica. O mecanismo de ponderação (MP) utilizou-se da diferença entre a fração molar de CO_2 lida do ensaio, on line, no instante i , y_{CO_2i} , e a fração molar de CO_2 da entrada de cada rede do comitê, y_{CO_2j} , que integra a base de dados. A mesma concepção é utilizada para as velocidades de agitação, uma vez que existia controle automático do oxigênio dissolvido e, assim, a velocidade do impelidor é a variável relacionada com a razão de velocidades de transporte/consumo de oxigênio pelo microrganismo.

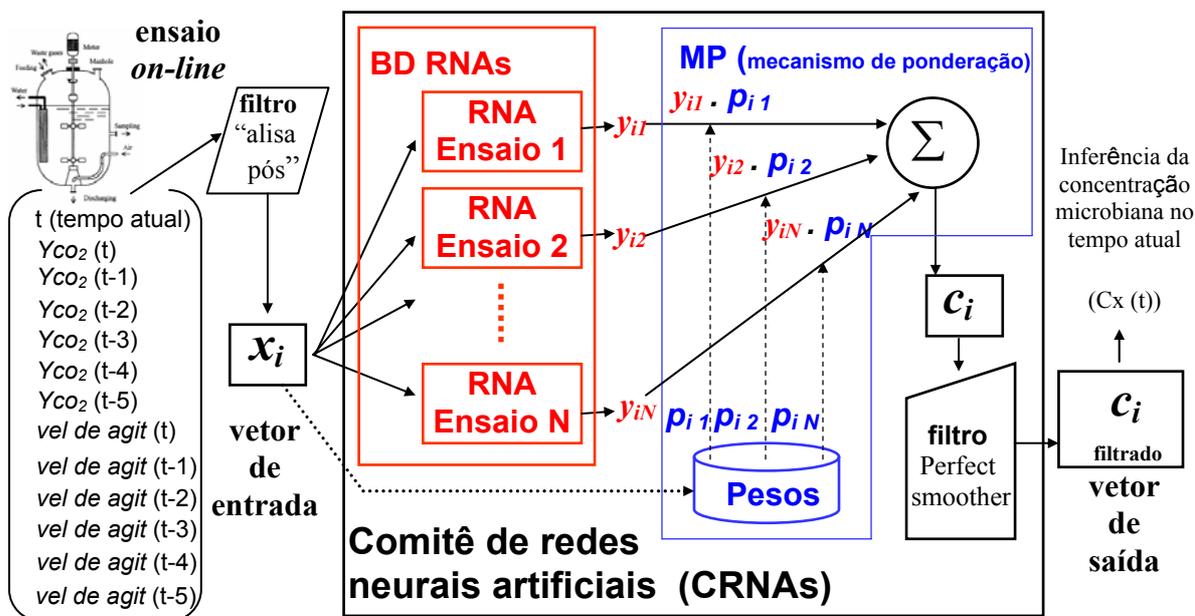


Figura 3.5 Estrutura dinâmica do comitê, com seus dois componentes (BD RNAs e MP), utilizado na inferência da concentração microbiana (C_x).

Os dados medidos de y_{CO_2} e velocidade de agitação foram suavizados, antes de entrarem no comitê, pelo filtro “alisa pós”, desenvolvido para este trabalho, e descrito no item 4.4.1, da seção 4.

Para um vetor de entrada x_i , qualquer do ensaio *on-line*, cada rede neural (RNA) individual (j) do comitê fornece uma resposta y_j (concentração microbiana). Essas respostas são a seguir ponderadas pelos pesos p_{ij} (peso da rede j do comitê, no instante i).

Assim, a resposta final do comitê, c_i , é obtida pela Equação 3.8.

$$c_i = \sum_{j=1}^N y_j \cdot p_{ij} \quad (3.8)$$

Note-se que os pesos do comitê, p_{ij} , se alteram com o tempo, o que é característico do comitê com estrutura dinâmica.

Para o cálculo do fator de ponderação, p_{ij} , veja Equação 3.9, foi utilizada a diferença em y_{CO_2} e entre velocidades de agitação, d_{Agit} . Para ponderar os desvios de y_{CO_2} e velocidade de agitação, foram utilizadas as normas 1 da distância em relação à média das duas medidas, s_{CO_2} e s_{Agit} , calculadas a cada instante de tempo i , utilizando a medida do experimento *on-line* e os valores empregados para treinamento, para todas as N redes. A média das diferenças de todas as redes, \bar{d}_i , era calculada a cada instante i .

$$p_{ij} = \frac{\frac{s_{CO_2i}}{d_{CO_2ij}} + \frac{s_{Agiti}}{d_{Agitij}}}{\sum_{j=1}^N \left(\frac{s_{CO_2i}}{d_{CO_2ij}} + \frac{s_{Agiti}}{d_{Agitij}} \right)} \quad (3.9)$$

$$\text{onde } d_{CO_2ij} = y_{CO_2i} - y_{CO_2j} \quad ; \quad d_{Agitij} = Agit_i - Agit_j \quad ;$$

$$s_{(\bullet)i} = \frac{1}{N-1} \sum_{j=1}^N |d_{(\bullet)ij} - \bar{d}_{(\bullet)i}| \quad \text{e} \quad \bar{d}_{(\bullet)i} = \frac{\sum_{j=1}^N (d_{(\bullet)ij})}{N}.$$

Para suavizar os valores de saída do comitê, foi empregado um filtro, o algoritmo denominado “alisador perfeito” (*perfect smoother*), desenvolvido por EILERS, 2003. A estratégia utilizada foi a seguinte: os valores de saída do comitê (c_i) eram armazenados em um vetor auxiliar. A cada instante de tempo i (no caso um minuto), o vetor auxiliar aumentava de um novo elemento, ou seja, com uma nova saída do comitê. Este vetor auxiliar, então, era submetido ao algoritmo *perfect smoother*, que retornava um vetor alisado. Apenas o último valor, do vetor alisado, foi armazenado no vetor de saída final do comitê (c_i filtrado).

3.2.5 Metodologia do sistema *fuzzy* baseado em regras (SFBR)

O sistema *fuzzy* baseado em regras, denominado nesta dissertação de sistema *fuzzy* ou SFBR, foi composto por dois componentes principais: a base de conhecimento *fuzzy* e o mecanismo de inferência. O mecanismo de inferência é responsável pelo processamento das regras, o qual é realizado por um método de raciocínio *fuzzy* (PEDRYCZ E GOMIDE, 1998).

3.2.5.1 Base de conhecimento *fuzzy*

A base de conhecimento *fuzzy*, foi composta pela base de dados dos conjuntos *fuzzy*, denominado de BD CFs, e contém as definições dos conjuntos *fuzzy* relacionados aos termos lingüísticos usados nas regras *fuzzy*, e pela base de regras, denominado de BD RFs, e armazena o conjunto de regras *fuzzy* que modelam conhecimento sobre um determinado problema.

Base de dados dos conjuntos *fuzzy* (BD CFs)

Foram utilizadas duas variáveis de entrada, a fração molar de CO_2 (y_{CO_2}) e quociente respiratório (RQ), que é um parâmetro importante em cultivos aeróbios, definido como a razão da velocidade de formação de gás carbônico pela velocidade de consumo de oxigênio. A variável de saída é o valor da correção (negativa ou positiva) a ser feita na predição feita pelo comitê (CRNAs), da concentração microbiana (Cx) do ensaio.

1ª Partição de entrada - a fração molar de CO_2 (y_{CO_2})

Esta partição mostrada na Figura 3.6, foi construída tendo como base o menor e maior valor de y_{CO_2} , encontrado a cada instante de tempo i , na base de dados de ensaios selecionados (BD ES), descrito no APÊNDICE A desta dissertação.

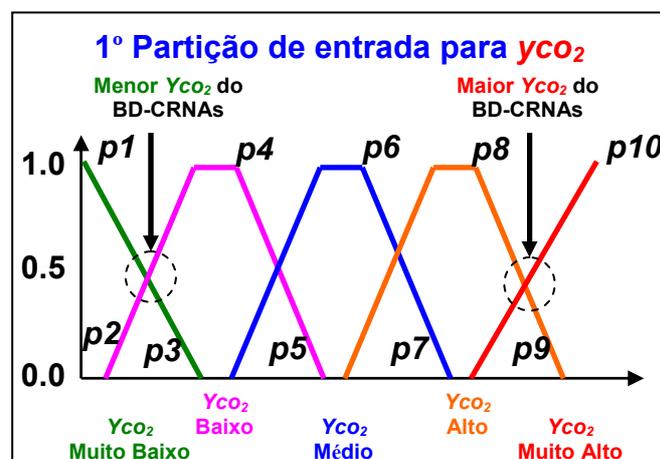


Figura 3.6 Formato da 1ª partição de entrada *fuzzy* para a fração molar de CO_2 (y_{CO_2}).

Os valores $p1, p2, p3, p4, p5, p6, p7, p8, p9$ e $p10$ descrevem os cinco conjuntos *fuzzy* (no formato trapezoidal), cujos rótulos lingüísticos são respectivamente Muito Baixo ($p1$ e $p3$), Baixo ($p2, p3, p4$ e $p5$), Médio ($p4, p5, p6$ e $p7$), Alto ($p6, p7, p8$ e $p9$), Muito Alto ($p8$ e $p10$). Estes valores se alteram com o tempo, caracterizando o conjunto de dinâmico. Os valores de $p1$ a $p10$ são encontrados da seguinte forma: para cada instante de tempo i foi pesquisado o menor y_{CO2} e o maior y_{CO2} , e então foram realizados os seguintes cálculos:

$$a = (\text{maior } y_{CO2} - \text{menor } y_{CO2}) / 6;$$

$$b = a / 2;$$

$$p2 = \text{menor } y_{CO2} - b;$$

$$p1 = p2 - 0,2;$$

$$p3 = \text{menor } y_{CO2} + b;$$

$$pk+1 = pk + c; k = 3..8;$$

$$p10 = p9 + 0,2.$$

2ª Partição de entrada – o quociente respiratório (RQ)

Esta partição mostrada na Figura 3.7, foi construída tendo como base o menor e maior valor do quociente respiratório (RQ), encontrado a cada instante de tempo i , na base de dados de ensaios selecionados (BD ES), descrito no APÊNDICE A desta dissertação.

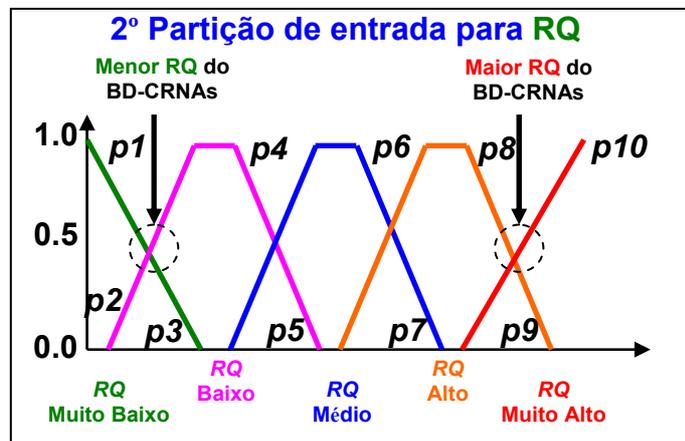


Figura 3.7 Formato da 2ª partição de entrada *fuzzy* para o quociente respiratório (RQ).

A descrição, os cálculos dos valores de $p1$ a $p10$, os conjuntos *fuzzy*, os rótulos lingüísticos e característica de conjunto dinâmico são definidos, de forma análoga, ao 1º conjunto de entrada.

Partição de saída – correção para o valor predito pelo comitê (CRNAs)

Esta partição é composta de conjuntos fixos, caracterizando a de estática, veja a Figura 3.8, e válido para todas as inferências. Os rótulos lingüísticos e os valores, que descrevem os cinco conjuntos *fuzzy* dessa partição, de formato triangular, são: Corrige Muito Baixo (-0,005 -0,001), Corrige Baixo (-0,003 zero), Não Corrige (-0,001 0,002), Corrige Alto (0 0,004) e Corrige Muito Alto (0,002 0,006). Os valores que descrevem os conjuntos *fuzzy* foram encontrados através de tentativa e erro. O valor inferido por este conjunto de saída (positivo ou negativo) será somado ao valor da concentração microbiana (Cx) inferido pelo comitê (CRNAs).

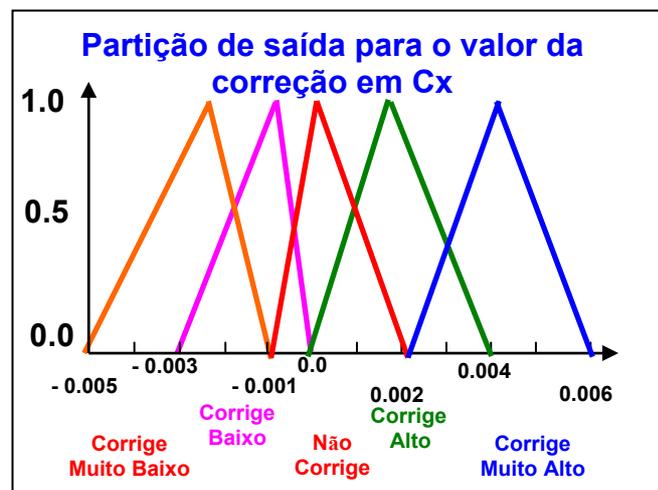


Figura 3.8 Formato da partição de saída dos conjuntos *fuzzy* para correção do valor predito de crescimento microbiano (Cx), pelo comitê (CRNAs).

Base de dados de regras *fuzzy* (BD RFs)

Os conhecimentos extraídos, para correção de Cx inferido pelo comitê (CRNAs), foram sintetizados em cinco regras:

- R1: Se y_{co2} é MUITO ALTO
então **Corrige Cx** para MUITO ALTO.
- R2: Se y_{co2} é MUITO ALTO
E **RQ** é ALTO
então **NÃO Corrige Cx**.
- R3: Se y_{co2} é MUITO BAIXO
então **Corrige Cx** para MUITO BAIXO.
- R4: Se y_{co2} é MUITO BAIXO
E **RQ** é MUITO BAIXO
então **Corrige Cx** para MUITO ALTO.
- R5: Se y_{co2} é MUITO BAIXO
E **RQ** é BAIXO
então **NÃO Corrige Cx**.

O raciocínio usado, nas regras R1 e R3, que só utiliza o valor da y_{CO_2} , são de que a produção de CO_2 reflete o consumo de aminoácidos para crescimento celular. Já as regras R2, R4 e R5, que utilizam também RQ, suavizam esta correção. O SFBR tem o seu caso mais extremo na contradição entre as regras R3 e R4, onde a correção pode ser muito próxima de zero. O raciocínio usado na regra R4 (e de forma similar nas regras R2 e R5) é de que se y_{CO_2} é muito baixo e RQ também é muito baixo então o carbono proveniente do substrato (aminoácidos) estaria sendo incorporado como massa celular, evidenciando assim uma região de alta velocidade de crescimento celular.

3.2.5.2 Mecanismo de inferência

Da mesma forma do comitê (CRNAs), os dados do vetor de entrada, y_{CO_2} e RQ foram suavizados, antes de entrarem no mecanismo de inferência, no SFBR, pelo filtro “alisa pós”, desenvolvido para este trabalho, e descrito no item 4.4.1, da seção 4. O mecanismo de inferência do SFBR utilizou o modelo de Mamdani. Uma simulação verdadeira do mecanismo de inferência, em instante de tempo i , do ensaio 2, é mostrada na Figura 3.9. Os retângulos vermelhos representam cada regra *fuzzy*, descritas na BD RFs. E seus respectivos conjuntos *fuzzy*, descrito na BD CFs. O retângulo azul representa a etapa de agregação. Neste instante ocorre o seu caso mais extremo. A regra R3 entra em contradição com a regra R4, suavizando a correção, onde fica muito próxima de zero.

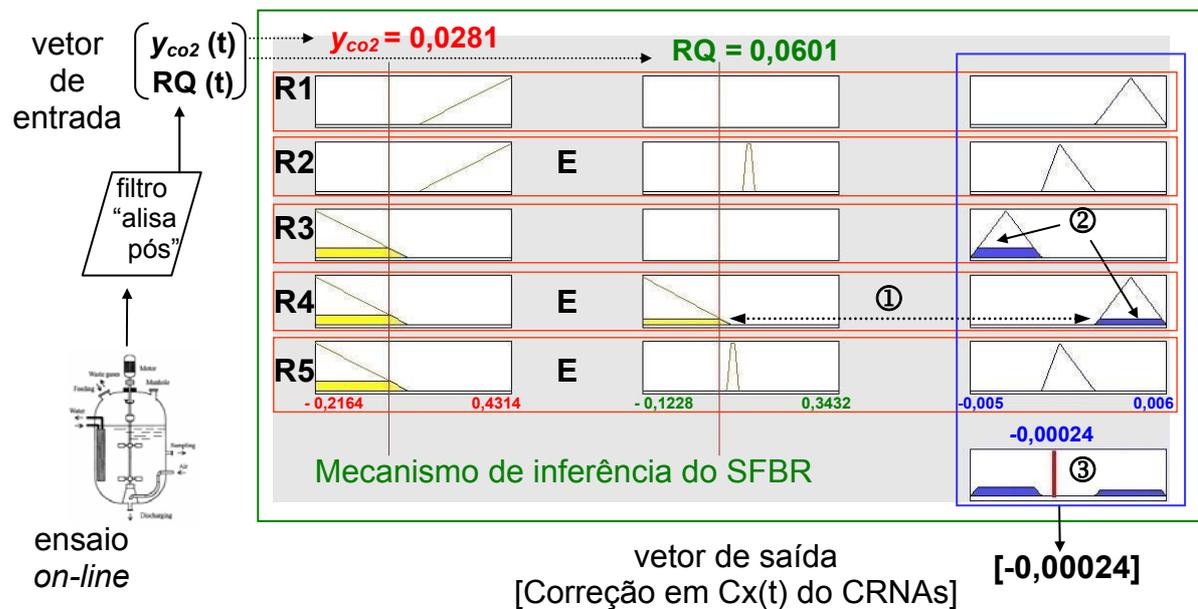


Figura 3.9 Resultado da simulação, feita no instante i , do ensaio2, mostrando o mecanismo de inferência do sistema *fuzzy* (SFBR).

A conjunção (E), utilizada nas regras R2, R4 e R5, utilizou o método min (mínimo). Um exemplo do resultado deste método pode ser visto em ①, na Figura 3.9, onde a

linha pontilhada mostra a escolha da partição de RQ, mínimo, para definir o tamanho da partição do conjunto de saída. A etapa de agregação utilizou o método max (máximo). Um exemplo deste método pode ser visto em ②, na Figura 3.9, as setas mostram as duas partições dos conjuntos *fuzzy* de saída, definidos pelas regras R3 e R4, e depois agregadas em ③, na Figura 3.9. O método de defuzificação escolhido foi o centróide, veja em ③, na Figura 3.9, traço vermelho indica o local do centróide.

3.2.6 Método de validação cruzada

Para comparar os resultados das metodologias foi utilizado o método de validação cruzada. O esquema do método utilizado nesta dissertação é mostrado na Figura 3.10.

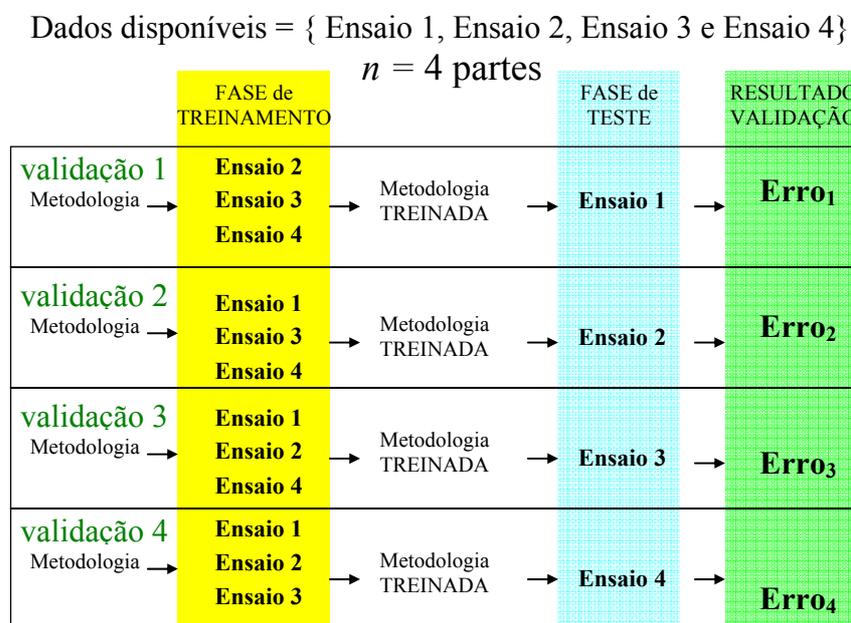


Figura 3.10 Esquema do método de validação cruzada utilizada nesta dissertação.

O método consiste em dividir todos os dados disponíveis do processo, em n partes. A partir destas n partes, dois conjuntos são formados. O primeiro conjunto, com $n-1$ partes, o de maior tamanho, representado de amarelo na Figura, tem a função de treinar a metodologia. O segundo conjunto, composto de uma parte que não foi utilizada no treinamento, representada de azul na Figura, tem a função de testar a metodologia treinada. O resultado da validação é a medida do erro, diferença entre a concentração microbiana (C_x) experimental, medida *off-line* do ensaio, e a concentração microbiana (C_x) inferida pela metodologia na fase de teste. Todos esses procedimentos anteriores são repetidos, para todas as combinações, com as n partes (NELLES, 2001).

3.2.7 Metodologia do sistema híbrido inteligente (SHI)

O sistema híbrido inteligente, denominado nesta dissertação de SHI, foi composto por duas metodologias: o comitê de redes neurais artificiais (CRNAs) e o sistema *fuzzy* baseado em regras (SFBR).

A escolha desta arquitetura inicia-se com o trabalho feito por FERREIRA *et al*, 2007, que constrói o comitê (CRNAs) tendo como objetivo principal a predição da concentração microbiana (C_x), a partir do conhecimento adquirido em ensaios passados, de modo empírico.

Os resultados, obtidos pela validação cruzada, foram de boa qualidade, principalmente até o final da fase exponencial. Na fase estacionária, apesar das variações, o valor inferido da concentração celular, acompanhou a tendência.

Os erros maiores na inferência de C_x , pelo comitê (CRNAs), foram encontrados nas validações 1 e 2, mostradas na Figura 3.10, onde os ensaios 1 e 2 foram o conjunto de teste, respectivamente. Foi observado, nesses ensaios, o fato do valor atual de y_{CO_2} , representado nos gráficos pela cor azul na Figura 3.11, extrapolar os valores dos ensaios treinados pelas redes neurais especialistas (RNAs especialistas) do comitê (CRNAs), representados nos gráficos pela cor vermelha. No caso do ensaio 1, Figura 3.11a, extrapola-se para maior, e no caso do ensaio 2, Figura 3.11b, extrapola-se para menor.

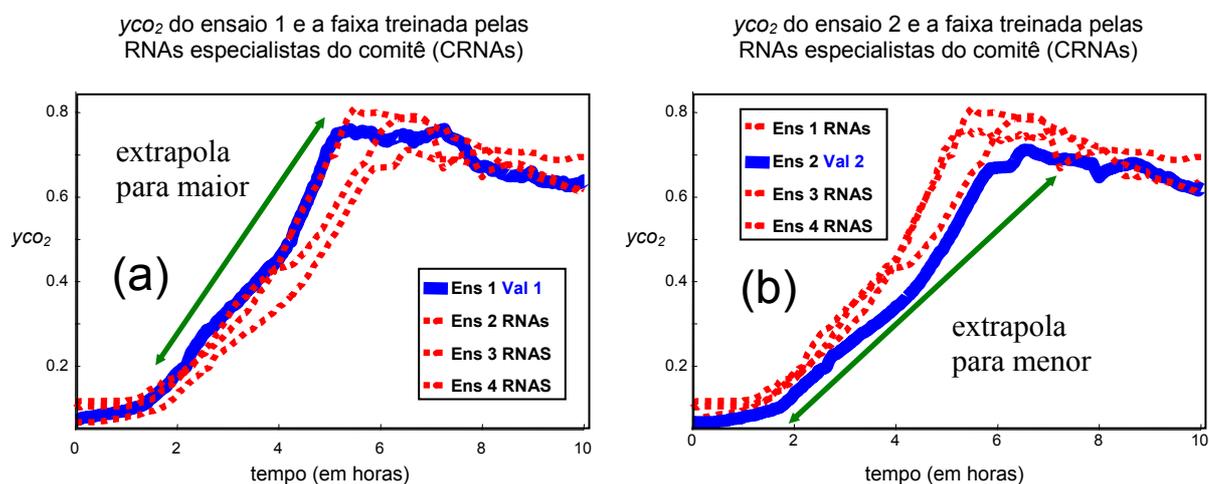


Figura 3.11 Gráficos de y_{CO_2} versus tempo, para ensaio validado, cor azul, e para os ensaios treinados pelas RNAs especialistas do comitê (CRNAs), cor vermelha: (a) do ensaio um e (b) do ensaio dois. A seta de cor verde, nos gráficos, indica regiões de extrapolação.

Ao encontrar essas regiões, o comitê (CRNAs) aumentou o erro em relação à predição correta. Observou-se, também, que a resposta do comitê tinha uma tendência final correta. As regiões de extrapolação foram responsáveis pelos desvios. Note-se que o comitê (CRNAs) não foi treinado nessas regiões, e esperava-se dele certo grau de extrapolação, o que

não ocorreu.

Para tratar os erros que causavam desvios na inferência do crescimento celular (Cx) pelo comitê (CRNAs), nas regiões de extrapolação, foi escolhido o raciocínio lógico, encontrado nas regras do sistema *fuzzy* (SFBR).

Assim, o sistema híbrido inteligente (SHI) foi escolhido para ser utilizado como *softsensor*, para inferir concentração microbiana (Cx), captando tanto conhecimento empírico, pelo comitê (CRNAs), quanto conhecimento fenomenológico (baseado no fenômeno), pelo sistema *fuzzy* (SFBR).

4. ARQUITETURA DO SISTEMA DE SUPERVISÃO E CONTROLE PROPOSTO

4.1 Introdução

Uma abordagem tecnológica foi desenvolvida nesta dissertação de mestrado, utilizando metodologias da Inteligência Computacional (IC), de forma simples e híbrida (duas ou mais metodologias utilizadas ao mesmo tempo). Elas foram usadas para realizar a modelagem do processo de produção de PGA, por microrganismo selvagem *Bacillus megaterium*. A implementação de estratégias (por vezes automáticas), de monitoração e controle da operação do biorreator, com o intuito de se assegurar de modo estável e reprodutível, uma produtividade máxima com um mínimo de custos, foi o objetivo principal deste trabalho.

Para melhor entendimento, do momento da aplicação da abordagem acima, temos que enfatizar que a abordagem tecnológica é normalmente utilizada após a concepção e otimização do processo, sob o ponto de vista das abordagens biológica e de engenharia. A abordagem biológica ou bioquímica procura melhorar o desempenho do processo e dos próprios catalisadores pela seleção de microrganismos e de meios de cultura apropriados ou por manipulações genéticas (mutagênese induzida, recombinação genética baseada em fusão de protoplastos,...) que são objetivos da Engenharia Genética e da Microbiologia. A abordagem de engenharia se concentra no desenvolvimento de melhores tecnologias e condições operacionais (regime de funcionamento, novas concepções de reatores, etc.) que são objetivos da Engenharia Biológica e da Engenharia Bioquímica. O que se denomina, nesta dissertação, abordagem tecnológica, é um subconjunto da de engenharia, permitindo racionalizar a busca por processos industriais controlados (reprodutíveis) e otimizados.

Sabe-se que uma abordagem tecnológica para fins de controle e monitoramento deveria ser aplicada somente depois de dominadas as variáveis operacionais, o que não foi possível pela exigüidade de tempo – ensaios foram realizados em paralelo ao desenvolvimento dos algoritmos desta dissertação. Ainda assim, mesmo trabalhando com ensaios onde as condições operacionais mudavam, os resultados obtidos apontam que o projeto e suas ferramentas têm consistência.

4.2 Descrição do controlador dinâmico e adaptativo

Controle adaptativo e otimização de processos são usualmente baseados simultaneamente em identificação de modelo e leis de controle que, para serem adaptativos,

requerem atualização *on-line* dos parâmetros do modelo. Nesta dissertação foi proposto um controlador, dinâmico e adaptativo, para monitoração e controle do biorreator, operando em batelada alimentada (“fed-batch”), no processo de produção de PGA, por microrganismo selvagem *Bacillus megaterium*. O esquema geral do funcionamento desse controlador, composto de quatro módulos e uma *Base de Conhecimento*, é mostrado na Figura 4.1. É importante ressaltar, que esta dissertação, focou a implementação dos módulos de 1 a 3, deixando o módulo 4, como proposta para continuidade desta linha de pesquisa.

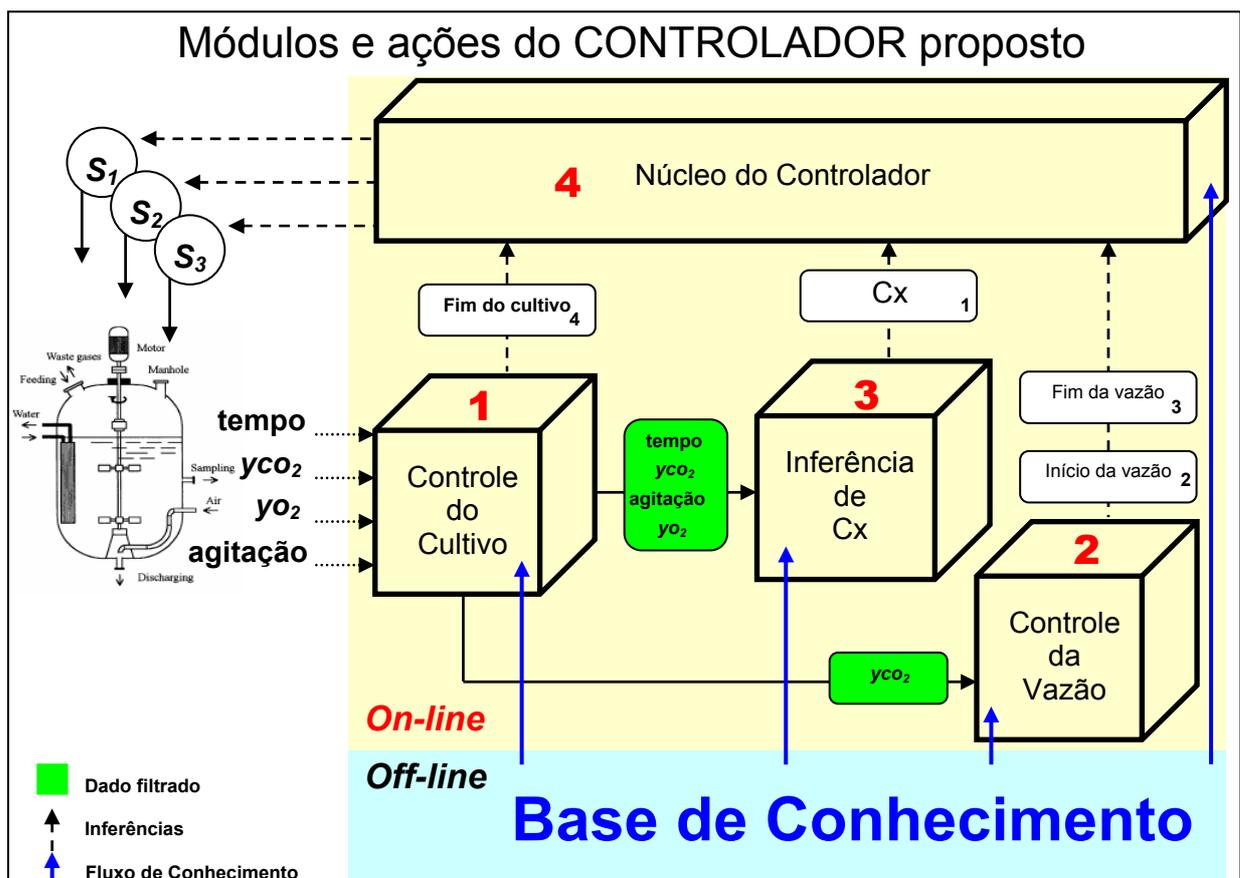


Figura 4.1 Esquema geral do funcionamento do controlador proposto.

Cada módulo foi responsável pelo controle de procedimentos distintos, durante o cultivo. O motivo de dividir o controlador em módulos foi a idéia de que uma tarefa complexa pode ser resolvida mais eficientemente se for dividida em tarefas mais simples, tendo suas respostas combinadas posteriormente. Além disso, esta estrutura pode beneficiar-se de vantagens, oferecidas no nível de processamento, pelas novas tecnologias de *software* e *hardware*. Estas novas tecnologias permitem o processamento paralelo, de softwares (módulos), em núcleos diferentes do processador. Isto é propiciado pela nova versão do software que controla a aquisição de dados (LabView 8 da National Instruments) e pelos novos processadores, que pode conter vários núcleos.

Para melhor entendimento da proposta de arquitetura para o controlador, esta seção foi dividida da seguinte forma: item 4.2.1, descreve o funcionamento geral do controlador; item 4.2.2 descreve a lógica proposta para a vazão de alimentação do meio suplementar no bioprocessamento, em tempo real. Em seguida foram descritos o funcionamento específico de cada módulo, dessa arquitetura, iniciando pela *Base de Conhecimento*, uma vez que todos os módulos utilizaram informações contidas nessa base de dados, para realizarem suas funções específicas.

4.2.1 Funcionamento geral do controlador

O módulo que deve iniciar esta arquitetura, o *Núcleo Controlador*, módulo que está sendo proposto para uma continuidade desta linha de pesquisa, será o responsável por coordenar e integrar as ações entre os módulos e enviar as ações de monitoramento e controle ao bioprocessamento. Ele terá como primeira ação, o início do módulo *Controle do Cultivo*. As funções desempenhadas pelo *Controle do Cultivo* são: enviar a decisão do término do cultivo para o *Núcleo Controlador*, e suavizar os dados ruidosos de entrada, preparando os dados de entrada para uma utilização correta dessa arquitetura. O término do cultivo pode ocorrer em dois momentos: duas horas após o início de cultivo ou uma hora após o término da vazão de alimentação do meio suplementar. Para suavizar os dados ruidosos dos sensores, um filtro foi desenvolvido, denominado nesta dissertação de “alisa pós”. Esse filtro, importante para esta arquitetura, permitiu um uso correto das metodologias de inteligência computacional (IC), e desses dados nos outros módulos.

Se o término do cultivo não ocorresse às duas horas, dois outros módulos eram iniciados: *Controle da Vazão* e *Inferência de Cx*.

O módulo *Controle da Vazão* foi o responsável por inferir o início, o fim e as condições da vazão de alimentação do meio suplementar. Para inferir estes momentos, o início e o fim da vazão de alimentação, foi utilizada somente a noção de conjuntos *fuzzy*, e de outro filtro, denominado nesta dissertação de “escada”, e construído a partir da derivada do valor alisado, pelo filtro “alisa pós”, da y_{CO_2} (fração molar de CO_2). O *Núcleo Controlador*, ao receber o sinal de início da vazão de alimentação do módulo *Controle da Vazão*, deverá comparar o crescimento celular inferido pelo módulo *Inferência de Cx* com três propostas de modelos de crescimento celular e que estão descritos no próximo item (4.2.2). O modelo mais próximo do crescimento atual inferido será escolhido para definir as condições da vazão de alimentação, uma vez que vazões de alimentação já teriam sido treinadas (*off-line*) para os modelos propostos.

O módulo *Inferência de Cx* utilizou-se de um *softsensor*, de estrutura não convencional, para inferir *on-line* a concentração microbiana (Cx). Este *softsensor* buscou capturar conhecimento

do processo, tanto empírico, pelo comitê de redes neurais (CRNAs), quanto pelo raciocínio lógico com base fenomenológica, pelo sistema *fuzzy* (SFBR).

4.2.2 Lógica da vazão de alimentação do meio suplementar

A lógica escolhida para a vazão de alimentação do meio suplementar para esse controlador, teve como base os conhecimentos sobre este bioprocessamento, obtidos pelo grupo de pesquisas da UFSCar.

Um ponto importante é a fonte de carbono utilizada para a produção da enzima, sendo reportado que glicose atua como repressora na síntese de PGA (SAVIDGE, 1984; ILLANES *et al.*, 1994; PINOTTI, 2003). A fonte de carbono nos ensaios empregados neste trabalho foi um *pool* de aminoácidos, além do indutor para produção de PGA, o ácido fenilacético (AFA). SILVA (2003) propôs modelo cinético para o crescimento de *B. megaterium* nesse meio, baseado nas seguintes hipóteses: a produção da PGA está associada ao crescimento do microrganismo. O valor máximo de enzima coincide, com pequeno atraso, com a máxima concentração celular; o consumo de AFA e de aminoácidos preferenciais ocorre quase simultaneamente e existe um grupo de aminoácidos consumidos mais lentamente.

Levando-se em conta a presença de um ou mais substrato(s) limitante(s), foram propostos por SILVA (2003) três modelos para descrever o crescimento do *Bacillus megaterium*.

Proposta 1: O AFA, somado à concentração de aminoácidos totais, representando apenas um substrato limitante.

Proposta 2: Os aminoácidos totais e o AFA foram separados, como dois substratos limitantes.

Proposta 3: Os aminoácidos mais rapidamente consumidos (aa+), os aminoácidos mais lentamente consumidos (aa-) e o AFA eram três substratos limitantes distintos.

Propõe-se que o crescimento celular e o consumo de substrato sigam a cinética de crescimento de Monod (equação 2.16), para múltiplos substratos, na forma aditiva, pois o crescimento continua mesmo quando o AFA ou algum grupo de aminoácidos se esgota, o que não poderia ocorrer se a cinética multiplicativa fosse escolhida (BLANCH e CLARK, 1997):

$$\frac{dC_X}{dt} = \sum_{i=1}^{n^o S} \left(\frac{\mu_{\max i} \cdot S_i}{K_i + S_i} \right) \cdot C_X \quad (4.1)$$

$$\frac{dS_i}{dt} = -\frac{1}{Y_{X/S_i}} \left(\frac{\mu_{\max i} \cdot S_i}{K_i + S_i} \right) \cdot C_X \quad (4.2)$$

$$\frac{dP}{dt} = \frac{1}{Y_{X/P}} \sum_{i=1}^{n^o S} \left(\frac{\mu_{\max i} \cdot S_i}{K_i + S_i} \right) \cdot C_X \quad (4.3)$$

Através de algoritmos de estimação de parâmetros das equações de Monod, os modelos são ajustados aos dados dos ensaios da BD Seleção de Ensaios (BD SE), descrito na *Base de Conhecimento*, para os modelos cinéticos escolhidos acima. Esses algoritmos de busca são tanto globais (como Simulated Annealing) como locais (como Levenberg-Marquardt), ou híbridos. Por concisão, não são descritos aqui. Basta, para os nossos fins, saber que esses parâmetros estão também disponíveis na BD Modelos Ajustados (BD MA), descrito na *Base de Conhecimento*. Com estes parâmetros armazenados será possível calcular a velocidade específica de crescimento (μ) para cada modelo proposto acima. Este μ calculado para os modelos será comparado com o μ inferido *on-line*, e aquele que possuir mais similaridade com o μ *on-line* no momento de iniciar a vazão de alimentação será escolhido como o modelo para definir a política de vazão de alimentação do meio suplementar no biorreator. Essa decisão será reavaliada, dinamicamente, em tempo real, inclusive com a possibilidade de se interpolar entre diferentes modelos. Esta etapa, entretanto, extrapola o escopo deste mestrado, ficando como sugestão para continuidade dos trabalhos.

A política escolhida para vazão de alimentação do meio suplementar foi a de manter a concentração de substrato (C_s) (ou substratos, $C_{s(i)}$, caso se escolha um modelo com mais de um substrato limitante) inalterada, tendo como base a concentração do substrato no momento de início da vazão de alimentação (C_{s_0} ou $C_{s_0(i)}$). A razão desta política tem aplicação prática óbvia. Quando a concentração de substrato (C_s ou $C_{s(i)}$) é mantida em um nível baixo no qual parte do substrato adicionado é convertido em produtos indesejados, uma grande quantidade de biomassa (com uma proteína associada, o produto desejado, no caso a enzima PGA) é produzida em uma velocidade razoavelmente alta e, assim, uma concentração final alta de biomassa e de produto pode ser obtida (NIELSEN, VILLADSEN e LIDÉN, 2002).

O conhecimento da vazão de alimentação, obtida com esta política (que resulta em um perfil exponencial), para cada um dos modelos cinéticos escolhidos no item anterior, será armazenado na base de dados de estratégias para vazão de alimentação (BD EVA), descrito na *Base de Conhecimento*.

4.3 Base de Conhecimento

A *Base de Conhecimento* foi constituída de informação adquirida e/ou processada *off-line*. Foram utilizados dois tipos de conjuntos de informação: dados experimentais medidos durante ensaios passados, e informações das estruturas usadas pelas metodologias de inteligência computacional (IC), cujos parâmetros foram obtidos a partir dos dados experimentais. O conhecimento está distribuído na *Base de Conhecimento* em estrutura de uma base de dados, abreviado aqui pelas iniciais BD ou na forma de parâmetros, para facilitar os desenhos dos diagramas de fluxo de processos. A composição de cada base de dados e os conhecimentos que representam são descritos a seguir.

BD Seleção de Ensaios (BD SE)

O conhecimento sobre ensaios considerados significativos, ou seja, com resultados consistentes para o bioprocessamento, foi representado fisicamente por esta base de dados. Note-se que a exclusão de ensaios considerados ruins, que não possuem significância para o bioprocessamento, foi um dos diferenciais deste trabalho. Os dados dos ensaios incluem: tempo, fração molar de CO₂ (y_{CO_2}) e de O₂ (y_{O_2}) nos gases efluentes, velocidade de agitação e RQ (quociente respiratório), atualizados a cada intervalo de amostragem (1 min, aqui). No APÊNDICE A desta dissertação, foi apresentada uma breve descrição dos ensaios utilizados neste trabalho. Os dados de tempo, fração molar de CO₂ (y_{CO_2}) e de O₂ (y_{O_2}) nos gases efluentes, velocidade de agitação e RQ (quociente respiratório) estão arquivados em meio digital.

BD do Comitê de RNAs Especialistas (BD CRNAs)

Constituído das estruturas (camadas ocultas, camadas de saída, pesos internos etc) das redes neurais artificiais (BD RNAs) especialistas, que são utilizadas pelo *comitê de redes neurais* (CRNAs), descrito no item 3.2.4.1, mais o mecanismo de ponderação (MP CRNAs), descrito no item 3.2.4.2.

BD do Sistema *Fuzzy* Baseado em Regras (BD SFBR)

Constituído pelo conhecimento de como construir os conjuntos *fuzzy* (BD CFs) e das regras *fuzzy* (BD RFs), descritos no item 3.2.5.1, e do mecanismo de inferência do sistema *fuzzy* (SFBR), descrito no 3.2.5.2.

BD Modelos Ajustados (BD MA)

Constituído pelos modelos ajustados e dos seus parâmetros. A forma como estes modelos foram ajustados e os parâmetros encontrados já foram explicados no item 4.2.1, desta seção.

BD Estratégias para Vazão de Alimentação (BD EVA)

Constituído pelo conhecimento das vazões de alimentação do meio suplementar, para cada um dos modelos ajustados no item anterior, BD MA, utilizando-se da política explicada descrita no item 4.2.1, desta seção.

Parâmetro mínimo para continuar o cultivo

Este parâmetro é calculado com base no BD Seleção de Ensaio, usando a variação de y_{CO_2} (Fração molar de CO_2), calculada no intervalo de tempo de duas horas após o início do cultivo. A menor variação dentre os ensaios selecionados é o parâmetro mínimo de aceitação para que o cultivo atual continue, uma vez que estes ensaios selecionados foram escolhidos como “especialistas” para este bioprocessamento.

Parâmetro *fuzzy* para o início de vazão de alimentação

Este parâmetro foi encontrado através de um algoritmo que determinou o início e o fim de vazão de alimentação, e será descrito no item 4.5.2, desta seção. Atualmente ele tem o valor de 0,7, e significa a pertinência deste valor ao conjunto *fuzzy* intermediário.

Parâmetro *fuzzy* para o fim de vazão de alimentação

Vale aqui o mesmo do item anterior. O procedimento para definição deste parâmetro será descrito no item 4.5.2., desta seção. Atualmente, tem o valor de 1,0, e significa a pertinência deste valor ao conjunto *fuzzy* grande.

4.4 Módulo 1 – Controle do Cultivo

Tomar a decisão de finalizar o cultivo, e informar ao *Núcleo Controlador* foi uma das funções desse módulo. O final do cultivo pode ocorrer em dois momentos. Com o objetivo de abortar um cultivo inviável, o primeiro momento ocorre pontualmente duas horas após o início de cultivo. Nesse momento, é calculado o valor da variação de y_{CO_2} , a partir do valor alisado, pelo filtro “alisa pós” e descrito logo abaixo, do sinal do sensor no intervalo de tempo de zero a duas horas, e comparado com *parâmetro mínimo para continuar o cultivo*, que foi explicado no item 4.3 desta seção. O segundo momento ocorre uma hora após o término da vazão de alimentação do meio suplementar, seguindo considerações feitas por SILVA (2003), que se baseou na hipótese de que a produção da PGA está associada ao crescimento do microrganismo, e que o valor máximo de enzima coincide, com pequeno atraso, com a máxima concentração celular.

Outra função deste módulo, muito importante para o funcionamento desta arquitetura, está no alisamento dos dados de entrada (y_{CO_2} , y_{O_2} , agitação). Dados de entrada ruidosos podem prejudicar o monitoramento e possíveis tomadas de decisões. Para suavizar os

dados ruidosos de entrada, uma estratégia foi desenvolvida, onde os ruídos, nos dados, foram suavizados após a detecção de alguns eventos, e por isso a denominação de “alisa pós”. A importância desse filtro reside no tratamento aplicado aos dados ruidosos de entrada, tratamento este que deve ser consistente com a lógica geral da arquitetura proposta, permitindo a utilização correta desses dados, em todos os outros módulos e principalmente pelas metodologias de inteligência computacional: o comitê de redes neurais (CRNAs) e o sistema *fuzzy* (SFBR).

A Figura 4.2 mostra o esquema de funcionamento do módulo *Controle do Cultivo*.

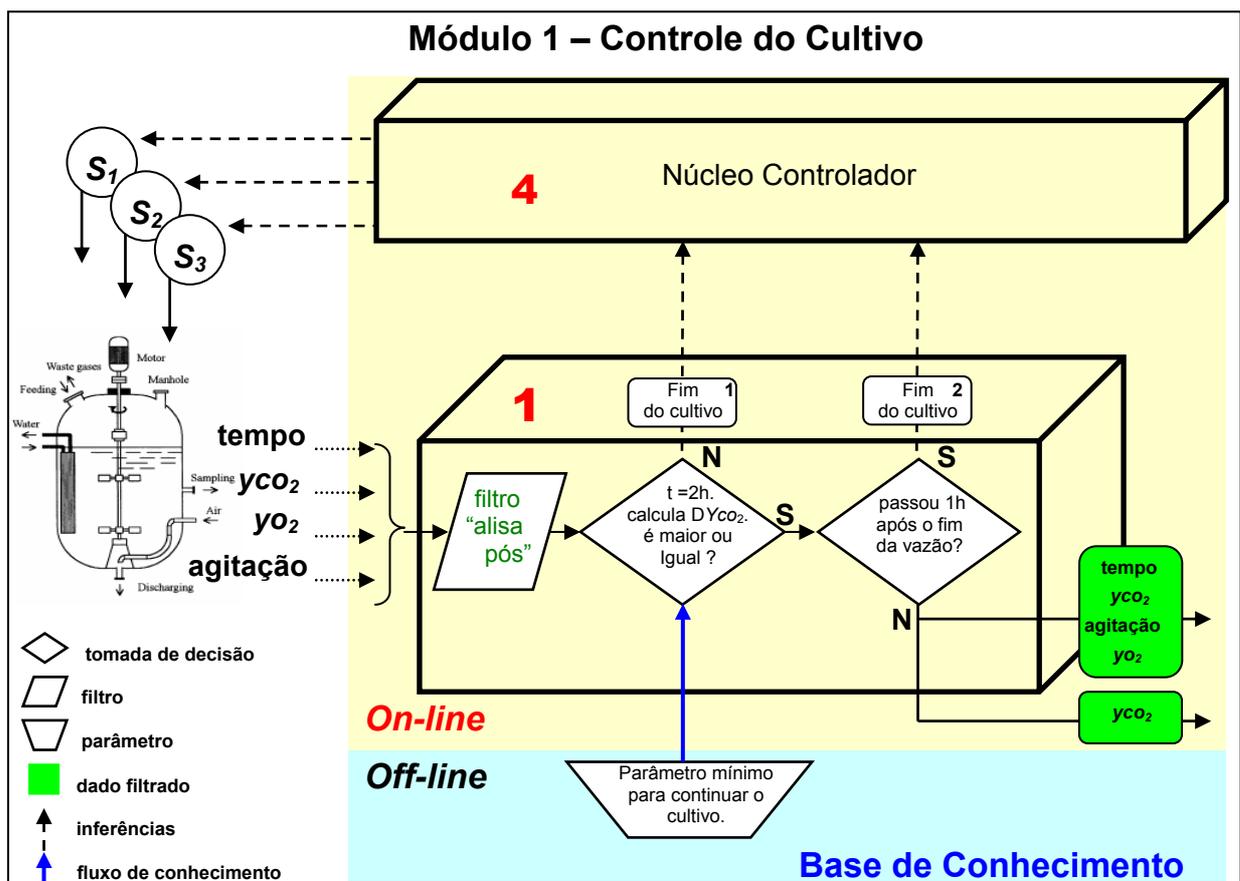


Figura 4.2 Esquema de funcionamento do módulo *Controle do Cultivo*.

4.4.1 Filtro “alisa pós”

Os dois objetivos principais do filtro “alisa pós”, foram: descrever melhor o comportamento da variável monitorada, e eliminar as distorções conhecidas dos dados do sensor (ruído identificado), preparando essa variável para uso nas próximas etapas da modelagem. Para entender melhor o comportamento da variável foi realizado um estudo, da distribuição da variação de y_{CO_2} , nos quatro ensaios considerados significativos para este bioprocessamento, nas primeiras 20 horas de cultivo. Foi definido para cada ensaio um intervalo de amostragem de um minuto. Assim, durante o ensaio tem-se 1200 dados, mais o dado inicial do tempo igual a zero, perfazendo 1201 dados e assim 1200 aproximações das derivadas de y_{CO_2}

(por diferenças finitas para trás). O motivo do estudo foi identificar se a distribuição das derivadas, nos quatro ensaios, era simétrica no tempo. A importância da distribuição simétrica (tipicamente, uma distribuição normal ou de Gauss), da derivada de y_{CO_2} , é que ela embasa o uso da metodologia estatística paramétrica, uma vez que, salvo raras exceções, resultados de análises estatísticas efetuadas através de métodos paramétricos não são confiáveis se os dados não pertencem a amostras extraídas de populações normais (RODRIGUES e IEMMA, 2005). Como pode ser visto na Figura 4.3, as distribuições dos quatro ensaios foram simétricas.

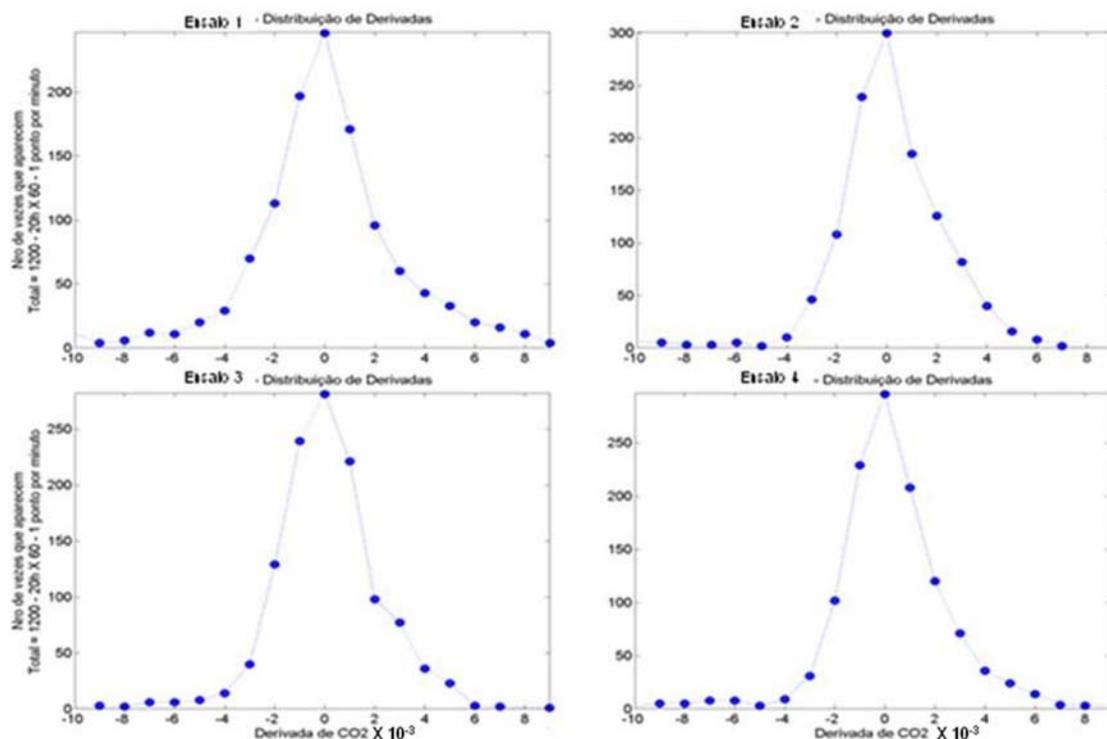


Figura 4.3 Gráficos de distribuição da derivada de y_{CO_2} nos ensaios um, dois, três e quatro.

O próximo passo foi realizar um estudo mais detalhado, da distribuição da derivada de y_{CO_2} , nos ensaios, tendo como base não mais o ensaio todo (20 horas), e sim a sua distribuição em intervalos de tempo. Esta estratégia tinha a motivação de encontrar fases similares nos ensaios. Nessa nova distribuição, foram selecionados os valores incluídos em oitenta por cento (80%) da área sob a curva normalizada, para separar os dados que podiam ser ruído. Notaram-se comportamentos repetidos, nos quatro ensaios, com pequenas variações. Estes comportamentos, mostrado para o ensaio um na Figura 4.4. da próxima página, são: (a) na 1^a hora do ensaio a distribuição central (ou seja, a área sob a curva de distribuição, ao redor da média, igual ou superior a 0,8, ou 80%) ficou no meio do eixo; (b) da 2^a à 5^a horas a distribuição central desloca se para o eixo positivo com grande intensidade; (c) da 6^a à 10^a horas a distribuição central volta a ficar no meio do eixo, com mais proximidade do eixo negativo e (d) da 11^a à 20^a horas a distribuição central desloca se para o eixo negativo com grande intensidade

no início e retornando aos poucos para o meio do eixo. O reconhecimento desses comportamentos, observados na distribuição da derivada de y_{CO_2} da população alvo, realizada em intervalos de tempo, direcionou a pesquisa no sentido de relacionar esses comportamentos às várias fases de crescimento da massa celular do microrganismo (fase lag, aceleração, de crescimento exponencial, de desaceleração e finalmente a estacionária). Essa identificação das fases de crescimento trouxe vários benefícios para este trabalho, entre os quais: a identificação de limites mínimo e máximo de tolerância para aceitar ou rejeitar os sinais do sensor de y_{CO_2} na construção do filtro; a identificação do início do estado de atenção, quando se processa a verificação sistemática da conveniência de se iniciar a vazão de alimentação do meio suplementar; e a definição de pesos diferentes para as regras *fuzzy* no módulo *Inferência de Cx*.

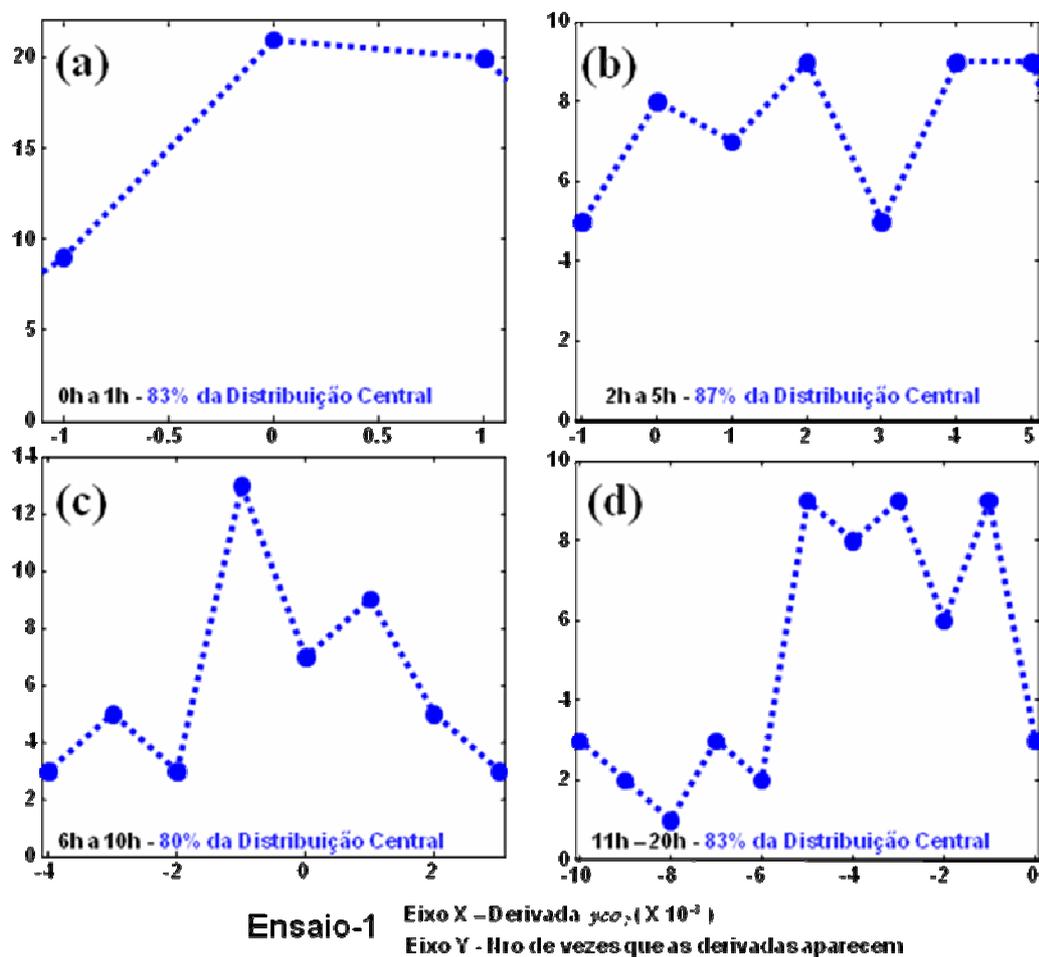


Figura 4.4 Distribuição central, da derivada de y_{CO_2} do ensaio um: (a) da 1ª hora, (b) da 2ª a 5ª hora, (c) da 6ª a 10ª hora e (d) da 11ª a 20ª hora.

4.4.1.1 Etapas de construção do filtro

A lógica utilizada para a construção do filtro foi dividida em duas etapas, descritas a seguir:

1ª etapa – tratamento da 1ª hora do ensaio

O início de um cultivo, em um biorreator, exige do operador muito trabalho e atenção, já que várias operações são realizadas ao mesmo tempo. Inocular; retirar amostra do tempo zero; conferir se os valores da vazão de ar, temperatura, agitação e O_2 dissolvido estão sendo mantidos, nos valores estabelecidos para o experimento; e também, verificar se todos os sensores do Sistema de Aquisição de Dados (SAD) estão registrando os dados. Também foi observado que os primeiros dados dos ensaios são muito instáveis, inclusive em razão da pequena massa de células existentes no biorreator. Diante dessas observações, foi construído um tratamento diferenciado para 1ª hora do ensaio. Nesta primeira hora do ensaio o tratamento foi *off-line*. O algoritmo do filtro tratou a 1ª hora de dados (61 valores de $y_{CO_2} - 1$ por minuto, mais o valor no tempo igual a zero) a partir do sexagésimo primeiro minuto, ou seja, esperou a 1ª hora passar para tratá-los de uma única vez. Do conjunto inicial (61 dados) o algoritmo calculou a derivada, minuto a minuto, e construiu subconjuntos. Estes subconjuntos de y_{CO_2} foram construídos a partir da verificação da obediência das derivadas de y_{CO_2} ao limite mínimo e máximo da distribuição da região central, onde estes limites foram aprendidos a partir do estudo da distribuição da derivada de y_{CO_2} dos ensaios anteriores (ver figura 4.4). A separação de um subconjunto do outro foi feita quando uma derivada não respeitava os limites estabelecidos. Então o subconjunto atual era encerrado e este valor de y_{CO_2} , que interrompia o subconjunto anterior, era o primeiro elemento do novo subconjunto. Uma vez construídos os subconjuntos confiáveis de y_{CO_2} (intervalos que respeitavam os limites estabelecidos para a fase lag) o algoritmo do filtro varreu os subconjuntos, iniciando a pesquisa do subconjunto final para o inicial, obedecendo a lógica que os valores de y_{CO_2} ficou constante ou cresceu pouco, uma vez que o ensaio está na fase lag. Para esta lógica, foi atribuído maior valor aos subconjuntos do fim e menor valor para os subconjuntos do início, devido aos problemas já mencionados acima sobre o operador e a pequena massa de células. Essa estratégia, do algoritmo do filtro, de pesquisar do fim para o começo, evitou grandes distorções, nos valores de y_{CO_2} filtrados, no início do cultivo e que poderiam comprometer a decisão de continuar ou não este ensaio, feita pontualmente em duas horas de cultivo.

2ª etapa – tratamento dos dados restantes com identificação das fases

A partir do sexagésimo primeiro minuto, o algoritmo do filtro, iniciou o tratamento *on-line*, dos dados do sensor. Nessa etapa, para cada valor lido da y_{CO_2} , do sensor, calculou se a derivada, e verificou se este respeitava os limites, mínimo e máximo, aprendidos no estudo da distribuição da derivada de y_{CO_2} dos ensaios anteriores. Caso a derivada não respeitasse os

limites, o algoritmo do filtro esperava pelo próximo valor da y_{CO_2} , lido do sensor, para tomar uma decisão da melhor forma, para alisar, estes valores ruidosos. Essa decisão, melhor forma de alisar os valores ruidosos de y_{CO_2} do sensor, dependia necessariamente de qual fase de crescimento, a população microbiana se encontrava. Esse foi mais um motivo para detectar as fases *on-line*. Nas fases lag, de aceleração e exponencial, o algoritmo do filtro, também utilizou a seguinte regra: o último valor armazenado de y_{CO_2} , no filtro, nunca devia ser menor do que o valor de y_{CO_2} , recebido *on-line* do sensor. Essa regra foi amparada no seguinte fato relacionada ao crescimento celular: na fase lag, as células crescem pouco ou não crescem e nas fases de aceleração e exponencial, as células estão crescendo. Nessas fases é desprezível a morte celular. Quando isto acontecia, o valor recebido do sensor de y_{CO_2} ser menor do que o último valor armazenado de y_{CO_2} no filtro, o algoritmo do filtro esperava um valor maior de y_{CO_2} do sensor. Uma vez encontrado este y_{CO_2} maior, foi calculado entre os dois dados, uma derivada constante, linear com o tempo e então, o filtro interpolava entre esses dois dados. Uma das vantagens desse procedimento foi o de não perder o controle automático do ensaio, devido, por exemplo, a uma falha no funcionamento do sensor como mostra a Figura 4.5, onde se perdeu o sinal por alguns momentos devido a uma falha (queda de tensão na linha de energia). Este ruído prejudicaria muito o resultado final deste ensaio se não houvesse filtro.

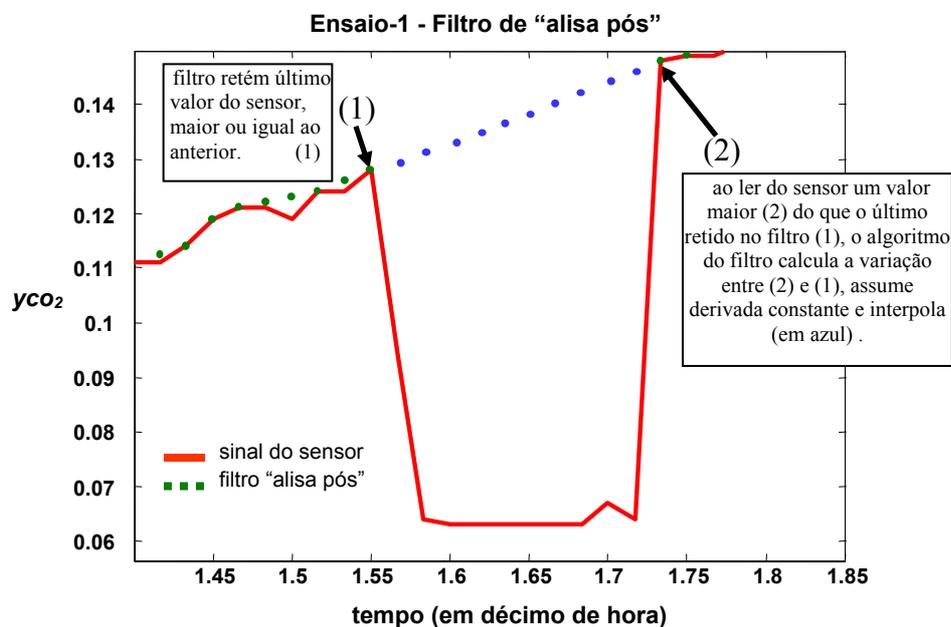


Figura 4.5 Gráfico da y_{CO_2} , do tratamento dado pelo filtro "alisa pós", a uma perda de sinal do sensor, ocorrido no ensaio um.

A lógica de construção desse filtro, descrita nesse item, abordou somente a fração molar de CO_2 (y_{CO_2}), por motivo de concisão, mas o filtro "alisa pós" foi utilizada também para a fração molar de O_2 (y_{O_2}) e velocidade de agitação. A lógica foi similar, utilizando apenas algumas adaptações necessárias, específicas para cada variável.

4.4.1.2 Identificação das fases de crescimento celular (Cx)

As regras que identificaram as fases de crescimento celular foram calculadas sobre os valores de y_{CO_2} alisados pelo filtro “alisa pós”.

1. Identificação do início da fase de aceleração

O início da fase de aceleração foi identificada fazendo a somatória das vinte (20) últimas derivadas de y_{CO_2} . A identificação ocorria se por vinte (20) minutos seguidos, a somatória dessas vinte (20) últimas derivadas fosse maior que $0,01 \text{ h}^{-1}$.

2. Identificação do início da fase Exponencial

O início da fase exponencial foi identificada, de forma similar ao início da fase de aceleração, fazendo a somatória das vinte (20) últimas derivadas de y_{CO_2} . A identificação ocorria se por vinte (20) minutos seguidos, a somatória dessas vinte (20) últimas derivadas fosse maior que $0,02 \text{ h}^{-1}$.

3. Identificação do maior valor de y_{CO_2}

A identificação deste valor, que caracteriza a região de inflexão do crescimento celular, é importante para a estratégia de cultivo: ele identifica o fim da fase de crescimento exponencial. O início de uma vazão de alimentação de meio suplementar terá com certeza este ponto como referência. Além disto, os valores de y_{CO_2} dos ensaios a partir deste ponto não apresentam estabilidade, o que dificultaria a aplicação de comparações entre ensaios do banco de dados. Com a identificação deste ponto, foi possível diminuir a importância dessa comparação, após este ponto. Isto ocorreu no momento que o sistema *fuzzy* (SFBR) corrigia o valor de Cx, inferido pelo comitê (CRNAs). Até este ponto, o peso da correção foi multiplicado por um (1) e depois deste ponto a correção foi multiplicado por 0,1.

4. Identificação do início da fase estacionária

Esta fase foi identificada pelo filtro “escada”, do módulo *Controle da Vazão*. O momento de terminar a vazão de alimentação do meio suplementar foi escolhido neste trabalho como sendo o da identificação do início da fase de desaceleração.

4.5 Módulo 2 – *Controle da Vazão de meio suplementar*

Um dos objetivos desse módulo foi inferir com acuidade dois momentos importantes do bioprocesso: o início e o fim da vazão de alimentação do meio suplementar. Para alcançar este objetivo foi construído um segundo filtro, chamado “filtro escada”. O filtro “escada” utiliza-se das derivadas dos valores alisados da y_{CO_2} , pelo filtro “suaviza pós”. Outra importante função desse módulo foi o de definir as condições da vazão de alimentação. Esta definição foi

feita no momento seguinte ao comando “iniciar a vazão suplementar”. O algoritmo compara o crescimento celular inferido, pelo módulo *Inferência de Cx*, para o cultivo atual, com três propostas de modelos de crescimento celular. O modelo mais próximo do crescimento atual inferido é escolhido para definir a vazão de alimentação, uma vez que várias vazões de alimentação em função do tempo haviam sido determinadas *off-line*, para esses modelos. A lógica da vazão de alimentação do meio suplementar foi explicada no item 4.2.2 desta seção. A Figura 4.6 mostra o esquema de funcionamento do módulo *Controle da Vazão*.

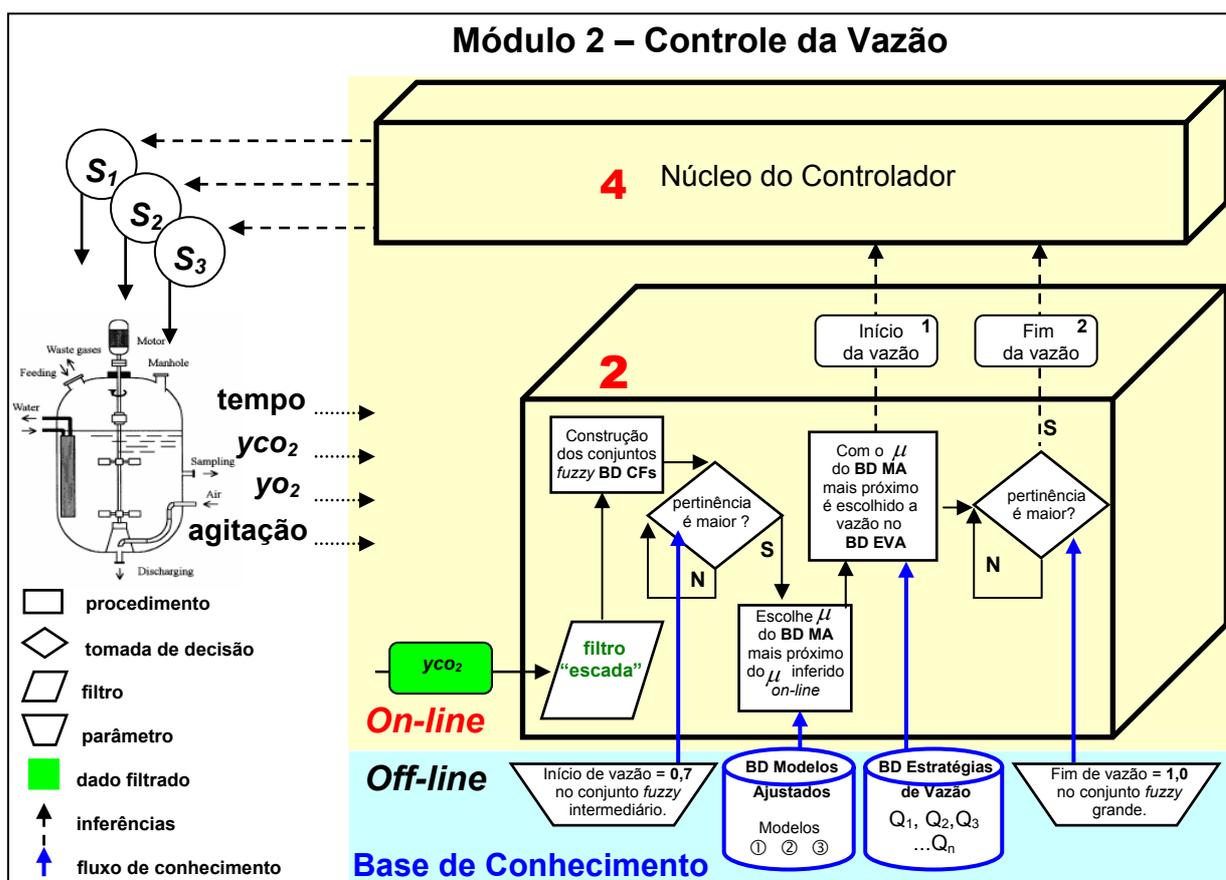


Figura 4.6 Esquema de funcionamento do módulo *Controle da Vazão*.

4.5.1 Filtro “escada”

Após várias tentativas, sem sucesso, de encontrar uma maneira de definir *on-line* o início e o término da vazão de alimentação do meio suplementar, surgiu a idéia da construção desse filtro. O objetivo principal deste filtro foi o de preparar a derivada do valor de y_{CO_2} , alisado pelo filtro “alisa pós”, para uma solução adaptativa, onde o sinal, do início e término da vazão de alimentação, pudesse ser encontrado em uma região imprecisa, e não determinada por um valor fixo. Para tratamento desta imprecisão foi utilizada uma lógica utilizando conjuntos *fuzzy*. A heurística empregada para construção do filtro “escada” foi a de calcular a derivada de y_{CO_2} do filtro “alisa pós”. Se a derivada de y_{CO_2} fosse maior que a anterior seu

valor era armazenado; caso contrário, descartado. Após a identificação da maior derivada de y_{CO_2} ($Dy_{CO_2_MAX}$, Figura 4.7) o procedimento foi o inverso. As derivadas foram comparadas e o menor valor foi armazenado. Assim, o tamanho da “escada”, $Dy_{CO_2_MAX}$, foi utilizado como parâmetro para definição da base do conjunto *fuzzy* de entrada. Dessa forma essa definição é adaptativa, já que cada ensaio tem um $Dy_{CO_2_MAX}$ diferente. O resultado do filtro “escada”, para o ensaio cinco, é mostrado na Figura 4.7, e como pode ser notado, seu formato é de uma escada.

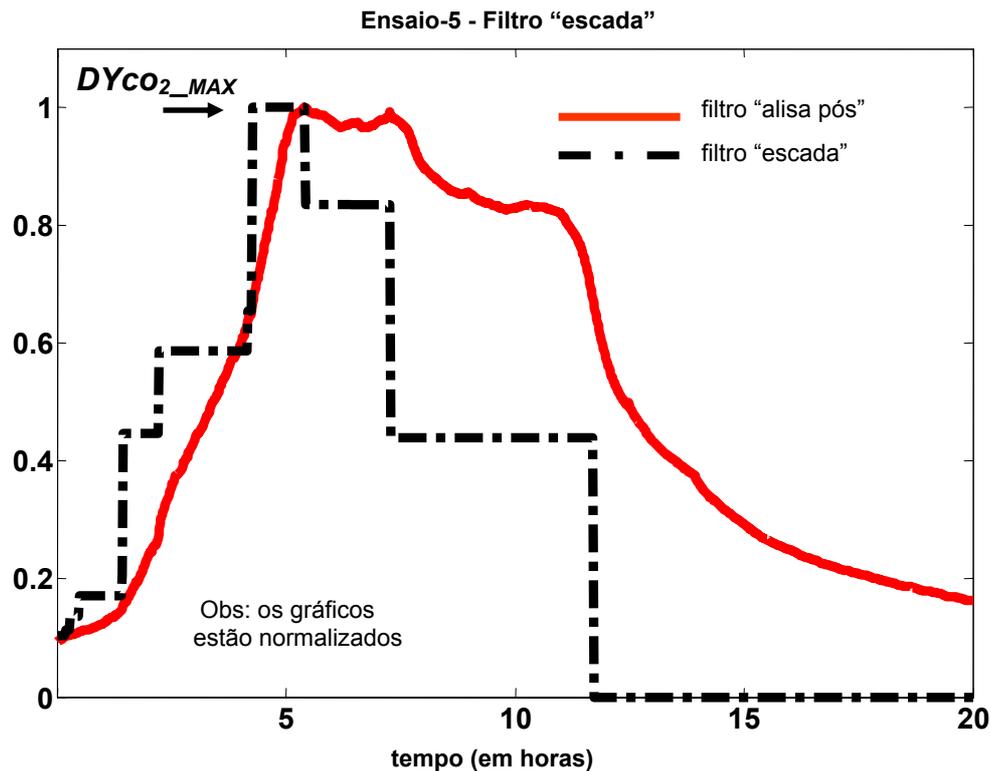


Figura 4.7 Gráfico dos filtros “alisa pós”, de cor vermelha, e “escada”, de cor preta.

4.5.2 Inferência *fuzzy* para definição dos momentos de início e fim de vazão de alimentação

Ao receber do sistema de aquisição de dados (SAD), o valor da y_{CO_2} , medido pelo sensor, este valor foi filtrado pelo filtro “alisa pós”. A derivada do valor alisado (Dy_{CO_2}) foi a entrada para o filtro “escada”. O algoritmo espera a identificação do maior valor da derivada de y_{CO_2} ($Dy_{CO_2_MAX}$), mostrado na Figura 4.7. Esta identificação foi explicada na construção do filtro “alisa pós”. Com $Dy_{CO_2_MAX}$, os pesos internos do algoritmo *fuzzy* são automaticamente modificados para se adequarem ao ensaio em questão, isto é, são redefinidos os limites dos rótulos lingüísticos que descrevem a magnitude do módulo da derivada de y_{CO_2} (pequena, intermediária ou grande). Através de simulações foi encontrado o valor que corresponde a $Dy_{CO_2_MAX}/3$. Assim, as funções trapezoidais de pertinência são definidas em relação à terça

parte do valor $Dy_{CO_2_MAX}$ de cada ensaio. Esta heurística é adaptativa, pois para cada ensaio o valor de $Dy_{CO_2_MAX}$ é diferente.

A partir da identificação do $Dy_{CO_2_MAX}$, e com os conjuntos *fuzzy* construídos, cada alteração da derivada de y_{CO_2} (negativa, mas tomada em seu valor absoluto pelo algoritmo *fuzzy*), ou seja, a cada tamanho do degrau na descida da “escada”, foi submetida ao conjunto *fuzzy*. Estudos mostraram que quando a saída do conjunto *fuzzy* possuía um grau de pertinência maior que 0.7, no conjunto Intermediário, o controlador devia acionar estado de atenção para ligar bomba de meio suplementar. Como seria de se esperar, o início da fase estacionária corresponde a um declínio sistemático de y_{CO_2} . A velocidade desta diminuição, assim, pode ser um indicativo do momento em que a fase de crescimento se esgota. O início da fase estacionária foi identificado com grau de pertinência igual a 1.0, no conjunto Grande. Os resultados e discussões sobre este módulo será realizado de forma mais ampla no item 5.1, da seção 5.

4.6 Módulo 3 – Inferência de Cx

A Figura 4.8 mostra o esquema de funcionamento do módulo *Inferência de Cx*.

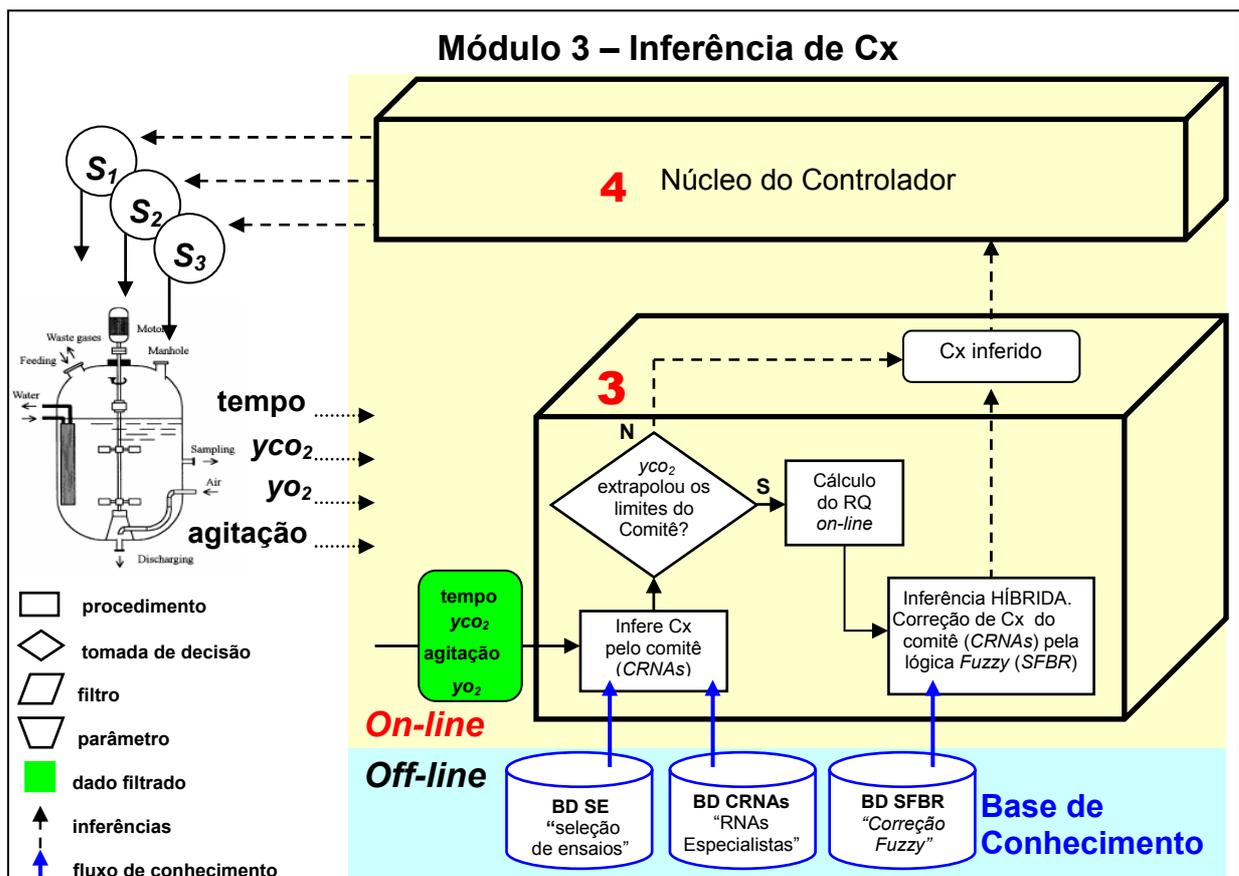


Figura 4.8 Esquema de funcionamento do módulo *Inferência de Cx*.

Inferir o crescimento celular (C_x) com acuidade, tarefa deste módulo, é muito importante para essa arquitetura proposta. Um *softsensor* foi desenvolvido. Ele se utilizou de um sistema híbrido inteligente (SHI). Esse sistema híbrido inferiu a concentração celular (C_x) e a velocidade específica de crescimento (μ) a partir do tempo, das frações molares de CO_2 (y_{CO_2}) e O_2 (y_{O_2}) nos gases efluentes, do quociente respiratório (RQ) e da velocidade de agitação durante o cultivo. Lembrando que os valores de y_{CO_2} , y_{O_2} e velocidade de agitação foram suavizados pelo filtro “alisa pós”. A inferência tinha como base um *comitê de redes neurais* (CRNAs), como descrito no item 3.2.4, da seção 3. Esta inferência provou ser eficiente, quando a fração molar de CO_2 (y_{CO_2}) *on-line* estava entre o intervalo definido pelo menor até o maior valor, da y_{CO_2} da base de dados Seleção de Ensaio (BD SE). No momento em que y_{CO_2} *on-line* ficava fora destes limites, foi adotada uma solução híbrida, já que nestas regiões de extrapolação o comitê não foi capaz de inferir crescimento celular (C_x) com acuidade. Note-se que o comitê (CRNAs) não foi treinado nestas regiões, e esperava-se dele, um certo grau de extrapolação, o que não ocorreu. Como a solução do comitê (CRNAs) baseava no conhecimento do processo de forma empírica, em situações ocorridas em ensaios passados, pensou-se para sua correção, uma base fenomenológica. Foi escolhida como base a evolução do teor de CO_2 nos gases efluentes (y_{CO_2}) e o quociente respiratório (RQ). A escolha, destas duas variáveis, foi pelo fato de estarem intimamente ligada à dinâmica de crescimento celular, sendo assim uma boa base fenomenológica para utilizar uma lógica que mensure o crescimento celular (C_x). A metodologia escolhida para tratar as regiões de extrapolação foi um sistema *fuzzy*, devido as imprecisões e incertezas dos dados de entrada nesse momento. Este sistema *fuzzy* possuía como entrada os valores da y_{CO_2} e o quociente respiratório (RQ), *on-line*, e como saída o valor da correção na concentração celular (C_x), inferida pelo comitê (CRNAs), como foi descrito no item 3.2.5. O sistema híbrido inteligente (SHI) corrigia, nas regiões de extrapolação, o valor inferido de crescimento celular (C_x), do comitê de redes neurais (CRNAs), por um valor calculado pelo sistema *fuzzy* (SFBR). Os resultados e discussões sobre este módulo será realizado de forma mais ampla no item 5.2, da seção 5.

4.7 Módulo 4 – Núcleo Controlador

Este módulo, como já dito anteriormente, não foi implementado, ficando como proposta para continuidade desta linha de pesquisa. Seus principais objetivos são de coordenar e integrar as ações dos outros módulos, assim como de realizar a interface de controle com o sistema de aquisição de dados, enviando as ações de monitoramento/controle.

O objetivo principal deste controlador é contribuir para que o cultivo microbiano alcance a sua máxima produtividade. Para atingir essa meta, dois desafios devem ser superados: identificar com acuidade o comportamento do crescimento microbiano (C_x) e, a seguir, encontrar na base de dados dos modelos ajustados (BD MA), o modelo mais similar/próximo deste comportamento de crescimento que estaria ocorrendo no bioprocesso. Essa decisão poderá ser reavaliada, dinamicamente, em tempo real, inclusive com a possibilidade de se interpolar entre diferentes modelos. Para cada modelo similar que foi armazenado na base de dados dos modelos ajustados (BD MA), existe uma condição de vazão de alimentação, definida em função do tempo, e treinada *off-line*, com os ensaios da base de dados de seleção de ensaios (BD SE). Essas condições de alimentação estão carregadas na base de dados de estratégias de vazão de alimentação (BD EVA). Esta estratégia de vazão de alimentação escolhida poderá sofrer uma correção final, através de uma comparação entre μ inferido *on-line* e o μ do modelo escolhido como o mais similar. A Figura 4.9 mostra o esquema de funcionamento do módulo *Núcleo Controlador*.

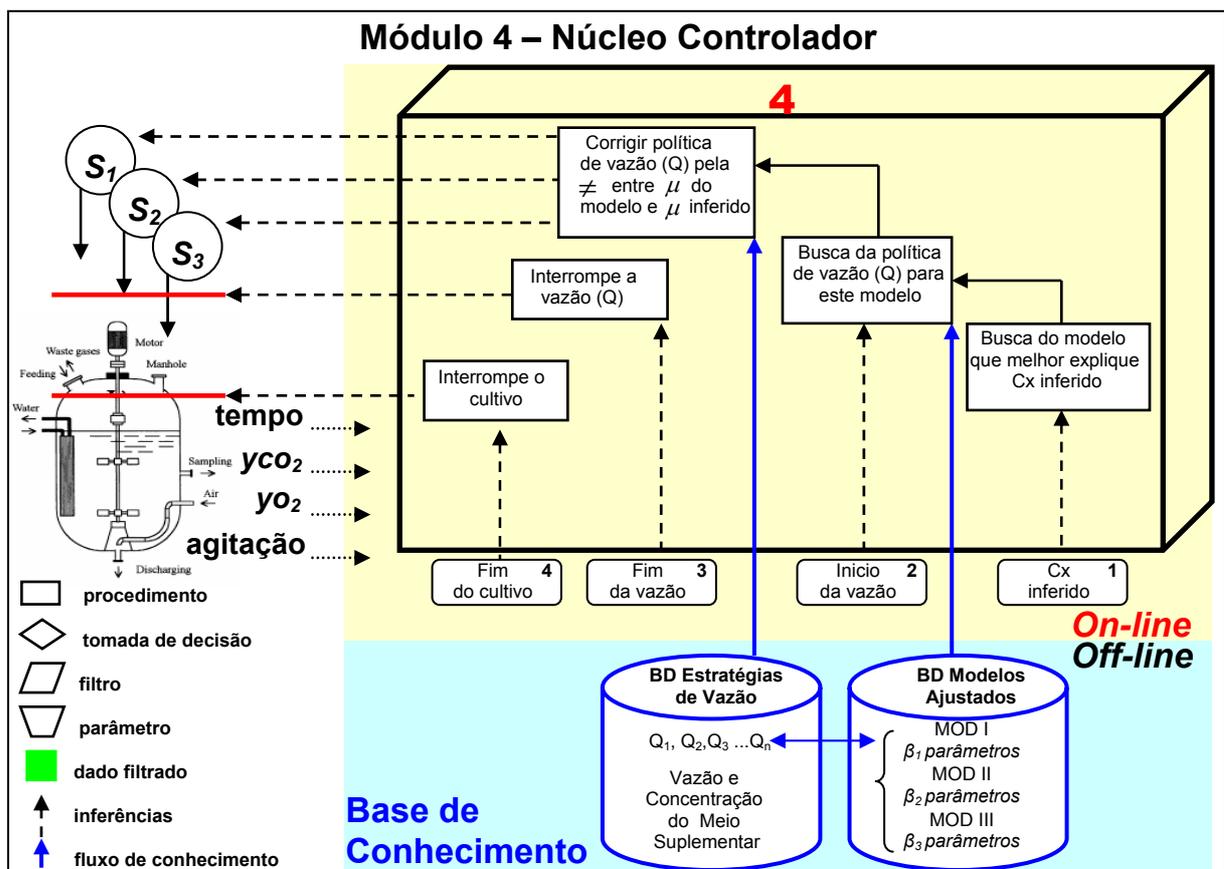


Figura 4.9 Esquema de funcionamento do módulo *Núcleo Controlador*.

5. RESULTADOS E DISCUSSÃO

Apresentam-se aqui os resultados do sistema proposto para monitoramento e controle de um biorreator (tipo SBTR), operado inicialmente em batelada e depois em batelada alimentada (“fed-batch”), para o bioprocesso de produção da enzima PGA pelo microorganismo *Bacillus megaterium*. Serão apresentados os resultados dos seguintes módulos: *Controle do Cultivo*, *Controle da Vazão* e *Inferência de Cx*.

5.1 Resultados do módulo *Controle do Cultivo*

Este módulo utilizou o filtro “alisa pós” que alisou os dados de entrada (y_{CO_2} , y_{O_2} e velocidade de agitação). Os seus dois objetivos principais foram descrever melhor o comportamento da variável monitorada, e eliminar as distorções conhecidas dos dados do sensor (ruído identificado), preparando essa variável para utilização das metodologias de inteligência computacional (IC). Os resultados serão apresentados para a variável y_{CO_2} , e em duas partes seguindo a mesma divisão de como foi construído. O resultado do filtro “alisa pós” utilizou um tratamento especial para a primeira hora do ensaio devido a uma sobrecarga de tarefas que o operador possui no início da operação de cultivo, onde várias tarefas eram realizadas ao mesmo tempo, tendo como consequência um sinal muito ruidoso principalmente na primeira meia hora. Este ruído, se não fosse tratado, causaria problemas para a lógica do primeiro momento de finalizar o cultivo, realizado por este módulo. A Figura 5.1a mostra o gráfico de y_{CO_2} versus tempo do sinal do sensor e do filtro “alisa pós” para o tratamento da primeira hora de cultivo para o ensaio cinco e na Figura 5.1b as mesmas informações para o ensaio seis.

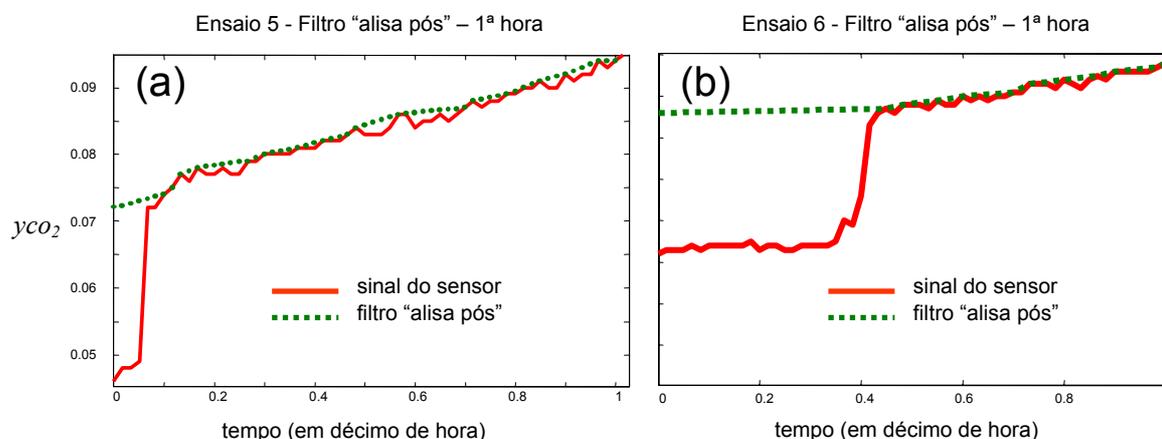


Figura 5.1 Gráfico da y_{CO_2} versus tempo do sinal do sensor e do filtro “alisa pós”, para o tratamento da 1ª hora: (a) do ensaio cinco e (b) do ensaio seis.

O filtro “alisa pós” iniciou o tratamento *on-line* a partir do sexagésimo primeiro minuto dos dados do sensor. A Figura 5.2a mostra o gráfico de y_{CO_2} versus tempo do sinal do sensor e do filtro “alisa pós” para o tratamento das vinte horas de cultivo, inclusive com o tratamento da primeira hora, para o ensaio cinco e na Figura 5.2b as mesmas informações para o ensaio seis.

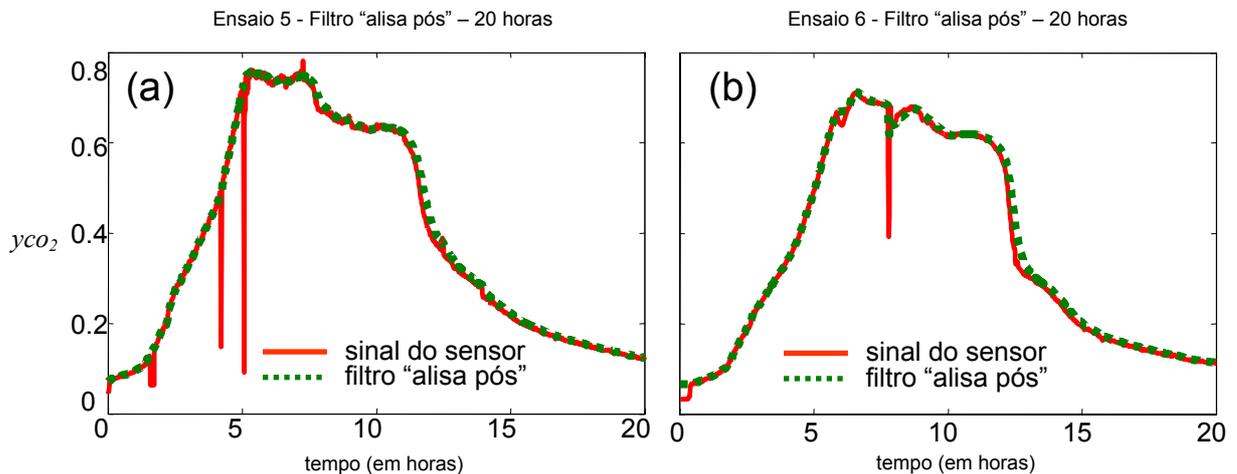


Figura 5.2 Gráfico da y_{CO_2} versus tempo do sinal do sensor e do filtro “alisa pós”, para o tratamento de vinte horas do cultivo: (a) do ensaio cinco e (b) do ensaio seis.

A Figura 5.3a mostra no gráfico de y_{CO_2} versus tempo do sinal do filtro “alisa pós” o resultado da identificação de fases do crescimento celular (C_x) *on-line*, para o ensaio cinco e na Figura 5.3b as mesmas informações para o ensaio seis.

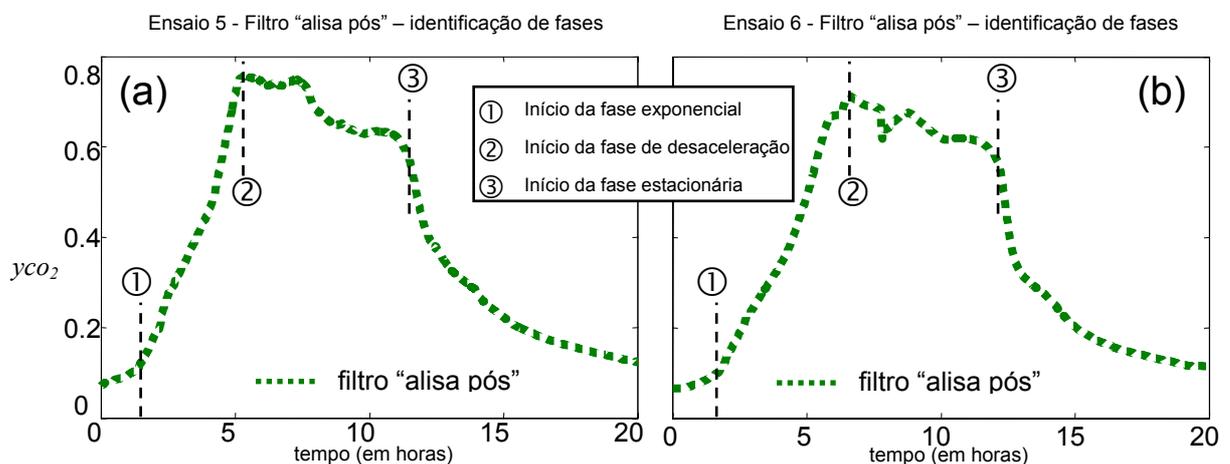


Figura 5.3 Gráfico da y_{CO_2} versus tempo do filtro “alisa pós”, para a identificação das fases do crescimento celular (C_x): (a) do ensaio cinco e (b) do ensaio seis.

5.1.1 Discussões dos resultados

Como pode ser visto na Figura 5.1a do ensaio cinco e Figura 5.1b do ensaio seis, nos momentos iniciais do cultivo a diferença entre o sinal do sensor e o valor alisado pelo filtro “alisa pós” é muito grande. Também pode ser percebido o porquê da estratégia de atribuir maior valor aos subconjuntos do fim e menor valor para os subconjuntos do início na rotina que tratava esta primeira hora. Sem esta estratégia não seria possível manter a lógica do primeiro

momento de finalizar o cultivo, na qual é calculado a variação de y_{CO_2} no intervalo de tempo de zero a duas horas, e comparado com *parâmetro mínimo para continuar o cultivo*.

Os resultados do filtro “alisa pós”, para as vinte primeiras horas do cultivo, mostrados na Figura 5.2a para ensaio cinco e Figura 5.2b para ensaio seis, conseguiram atingir seus dois objetivos propostos.

Os resultados da identificação de fases, mostrados na Figura 5.3a para ensaio cinco e Figura 5.3b para ensaio seis, foram de forma satisfatórios, e os atrasos ocorridos para identificar as fases não atrapalharam uma série de benefícios que este algoritmo trouxe para este trabalho, e que estão descritos no item 4.4.1.

5.2 Resultados do módulo *Controle da Vazão de meio suplementar*

A Figura 5.4 mostra a identificação do valor máximo da derivada da fração molar de CO_2 (y_{CO_2}) em relação ao tempo ($Dy_{CO_2_MAX}$), a partir dos valores do filtro “escada” do ensaio cinco.

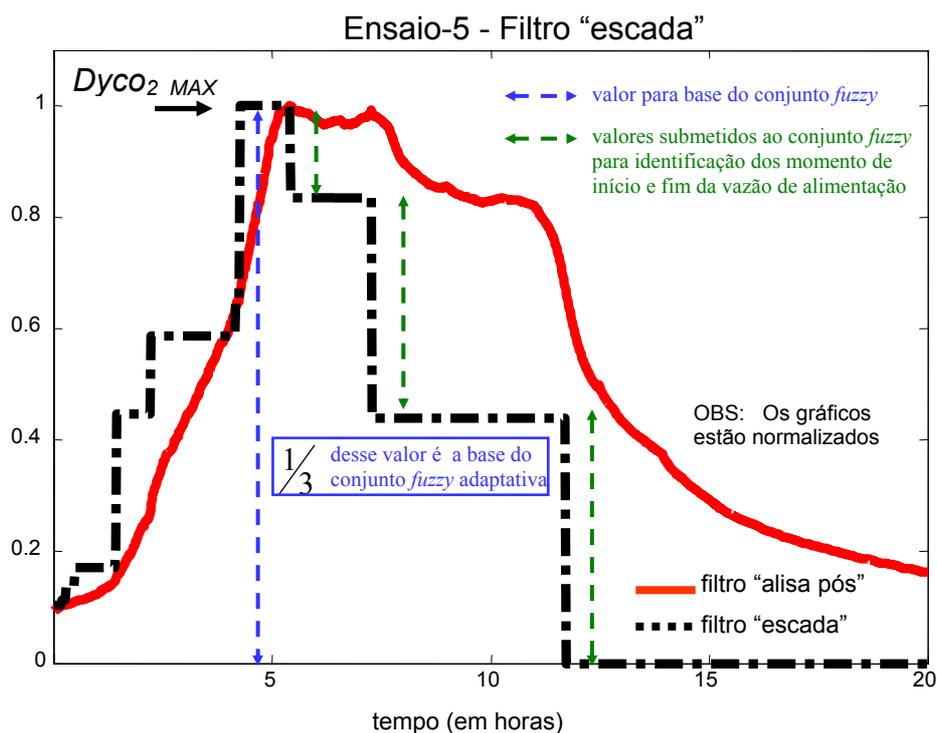


Figura 5.4 Gráfico dos filtros “alisa pós” e “escada”, e os valores utilizados pela lógica *fuzzy* para inferir os momentos de início e fim da vazão de alimentação de meio suplementar.

Com um terço do valor de $Dy_{CO_2_MAX}$, o domínio da magnitude do módulo da derivada de y_{CO_2} (Dy_{CO_2}), representado pelos conjuntos *fuzzy* pequena, intermediária e grande, são automaticamente adaptados para se adequarem ao ensaio em questão, isto é, são redefinidos os limites dos conjuntos *fuzzy*. Isto pode ser visto observando as diferenças das

bases dos conjuntos *fuzzy* pequena, intermediária e grande do ensaio cinco, Figura 5.5a, comparadas as mesmas bases dos conjuntos *fuzzy* do ensaio seis, Figura 5.5b.

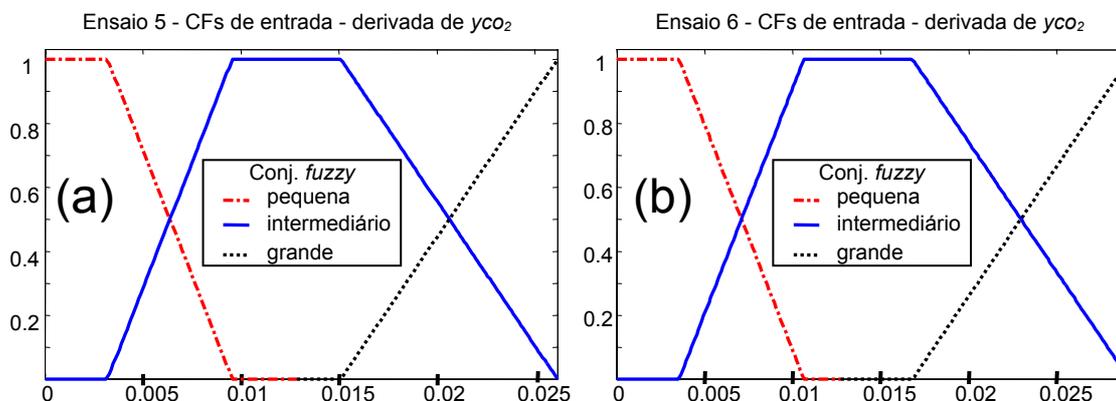


Figura 5.5 Domínio da magnitude do módulo da derivada de y_{CO_2} ($D_{y_{CO_2}}$) e seus conjuntos *fuzzy*: (a) do ensaio cinco e (b) do ensaio seis.

Os resultados obtidos pela inferência *fuzzy* para indicar o momento de início e fim para a vazão de alimentação de meio suplementar do ensaio cinco é mostrado junto com os gráficos dos filtros “alisa pós” e “escada” normalizados na Figura 5.6a, e a mesma inferência *fuzzy* para o mesmo ensaio é mostrada junto aos dados experimentais de crescimento microbiano (C_x) e consumo de substrato (C_s), em gramas por litro (g/L), na Figura 5.6b. A forma de operação do biorreator para o ensaio cinco foi batelada.

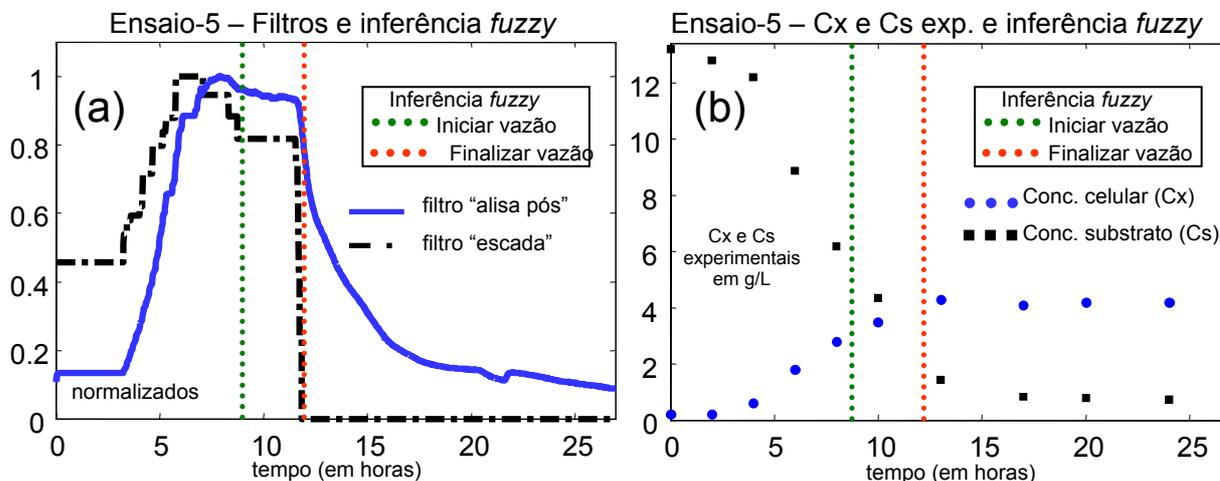


Figura 5.6 Visualização da inferência *fuzzy* para os momentos de início e fim de vazão de alimentação do ensaio cinco: (a) com os filtros “alisa pós” e “escada” e (b) com os dados experimentais de crescimento celular (C_x) e consumo de substrato (C_s).

Os resultados obtidos pela inferência *fuzzy* para indicar o momento de início e fim para a vazão de alimentação de meio suplementar do ensaio seis é mostrado junto com os gráficos dos filtros “alisa pós” e “escada” normalizados na Figura 5.7a, e a mesma inferência *fuzzy* para o mesmo ensaio é mostrada junto aos dados experimentais de crescimento

microbiano (C_x) e consumo de substrato (C_s), em gramas por litro (g/L), na Figura 5.7b. A forma de operação do biorreator para o ensaio seis foi batelada alimentada (“fed-batch”).

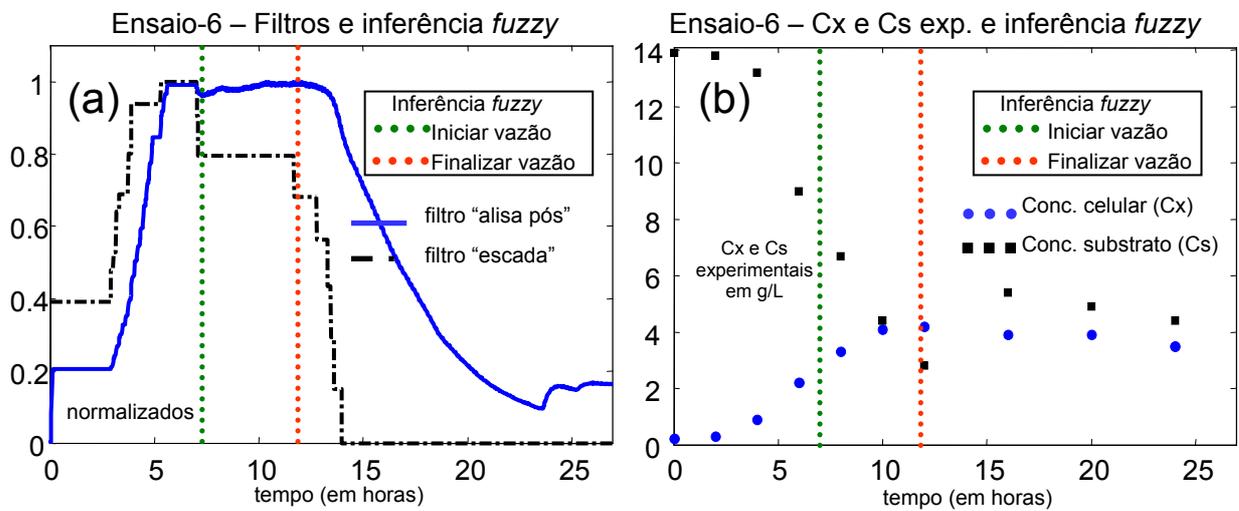


Figura 5.7 Visualização da inferência *fuzzy* para os momentos de início e fim de vazão de alimentação do ensaio seis: (a) com os filtros “alisa pós” e “escada” e (b) com os dados experimentais de crescimento celular (C_x) e consumo de substrato (C_s).

5.2.1 Discussões dos resultados

Fazendo uso apenas do sinal de fração molar de CO_2 (y_{CO_2}) na composição do gás efluente de um biorreator, o algoritmo de inferência *fuzzy* desenvolvido neste trabalho mostrou-se capaz de indicar o momento para início de vazão de alimentação de meio suplementar no biorreator nas regiões esperadas. É bom lembrar que a política escolhida para vazão de alimentação do meio suplementar é a de manter a concentração de substrato (C_s) (ou substratos, $C_{s(i)}$, caso se escolha um modelo com mais de um substrato limitante) inalterada, tendo como base a concentração do substrato (C_s) no momento de início da vazão de alimentação, representado nas Figuras por pontos verdes. O momento de parada de vazão de alimentação de nutrientes foi, também, previsto com acuidade. Deve-se atentar para a capacidade do algoritmo em se adaptar para o ensaio no qual esta sendo utilizado, isto é, existe uma adaptação do domínio da magnitude do módulo da derivada de y_{CO_2} (Dy_{CO_2}), representado pelos conjuntos *fuzzy* pequena, intermediária e grande, em função do valor de um terço de $Dy_{\text{CO}_2_MAX}$, valor este diferente para cada ensaio. Os ensaios utilizados para validação deste algoritmo utilizaram formas de operações do biorreator distintas, como os ensaios cinco e seis, mostrados nas Figuras 5.6 e 5.7, e os seus resultados mostram a robustez do algoritmo de inferência *fuzzy*. É bom lembrar que este algoritmo não utilizou toda a potencialidade de um sistema *fuzzy*, mas a capacidade de representação de categorias por meio de conjuntos *fuzzy* teve um papel fundamental para essa heurística de identificação dos pontos-chave de operação.

5.3 Resultados do módulo *Inferência de Cx*

Os resultados do *softsensor* para a inferência da concentração celular (C_x) utilizando somente o comitê de redes neurais (CRNAs) são visualizados no gráfico juntos dos dados experimentais na Figura 5.8a para o ensaio um, na Figura 5.8b para o ensaio dois, na Figura 5.9a para o ensaio três e na Figura 5.9b para o ensaio quatro.

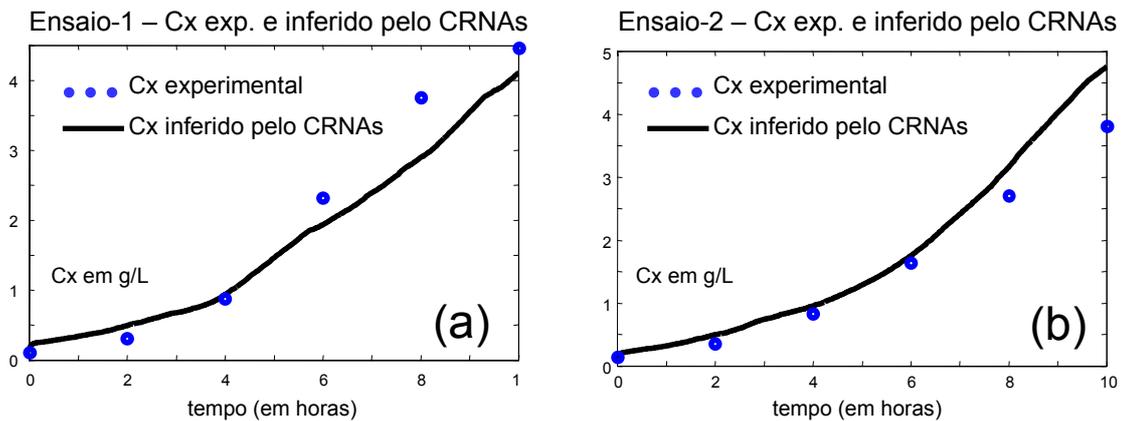


Figura 5.8 Visualização do resultado da inferência da concentração celular (C_x) do comitê (CRNAs) juntos dos dados experimentais: (a) do ensaios um e (b) do ensaio dois.

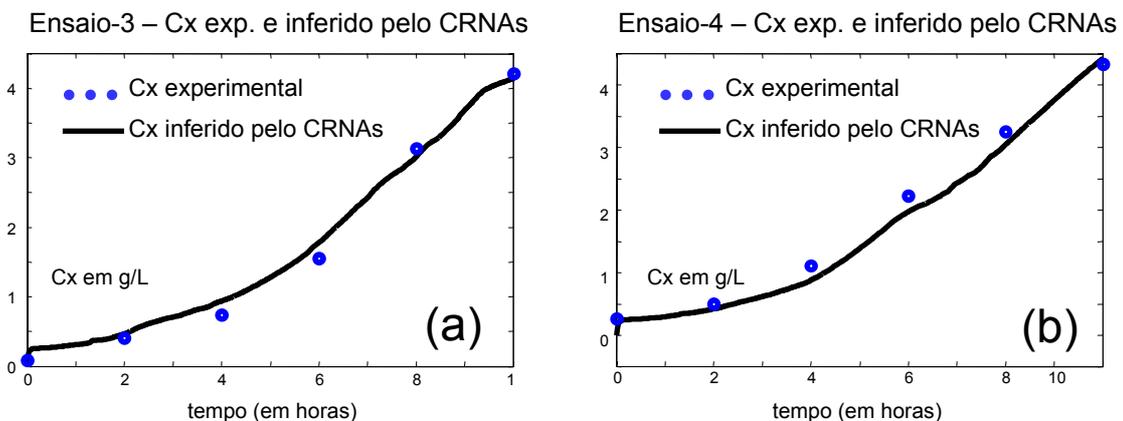


Figura 5.9 Visualização do resultado da inferência da concentração celular (C_x) do comitê (CRNAs) juntos dos dados experimentais: (a) do ensaios três e (b) do ensaio quatro.

O comitê de redes neurais (CRNAs) mostrou-se eficiente na inferência da concentração celular (C_x) nos ensaios três e quatro (Figuras 5.9a e 5.9b), mas nos ensaios um e dois (Figuras 5.8a e 5.8b) os erros foram maiores. Foi observado, no ensaio um, o fato do valor atual de $y_{CO_2, on-line}$, extrapolava para maior os valores da faixa treinada pelo comitê (CRNAs). No caso do ensaio dois, extrapolava para menor os valores da faixa treinada pelo comitê (CRNAs). Isto já foi explicado no item 3.2.7 da seção 3, inclusive é mostrado na Figura 3.11, nos dois ensaios as regiões onde ocorreram as extrapolações nos ensaios um e dois. Nos ensaios três e quatro não houve nenhuma ocorrência do valor atual de $y_{CO_2, on-line}$, ficar fora da faixa treinada pelo comitê (CRNAs), e por isso, suas inferências foram boas, uma

vez que o comitê foi treinado nestas regiões, que vai do menor até o maior valor, da y_{CO_2} da base de dados Seleção de Ensaio (BD SE).

As observações de que a resposta da inferência do comitê (CRNAs) tinha uma tendência final correta, necessitando apenas de *aproximações* do valor inferido ao valor medido da concentração celular (Cx), e que existiam *incertezas* nos limites de detecção das regiões de extrapolações levou a escolha de um sistema *fuzzy* baseado em regras (SFBR). Como o comitê (CRNAs) baseava no conhecimento do processo de forma empírica, em situações ocorridas em ensaios passados, optou-se pela utilização de um raciocínio lógico, nas regiões de extrapolação, baseada na estequiometria de balanços de espécies atômicas dos elementos do ensaio *on-line*, tendo assim uma base fenomenológica. O lógico teve como base o balanço de espécies atômicas para a forma geral da reação de crescimento celular e formação de produto. A Equação 2.13 (item 2.2.2.3) representa a estequiometria que simplifica o complexo fenômeno, repetida neste ponto da dissertação para facilitar o entendimento do raciocínio e sua utilização.



onde: $CH_a O_b N_c$ é a composição elementar da célula

$CH_l O_m N_n$ é a composição elementar da fonte de carbono (aminoácidos)

$CH_p O_q N_r$ é a composição elementar do produto

O mecanismo de inferência do sistema *fuzzy* (SFBR) foi constituído de cinco regras que utilizou os seguintes os seguintes elementos: o teor de CO_2 nos gases efluentes ($K y_{CO_2}$), o quociente respiratório (RQ) que é a razão do teor de CO_2 ($K y_{CO_2}$) pela concentração do oxigênio consumido pelas células (γy_{O_2}) e a composição elementar da célula ($\alpha CH_a O_b N_c$). O sistema *fuzzy* (SFBR) tiveram duas partições de entrada dinâmicas (muda com o tempo), com conjuntos *fuzzy* dos componentes da evolução da fração molar de CO_2 (y_{CO_2}) nos gases efluentes e do quociente respiratório (RQ), e como única partição de saída estática (não muda com o tempo) o valor da correção em Cx inferido pelo comitê (CRNAs). A escolha dos componentes de entrada do sistema *fuzzy* (SFBR) foi o fato de estarem intimamente ligada à dinâmica de crescimento celular, sendo assim, uma boa base fenomenológica para ser utilizada para medir uma variação no crescimento celular (Cx).

Assim foi feita a composição do sistema híbrido inteligente (SHI), que inferiu o valor do crescimento celular (Cx) por um comitê (CRNAs) de forma empírica e nas regiões de extrapolação, o valor inferido pelo comitê recebeu uma correção do sistema *fuzzy* (SFBR), pelo raciocínio lógico com base fenomenológica.

A comparação dos erros, medidos entre os valores dos dados experimentais da concentração celular (C_x) para os valores inferidos de C_x , tanto pelo comitê (CRNAs), como para o sistema híbrido (SHI), são mostrados para o ensaio um na Figura 5.10a e para o ensaio 2 na Figura 5.10b.

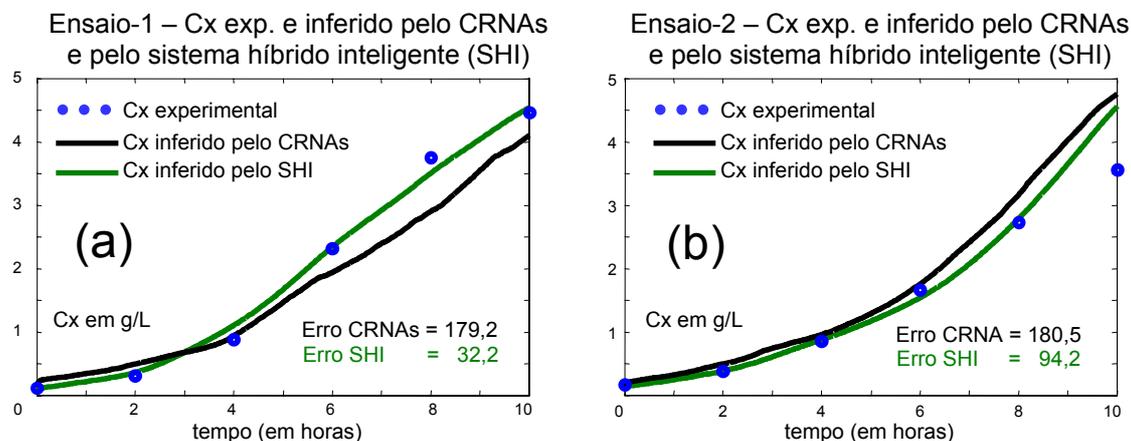


Figura 5.10 Comparação dos erros, medidos pela diferença entre os dados experimentais da concentração celular (C_x) e os C_x inferidos pelo comitê (CRNAs) e pelo sistema híbrido (SHI): (a) para o ensaios um e (b) para o ensaio dois.

5.3.1 Discussões dos resultados

A utilização do comitê de redes neurais artificiais (CRNAs) mostrou-se eficiente para a inferência da concentração microbiana (C_x) durante o cultivo aeróbio de *B. megaterium* para produção da enzima penicilina G acilase. Contudo, suas previsões apresentavam-se tendenciosas quando os dados de fração molar de CO_2 (y_{CO_2}) nos gases efluentes medidos on-line extrapolavam os limites que vai do menor até o maior valor, da y_{CO_2} da base de dados Seleção de Ensaio (BD SE). Note-se que o comitê (CRNAs) não foi treinado nestas regiões, e esperava-se dele, um certo grau de extrapolação, o que não ocorreu. A utilização de uma correção por um sistema *fuzzy* (SFBR) para as previsões do comitê (CRNAs), empregando o quociente respiratório (RQ) com base em considerações fenomenológicas, conseguiu superar essa restrição. O algoritmo resultante, sistema híbrido inteligente (SHI), mostrou-se mais eficiente e robusto, suportando inclusive algum grau de extrapolação em seus dados de entrada, medidos em tempo real.

6. CONCLUSÕES E SUGESTÕES.

As principais dificuldades encontradas no controle para o processo de produção de PGA, por microrganismo selvagem *Bacillus megaterium* e dos processos biológicos em geral, são a variabilidade dos parâmetros cinéticos e a limitada disponibilidade de informação *on-line*. Os resultados da arquitetura do sistema de supervisão e controle proposto nesta dissertação, utilizando um controlador adaptativo e dinâmico, mostraram que essas dificuldades podem ser superadas.

As metodologias da Inteligência Computacional (IC) utilizadas nessa abordagem cumpriram com os objetivos determinados.

O comitê de redes neurais (CRNAs), composto de redes neurais “especialistas” sobre ensaios considerados significativos e consistentes para o bioprocessos, e de um mecanismo de ponderação, capturou o conhecimento empírico do processo de forma adequada. Assim, os resultados da inferência de crescimento celular (Cx) pelo comitê (CRNAs) foram de boa qualidade, principalmente até o final da fase exponencial. Na fase estacionária, apesar das variações, o valor inferido da concentração celular acompanhou a tendência. Note-se que a inferência de crescimento celular (Cx), ou a velocidade específica de crescimento (μ), foi utilizada neste trabalho para definir os momentos de início e fim da alimentação de meio suplementar, e das condições dessa vazão de alimentação em função do tempo. Assim, um erro pequeno na inferência do crescimento celular (Cx) não comprometeu a qualidade do desempenho da arquitetura proposta.

O sistema *fuzzy* baseado em regras (SFBR), que tratou os erros que causavam desvios na inferência de Cx pelo comitê (CRNAs) nas regiões de extrapolação de fração molar de CO₂ (y_{CO_2}) capturou conhecimento fenomenológico do processo também de forma adequada. Os resultados da correção pelo sistema *fuzzy* (SFBR) na inferência do crescimento celular (Cx) feita pelo comitê (CRNAs) mostraram ser consistentes.

O sistema híbrido inteligente (SHI), composto das metodologias de inteligência computacional (IC) descritas acima mostrou sua importância ao capturar conhecimentos distintos do processo. Neste trabalho os conhecimentos distintos foram utilizados de forma complementar, e permitindo um melhor desempenho, com eventuais deficiências de uma metodologia (CRNAs) sendo complementadas por outra (SFBR). Em trabalhos futuros, outras formas podem ser experimentadas: conjuntas, excludentes e/ou mistas.

A aplicação de um *softsensor* baseado em metodologias de inteligência computacional, como o utilizado neste trabalho, é uma boa opção para a inexistência, custo

proibitivo ou inadequação de sensores para a medição em linha de algumas das principais variáveis de estado. Sendo utilizada de forma adequada, pode assegurar uma monitoração e controle estáveis e reprodutíveis, da operação do biorreator, aumentando a competitividade do processo.

As variáveis frações molares de CO_2 (y_{CO_2}) e O_2 (y_{O_2}) nos gases efluentes, velocidade de agitação e quociente respiratório (RQ), utilizadas na arquitetura proposta, mostraram ser escolhas certas porque estão intimamente ligadas à dinâmica de crescimento celular (Cx). Nesse ponto é preciso ressaltar a importância dos dois filtros desenvolvidos neste trabalho. O filtro “alisa pós” suavizou os sinais ruidosos dos sensores das variáveis descritas acima, e permitiu a utilização correta dos mesmos pelas metodologias de inteligência computacional (IC). O filtro “escada” usou a derivada de y_{CO_2} , alisada pelo filtro “alisa pós”, do ensaio *on-line*, para construir uma partição *fuzzy* de entrada adaptativa. A partir de um conjunto de entrada, e utilizando lógica *fuzzy*, puderam-se definir os momentos de início e fim de vazão alimentação do meio suplementar. Assim, em trabalhos futuros, filtros podem ser uma opção tanto para fins de descrever melhor o comportamento da variável monitorada e eliminar as distorções conhecidas dos dados do sensor (ruído identificado), como para preparar variáveis para serem utilizadas por metodologias de inteligência computacional (IC).

Os ganhos que uma abordagem tecnológica, como a desenvolvida nesta dissertação, pode não ser tão significativos quanto os das abordagens biológica ou de engenharia. A abordagem biológica ou bioquímica pode melhorar significativamente o rendimento na produção de metabólitos secundários. A engenharia genética pode introduzir incrementos de 50 vezes em certos processos. A abordagem de engenharia pode conseguir melhoramentos na ordem dos 50% (MORGAN, 1992). Os ganhos de desempenho introduzidos pela abordagem tecnológica são complementares, sem dúvida, mas podem ser o diferencial em um ambiente competitivo, como o da indústria de base biotecnológica (a chamada Biotecnologia branca).

A sugestão básica, suporte para as outras sugestões, para a continuidade desta linha de pesquisa e trabalho é a escolha de um processo biológico com variáveis operacionais conhecidas e com necessidades de operação em batelada alimentada (“*fed-batch*”) do biorreator. Pode ser o processo de produção de PGA por microrganismo clonado de *Bacillus megaterium*, um dos projetos do grupo de pesquisa da UFSCar, desde que as variáveis operacionais desse processo tenham sido dominadas. A importância do conhecimento das variáveis operacionais decorre da necessidade de uma comparação limpa entre as metodologias escolhidas para representar o processo, impedindo, nesta comparação, uma possível contaminação dos

resultados por uma mudança nas condições operacionais. Outras sugestões, além das já mencionadas nesta seção, são:

- construir o módulo *Núcleo Controlador*;
- avaliar outras variáveis de entrada, como pH, verificando as regiões onde estas variáveis poderiam trazer informações adicionais ao processo; e
- utilizar outras metodologias da inteligência computacional (computação evolucionária, computação quântica e etc.), a fim de obter o máximo de suas potencialidades.

Referências Bibliográficas

- AGRAWAL R.; DEEPIKA, N. U.; JOSEPH R. Strain improvement of *Aspergillus* sp. and *Penicillium* spp. by induced mutation for biotransformation of pinene to verbenol. Biotechnol. Bioeng., v.63, 1999.
- ARAÚJO, M.L.G.C. Estudo cinético do processo de produção de Cefalosporina C com células livres e imobilizadas de Cefalosporina acremonium ATCC 48272. São Carlos, Depto Engenharia Química, UFSCar, 1996.
- BAILEY, J. E.; OLLIS, D. F. Biochemical engineering fundamentals. New York, USA: McGraw-Hill Book Company, 1986.
- BAGGALEY, K. H. *et al.* Chemistry and biosynthesis of clavulanic acid and another clavans. Natural Product Report, v.14, n. 4, 1997.
- BASTIN, G.; MORRIS, A.J. Proposal for a Research Programme in Bioprocess Science, Part 2: Modelling and Control, Biotech Forum Europe, v. 9, n. 11, 1991.
- BERAZAÍN, J. G. H. Q. Estudo sobre a utilização do soro de queijo na produção de penicilina G acilase por *Bacillus megaterium* ATCC 14945. 97f. Dissertação de Mestrado em Engenharia Química, UFSCar, São Carlos, 1997.
- BERGEY; S. Manual of Determinative Bacteriology, USA, 1974.
- BEZDEK, J. What is computational intelligence, in J.M. Zurada, R. J.; Marks II, C.; Robinson, J. Computational intelligence: imitating life, New York: IEEE Press, 1994.
- BONISSONE, P. P. *et al.* Hybrid soft computing systems: industrial and commercial applications, Proceedings of the IEEE, v. 87, n. 9, 1999.
- BORDENEUVE-GUIBE, J.; VAUCORET, C. Robust multivariable predictive control: an application to an industrial test stand, IEEE Control Systems Magazine, v. 21, n. 2, 2001.
- BLANCH, H.W.; CLARK, D.S. Biochemical Engineering. New York, USA, Marcel Dekker, 1997.
- BRUGGINK, A. *et al.* Penicillin Acylase in the Production of β -Lactam Antibiotics. Organic Process Research & Development, v. 2, n. 2, 1998.
- CARLEYSMITH, S. W. Advances in Biotechnological Processes, v. 3, New York, 1984.
- CHIPPERFIELD, A.; FLEMING, P. Multiobjective gas turbine engine controller design using genetic algorithms, IEEE Transactions on Industrial Electronics, v. 43, n. 5, 1996.
- COELHO, L. S. Metodologias de inteligência artificial aplicadas ao projeto de sistemas de controle e automação industrial. Anais do ISA SHOW BRASIL, São Paulo, SP, 2001.
- CRAVEN, P. J. *et al.* Multivariable neurofuzzy control of an autonomous underwater vehicle, Integrated Computer-Aided Engineering, v. 6, n. 4, 1999.
- CRUZ, A. J. G. Estudo de produção de cefalosporina C com acompanhamento em linha de variáveis do processo. 2000. 263f. Tese de Doutorado em Engenharia Química, UFSCar, São Carlos, 2000.
- DOCHAIN, D. Design of adaptive linearizing controllers for non-isothermal reactors. International Journal of Control, v. 59, 1994.
- DUCH, W. Towards comprehensive foundations of computational intelligence. Duch, W; MANDZIUK, J. Challenges for computational intelligence, New York: Springer, 2007.

- DUMITRESCU, D. *et al.* Fuzzy sets and their application to clustering and training. Boca Raton: CRC Press (International Series on Computational Intelligence), 2000.
- EILERS, P. H. C. A perfect smoother. Analytical Chemistry, v. 75, 2003.
- FERREIRA, E.M.F.C. Identificação e controle adaptativo de processos biotecnológicos. 299f. Tese de Doutorado (Engenharia Química), FEUP, Portugal, 1995.
- FERREIRA, D.S. *et al.* Inferência de estado durante o cultivo de *Bacillus megaterium* para produção de Penicilina G Acilase através de consórcio de redes neurais, Anais do XVI Simpósio Nacional de Bioprocessos, Curitiba, 2007.
- GE, S. S. *et al.* Adaptive control of partially known nonlinear multivariable systems using neural networks, Proceedings of the IEEE International Symposium on Intelligent Control, Mexico City, Mexico, 2001.
- GERSON, D. *et al.* Substrate Concentration Control in Bioreactors. Biotechnol. Gen. Eng. Rev., v. 6, n. 67, 1988.
- GOLEA, N.; GOLEA, A. Fuzzy adaptive control of multivariable nonlinear systems, Proceedings of the IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2002, v. 1, Honolulu, HI, USA, 2002.
- HADDAD, W. M. *et al.* Direct adaptive control for discrete-time nonlinear uncertain dynamical systems, Proceedings of the American Control Conference, Anchorage, AK, USA, v. 3, 2002.
- HAYKIN, S. Redes neurais: princípios e prática. Brasil: Bookman, 2001.
- HIMMELBLAU, D.M. Process analysis by statistical methods. New York: John Wiley & Sons, 1970.
- HOJO, O. Estudo sobre a produção de penicilina G acilase por *Bacillus megaterium* ATCC 14945. 1997. 123f. Tese de Doutorado em Engenharia Química, UFSCar, 1997.
- HOU, J.P.; POOLE, J.W. The amino acid nature of ampicillin and related penicillins. J. Pharm. Sci., v. 58, 1969.
- ILLANES, A. Biotecnología de Enzimas. Valparaíso, Chile, Ediciones Universitarias de Valparaíso, 1994.
- ILLANES, A. *et al.* Production of penicillin acylase from *Bacillus megaterium* in complex and defined media. Process Biochemistry, v.29, 1994.
- ISAACS, S. *et al.* Use of orthogonal collocation approximation for parameter identification and control optimization of a distributed parameter biological reactor. Biotechnol. Prog., v. 3, n. 4, 1987.
- JAFAROV, E. M. Design of sliding mode control for multi-input systems with multiple state delays, Proceedings of the American Control Conference, Chicago, IL, USA, v. 2, 2000.
- JAMSHIDI, M. *et al.* Robust control with genetic algorithms, Boca Raton, CRC Press, 2002.
- KAFAROW, W.W. *et al.* Modelling of biochemical reactors. Moskow: Lesnaja Promyshlenost, 1979.
- KANG, J.H. *et al.* Expression of Penicillin G Acylase gene from *Bacillus Megaterium* ATCC 14945 in *Escherichia coli* and *Bacillus Subtilis*. Journal of Biotechnology, v. 17, 1991.

- KAWANAKA, H. *et al.* Acquisition of fuzzy control rules for a mobile robot using genetic algorithm, Proceedings of the 6th International Workshop on Advanced Motion Control, Nagoya, Japan, 2000.
- KING, R. E. Computational intelligence in control engineering. New York: M. Dekker, 1999.
- KLIR, G.; YUAN, B. “Fuzzy sets and fuzzy logic : theory and applications”. Prentice-Hall, 1995.
- LARSEN, P. M. Industrial applications of fuzzy logic control. International Journal of Man-Machine Studies, v. 12, n. 1, 1980.
- LEHNINGER, A.L. *et al.* Princípios de Bioquímica, 4ª ed., Editora Sarvier, 2006.
- LUGER, G. Artificial intelligence: structures and strategies for complex problem solving. Addison Wesley Longman, 2002.
- LUKASIEWICZ, J. Philosophical remarks on many-valued systems of propositional logic. (Studies in Logic and the Foundations of Mathematics), North Holland, 1970.
- MANDANI, E. H.; ASSILIAN, S., An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies, v.7, 1975.
- MAMDANI, E. H. Application of fuzzy logic to approximate reasoning using linguist systems. Fuzzy Sets and Systems, v. 26, 1977.
- MAVROVOUNIOTIS, G. M. *et al.* Transport of interfaces involving surface-active substances. Tese de PhD, Massachusetts Institute of Technology, Cambridge, MA, USA, 1989.
- MCCULLOCH, W.S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity, Bullwtn of Mathematical Biophysics, v. 5, 1973.
- MITCHELL, T. M. Toward combining empirical and analytical methods for inferring heuristics. Technical Report LCSR-TR-27, Rutgers University: Laboratory for computer science research, 1982.
- MORGAN, M. J. Spatial filtering precedes motion detection. Nature (London), 1992.
- MOSER, A. Bioprozeßtechnik. New York: Springer Verlag, 1981.
- MURAO, S. *et al.* Process for preparing 6-aminopenicilanic acid by Bacillus megaterium. U.S. Patent 3, 1964.
- NELLES, O. Nonlinear System Identification – Germany: Springer, 2001.
- NIELSEN, J.; VILLADSEN, J. Bioreaction engineering principles. New York: Plenum Press, 1994.
- NIELSEN, J.; VILLADSEN, J.; LIDÉN, G. Bioreaction engineering principles. New York: Plenum Press, 2002.
- NUCCI, E. R. *et al.* Comparing the performance of multi layer perceptrons networks and neuro-fuzzy systems for on-line inference of *Bacillus megaterium* cellular concentrations. Bioprocess and Biosystem Engineering, v. 30, 2007.
- OLIVEIRA, L. L.; CARRAPICO, O. L. Multivariable control increases coke unit, refinery throughput, Hydrocarbon Processing, v. 77, n. 7, 1998.
- OLIVEIRA, J. A. Modelagem matemática do processo de produção de leveduras de panificação – Um estudo de caso. 103p. Dissertação de Mestrado em Engenharia Química, UFSCar, São Carlos, 2006.

- PARMAR, A. *et al.* Advances in Enzymatic Transformation of Penicillins to 6-Aminopenicillanic Acid (6-APA). Biotechnology Advances, v. 18, 2000.
- PAUL, G. C.; KENT, C. A.; THOMAS, C. R. Viability testing and characterisation of germination of fungal spores by automatic image analysis. Biotechnol. Bioeng. v. 42, 1993.
- PEDRYCZ, W. Fuzzy modelling: paradigms and practice. Kluwer: Academic Press, 1996.
- PEDRYCZ, W.; GOMIDE, F. An introduction to fuzzy sets: Analysis and design. Cambridge: MIT Press, 1999.
- PINOTTI, L. M. Produção e purificação de Penicilina G Acilase. 2003. Tese de Doutorado em Engenharia Química, UFSCar, São Carlos, 2003.
- PONCIA, G.; BITTANTI, S. Multivariable model predictive control of a thermal power plant with built-in classical regulation, International Journal of Control, v. 74, n. 11, 2001.
- RAJENDHRAN, J.; GUNASEKARAN, P. Recent biotechnological interventions for developing improved Penicilin G Acylase. Journal of Bioscience and Bioengineering, v. 97, n. 1, 2004.
- REHM, H.-J; REED, G. Biotechnology, Fundamentals of Biochemical Engineering, Editor H. Brauer , v.2, 1985.
- REZENDE, S.O. Sistemas inteligentes: Fundamentos e aplicações. Barueri-SP: Ed. Manole, 2005.
- RODRIGUES, M. I.; IEMMA, A. F. Planejamento de experimentos e otimização de processos: uma estratégia seqüencial de planejamentos. Campinas-SP: Casa do Pão Editora, 2005.
- ROELS, J.A. Biotechnol. Bioeng., 1980.
- ROELS, J.A.; KOSSEN, N.W.F. On the modeling of microbial metabolism. In: BULL, M.J. Progress in industrial microbiology, Amsterdam: Elsevier., v. 14, 1978.
- ROYCE, P.N. A discussion of recent developments in fermentation monitoring and control from a practical perspective, Critical Reviews in Biotechnology, v. 13, n. 2, 1993.
- SAAJ, C. M.; BANDYOPADHYAY, B. Output feedback sliding mode control for MIMO discrete time systems, Proceedings of the 15th IFAC World Congress, Barcelona, Spain, 2002.
- SANA, S.; RAO, V. S. Robust control of input limited smart structural systems, IEEE Transactions on Control Systems Technology, v. 9, n. 1, 2001.
- SAVIDGE, T. A.; COLE, M. Penicillin Acylase (bacterial). Methods in Enzymology, v. 43, 1975.
- SAVIDGE, T.A. Enzymatic conversions used in the production of Penicillins and Cephalosporins. In: VANDAME, E.J. (Ed). Biotechnology of Industrial Antibiotics. USA: Marcel Dekker, 1984.
- SENG, T. L. *et al.* Tuning of a neuro-fuzzy controller by genetic algorithm, IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics, v. 29, n. 2, 1999.
- SHIOYA, S., SHIMIZU, K., & YOSHIDA, T. Knowledge based design and optimization of bioprocess systems. Journal of Bioscience and Bioengineering, v. 87, 1999.

- SILVA, R. G. Inferência de variáveis do processo de produção de penicilina G acilase por *Bacillus megaterium* ATCC 14945. 2003. 266f. Tese em Doutorado em Engenharia Química, UFSCar, São Carlos, 2003.
- SONNLEITNER, B. On line measurement and validation in biotechnological processes. On Bioprocess Modelling and Control, Newcastle upon Tyne, 1991.
- TAKAGI, T.; SUGENO, M. Derivation of fuzzy control rules from human operator's control actions. In Proceedings of the IFAC Symp. on Fuzzy Information, Knowledge Representation and Decision Analysis, Marseilles, France, 1983.
- TAKAGI, T.; SUGENO, M., Fuzzy Identification of Systems and its application to modeling and control. IEEE Transaction on Systems, Man and Cybernetics, v. 15, 1985.
- TAN Z-J *et al.* Pattern formation on nonuniform surfaces by correlated random sequential absorptions, Physical Review, v. 65, 2002.
- TANAKA, K.; IKEDA, T. Absolute stability conditions in a fuzzy phase-lead compensation and their extension to MIMO systems, IEEE Transactions on Industrial Electronics, v. 45, n. 2, 1998.
- VARY, P. S. Prime time for *Bacillus megaterium*. Microbiology, v.140, 1994.
- VARY, P. S. *et al.* *Bacillus megaterium*- from simple soil bacterium to industrial protein production host. Applied Microbiology and Biotechnology, v. 76, 2007.
- VOTRUBA, J. Acta Biotechnol., v. 2, n. 2, 1982.
- YU, A. G. A.; ORLOV, F. Frequency adaptive control of multivariable plants. Proceedings of the 15th IFAC World Congress, Barcelona, Spain, 2002.
- YU; D. L. *et al.* Model predictive control of a chemical process based on an adaptive neural network, Proceedings of the 15th IFAC World Congress, Barcelona, Spain, 2002.
- WERNER, R. G.; LANGLOUIS-GAU, H. Arzneim-Forsch, v. 39, 1989.
- ZADEH, L. A. Fuzzy sets. Information and Control, v.8, 1965.
- ZADEH, L. A. Outline of a new approach to the analysis of complex system and decision process. IEEE Trans. on Systems, Man, and Cybernetics, v. 3, 1973.
- ZADEH, L. A. Soft computing and fuzzy logic, IEEE Software, v. 11, n. 6, 1994.
- ZHANG, X.-C. *et al.* Functional state modeling and fuzzy control of fed-batch aerobic baker's yeast process. J. Biotechnol., v. 37, 1994.
- ZHANG, M. *et al.* Genetic algorithm for input/output selection in MIMO systems based on controllability and observability indices, Electronics Letters, v. 38, n. 19, 2002.
- ZIMMERMANN, H. J. Fuzzy set theory and its applications. Kluwer: Academic Publishers, 1991.

APÊNDICE A – Descrição e dados dos ensaios

a) Ensaios utilizados para validação do módulo Inferência de Cx

Os ensaios identificados por 1, 2, 3 e 4 foram realizados no biorreator de bancada marca Applikon, tipo tanque agitado e aerado (STR) de 2 litros de volume útil. Nos ensaios 1,2,3 foi adicionado ao meio, soro hidrolisado por tripsina imobilizada em quitosana, e no ensaio 4, soro de queijo.

Ensaio um

Ensaio realizado em 06 de dezembro de 2006, e a forma de operação do biorreator foi batelada. O meio continha uma solução de aminoácidos (10,09 g/L). O oxigênio dissolvido (OD) foi controlado em 20% durante todo cultivo.

t	Cx	AE
0	0,11	-
2	0,31	-
4	0,88	-
6	2,32	63 ± 3
8	3,75	171 ± 8
10	4,46	274 ± 14
12	5,28	276 ± 2
14	5,25	283 ± 5
16	5,02	258 ± 3
18	4,76	270 ± 3
20	4,64	293 ± 14
22	3,95	281 ± 14
24	3,88	279 ± 14

(a)

Unidades: t em h
Cx em g/L
AE em UI/L

t	aa+					aa-		aa	AFA
	Ala	Arg	Lys	Ser	Thr	Asp	Glu	Totais	
0	0,70	0,79	0,89	0,82	0,33	2,38	4,18	10,09	3,69
2	0,66	0,84	0,86	0,73	0,31	2,32	4,18	9,90	3,69
4	0,48	0,69	0,82	0,37	0,22	2,24	3,98	8,79	3,67
6	0,05	0,20	0,56	0,00	0,00	2,26	3,62	6,68	3,32
8	0,00	0,00	0,07	0,00	0,00	1,65	3,43	5,15	2,55
12	0,00	0,00	0,00	0,00	0,00	1,59	3,40	4,99	0,06
14	0,00	0,00	0,00	0,00	0,00	1,08	2,32	3,40	0,00
18	0,00	0,00	0,00	0,00	0,00	1,24	2,76	3,10	0,00
24	0,00	0,00	0,00	0,00	0,00	1,24	1,12	2,36	0,00

(b)

Unidades: t em h
aa em g/L
AFA em g/L

Tabela 1 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA).

Ensaio dois

Ensaio realizado em 12 de dezembro de 2006, e a forma de operação do biorreator foi batelada. O meio continha uma solução de aminoácidos (8,31 g/L). O oxigênio dissolvido (OD) foi controlado em 20% até as oito horas e após este horário foi controlado em 5% até o fim do cultivo.

t	Cx	AE
0	0,14	-
2	0,35	-
4	0,83	-
6	1,64	48 ± 2
8	2,70	92 ± 5
10	3,44	186 ± 5
12	4,24	266 ± 3
14	4,11	293 ± 4
16	3,81	284 ± 14
18	3,72	271 ± 13
20	3,41	268 ± 13
22	3,41	262 ± 13
24	3,35	240 ± 12

(a) Unidades: t em h
Cx em g/L
AE em UI/L

t	aa+					aa-		aa	AFA
	Ala	Arg	Lys	Ser	Thr	Asp	Glu	Totais	
0	0,61	0,71	0,67	0,64	0,26	2,19	3,27	8,35	3,11
2	0,57	0,72	0,70	0,65	0,26	2,21	3,39	8,50	3,21
4	0,58	0,64	0,64	0,59	0,24	2,12	3,22	8,03	3,08
6	0,52	0,60	0,66	0,55	0,21	2,15	3,24	7,93	2,58
8	0,48	0,09	0,63	0,01	0,00	2,07	3,49	6,77	1,96
12	0,16	0,05	0,30	0,00	0,00	1,28	2,21	4,00	1,71
18	0,00	0,00	0,10	0,00	0,00	1,87	1,97	3,94	1,28

(b) Unidades: t em h
aa em g/L
AFA em g/L

Tabela 2 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA).

Ensaio três

Ensaio realizado em 19 de dezembro de 2006, e a forma de operação do biorreator foi batelada. O meio continha uma solução de aminoácidos (9,06 g/L). O oxigênio dissolvido (OD) foi controlado em 10% durante todo cultivo.

t	Cx	AE
0	0,46	-
2	0,99	-
4	1,89	-
6	3,23	43 ± 2
8	3,78	119 ± 6
10	4,70	215 ± 5
12	4,73	265 ± 5
14	4,24	283 ± 4
16	4,23	297 ± 4
18	4,05	288 ± 3
20	3,74	271 ± 1
22	3,69	243 ± 12
24	3,81	236 ± 11

(a) Unidades: t em h
Cx em g/L
AE em UI/L

t	aa+					aa-		aa	AFA
	Ala	Arg	Lys	Ser	Thr	Asp	Glu	Totais	
0	0,64	0,71	0,76	0,72	0,28	2,41	3,54	9,06	3,57
2	0,57	0,63	0,72	0,66	0,25	2,32	3,29	8,44	3,13
4	0,51	0,61	0,68	0,60	0,23	2,24	3,14	8,02	3,10
6	0,48	0,51	0,70	0,57	0,17	2,24	3,30	7,96	2,97
8	0,45	0,25	0,76	0,42	0,00	2,53	3,81	8,23	2,73
12	0,12	0,02	0,56	0,05	0,00	2,20	3,51	6,45	1,80
14	0,01	0,00	0,26	0,00	0,00	1,90	3,08	5,25	0,77
20	0,00	0,00	0,00	0,00	0,00	1,68	2,77	4,45	0,00

(b) Unidades: t em h
aa em g/L
AFA em g/L

Tabela 3 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA).

Ensaio quatro

Ensaio realizado em 6 de março de 2007, e a forma de operação do biorreator foi batelada. O meio continha uma solução de aminoácidos (9,26 g/L). O oxigênio dissolvido (OD) foi controlado em 20% até as oito horas e após este horário foi controlado em 5% até o fim do cultivo.

t	Cx	AE
0	0,26	-
2	0,50	-
4	1,11	23 ± 2
6	2,22	109 ± 6
8	3,25	131 ± 6
10	4,33	285 ± 11
12	4,92	346 ± 10
14	4,14	327 ± 9
16	4,24	319 ± 9
18	4,12	292 ± 9
20	3,9	279 ± 8
22	3,92	261 ± 8
24	3,88	269 ± 8

(a)

Unidades: t em h
Cx em g/L
AE em UI/L

t	aa+					aa-		aa	AFA
	Ala	Arg	Lys	Ser	Thr	Asp	Glu	Totais	
0	0,61	0,79	0,75	0,69	0,29	2,26	3,87	9,26	3,29
2	0,55	0,76	0,76	0,59	0,27	2,27	3,89	9,09	3,13
4	0,37	0,61	0,69	0,21	0,17	2,32	3,86	8,24	2,88
6	0,01	0,11	0,46	0,00	0,00	1,79	3,55	5,93	2,31
8	0,00	0,00	0,00	0,00	0,00	1,70	3,52	5,23	1,79
12	0,00	0,00	0,00	0,00	0,00	1,32	2,51	3,83	0,68
14	0,00	0,00	0,00	0,00	0,00	1,14	2,12	3,26	0,00
16	0,00	0,00	0,00	0,00	0,00	1,07	2,13	3,20	0,00
18	0,00	0,00	0,00	0,00	0,00	1,18	2,18	3,36	0,00
24	0,00	0,00	0,00	0,00	0,00	1,14	1,94	3,08	0,00

(b)

Unidades: t em h
aa em g/L
AFA em g/L

Tabela 4 Dados experimentais de: (a) crescimento microbiano (Cx) e produção de PGA (AE) e (b) consumo de aminoácidos (aa+,aa- e aa Totais) e de ácido fenilacético(AFA).

b) Ensaio utilizado para validação do módulo Controle da Vazão

Os ensaios identificados por 5, 6, 7, 8 e 9 foram realizados no biorreator de bancada New Brunswick BIOFLO II-C, tipo tanque agitado e aerado (STR) de 5 litros de volume útil. A aeração foi de 2,0 L/min ao longo do cultivo. O oxigênio dissolvido (OD) foi controlado em 20% durante ao longo do cultivo.

Ensaio cinco

A forma de operação do biorreator foi batelada. Foi adicionada ao meio uma solução de aminoácidos (10,6 g/L) mais de ácido fenilacético (2,8 g/L).

t	Cx	AE	Cs		
			aa	AFA	Totais
0	0,28	0	10,60	2,80	13,40
2	0,30	0	10,00	2,80	12,80
4	0,70	14	9,50	2,69	12,19
6	2,00	19	6,20	2,67	8,87
8	3,00	61	4,10	2,07	6,17
10	3,70	127	3,10	1,24	4,34
13	4,50	175	1,30	0,13	1,43
17	4,30	158	0,72	0,13	0,85
20	4,30	164	0,70	0,13	0,83

Unidades: t em h
Cx em g/L
AE em UI/L
aa em g/L
AFA em g/L
Cs em g/L

Tabela 5 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais.

Ensaio seis

A forma de operação do biorreator foi batelada alimentada. Foi adicionada no início uma solução de aminoácidos (11,5 g/L) mais de ácido fenilacético (2,8 g/L). Após oito horas de cultivo, uma solução de aminoácidos de concentração 40 g/L, foi adicionada a uma vazão de alimentação de 1,76 mL/min.

t	Cx	AE	Cs		
			aa	AFA	Totais
0	0,22	0	11,50	2,80	14,30
2	0,30	0	11,10	2,70	13,80
4	0,90	14	10,60	2,62	13,22
6	2,20	25	6,50	2,49	8,99
8	3,30	75	4,70	1,99	6,69
10	4,10	136	3,60	0,46	4,06
12	4,20	166	2,30	0,50	2,80
16	3,90	207	5,20	0,18	5,38
20	3,90	203	4,76	0,20	4,96

Unidades: t em h
 Cx em g/L
 AE em UI/L
 aa em g/L
 AFA em g/L
 Cs em g/L

Tabela 6 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais.

Ensaio sete

A forma de operação do biorreator foi batelada alimentada. No início, foi adicionada uma solução de aminoácidos (10,4 g/L) mais de ácido fenilacético (2,8 g/L). Após cinco horas de cultivo, uma solução de aminoácidos de concentração de 40 g/L, foi adicionada a uma vazão de alimentação de 1,76 mL/min.

t	Cx	AE	Cs		
			aa	AFA	Totais
0	0,37	0	10,80	2,80	13,60
2	0,33	0	10,10	2,70	12,80
4	0,78	22	7,94	2,50	10,44
6	1,77	22	5,93	1,82	7,75
8	2,66	41	6,35	0,90	7,25
10	3,25	130	4,92	0,40	5,32
12	3,20	137	4,27	0,00	4,27
16	3,57	191	3,74	0,00	3,74
21	3,93	175	4,10	0,00	4,10

Unidades: t em h
 Cx em g/L
 AE em UI/L
 aa em g/L
 AFA em g/L
 Cs em g/L

Tabela 7 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais.

Ensaio oito

A forma de operação do biorreator foi batelada alimentada. No início, foi adicionada uma solução de aminoácidos (10,8 g/L) e de ácido fenilacético (2,8 g/L). Após cinco horas de cultivo, uma solução de aminoácidos de concentração de 40 g/L, foi adicionada a uma vazão de alimentação de 1,76 mL/min. Após oito horas de cultivo, soro de queijo com concentração de 176,8 g/L, foi adicionada a uma vazão de alimentação de 1,76 mL/min.

t	Cx	AE	Cs		
			aa	AFA	Totais
0	0,40	0	10,80	2,80	13,60
2	0,50	0	10,10	2,70	12,80
4	0,98	0	9,00	2,65	11,65
6	2,40	23	7,94	2,50	10,44
8	3,00	60	7,90	2,00	9,90
10	4,70	121,8	5,93	1,36	7,29
14	5,90	184	6,35	0,17	6,52
16	5,80	199	4,99	0,17	5,16
20	5,80	215,4	4,10	0,17	4,27

Unidades: t em h
 Cx em g/L
 AE em UI/L
 aa em g/L
 AFA em g/L
 Cs em g/L

Tabela 8 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais.

Ensaio nove

A forma de operação do biorreator foi batelada. No início, foi adicionada ao meio uma solução de aminoácidos (9,20 g/L) e de ácido fenilacético (2,8 g/L).

t	Cx	AE	Cs		
			aa	AFA	Totais
0	0,36	0	9,20	2,80	12,00
2	0,34	0	8,00	2,70	10,70
4	0,74	0	7,30	2,65	9,95
6	2,00	14	5,00	2,50	7,50
8	3,20	48	3,10	2,00	5,10
10	4,00	123	2,50	1,36	3,86
12	4,50	143	2,00	0,40	2,40
16	4,80	157	1,00	0,17	1,17
20	4,50	180	0,70	0,17	0,87

Unidades: t em h
 Cx em g/L
 AE em UI/L
 aa em g/L
 AFA em g/L
 Cs em g/L

Tabela 9 Dados experimentais de: crescimento microbiano (Cx) e produção de PGA (AE) e consumo de aminoácidos (aa), de ácido fenilacético(AFA) e de substrato (Cs) totais.