

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

Inteligência computacional aplicada à modelagem e otimização de bioprocessos

PEDRO LUIZ DA MOTA E AQUINO

SÃO CARLOS - SP

2016

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

Inteligência computacional aplicada à modelagem e otimização de bioprocessos

PEDRO LUIZ DA MOTA E AQUINO

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Química da Universidade Federal de São Carlos como parte dos requisitos à obtenção do título de Doutor em Engenharia Química, área de concentração em Pesquisa e Desenvolvimento de Processos Químicos.

Orientador: Prof. Prof. Dr. Ruy de Sousa Junior

Co-orientador: Dr. Roberto de Campos Giordano

SÃO CARLOS - SP

2016

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária UFSCar
Processamento Técnico
com os dados fornecidos pelo(a) autor(a)

A657i Aquino, Pedro Luiz da Mota e
Inteligência computacional aplicada à modelagem e
otimização de bioprocessos / Pedro Luiz da Mota e
Aquino. -- São Carlos : UFSCar, 2016.
160 p.

Tese (Doutorado) -- Universidade Federal de São
Carlos, 2016.

1. Inteligência computacional. 2. Lógica fuzzy. 3.
Simulated annealing. 4. Redes neurais. 5. Neuro-
fuzzy. I. Título.

MEMBROS DA BANCA EXAMINADORA DA DEFESA DE TESE DE PEDRO LUIZ DA MOTA E AQUINO APRESENTADA AO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA DA UNIVERSIDADE FEDERAL DE SÃO CARLOS, EM 29 DE ABRIL DE 2016.

BANCA EXAMINADORA:




Ruy de Sousa Júnior
Orientador, UFSCar



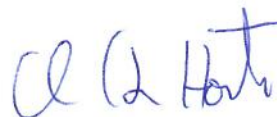
Roberto de Campos Giordano
Coorientador, UFSCar



Otto Domenici Mozzer
VALLÉE



Edson Romano Nucci
UFSJ



Antonio Carlos Luperni Horta
UFSCar

A minha família e esposa eu dedico este trabalho.

Agradecimentos

Aos meus orientadores Ruy de Sousa Junior e Roberto de Campos Giordano por tudo que me ensinaram, pelas criticas que aceleraram meu desenvolvimento e pelo apoio e compreensão nas dificuldades impostas pela vida.

Agradeço aos meus pais, Jairo e Marta, pelo apoio. À minha esposa, Amanda, pelo companheirismo e amor. Aos meus tios, Alice e José, por todo carinho e cumplicidade.

Agradeço colegas de Vallée, Otto, Ricardo, Felipe, Cássia e Donizete pela amizade e apoio ao desenvolvimento do projeto.

Aos colegas de Laboratório por tornar o tempo de trabalho mais agradável.

*“Se não puder voar, corra. Se não puder correr, ande.
Se não puder andar, rasteje, mas continue em frente
de qualquer jeito”.*

Martin Luther King

RESUMO

Este trabalho aborda aplicações de modelagem, metodologias sistemáticas e confiáveis de otimização por busca global, além de outras ferramentas computacionais. Espera-se que métodos de inteligência computacional existentes, codificados em uma ferramenta apropriada para a aplicação da engenharia de processos assistida por computador, resultem em resultados numéricos úteis para a modelagem e otimização de diferentes processos, incluindo-se os processos biotecnológicos (foco deste trabalho). Assim, diferentes tipos de metodologias, apropriadas para aplicações em computador, foram aqui estudadas. Os métodos propostos foram aplicados e avaliados ao desenvolvimento e otimização de meios de cultura para o processo fermentativo do microrganismo *Clostridium novyi* tipo B, além dos processos de fermentação alcoólica e hidrólise enzimática de bagaço de cana, associados à produção de bioetanol (1G e 2G). Desta forma, foi avaliado o potencial de aplicação destas técnicas computacionais aos sistemas biotecnológicos, em diversas abordagens. Mais especificamente, foram realizadas: classificação (“clustering”) de sistemas em regiões cineticamente semelhantes para a produção de etanol celulósico (Etanol 2G) utilizando lógica Fuzzy; estimação por busca global de parâmetros cinéticos do modelo para uma fermentação alcoólica utilizando o algoritmo Simulated Annealing (SA) (Contribuições ao projeto temático FAPESP 2011/51902-9); formulação e otimização do meio de cultura economicamente viável para o *Clostridium novyi* tipo B utilizando a modelagem de dados por neuro-fuzzy seguido de busca global da composição de meio que maximize a produtividade utilizando também o algoritmo SA como ferramenta de busca global (esta etapa do projeto foi realizado em parceria com a empresa farmacêutica veterinária Vallée S.A). As ferramentas computacionais apresentadas neste trabalho se mostraram altamente efetivas para a modelagem e otimização dos bioprocessos estudados."

Palavras-Chave: Inteligência computacional, lógica *Fuzzy*, *Simulated annealing*, redes neurais, *neuro-Fuzzy*, bioprocessos, estimativa de parâmetros, modelagem.

ABSTRACT

This work deals with modeling applications, systematic and reliable optimization methodologies of global search, and other computational tools. It is expected that existing computational intelligence methods, encoded in an appropriate tool for the application of process engineering assisted by computer, can lead to useful numerical results for the modeling and optimization of different processes, including biotechnological processes (focus of this work). Thus, different types of methodologies suitable for computer applications, were studied here. The proposed methodologies were implemented and evaluated for the development and optimization of culture media for the fermentation process of *Clostridium novyi* type B, besides the fermentation process and enzymatic hydrolysis of bagasse associated with the production of bioethanol (1G and 2G). Thus, the potential application of these computational techniques was evaluated to biotechnological systems in different approaches. More specifically, it was performed: Classification of biotechnological systems ("clustering") in kinetically similar regions to produce cellulosic ethanol (2G ethanol) using fuzzy logic; estimation by global search of kinetic parameters to an alcoholic fermentation model using Simulated Annealing algorithm (SA) (Contributions to the thematic project FAPESP 2011 / 51902-9); formulation and optimization of economically viable culture media for *Clostridium novyi* type B using neuro-fuzzy data modeling followed by global search which maximize productivity, also utilizing SA algorithm as a search engine (this step of the project was conducted in partnership with the veterinary pharmaceutical company Vallée SA). The computational tools presented in this work were highly effective for modeling and optimization of the bioprocesses studied.

Keywords: Computational intelligence, Fuzzy logic, Simulated annealing, neural networks, neuro-Fuzzy, bioprocesses, parameter estimation, modeling.

SUMÁRIO

1.	CAPÍTULO 01: INTRODUÇÃO	29
2.	CAPÍTULO 02: REVISÃO BIBLIOGRÁFICA.....	33
2.1.	TÉCNICAS COMPUTACIONAIS	34
2.1.1.	Método de Monte Carlo.....	34
2.1.2.	Método simulated annealing	35
2.1.3.	Redes neurais artificiais.....	38
2.1.4.	Logica Fuzzy	41
2.1.1.	Redes Neuro-Fuzzy	46
2.1.	SISTEMAS BIOTECNOLÓGICOS EM ESTUDO	51
2.1.1.	Fermentação alcoólica.....	51
2.1.2.	Micro-organismos: matérias-primas e processos	58
2.1.3.	Produção de etanol no Brasil	58
2.1.4.	Clostridium novyi.....	60
3.	CAPÍTULO 03: CLASSIFICAÇÃO DE DADOS DO PROCESSO DE HIDRÓLISE ENZIMÁTICA DE BAGAÇO DE CANA-DE-AÇÚCAR: UMA ANÁLISE DE “CLUSTERS” NEBULOSOS.....	62
3.1.	INTRODUÇÃO	65
3.2.	MATERIAIS E MÉTODOS.....	66
3.3.	RESULTADOS E DISCUSSÃO.....	67
3.4.	CONCLUSÃO.....	70
4.	CAPÍTULO 04: ESTIMATIVA DE PARÂMETROS CINÉTICOS DE FERMENTAÇÃO POR S. CEREVISIAE VISANDO À PRODUÇÃO INDUSTRIAL DE ETANOL 1G	71
	ABSTRACT	73
4.1.	INTRODUCTION.....	73
4.1.1.	Production of ethanol in Brazil	73
4.1.2.	Mathematical modeling of bioprocesses	74
4.1.3.	Estimation of model kinetic parameters	80
4.2.	MATERIALS AND METHODS.....	83
4.2.1.	Model	83
4.2.2.	Simulated annealing algorithm	83
4.3.	RESULTS AND DISCUSSION	84

4.4.	CONCLUSIONS.....	90
4.5.	REFERENCES.....	90
5.	CAPÍTULO 05: OPTIMIZATION OF THE PRODUCTION OF INACTIVATED <i>CLOSTRIDIUM NOVYI</i> TYPE B VACCINE USING COMPUTATIONAL INTELLIGENCE TECHNIQUES.....	93
	ABSTRACT.....	96
5.1.	INTRODUCTION.....	97
5.2.	OBJECTIVES.....	99
5.2.1.	General objective.....	99
5.2.2.	Specific objective.....	99
5.3.	NEURO-FUZZY SYSTEMS.....	99
5.3.1.	Fuzzy logic and neural networks.....	100
5.3.2.	Neuro-fuzzy hybrid systems.....	101
5.3.3.	Simulated annealing.....	103
5.4.	METHODOLOGY.....	105
5.4.1.	Microbial cultures.....	105
5.4.2.	Quantification of toxin.....	106
5.4.3.	Medium optimization procedure.....	106
5.4.4.	Validation of the optimization.....	111
5.5.	RESULTS.....	111
5.5.1.	Constructing the data matrix.....	111
5.5.2.	Optimizing toxin productivity.....	114
5.6.	DISCUSSION.....	116
5.7.	CONCLUSIONS.....	117
5.9.	REFERENCES.....	118
6.	APÊNDICE 1: INTEGRAÇÃO COM O SOFTWARE EMSO.....	125
6.1.	INTRODUÇÃO.....	126
6.2.	OBJETIVOS.....	127
6.3.	MATERIAIS E MÉTODOS.....	127
6.3.1.	Simulador.....	127
6.3.2.	Interface Emso MATLAB.....	128
6.4.	RESULTADOS.....	129

7. APÊNDICE 2: Modeling the Kinetics of Complex Systems: Enzymatic Hydrolysis of Lignocellulosic Substrates	131
7.1. REFERÊNCIAS	145

ÍNDICE DE FIGURAS

Figura 1: O Particle swarm optimization (PSO) é um dos exemplos de algoritmo baseado em fenômenos naturais	30
Figura 2: Figura esquemática do neurônio humano.	39
Figura 3: Representação das entradas e saídas das RNA's.	41
Figura 4: Fluxograma simplificado do algoritmo FCM.	46
Figura 5: Sistema Híbrido Incorporado (Mitra e Hayashi, 2000).....	47
Figura 6: Sistema Híbrido Auxiliar (Mitra e Hayashi, 2000)	48
Figura 7: Sistema Híbrido Sequencial (Mitra e Hayashi, 2000).	48
Figura 8: Arquitetura básica de um sistema Neuro-Fuzzy. Fonte: (Shi e Shimizu, 1992).....	49
Figura 9: Rota fermentativa para a produção de etanol.....	54
Figura 10: Rota de produção de etanol por sacarose	55
Figura 11: Rota de produção de etanol por amido	56
Figura 12: Rota de produção de etanol por fonte celulósica	57
Figura 13: Bastonetes de <i>Clostridium novyi</i>	60
Figura 14: Colônia de <i>Clostridium novyi</i> em meio ágar.	61
Figura 15: Dados de conversão de celulose pela hidrólise enzimática (1. Papel filtro, para fins de comparação; 2. Bagaço de cana explodido a vapor e tratado com NaOH, com baixa carga de celulose no reator; 3. Bagaço de cana explodido a vapor e tratado com NaOH, com alta carga de celulose; 4. Bagaço de cana tratado com H ₂ SO ₄ e NaOH, com baixa carga de celulose no reator; 5. Bagaço de cana tratado com H ₂ SO ₄ e NaOH, com alta carga de celulose).....	68
Figura 16: Função Objetivo	69
Figura 17: Dados agrupados pela proximidade com o centro do "cluster"	69
Figure 18 Culture P15.	84
Figure 19. Linear confidence region for culture P15.	85
Figure 20. Culture P55A.	85
Figure 21. Linear confidence region for culture P55A.	86
Figure 22. Culture P.7.5-7.5A.	86
Figure 23. Linear confidence region for culture P.7.5-7.5A.	87
Figure 24. Culture P.7.5-7.5B.	87
Figure 25. Linear confidence region for culture P.7.5-7.5B.	88
Figure 26. Boxplot of Mm, Kp, Kd, K1, K2, Ks; and n.....	89
Figure 27: Rods of <i>Clostridium novyi</i>	97
Figure 28: Colonies of <i>Clostridium novyi</i> on agar medium.....	98
Figure 29: Basic architecture of a neuro-fuzzy system.	102
Figure 30: Simulated annealing pseudocode.	105
Figure 31: Laboratory process for production of <i>Clostridium novyi</i>	108
Figure 32: Optimization route using DOE methodology to obtain the LTV29 medium.....	112
Figure 33: Cellular growth (OD600) and glucose consumption (mg/mL) curves for the culture performed with LTV29 medium.	112
Figure 34: Growth curves using the media obtained by computational optimization.	113
Figure 35: Microbial growth curves obtained using the media generated by computational optimization.	114
Figura 36: Interface do simulador EMSO	126

Figura 58: Resultado da simulação no EMSO para a produção de células CX1 e CX2, com parâmetros do modelo ajustados no MATLAB.	129
Figura 59: Resultado da simulação no EMSO para a produção de etanol, com parâmetros do modelo ajustados no MATLAB.	129
Figura 60: Resultado da simulação no EMSO para o consumo de substrato, com parâmetros do modelo ajustados no MATLAB.	130

ÍNDICE DE TABELAS

Tabela 1: Pseudocódigo do simulated annealing.....	37
Table 1. Simulated annealing pseudocode.	82
Table 3: Culture medium formulation input data matrix.....	107
Table 4: Output data matrix for toxin production (log L+/mL).	109
Table 5: Output data matrix for time (h).....	109
Table 6: Data matrix for productivity ((log L+/mL)/h).....	109
Table 7: Main identification parameters.....	111
Table 8: Media generated using the proposed optimization method.	115
Table 9: Experimental productivities obtained using the optimized media.	115
Tabela 10: Parâmetros ajustados no MATLAB.....	128

1. CAPÍTULO 01: INTRODUÇÃO

Problemas de elevado nível de complexidade computacional têm sido um desafio constante para pesquisadores dos mais diversos campos de estudo. Uma das áreas mais estudadas atualmente é a de sistemas computacionais inteligentes. As pesquisas sobre sistemas computacionais inteligentes visam à resolução de problemas combinatórios, estocasticamente¹ ou não, e têm se caracterizado, nos últimos anos, pela tendência em buscar inspiração na natureza, onde existem inúmeros exemplos de processos que podem ser ditos “inteligentes”. Embora não se possa afirmar que soluções tiradas desses processos sejam todas “ótimas”, não há a menor dúvida de que os processos naturais, em particular os relacionados diretamente com os seres vivos, são bem concebidos e adequados ao nosso mundo (Chen et al., 2015).

Figura 1: O Particle swarm optimization (PSO) é um dos exemplos de algoritmo baseado em fenômenos naturais



O algoritmo Recozimento Simulado, mais conhecido na literatura como Simulated Annealing (Kirkpatrick *et al.*, 1983), é também um desses algoritmos. Seu objetivo principal é encontrar o mínimo global de uma função, que caracteriza sistemas grandes e complexos, e é motivado por uma ideia simples: “para otimizar um sistema muito

¹ De modo aleatório e não determinístico.

grande e complexo, ao invés de sempre avançar no sentido descendente, procure prosseguir no sentido descendente na maior parte do tempo” (Musharavati e Hamouda, 2012). Ele foi idealizado a partir do algoritmo (Metropolis et al., 1953) que, por sua vez, usa técnicas de Monte Carlo para simular o processo inerente a esse algoritmo.

Outro exemplo são as redes neurais, que são baseadas no cérebro humano, o qual é composto por cerca de 100 bilhões de neurônios, que são conectados entre si de maneira complexa através de sinapses. Essa complexa conectividade é responsável pelas características atribuídas à inteligência. Com o desenvolvimento da Inteligência Artificial (AI) surgiu a ideia de representar, por meio de determinados programas, o funcionamento do processo de aprendizagem do cérebro humano. A tentativa de simular a rede neural do cérebro deu origem à chamada Rede Neural Artificial (RNA).

Assim, é imprescindível o aprimoramento na utilização de métodos computacionais que resolvam problemas complexos e em larga escala relacionados à Engenharia Química e demais áreas. Desde muito tempo, tais ferramentas são usadas na rotina de empresas, centros de pesquisa e universidades para automatizar os cálculos e fornecer um maior número de resultados em menos tempo. Quando se trata de milhares de contas e inúmeros tratamentos estatísticos não há como se usar lápis e papel. Qualquer erro, como uma simples soma errada, ou um expoente que falta, leva à decadência de todo o resultado final.

Dessa forma, a finalidade da realização deste trabalho é demonstrar, de maneira prática, que a utilização de modelagem de processos e buscas globais por ferramentas heurísticas podem proporcionar resultados que venham a aperfeiçoar os processos conhecidos.

Portanto, foi avaliado o potencial dessas técnicas computacionais nas diversas abordagens que abrangem os sistemas biotecnológicos. Os sistemas e técnicas avaliados foram: (i) realização de clusterização de sistemas em regiões cineticamente semelhantes para a produção de etanol celulósico (Etanol 2G), utilizando lógica *Fuzzy* (Contribuições ao projeto temático FAPESP/BIOEN 2008/56246-0); (ii) estimação de parâmetros cinéticos do modelo de equações de fermentação alcoólica (Etanol 1G) (Monod-Levenspiel) por busca global, utilizando o algoritmo Simulated annealing (SA) (Contribuições ao projeto temático FAPESP 2011/51902-9); (iii) formulação e otimização do meio de cultura economicamente viável para a produção de toxina do

Clostridium novyi tipo B, utilizando a modelagem de dados por *Neuro-Fuzzy* seguido de busca global da composição de meio que maximize a produtividade de toxinas pelo microrganismo (L+/mL). Essa etapa do projeto foi realizada em parceria com a empresa farmacêutica veterinária Vallée S.A.

Esta Tese de Doutorado esta estruturada da seguinte forma:

- Capítulo 01: INTRODUÇÃO (este capítulo);
- Capítulo 02: REVISÃO BIBLIOGRÁFICA GERAL: Aborda as técnicas de inteligência computacional e as aplicações de interesse neste trabalho, isto é, produção de etanol e vacina para uso veterinário;
- Capítulo 03: CLASSIFICAÇÃO DE DADOS DO PROCESSO DE HIDRÓLISE ENZIMÁTICA DE BAGAÇO DE CANA-DE-AÇÚCAR: UMA ANÁLISE DE “CLUSTERS” NEBULOSOS (artigo publicado no SINAFERM 2013);
- Capítulo 04: ESTIMATIVA DE PARÂMETROS CINÉTICOS DE FERMENTAÇÃO POR *S. Cerevisiae* VISANDO À PRODUÇÃO INDUSTRIAL DE ETANOL 1G (Artigo em preparação);
- Capítulo 05: INTELIGÊNCIA COMPUTACIONAL APLICADA À OTIMIZAÇÃO DO PROCESSO DE PRODUÇÃO DE VACINA INATIVADA DE *CLOSTRIDIUM NOVYI* TIPO B, UTILIZANDO SIMULTANEAMENTE NEUROFUZZY E SIMULATED ANNEALING. (Artigo publicado pela Applied Biochemistry and Biotechnology ISSN 0273-2289 – A1).
- APÊNDICE 1: Integração EMSO/MATLAB: utilização pelo software EMSO do algoritmo Simulated Annealing em Matlab, aplicado à fermentação de etanol 1G.
- APÊNDICE 2: Modeling the Kinetics of Complex Systems. Artigo sobre a aplicação da classificação de dados do processo de hidrólise (apresentada no capítulo 3) à modelagem da hidrólise enzimática do bagaço de cana-de-açúcar.

2.CAPÍTULO 02: REVISÃO BIBLIOGRÁFICA

2.1.TÉCNICAS COMPUTACIONAIS

2.1.1. Método de Monte Carlo

O Método de Monte Carlo (MMC) teve sua denominação originada do uso da aleatoriedade e da natureza repetitiva das atividades realizadas nos cassinos de Monte Carlo, em Mônaco (Rai *et al.*, 2009), onde a roleta era um gerador de números aleatórios. O nome MMC surgiu em 1940, durante o projeto Manhattan (Rezvani e Bolduc, 2014), quando Jon Von Neuman e S.M. Ulam consideraram utilizar o método em pesquisa para construção da bomba atômica, durante a Segunda Guerra Mundial (Atanassov e Dimov, 2008). O MMC é uma técnica baseada no uso de números aleatórios e estatística para resolução de problemas de diversas áreas, que podem variar desde a pesquisa operacional, simulações econômicas e previsão de comportamento de sistemas biológicos (Sminchisescu e Welling, 2011).

Essa é uma técnica matemática computadorizada que possibilita levar em conta o risco em análises quantitativas e tomadas de decisão (Gariazzo *et al.*, 2015). Comumente ela é usada por profissionais de uma grande variedade de campos, como finanças, gerenciamento de projetos, energia, indústrias, engenharia, pesquisa e desenvolvimento, seguros, petróleo e gás, transportes, meio ambiente e biotecnologia (Rezvani e Bolduc, 2014).

O MMC fornece ao tomador de decisão uma gama de resultados possíveis e as probabilidades de ocorrências desses resultados de acordo com a ação escolhida como decisão (Tempo e Ishii, 2007). Ela mostra as possibilidades extremas, os resultados das decisões mais ousadas e das mais conservadoras e todas as possíveis consequências das decisões mais moderadas (Sminchisescu e Welling, 2011).

Desde a sua introdução, na época da Segunda Guerra Mundial, a simulação de Monte Carlo tem sido usada para modelar uma variedade de sistemas físicos e conceituais (Schadd *et al.*, 2012).

Entre os métodos probabilísticos baseados em MMC, um algoritmo que se destaca é conhecido como *Simulated Annealing* (KIRKPATRICK, et al. 1983), que, na verdade, é

uma modificação do algoritmo de Metropolis-Hastings (1970), transformando-o em um algoritmo de otimização.

2.1.2. Método simulated annealing

Nesta seção será apresentada a meta-heurística *Simulated Annealing* (SA) com maiores detalhes, uma vez que essa foi à técnica, baseada no método de Monte Carlo, utilizada para solucionar problemas em estudo neste trabalho. Conforme apresentado por Dolan *et al.* (1990), o método pertence à classe de algoritmos não especializados, aplicáveis na resolução de problemas de otimização combinatória.

A meta-heurística² *Simulated Annealing* é um algoritmo baseado nos processos de recozimento de sólidos, desenvolvido por Metropolis *et al.* (1953). A analogia com a otimização combinatória foi introduzida por Kirkpatrick *et al.* (1983) e aperfeiçoada por Cerny (1985).

No processamento tradicional de metais, um método para melhorar a qualidade do metal é aquecê-lo a altas temperaturas, depois, lentamente, esfriá-lo em ciclos repetitivos. Esse método é chamado de recozimento. O que fisicamente acontece é que quando esse processo é realizado, o recozimento remove os defeitos do cristal, pois a energia potencial média por átomo é diminuída durante o recozimento (Boykov *et al.*, 2001). Então o processo pode ser considerado um regime de minimização de energia macroscópica, o que levou a um processo análogo chamado *Simulated Annealing*, que quer dizer recozimento simulado (Yao *et al.*, 1999).

O SA pode ser utilizado na solução de diversos problemas de otimização combinatória. Muitos problemas de otimização de processos biotecnológicos são de natureza combinatória e, portanto, candidatos naturais à solução por essa técnica. Alguns exemplos incluem problemas de formulações de meios de cultura complexo para microrganismos e células. No entanto, existem poucos trabalhos a respeito da utilização dessa técnica na solução de tais problemas. Um exemplo é que em seu trabalho Subrahmanyam *et al.* (2001), buscaram resolver um problema de monômeros

² Uma meta-heurística é um método heurístico para resolver de forma genérica problemas de otimização (normalmente da área de otimização combinatória).

que dão o maior número de pontos de ligação usando *Simulated annealing*, de modo a imitar a complexação dos monômeros funcionais com molde na mistura de monômeros.

Outro trabalho em que houve utilização deste algoritmo foi aquele desenvolvido por Forbes *et al.* (2001), referente à predição de fluxos metabólicos microbiais, no qual foi proposto um método de solução particionada que tira proveito da relação quase linear entre fluxos e isotopômeros. *Simulated Annealing* e o método de Newton-Raphson foram usados para os componentes não lineares.

Como visto, esse algoritmo procura encontrar uma solução ótima para problemas de otimização combinatória explorando novas áreas no espaço de soluções do problema, iterativamente (Musharavati e Hamouda, 2012). O desempenho dessa técnica pode ser avaliado pelo valor encontrado para a função objetivo, observando o número de iterações necessárias para atingir um resultado ótimo global (Pinto *et al.*, 2009).

A Tabela 1 mostra a forma geral o pseudocódigo do método *Simulated annealing*, projetado para aplicações gerais, em que os principais parâmetros do método são as variáveis do modelo, X_j , a função objetivo, f , as temperaturas instantâneas e iniciais do sistema, T e T_0 , o número máximo de iterações do modelo, N_{max} , a função de geração de probabilidade, e os vizinhos, $R(x_i)$. As respostas do modelo são as funções f_b e x_b . A temperatura inicial e o número máximo de iterações são valores arbitrários definida antes do início das iterações. O fator temperatura de decomposição (α) é um valor constante ou uma função $\alpha = f(T)$ pertencente ao intervalo $(0; 1)$ ".

Com o uso do algoritmo citado, pode-se obter como resultado o valor mínimo ou máximo de uma função previamente modelada. Assim, por exemplo, se o sistema modelado descreve o crescimento de um microrganismo em função dos componentes do meio de cultura, podemos utilizar o algoritmo SA para maximizar a resposta de crescimento microbiano em função da composição do seu meio de cultura.

Tabela 1: Pseudocódigo do simulated annealing

```

 $f_b = +\infty$ ;
 $x_0 \rightarrow$  Solução inicial;
 $f_m = f(x_0)$ ;  $x_m = x_0$ 
I=0  $\rightarrow$  Número de Iterações;
 $T = T^0 \rightarrow$  Temperatura inicial do sistema;
    enquanto  $i \leq N_{\max}$  faça:
         $x_i = x_m$ 
         $x_j = R(x_i)$ ;
        Se  $f(x_j) < f(x_i)$  então:
             $x_m = x_j$ ;  $f_m = f(x_j)$ ;
        termina se;
        se não:
             $\Delta f = f(x_j) - f(x_i)$ ;  $P(f) \rightarrow \text{unif.}[0:1]$ ;
            se  $P(f) \leq e^{(-\Delta f/T)}$  então:
                 $x_m = x_j$ ;  $f_m = f(x_j)$ ;
            termina se;
        termina se;
     $T = T \cdot \alpha$ ;  $i = i + 1$ 
    Se  $f_b > f_m$  então:
         $f_b = f_m$ ;  $x_b = x_m$ ;
    termina se;
termina enquanto;
```

Fonte: (Zheng *et al.*, 2006)

2.1.3. Redes neurais artificiais

O cérebro humano é considerado o mais fascinante processador baseado em carbono existente, sendo composto por aproximadamente 10 bilhões de neurônios. Todas as funções e movimentos do organismo estão relacionados ao funcionamento dessas pequenas células (Swain *et al.*, 2014). Os neurônios estão conectados uns aos outros através de sinapses, e juntos formam uma grande rede, chamada REDE NEURAL (Kreutzberg, 1996). As sinapses transmitem estímulos através de diferentes concentrações de Na⁺ (Sódio) e K⁺ (Potássio), e o resultado disto pode ser estendido por todo o corpo humano (Swain *et al.*, 2014). Essa grande rede proporciona uma fabulosa capacidade de processamento e armazenamento de informação (Abbod *et al.*, 2001).

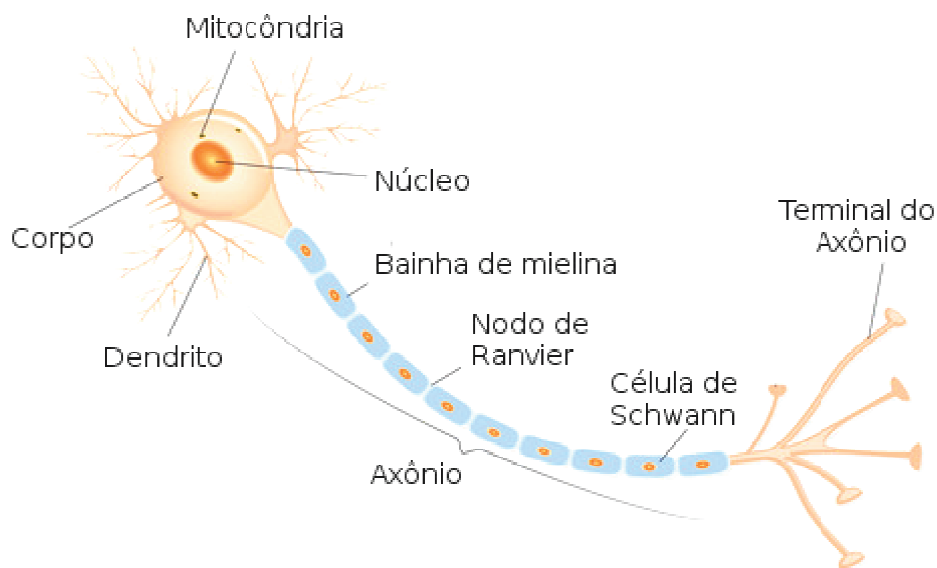
Esse sistema nervoso é formado por um conjunto muito complexo de neurônios (Figura 02). Nos neurônios, a comunicação é alcançada através de impulsos, quando um impulso é recebido, o neurônio o processa, e passado um limite de ação, dispara um segundo impulso que produz uma substância neurotransmissora que flui do corpo celular para o axônio (que, por sua vez, pode ou não estar conectado a um dendrito de outra célula) (Kordylewski *et al.*, 2001). O neurônio que transmite o pulso pode controlar a frequência de pulsos aumentando ou diminuindo a polaridade na membrana pós-sináptica. Eles têm um papel efetivo na determinação do funcionamento, comportamento e do raciocínio do ser humano. Ao contrário das redes neurais artificiais, redes neurais naturais não transmitem sinais negativos, sua ativação é medida pela frequência com que emite pulsos, frequência esta de pulsos contínuos e positivos (Mitra e Hayashi, 2000). As redes naturais não são uniformes como as redes artificiais, e apresentam uniformidade apenas em alguns pontos do organismo. Seus pulsos não são síncronos ou assíncronos, devido ao fato de não serem contínuos, o que a difere de redes artificiais (Rao e Kamat, 1996).

Os principais componentes dos neurônios são:

- Os dendritos, que tem por função receber os estímulos transmitidos pelos outros neurônios;

- O corpo de neurônio, também chamado de soma, que é responsável por coletar e combinar informações vindas de outros neurônios;
- E finalmente o axônio, que é constituído de uma fibra tubular que pode alcançar até alguns metros, e é responsável por transmitir os estímulos para outras células.

Figura 2: Figura esquemática do neurônio humano.



Assim, baseando-se neste sistema complexo humano, as redes neurais artificiais tiveram seus estudos iniciados, originalmente, na década de 40, pelo neurofisiologista Warren McCulloch, do MIT³, e pelo matemático Walter Pitts, da Universidade de Illinois, os quais, dentro do espírito cibernético, fizeram uma analogia entre células nervosas vivas e o processo eletrônico num trabalho publicado sobre "neurônios formais" (McCulloch e Pitts, 1990; Kay, 2001; Mira, 2008). O trabalho consistia num modelo de resistores variáveis e amplificadores representando conexões sinápticas de um neurônio biológico (McCulloch e Pitts, 1990).

Desde então, mais enfaticamente a partir da década de 80, diversos modelos de redes neurais artificiais têm surgido com o propósito de aperfeiçoar e aplicar essa

³ Massachusetts Institute of Technology (MIT): O Instituto de Tecnologia de Massachusetts é um centro universitário de educação e pesquisa privado localizado em Cambridge, Massachusetts, nos Estados Unidos.

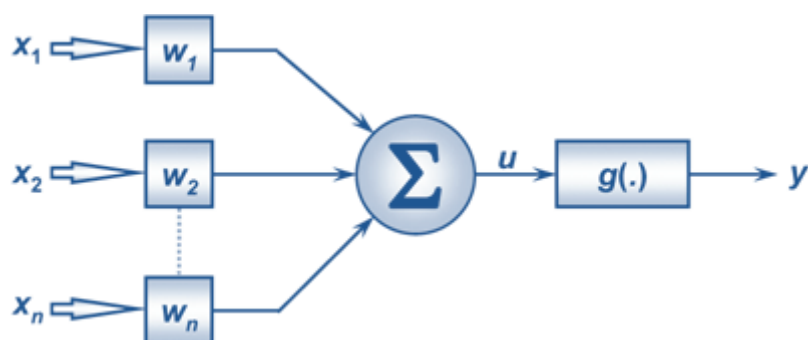
tecnologia. Algumas dessas propostas tendem a aperfeiçoar mecanismos internos da rede neural para aplicação na indústria e negócios, outras procuram aproximá-las ainda mais dos modelos biológicos originais.

Por exemplo, no trabalho desenvolvido por Riesenberg e Guthke (1999), foram modificadas estratégias tradicionais de controle do processo fermentativo por decisões tomadas a partir de aplicação do controle de estatística multivariada, redes neurais artificiais, controle *Fuzzy* e supervisão baseada no conhecimento (sistemas especialistas), para obtenção do cultivo de *Escherichia coli* em alta densidade celular.

Assim, as redes neurais artificiais (RNA's) são, portanto, um algoritmo computacional de uma rede de neurônios artificiais. Sendo assim, um neurônio artificial é, por outro lado, um modelo matemático inspirado em um neurônio real. (Passos, 2005).

Como visto, os neurônios biológicos estão conectados em uma rede complexa e dinâmica (Figura 03). Essas redes biológicas de interconexões apresentam características desejáveis em sistemas artificiais, como por exemplo: exibem capacidade de aprendizado; não precisam ser programadas; possuem capacidade de ação paralela, ou seja, capaz de realizar várias funções ao mesmo tempo; são pequenas, compactas e dissipam pouca potência (Zanetti *et al.*, 2008). Estas características podem ser desenvolvidas em algoritmos computacionais, utilizando-se de um modelo matemático que leva em conta os conexionismos para lidar com a manipulação de dados, resolvendo problemas complexos (Borg *et al.*, 2014).

Figura 3: Representação das entradas e saídas das RNA's.



Cada neurônio recebe suas entradas com os pesos associados, vindas de outros neurônios ou de estímulo externo. A função de ativação é usualmente um somatório agindo nas entradas da rede. Os valores bias⁴ são adicionados ao somatório das entradas com pesos. O estado de ativação do neurônio é determinado pela função de saída ou de transferência e geralmente é uma função sigmoideal. Uma regra de propagação consiste das funções de ativação e de transferência (saída). Esta saída do neurônio trafega pelas sinapses até a próxima célula. Portanto, uma rede neural artificial (RNA) é formada por um conjunto de neurônios dispostos em uma ou mais camadas, conectados por ligações de pesos variados (Passos, 2005).

2.1.4. Logica Fuzzy

Seres humanos são capazes de lidar com processos bastante complexos, baseados em informações imprecisas ou aproximadas. A estratégia adotada pelos operadores humanos é também de natureza imprecisa e geralmente possível de ser expressa em termos linguísticos. A Teoria de Conjuntos *Fuzzy* e os Conceitos de Lógica *Fuzzy* podem ser utilizados para traduzir, em termos matemáticos, a informação imprecisa expressa por um conjunto de regras linguísticas. Se um operador humano for capaz de articular sua estratégia de ação como um conjunto de regras da forma “se-então”, um algoritmo passível de ser implementado em computador pode ser construído como

⁴ - A camada de entrada deve possuir uma unidade especial conhecida como bias, usada para aumentar os graus de liberdade, permitindo uma melhor adaptação, por parte da rede neural, ao conhecimento a ela fornecido.

ferramenta de controle. O resultado é um sistema de inferência baseado em regras, no qual a Teoria de Conjuntos *Fuzzy* e Lógica *Fuzzy* fornecem as ferramentas matemáticas para se lidar com as tais regras linguísticas.

As primeiras noções da lógica dos conceitos “nebulosos” (*Fuzzy*) foi desenvolvida por um lógico polonês chamado Jan Lukasiewicz (1878-1956), que, em 1920, introduziu conjuntos com graus de pertinência (0, $\frac{1}{2}$ e 1) que, mais tarde, foram expandidos para um número infinito de valores entre 0 e 1 (Pearson, 1962). A primeira publicação sobre lógica *Fuzzy* foi em 1965, quando recebeu este nome (Zadeh, 1965). Seu autor foi Lotfi Asker Zadeh, docente em Berkeley, Universidade da Califórnia. Zadeh criou a lógica *Fuzzy*, concordando os fundamentos da lógica clássica e os conjuntos de Lukasiewicz, definindo os graus de pertinência. Ele observou que recursos tecnológicos, baseados na lógica booleana, não eram suficientes para automatizar atividades relacionadas a problemas de natureza industrial, biológica ou química. Entre 1970 e 1980, as aplicações industriais da lógica *Fuzzy* aconteceram com maior importância na Europa. Especificamente, em 1974, o Professor Ebrahim Mamdani (Mamdani e Assilian, 1975) conseguiu controlar uma máquina a vapor com tipos diferentes de controladores aplicando o raciocínio *Fuzzy*. E, após 1980, o Japão iniciou seu uso com aplicações na indústria. Algumas das primeiras aplicações foram em um tratamento de água feito pela Fuji Electric, em 1983, e pela Hitachi em um sistema de metrô inaugurado em 1987 (Hu *et al.*, 2001). Por volta de 1990 é que a lógica *Fuzzy* despertou um maior interesse em empresas dos Estados Unidos (Moran, 1992; Nass, 1992). Devido ao desenvolvimento e às inúmeras possibilidades práticas dos sistemas *Fuzzy* e ao grande sucesso comercial de suas aplicações, a lógica *Fuzzy* é considerada hoje uma técnica *standard*⁵ e tem uma ampla aceitação na área de controle de processos industriais.

A definição de um conjunto nebuloso (*Fuzzy*) A , definido no universo de discurso U , é caracterizada por uma função de pertinência μ_A , a qual mapeia os elementos de U para o intervalo $[0, 1]$ (Zadeh, 1965).

$$\mu_A:U \Rightarrow [0, 1]$$

Assim posto, a função de pertinência associa com cada elemento x pertencente a U um número real $\mu_A(x)$ no intervalo $[0,1]$, que representa o grau de possibilidade de que

⁵ Padrão

o elemento x venha a pertencer ao conjunto A , i. e., o quanto é possível para o elemento x pertencer ao conjunto A .

Um conceito relacionado com conjuntos nebulosos é o de variável⁶ linguística. Desse modo, uma variável linguística pode assumir um valor linguístico dentre vários outros em um conjunto de termos linguísticos (Zadeh, 1965).

Formalmente, uma variável linguística é caracterizada pela quintupla $(X, T(X), U, G, M)$, onde X é o nome do conjunto de termos, U o universo de discurso, G uma gramática para gerar os termos $T(X)$, e M o significado dos termos linguísticos, representados através de conjuntos nebulosos.

Variáveis linguísticas podem também conter modificadores (também linguísticos) que alteram seu valor. Exemplos de modificadores válidos são: “muito”, “pouco”, “não muito”, “mais” ou “menos”. Assim, um valor válido para a variável linguística peso seria não muito pesado e não muito leve. Existem também conectivos que podem ser aplicados a essas variáveis. Os conectivos *and* (e) e *or* (ou) são equivalentes a operações de união e intersecção de conjuntos, respectivamente, podendo dar origem a conjuntos complexamente definidos, porém representados linguisticamente de maneira simples.

2.1.1.1. Agrupamento de dados

O agrupamento de dados numéricos (do inglês, *clustering of numerical data*) forma a base de muitos algoritmos de classificação e de modelagem de sistemas. O propósito é identificar agrupamentos naturais de um conjunto de dados para produzir uma representação concisa do comportamento do sistema. Uma técnica bastante atraente para o agrupamento de dados é aquela associada à lógica nebulosa (*Fuzzy*), chamada de técnica FCM (do inglês, *Fuzzy C-Means*).

Trata-se de uma técnica de agrupamento de dados em que cada ponto de dados pertence a um *cluster*⁷, até certo ponto, o que é especificado por um grau de pertinência. Tal técnica foi introduzida originalmente por Jim Bezdek (1981) como uma melhoria nos métodos de agrupamento usados antes desse (Nelles, 1998). Ele fornece um método que mostra como agrupar pontos de dados que povoam algum espaço multidimensional

⁶ Entende-se por variável um identificador que pode assumir um dentre vários valores.

⁷ Aglomeração.

em um número específico de diferentes *clusters*. Existem técnicas a serem utilizadas quando esse número de *clusters* não é conhecido (Cinque *et al.*, 2004).

A técnica consiste em minimizar uma função de resíduos de modo a agrupar, da melhor forma possível, o conjunto de dados segundo o grau de pertinência. O algoritmo FCM trabalha executando os passos apresentados a seguir (Ahmed *et al.*, 2002; Dembele e Kastner, 2003): São escolhidos os centros dos *clusters* aleatoriamente, e, para cada ponto do conjunto de dados, é calculada a distância a cada centro através da equação:

Equação I

$$D_{ij}^2 = (u(i) - c_j)^T \sum (u(i) - c_j)$$

Onde $u(i)$ é um ponto da amostra, j é o número de *clusters* e c é o centro escolhido. Depois disso, é calculado o grau de pertinência para cada ponto (em relação a cada cluster), através da equação:

Equação II

$$\mu_{ij} = \frac{1}{\sum_{l=1}^C (D_{ij}^2 / D_{il}^2)^{\frac{1}{v-1}}}$$

Sendo μ é o grau de pertinência, cuja somatória deve ser 1, v é o expoente de ponderação, que está entre 1 e ∞ , que, quando não se tem conhecimento prévio, 2 é uma boa estimativa, e C é o número de centros. Após essa etapa, é calculado o centroide de cada cluster, ajustando cada centro de cluster a esse centroide, que pode ser obtido através da equação:

Equação III

$$c_j = \frac{\sum_{i=1}^N \mu_{ij}^v u_i}{\sum_{i=1}^N \mu_{ij}^v}$$

Sendo N o número de dados existentes. Os centros são modificados até que a função objetivo seja minimizada. A função objetivo é representada por I na equação:

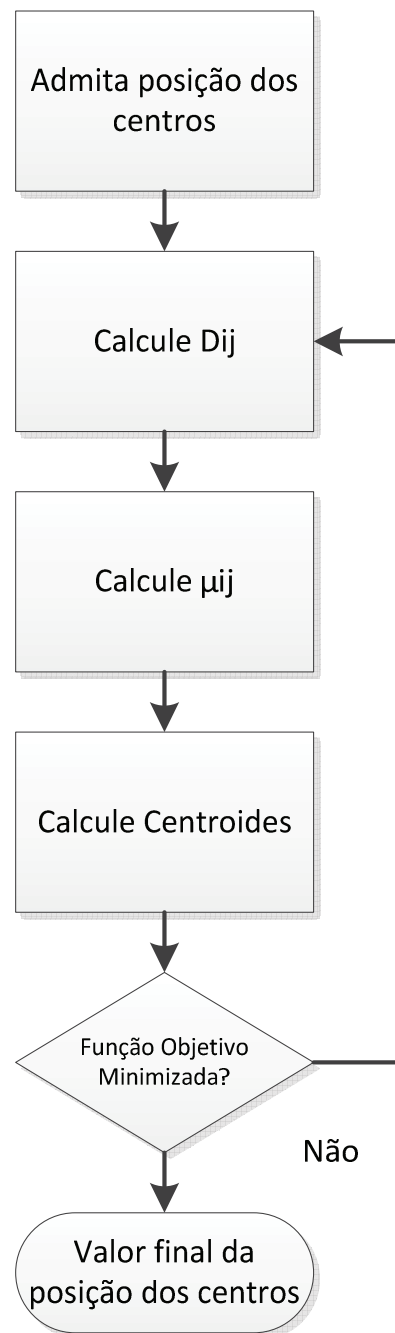
Equação IV

$$I = \sum_{j=1}^C \sum_{i=1}^N \mu_{ij}^v (u(i) - c_j)^T \sum (u(i) - c_j)$$

A Figura 04 resume esses passos em forma de algoritmo, indicando também a etapa em que ocorrem as iterações.

A execução desses passos é bastante extensa e mecânica, devido à quantidade de iterações que podem ocorrer. Assim, pacotes computacionais como o MATLAB oferecem funções com o algoritmo FCM acoplado, facilitando a utilização dessa técnica (MATHWORKS, 2012).

Figura 4: Fluxograma simplificado do algoritmo FCM.



Fonte: Acervo pessoal.

2.1.1. Redes Neuro-Fuzzy

O sistema *Neuro-Fuzzy* ajusta a teoria de redes neurais (Smith, 1992), (Lin e Lee, 1991) com a dos conjuntos *Fuzzy* (Zadeh, 1965), (Zimmermann, 1985). Os processos de treinamento de redes neurais permitem que esse sistema “aprenda” novas informações a

partir de um conjunto de dados. Por outro lado, a teoria de conjuntos *Fuzzy* admite que esses sistemas representem o conhecimento armazenado em uma forma acessível aos seres humanos.

Esses sistemas constituem-se mais comumente de sistemas híbridos incorporados, uma vez que essa classe de sistemas híbridos faz evidenciar mais ainda as vantagens de cada técnica utilizada (Lei et al., 2007). Existem três classes básicas de hibridização: incorporada, auxiliar e sequencial. As figuras 5, 6 e 7 exemplificam cada um dessas classes.

O sistema híbrido *Neuro-Fuzzy* apresenta-se como uma alternativa particularmente interessante na modelagem de sistemas, uma vez que consegue extrair o aprendizado das redes neurais com as vantagens do nível de interpretabilidade da lógica *Fuzzy* (Jiang e Adel, 2008).

Figura 5: Sistema Híbrido Incorporado (Mitra e Hayashi, 2000).

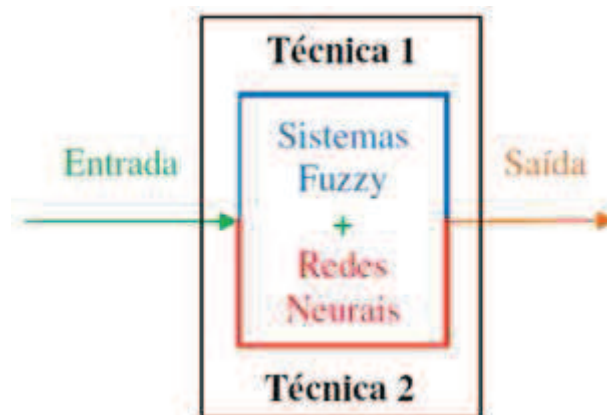


Figura 6: Sistema Híbrido Auxiliar (Mitra e Hayashi, 2000)

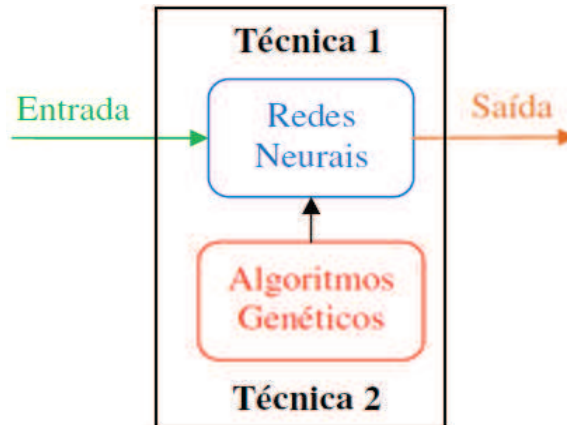
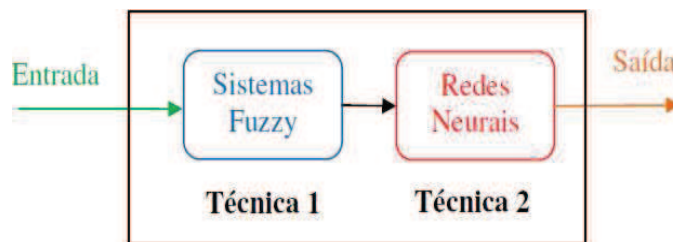


Figura 7: Sistema Híbrido Sequencial (Mitra e Hayashi, 2000).



É válido ressaltar que o único sistema híbrido em que não é possível a separação entre dois subsistemas é o incorporado, evidenciando, assim, um maior grau de hibridização (Kofman *et al.*).

A ideia básica é programar um sistema de inferência Fuzzy através de uma arquitetura distribuída de tal forma que os paradigmas de aprendizado, comuns às Redes Neurais, possam ser aproveitados nessa hierarquia híbrida.

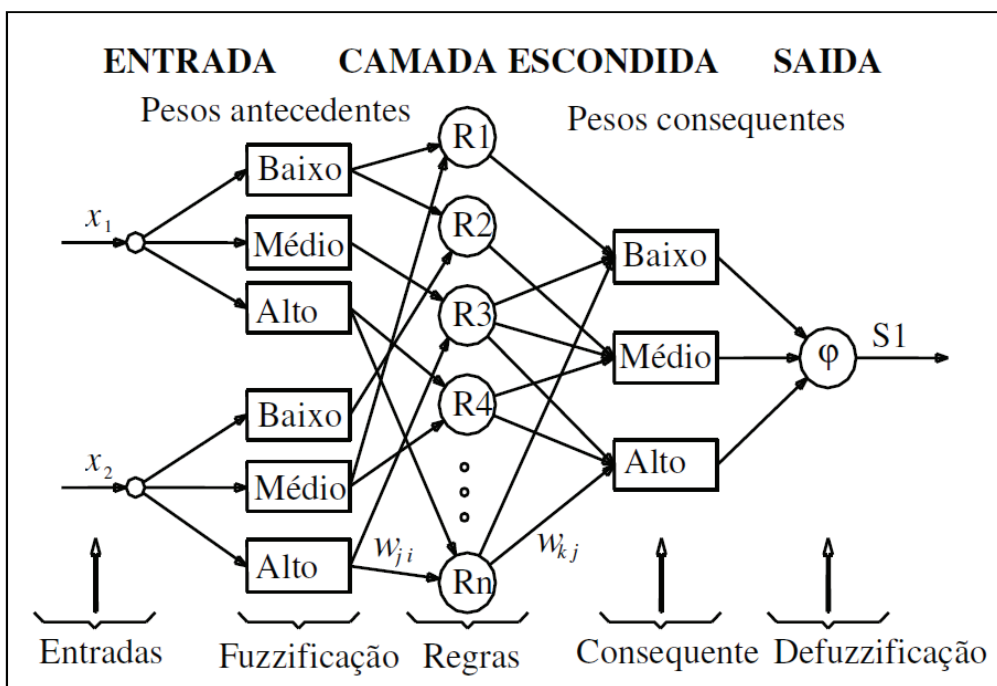
Como exemplo, temos o fato de que os sistemas *Fuzzy* se mostram ineficientes na determinação e composição das funções de pertinência por si só. Já redes neurais não aceitam inclusão de conhecimento prévio na sua inicialização, ou seja, torna-se difícil determinar qual estrutura obtém o melhor modelo para determinado problema (Nava e Ieee, 1998).

Dessa maneira, os sistemas *Neuro-Fuzzy* que aliam a capacidade de aprendizado das redes neurais artificiais com a facilidade de interpretação linguística dos sistemas *Fuzzy* estão sendo amplamente utilizados. Vale ressaltar algumas características importantes desses sistemas:

1. A possibilidade de realizar a integração de conhecimentos de maneira implícita e explicitamente;
2. A aplicabilidade de diversos algoritmos de aprendizado utilizados pelas Redes Neurais;
3. Facilita a interpretação e o entendimento do conhecimento, principalmente pelo fato de a extração do conhecimento dar-se através de regras *Fuzzy*.

A forma de implementação básica das redes *neuro-Fuzzy* é de que os algoritmos de aprendizado das redes neurais possam ser aproveitados nesses sistemas híbridos para ajustar os parâmetros do Sistema de Inferência *Fuzzy* (FIS) (Baharani *et al.*, 2014). Na figura 8, está representada a estrutura de um sistema de inferência *Neuro-Fuzzy*:

Figura 8: Arquitetura básica de um sistema *Neuro-Fuzzy*. Fonte: (Shi e Shimizu, 1992)



O método híbrido gera características fundamentais, vistas, em cada técnica geradora, em separado, aliando-se e adaptando-se à classe de problemas a serem resolvidos. As características de responsabilidade *Fuzzy* dos sistemas híbridos aqui estudados se dividem em:

- Modelo *Fuzzy*
- Formato das funções de pertinência
- Particionamento do Espaço do espaço de Entrada e Saída
- Método de Defuzzificação

Já tratando das Redes Neurais Artificiais, sabe-se que elas irão lidar com as seguintes características:

- Tipo de Aprendizado
- Identificação da estrutura
- Identificação dos parâmetros

Existem várias configurações para cada uma das características supracitadas (seja *Fuzzy* ou de Redes Neurais). O conjunto de métodos utilizados para formar a rede neuro-*Fuzzy* desejada irá gerar um tipo de sistema Neuro-*Fuzzy* diferente. Os mais conhecidos e utilizados são: ANFIS (Adaptative Network based *Fuzzy* Inference System) (Jang, 1993), FSOM (*Fuzzy* Self-Organized Map) (Vuorimaa, 1994), NEFCLASS (Neuro *Fuzzy* Classification), (Ichihashi *et al.*, 1996) NEFCON (Neuro-*Fuzzy* Control) (Rao e Kamat, 1996) e NEFPROX (Neuro-*Fuzzy* Function Aproximation) (Nauck e Kruse, 1999).

Segundo Shi e Shimizu (1992), vale evidenciar as vantagens e desvantagens da utilização dos sistemas Neuro-*Fuzzy*. Primeiramente, tratando das vantagens, podemos citar:

- A lógica *Fuzzy* provê uma interface de alto nível, de rápida computação e amigável de programar, permitindo que o projetista se concentre nos objetivos funcionais em vez dos detalhes matemáticos. As redes neurais artificiais são convenientes para a extração de conhecimento através do aprendizado.
- O desenvolvedor não precisa ter conhecimento prévio do processo, levando a uma fácil adaptação aos diferentes processos.
- Os modelos *Neuro-Fuzzy* têm a capacidade de lidar melhor com o problema de ruído nos dados do que as redes neurais artificiais puras.
- Possui capacidade de autoaprendizado, auto-organização e autodirecionamento, aproximando a capacidade humana do processo de tomada de decisão.

Em se tratando de desvantagens, podemos citar dois aspectos:

- O sistema deve trabalhar com um número reduzido de entradas, já que um número maior demanda um esforço computacional maior. Além disso, pode ocorrer uma explosão combinatória de regras geradas (alta quantidade de iterações necessárias, inviabilizando o cálculo ou o tempo necessário para o cálculo ser extremamente grande).
- Há limitação na capacidade de alterar sua estrutura pelo elevado número de regras.

2.1.SISTEMAS BIOTECNOLÓGICOS EM ESTUDO

2.1.1. Fermentação alcoólica

2.1.1.1. Aspectos gerais

Os primeiros usos práticos do etanol no Brasil deram-se entre meados dos anos 1920 e início dos anos 1930. Mas somente nos anos 1970, com a crise do petróleo, é que o Brasil passou a usar maciçamente o etanol como combustível, com a criação do programa Pró-Álcool (Dhere, 1989).

O Pró-Álcool foi um programa do governo brasileiro para substituição dos carros movidos por gasolina pelos movidos a etanol. Em 1973, o mundo passou por uma crise do petróleo, devido à Guerra do Yom Kipur⁸ (Rowley, 1989). O preço do barril do combustível natural subiu expressivamente, gerando grandes consequências na economia mundial. No Brasil, viviam-se os anos do Regime Militar, que começara em 1964 (Choi, 2010). Entre 1968 e 1973, o país passou por um tempo de grande crescimento econômico colocando-se entre as economias mais desenvolvidas do planeta. Esse período é chamado de Milagre Econômico (Rosillo-Calle, 1986; Veloso *et al.*, 2008).

Foi justamente com a crise mundial do petróleo, em 1973, que chegou ao fim essa fase de grande crescimento econômico (Veloso *et al.*, 2008). Assim, como os outros países, o Brasil também sofreu os efeitos da crise do petróleo (um problema que prejudicou a continuidade do crescimento econômico). Dessa forma, os governantes da época lançaram, em 14 de novembro de 1975, o Programa Nacional do Álcool, o Pró-Álcool (Veloso *et al.*, 2008; Cerqueira Leite *et al.*, 2009).

Criado pelos engenheiros Lamartine Navarro Júnior e Cícero Junqueira Franco e pelo empresário Maurílio Biagi, o programa contava com os projetos do físico José Walter Bautista Vidal e do engenheiro urbano Ernesto Stumpf, que desenvolveram o motor a álcool (Rosillo-Calle e Hall, 1987).

O objetivo era substituir gradativamente a frota de carros movida por combustíveis derivados do petróleo por motores que funcionavam com recursos naturais, caso do álcool. Preocupados com os impactos da primeira crise do petróleo e com possíveis novas crises, os idealizadores do Pró-Álcool acabaram por desenvolver uma tecnologia pioneira no Brasil. (Da Costa *et al.*, 2010).

O governo da época, por sua vez, incentivou a medida empenhando-se na produção de etanol no país e fornecendo subsídios. Assim, o Brasil ficou menos dependente da gasolina mundial e reduziu em 10 milhões o número de carros no país que eram movidos pelo derivado do petróleo. (De Oliveira, 1991).

⁸ Conhecida pelos nomes de “Quarta Guerra Árabe-Israelense”, “Guerra do Ramadã”, “Guerra de Outubro” e “Guerra Árabe-Israelense de 1973”, a Guerra do Yom Kippur ocorreu entre Israel e um grupo de nações sob a liderança da Síria e do Egito. O combate deu-se entre os dias 6 e 26 do mês de outubro de 1973.

Foi utilizada a cana de açúcar como substrato para a produção do etanol, no período, devido aos baixos preços com que o açúcar se encontrava. Dessa maneira, o Brasil conseguiu resguardar-se da nova crise do petróleo, que abalou o mundo, em 1979, devido ao corte na venda e distribuição desse item por parte do segundo maior produtor mundial, o Irã, em meio à revolução fundamentalista pela qual passou o país em 1979. (Rowley, 1989). O projeto entrou em fase de estagnação e recessão quando o preço do petróleo passou por processo inverso: queda no preço do barril de petróleo aliada à alta do açúcar no mercado exterior, tornando-se mais vantajoso às usinas a produção do açúcar. (Rosillo-Calle, 1986).

O ressurgimento da exploração do álcool como combustível deu-se em meados dos anos 2000 com o surgimento da tecnologia flex⁹, que é comum de se encontrar no mercado atualmente. (Cerqueira Leite *et al.*, 2009).

O veículo de combustível flex mais comum disponível no mercado mundial utiliza etanol como segundo combustível (Samanez *et al.*, 2014). Através de um sensor é detectada a combinação do combustível, e ajusta-se a injeção de acordo com a mistura (Oih Yu *et al.*, 2010). Dessa forma, pode-se usar tanto álcool quanto gasolina, ou uma mistura dos dois em qualquer proporção. No caso dos carros brasileiros, o ajuste da injeção é realizado através de um software automotivo, desenvolvido por engenheiros brasileiros (Oih Yu *et al.*, 2009).

2.1.1.2. Matéria-prima: substrato

O álcool etílico pode ser obtido a partir de vegetais ricos em açúcar, classificados como sacaríneos, amiláceos e celulósicos, sendo que, entre eles, estão a cana-de-açúcar, a beterraba, extratos da mandioca, do arroz, do milho, e da celulose proveniente dos eucaliptos. (Morimura *et al.*, 1997).

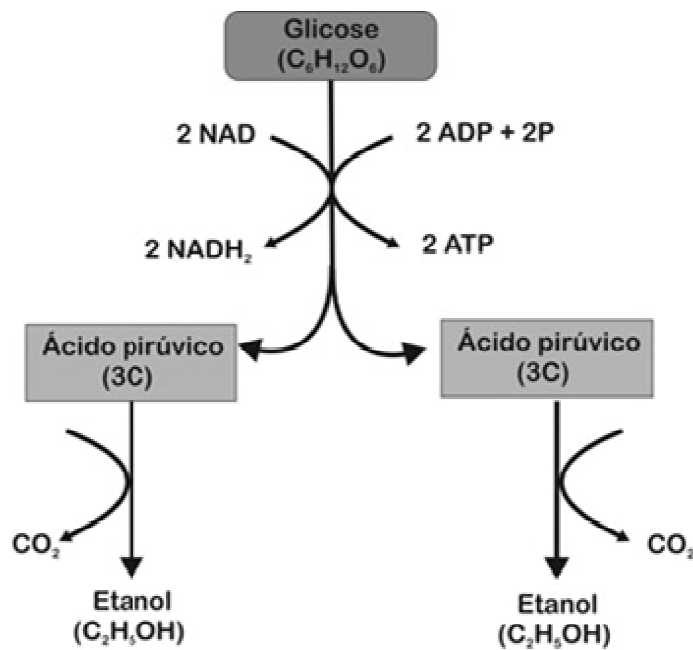
Grande parte do álcool produzido no Brasil é obtida através da cana-de-açúcar. Embora existam diferentes matérias-primas para a sua produção, todas elas derivam do mesmo processo, a fermentação alcoólica, que consiste numa reação anaeróbica,

⁹ Tecnologia que possibilita que os carros rodem com os dois tipos de combustíveis, etanol ou gasolina, independentemente da proporção da mistura.

realizada por micro-organismos que degradam parcialmente a glicose obtendo como produto final álcool etílico e dióxido de carbono (Figura 09).

Na fermentação alcoólica (ver figura 09), as duas moléculas de ácido pirúvico produzido são convertidas em álcool etílico (também chamado de etanol), com a liberação de duas moléculas de CO₂ e a formação de duas moléculas de ATP (Murai et al., 1998).

Figura 9: Rota fermentativa para a produção de etanol

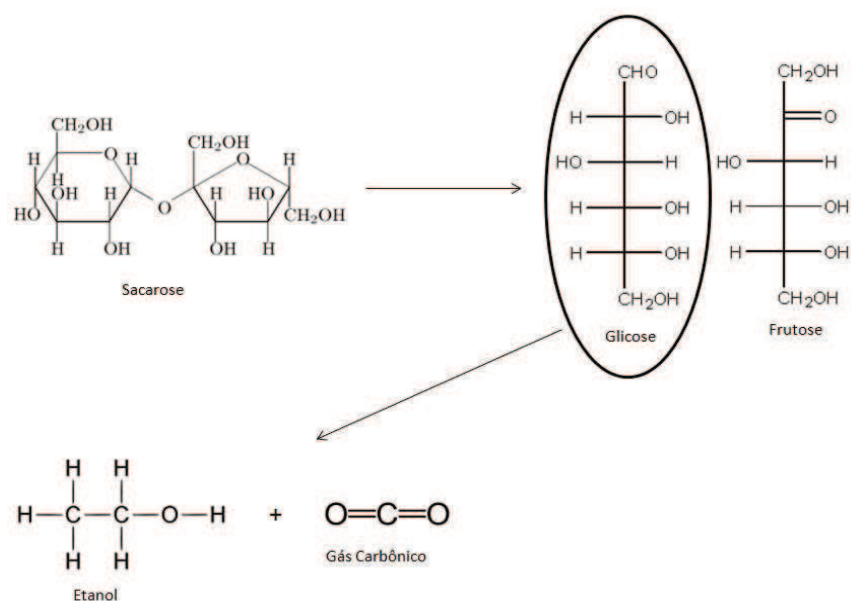


Esse tipo de fermentação é realizado por diversos micro-organismos, destacando-se os chamados “fungos de cerveja”, da espécie *Saccharomyces cerevisiae* (Adachi et al., 1998). O homem utiliza os dois produtos dessa fermentação: o álcool etílico empregado há milênios na fabricação de bebidas alcoólicas (vinhos, cervejas, cachaças etc.), e o gás carbônico importante na fabricação do pão, um dos mais tradicionais alimentos da humanidade. Mais recentemente, tratando-se de história da humanidade, tem-se utilizado esses fungos para a produção industrial de álcool combustível.

2.1.1.2.1. Matérias de origem sacarínica

Introduzida no período colonial, a cana-de-açúcar é a principal fonte de matéria prima sacarínica existente no Brasil, transformando-se, assim, em uma das principais culturas da economia brasileira. O Brasil não é apenas o maior produtor de cana, mas também o primeiro do mundo na produção de açúcar e o segundo maior produtor de etanol (perdendo apenas para os Estados Unidos), e conquista, cada vez mais, o mercado externo com o uso do biocombustível como alternativa energética. (Martinez et al., 2013). A figura 10 ilustra a rota de produção de etanol por via sacarínica.

Figura 10: Rota de produção de etanol por sacarose

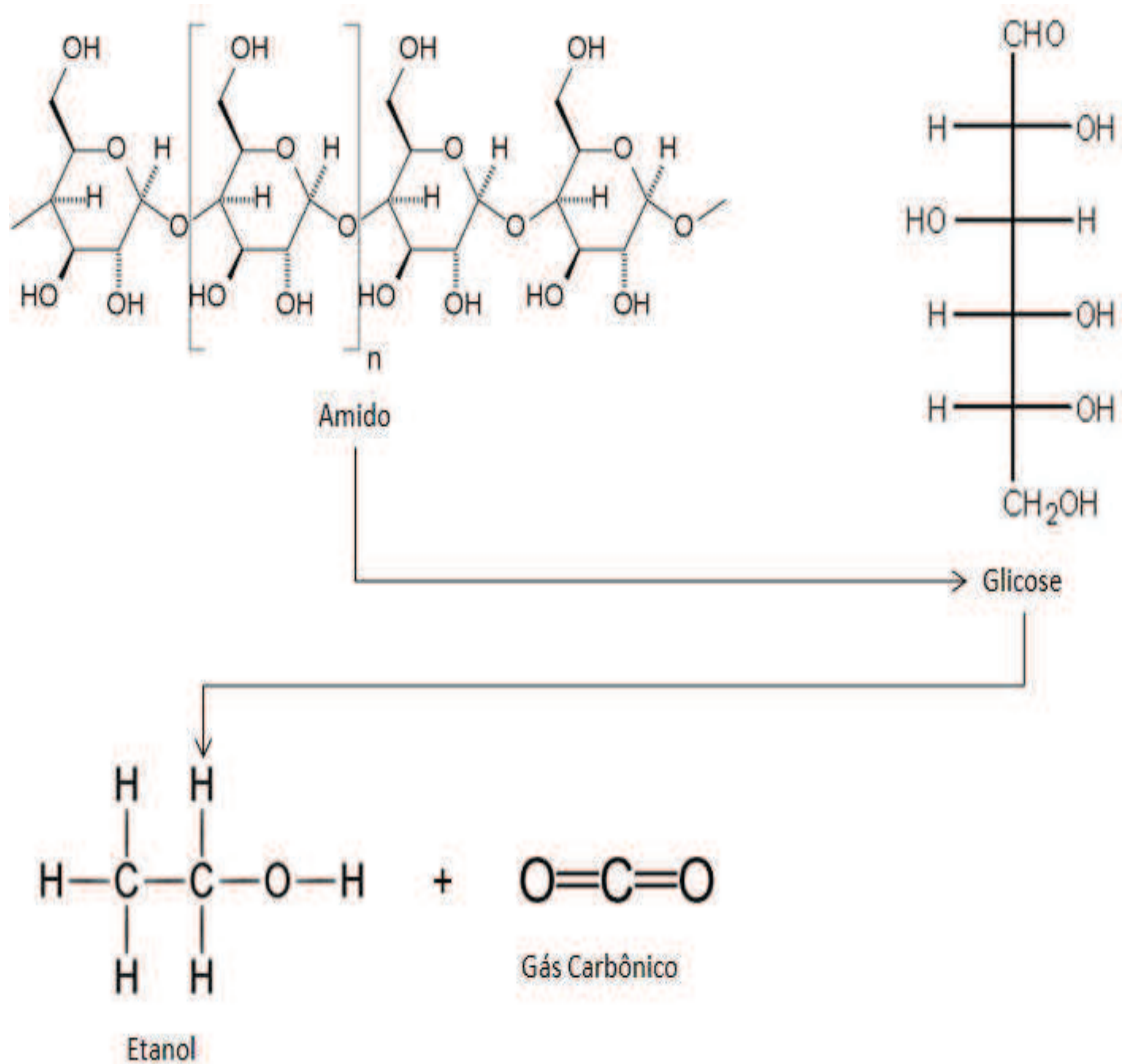


2.1.1.2.2. Matérias de origem amilácea

As matérias primas de origem amilácea são as que contêm o amido em sua composição. Sendo esse um polímero complexo, não apresenta sabor, sendo insípido ao paladar. Também é impossível fermentá-lo, mas é possível transformá-lo em açúcares e depois submetê-los à fermentação alcoólica. (Bannai e Oe, 2010).

Os amidos podem ser hidrolisados por via química ou física (ácidos, calor e pressão) ou por via enzimática. Os hidrolisados por enzimas são os mais importantes amidos modificados comerciais. Parte desses hidrolisados são dextrinas, ainda insípidas e pouco fermentescíveis, e parte são açúcares derivados de amido, como a maltose e a glicose (Nunez *et al.*, 2013). A figura 11 ilustra a rota de produção de etanol por via amilácea.

Figura 11: Rota de produção de etanol por amido



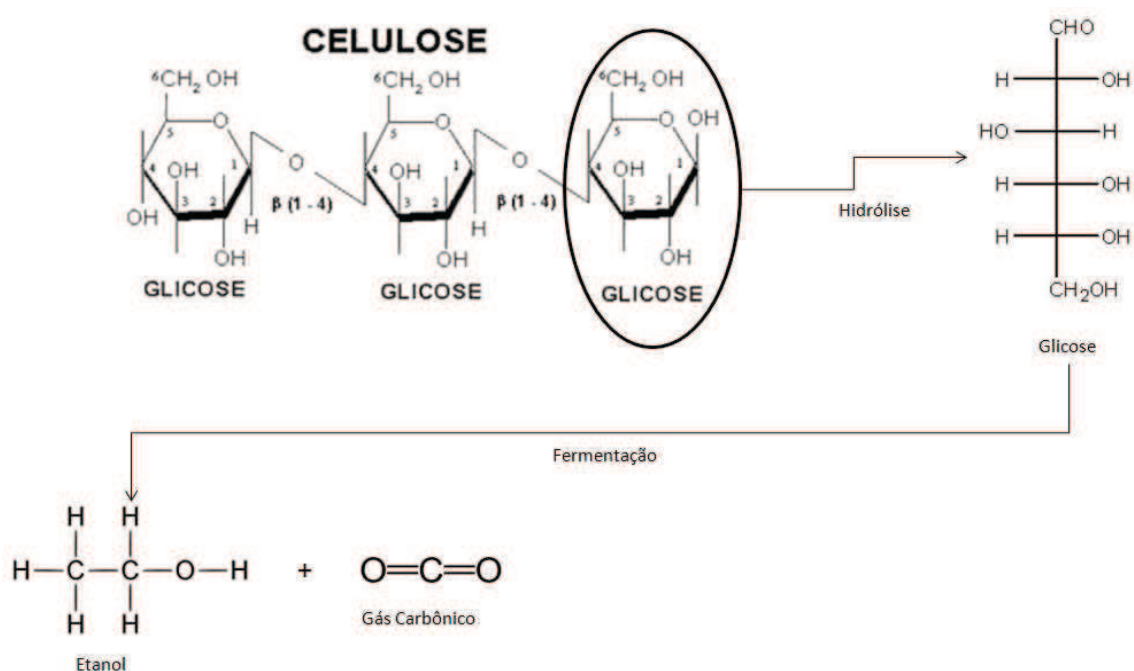
2.1.1.2.3. Matéria de origem celulósica

No Brasil, apesar de possuir-se uma grande produção de etanol a partir da sacarose de cana-de-açúcar, a produção de álcool de fontes alternativas de substrato pode ser interessante, principalmente se associada à indústria já existente. (Pandey, Soccol, Nigam e Soccol, 2000).

O etanol de segunda geração, como também é conhecido o etanol produzido através de matérias primas celulósicas, pode ser obtido não só do bagaço e da palha de cana-de-açúcar, mas também de palha de milho, palha de trigo, palha e casca de arroz, cavacos e restos de madeira, ou seja, quaisquer resíduos que contenham celulose, como papel, e ainda resíduos urbanos.

O bagaço de cana-de-açúcar pode ser convertido a etanol por meio do processo chamado de hidrólise enzimática, em que a celulose é convertida em açúcares fermentescíveis. Para que ocorra a hidrólise enzimática, primeiramente a biomassa lignocelulósica precisa ser previamente tratada, aumentando a acessibilidade ao ataque enzimático. (Pandey, Soccol, Nigam, Soccol, et al., 2000; Gírio et al., 2010). A figura 12 ilustra a rota de produção de etanol por rota celulósica.

Figura 12: Rota de produção de etanol por fonte celulósica



2.1.2. Micro-organismos: matérias-primas e processos

2.1.2.1. Saccharomyces cerevisiae

Os processos de fermentação já eram utilizados pelo homem há cerca de dez mil anos. Muitas bebidas eram fabricadas pelos antigos egípcios, germanos e israelitas. Embora as bebidas alcoólicas sejam diferenciadas por suas propriedades, tais como suas matérias primas e diferentes teores alcoólicos, todas elas têm uma origem básica comum, isto é, todas derivam de um processo bioquímico denominado fermentação alcoólica. A fermentação alcoólica é um tipo de reação química realizada pela ação de micro-organismos (leveduras) sobre os açúcares, produzindo etanol e gás carbônico. (Ferreira Filho e Horridge, 2014).

O micro-organismo do fermento, denominado *Saccharomyces cerevisiae*, é responsável pela produção das enzimas fundamentais para o processo de fermentação alcoólica. (Morimura *et al.*, 1997). A *Saccharomyces cerevisiae* é uma espécie de levedura (Lindegren *et al.*, 1945), e talvez a levedura mais útil, tendo sido fundamental para vinicultura, panificação e fabricação de cerveja, desde os tempos antigos.

Acredita-se que tenham sido originalmente isolados a partir da pele de uvas (pode-se ver a levedura como um componente da película branca fina sobre a pele de alguns frutos de cores escuras, tais como as ameixas). (Lindegren *et al.*, 1945). É um dos organismos eucariotas, modelo mais intensivamente estudado na biologia molecular e celular, assim como a *Escherichia coli* como o modelo de bactéria (procariota). É o micro-organismo que traz o tipo mais comum de fermentação. Células de *S. cerevisiae* são redondas a ovoides, com 5-10 μ m de diâmetro, as quais se reproduzem por um processo de divisão conhecido como brotação. (Swanson e Clifton, 1948).

2.1.3. Produção de etanol no Brasil

O etanol (C_2H_5OH) é uma substância pura, constituída por uma molécula formada por dois átomos de carbono, cinco de hidrogênio e um grupo hidroxila. Existem dois tipos de etanol: o anidro e o hidratado. A diferença entre os dois se deve à concentração

de água em sua composição. O anidro tem o teor de água equivalente a 0,5%, enquanto o hidratado apresenta teor de 5%. (Luchi e Trindade, 1982).

O processo industrial usado convencionalmente produz o etanol hidratado, que é vendido nos postos de combustíveis para ser usado em veículos automotores. Já a produção do etanol anidro requer um procedimento adicional e específico para a retirada do excesso de água. (Cerqueira Leite *et al.*, 2009).

O álcool anidro, entre outras aplicações, é usado para a produção de gasolina C, substância produzida a partir da mistura entre a gasolina A (gasolina pura) e de etanol. Nessa mistura, a quantidade de álcool anidro varia de 20% a 25%. Tanto a gasolina A quanto o álcool anidro não podem ser vendidos diretamente ao consumidor final. (Marjotta-Maistro e Barros, 2003).

Uma diversidade de espécies vegetais pode ser usada para a produção de etanol. No Brasil, o material mais utilizado é a cana-de-açúcar, apresentada também como uma alternativa para a geração de eletricidade a partir do resíduo em forma de bagaço. Cada tonelada de cana-de-açúcar produz 140 kg de bagaço de cana (Kiatkittipong *et al.*, 2009).

Segundo dados da Confederação da Agricultura e Pecuária do Brasil (CNA), o Brasil é o segundo maior produtor mundial de etanol, perdendo apenas para os Estados Unidos. A cana-de-açúcar é a espécie vegetal que sustenta a produção de etanol brasileiro, enquanto o etanol dos Estados Unidos é produzido a partir do milho (Nunez *et al.*, 2013).

O etanol tem algumas vantagens em comparação aos combustíveis fósseis, como o fato de poluir menos (Furtado *et al.*, 2011). Contudo, outras etapas da cadeia produtiva do etanol devem ser levadas em conta. Um exemplo é a produção do vinhoto, ou vinhaça, um subproduto do processo de produção do etanol (Tsao *et al.*, 2012).

Na produção de um litro de etanol, são usados 13 litros de água, dos quais 12 litros são vinhoto. Ao ser despejado nos corpos hídricos, o vinhoto torna a água imprópria para o consumo humano, além de afetar gravemente a fauna e a flora do ambiente aquático. Porém, o vinhoto pode ser usado para a produção de biogás, o que reduziria o impacto dessa etapa do ciclo produtivo do etanol. (Moreira Da Silva e Cardoso, 2013). Também o uso de vinhaça representa importantes aportes de matéria orgânica e de

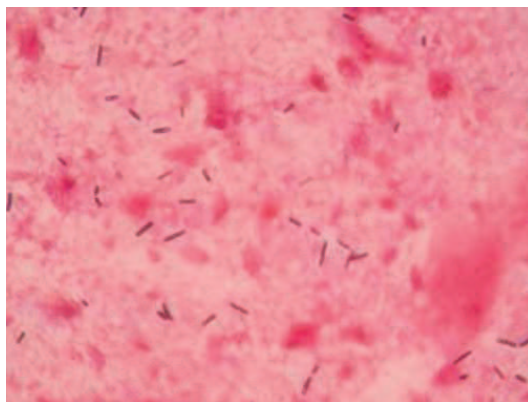
potássio e fósforo, respectivamente causando o aumento da fertilidade do solo reduzindo a captação de água para irrigação, reduzindo também o uso de fertilizantes químicos e, conseqüentemente, os custos decorrentes dessa prática. (Motizuki Lopes *et al.*, 2014).

2.1.4. *Clostridium novyi*

O *Clostridium novyi* é descrito como causador da hepatite necrótica dos ovinos e bovinos, conhecida como "black disease". É descrito também como causador de gangrena gasosa proveniente de ferimentos. (Amimoto *et al.*, 2007).

Segundo Skarin *et al.* (2011), as principais características desse micro-organismo são: bastonete grande (medindo 5 a 10 μm de comprimento e 1 a 1,5 μm de diâmetro), como pode ser observado na figura 13, forma esporo oval central ou subterminal, acapsulado, possui flagelos peritríquios (móveis sob condições de anaerobiose), gram positivo quando jovens, mas tornando-se gram variáveis em culturas após vários repiques ou longos períodos de cultivo e com distribuição mundial, estando presente principalmente no solo. (Belyi e Aktories, 2010).

Figura 13: Bastonetes de *Clostridium novyi*.



O *Clostridium novyi* é anaeróbico estrito, fastidioso, sendo difícil cultivá-lo. O crescimento pode ser melhorado com a adição de glicose, sangue ou cérebro no meio de cultura (Ziegler et al., 2008). As colônias são achatadas, com bordas irregulares que se espalham na superfície do ágar como um filme e possui crescimento transparente como demonstrado na figura 14 (Wei et al., 2008). Em crescimento em meios sólidos com sangue, há presença de hemólise, após 48 a 72 de incubação (Diaz et al., 2005). No meio contendo carne cozida, o crescimento é lento, e a carne torna-se rósea. A gelatina é liquefeita e o soro coagulado, mas não digerido. As colônias, quando crescidas em meio contendo gema de ovo, produzem uma zona de precipitação opalescente (Reação de Nagler), devido à produção de lecitinase (gama toxina) pelo *C. novyi* tipo A e devido à beta toxina do tipo B e D. A fermentação de carboidratos é variável. A grande maioria produz ácido e gás a partir da glicose, mas não fermentam a lactose. (Brett et al., 2005).

Figura 14: Colônia de *Clostridium novyi* em meio ágar.



O *C. novyi* está dividido em 4 tipos (A, B, C e D), sendo o tipo D reconhecido como *Clostridium haemolyticum* de acordo com a produção das toxinas (Alfa, Beta, Gama e Delta). (Batty e Walker, 1964). O *Clostridium novyi* Tipo B (tipo em estudo no presente trabalho) produz as toxinas alfa e beta (menor quantidade). Nesse micro-organismo, a produção da toxina alfa está ligada à presença de bacteriófagos específicos. Com a ausência desses bacteriófagos, apenas a toxina beta é produzida, semelhante ao *C. novyi* tipo D (*C. haemolyticum*). (Eklund et al., 1974).

A principal forma de combate às doenças causadas pelo *C. novyi* é a vacinação. Em geral as vacinas contra esse micro-organismo possuem em sua constituição toxóides ou toxóides e bacterinas (Amimoto et al., 2007).

**3. CAPÍTULO 03: CLASSIFICAÇÃO DE
DADOS DO PROCESSO DE
HIDRÓLISE ENZIMÁTICA DE
BAGAÇO DE CANA-DE-AÇÚCAR:
UMA ANÁLISE DE “CLUSTERS”
NEBULOSOS**

RESUMO

A modelagem da hidrólise enzimática de materiais lignocelulósicos é provavelmente um dos temas mais desafiadores em engenharia de (bio) reatores até o presente. A complexidade do sistema com relação ao substrato torna difícil o uso de um único modelo para todas as distintas condições de hidrólise. Pode-se pensar, neste ponto, na combinação dos atributos de diferentes modelos em um sistema “global” para cálculo da velocidade de hidrólise que incorpore os atributos de modelos ajustados para distintas condições de hidrólise. Para tanto, um passo inicial no sentido deste desenvolvimento é a análise de dados experimentais para se mapearem dados para distintas condições de hidrólise a classes pré-definidas, correspondendo a cada um dos modelos. O intuito principal deste trabalho consiste no tratamento dos dados experimentais de Carvalho (2011) de modo a se mapearem dados de porcentagem de lignina no substrato sólido e carga de celulose no reator a regiões cineticamente semelhantes. Utilizou-se a técnica associada à lógica nebulosa chamada técnica FCM (do inglês, “Fuzzy C-Means”).

Palavras-chave: bagaço de cana-de-açúcar, hidrólise, modelagem, “fuzzy”, classificação.

ABSTRACT

The modeling of the enzymatic hydrolysis of lignocellulosic materials is probably one of the most challenging issues in engineering (bio) reactors to the present. The complexity of the system with respect to the substrate makes it difficult to use a single model for all different hydrolysis conditions. One might think this point, the combination of attributes of different models in a "global" system to calculate the rate of hydrolysis which incorporates the attributes of adjusted models for different hydrolysis conditions. To this end, an initial step in this development is the analysis of experimental data to delineate data for different hydrolysis conditions pre-defined classes, corresponding to each of the models. The main purpose of this work consists in treating the experimental data Carvalho (2011) to delineate data lignin percentage in the solid substrate and ballast load on the cellulose to kinetically similar regions. We used the technique associated to the fuzzy logic technique called FCM (English, "Fuzzy C-Means").

Keywords: Sugarcane bagasse, hydrolysis, modeling, "fuzzy" classification.

3.1.INTRODUÇÃO

Há muito tempo se questiona até quando teremos matéria prima para suprir a demanda energética mundial, contudo ninguém chega a uma resposta definitiva. Enquanto isso, ela é consumida, e uma possível escassez acaba por se tornar um motivo de preocupação ascendente.

Para solucionar os problemas de demanda energética em uma sociedade que, cada vez mais, é dependente da tecnologia, buscam-se alternativas para as matérias-primas tradicionais, como o carvão, o petróleo e, inclusive, a sacarose para produção de etanol. Uma possibilidade, que há algumas décadas é pesquisada, é o etanol de segunda geração, como é chamado o etanol proveniente de fontes lignocelulósicas.

No caso brasileiro, o etanol de primeira geração é produzido a partir da sacarose extraída da cana-de-açúcar, que tem grande produtividade no território nacional (CONAB, 2012). Contudo, a produção gera muitos resíduos, tanto líquidos como sólidos. Entre os resíduos sólidos está o bagaço da cana-de-açúcar, que não é completamente consumido na integração energética da indústria sucroalcooleira. Sendo assim, uma das soluções para o grande volume de resíduos é a utilização do bagaço como matéria-prima para a produção de etanol de segunda geração. Com isso, aumenta-se também a produtividade, já que a produção de etanol será aumentada sem se incrementar a área plantada.

Assim, apesar de o país possuir uma grande produção de etanol a partir da sacarose de cana-de-açúcar, a produção de álcool de fontes alternativas de substrato pode ser interessante, principalmente se associada à indústria já existente. (Pandey, Soccol, Nigam e Soccol, 2000)

O bagaço de cana-de-açúcar pode ser convertido a etanol por meio do processo chamado de hidrólise enzimática. Para a hidrólise enzimática, primeiramente a biomassa lignocelulósica precisa ser previamente tratada, aumentando a acessibilidade ao ataque enzimático. (Hamelinck *et al.*, 2005; Gírio *et al.*, 2010)

Carvalho (2011) realizou o ajuste de modelos cinéticos de hidrólise enzimática para a obtenção de glicose a partir de diferentes substratos ligno-celulósicos, tendo, com isso,

definido equações de velocidade específicas para os diferentes tipos de substrato (equações do tipo Michaelis-Menten e Chrastil).

Devido à complexidade de um sistema desse tipo, torna-se difícil a utilização de um único modelo global para as diferentes condições de hidrólise. Uma ideia interessante é combinar os atributos dos diferentes modelos “locais” ajustados por Carvalho (2011) em um sistema global para cálculo da velocidade de hidrólise (incorporando os modelos ajustados para as distintas condições de hidrólise). Para tanto, um passo inicial no sentido desse desenvolvimento torna-se a análise dos dados experimentais de Carvalho (2011).

Assim, o objetivo deste trabalho foi efetuar a classificação em “clusters” dos dados experimentais de Carvalho (2011). Essa análise consistiu em agrupar entidades cujos valores dos principais atributos são próximos entre si (no presente caso, porcentagem de lignina no substrato sólido e carga de sólidos no reator). Essa classificação torna-se importante para determinar as regiões cineticamente semelhantes. O agrupamento de dados numéricos forma a base de muitos algoritmos de classificação e de modelagem de sistemas. Uma técnica bastante atraente para o agrupamento de dados é aquela associada à lógica nebulosa chamada de técnica FCM (do inglês, “Fuzzy C-Means”) (Bezdec, 1981).

3.2.MATERIAIS E MÉTODOS

Através dos dados experimentais obtidos por Carvalho (2011) de concentração de glicose em função do tempo, para diversos ensaios de hidrólise, com distintos substratos ligno-celulósicos (papel filtro; bagaço explodido tratado com NaOH 4%, baixa e alta carga; bagaço tratado com H₂SO₄ 1% e com NaOH 4%, baixa e alta carga), obtiveram-se dados de porcentagem de lignina no substrato sólido e carga de celulose no reator. Com esses dados em mãos, trabalhou-se com o agrupamento dos dados através da função FCM do “software” MATLAB (licença disponível no Departamento de Engenharia Química da UFSCar).

A função de linha de comando FCM começa com uma estimativa inicial para os centros de “cluster”, que são destinados a marcar a posição média de cada “cluster”. A estimativa inicial para esses centros de “cluster” é provavelmente incorreta, mas a função FCM iterativamente produz movimentos dos centros de “cluster” para o local correto dentro de um conjunto de dados. Essa iteração é baseada na minimização de uma função objetivo que representa a distância de qualquer ponto de dados fornecido a um centro de “cluster”, ponderada pelo grau de pertinência daquele ponto. O comando FCM gera uma lista de centros de “cluster” e vários graus de pertinência para cada ponto de dados.

3.3.RESULTADOS E DISCUSSÃO

Primeiramente foram plotados dados de porcentagem de celulose no reator em função do percentual de lignina, durante o processo de hidrólise (ao longo do tempo), para os ensaios realizados por Carvalho (2011). Os dados para os cinco diferentes tipos de substratos lignocelulósicos são apresentados numerados de um a cinco na figura 15.

Através da função FCM, calcularam-se os centros dos clusters. A figura 16 ilustra o processo de convergência associado à minimização da função objetivo representando a distância de qualquer ponto de dados fornecido a um centro de “cluster”.

A figura 17 mostra os centros de “cluster” obtidos e a pertinência preponderante de cada ponto em relação aos agrupamentos.

O que se observa é que os “clusters” agrupam, de forma coerente, os pontos de dados: parte dos dados referentes a substratos semelhantes (bagaço de cana explodido a vapor e tratado com NaOH, alta e baixa carga de celulose no reator) podem ser agrupados em um mesmo “cluster”, enquanto dados referentes ao final de experimentos para substratos inicialmente menos recalcitrantes (bagaço de cana explodido a vapor e tratado com NaOH) podem ser incorporados em “clusters” primordialmente associados a dados para substrato mais recalcitrante (bagaço de cana tratado com H₂SO₄ e NaOH, com baixa carga de celulose), porque se tornam cineticamente semelhantes.

As informações retornadas pelo FCM poderão ser usadas para gerar um sistema de inferência (nebuloso) que represente o comportamento do sistema (velocidade de hidrólise em função das condições de hidrólise) usando regras do tipo: se a carga é baixa ou alta e a lignina é pouca, a velocidade segue modelo cinético 1; se a carga é baixa e a lignina é média, a velocidade segue modelo cinético 2; se a carga é alta e a lignina é média, a velocidade segue modelo cinético 3; etc.

Figura 15: Dados de conversão de celulose pela hidrólise enzimática (1. Papel filtro, para fins de comparação; 2. Bagaço de cana explodido a vapor e tratado com NaOH, com baixa carga de celulose no reator; 3. Bagaço de cana explodido a vapor e tratado com NaOH, com alta carga de celulose; 4. Bagaço de cana tratado com H₂SO₄ e NaOH, com baixa carga de celulose no reator; 5. Bagaço de cana tratado com H₂SO₄ e NaOH, com alta carga de celulose)

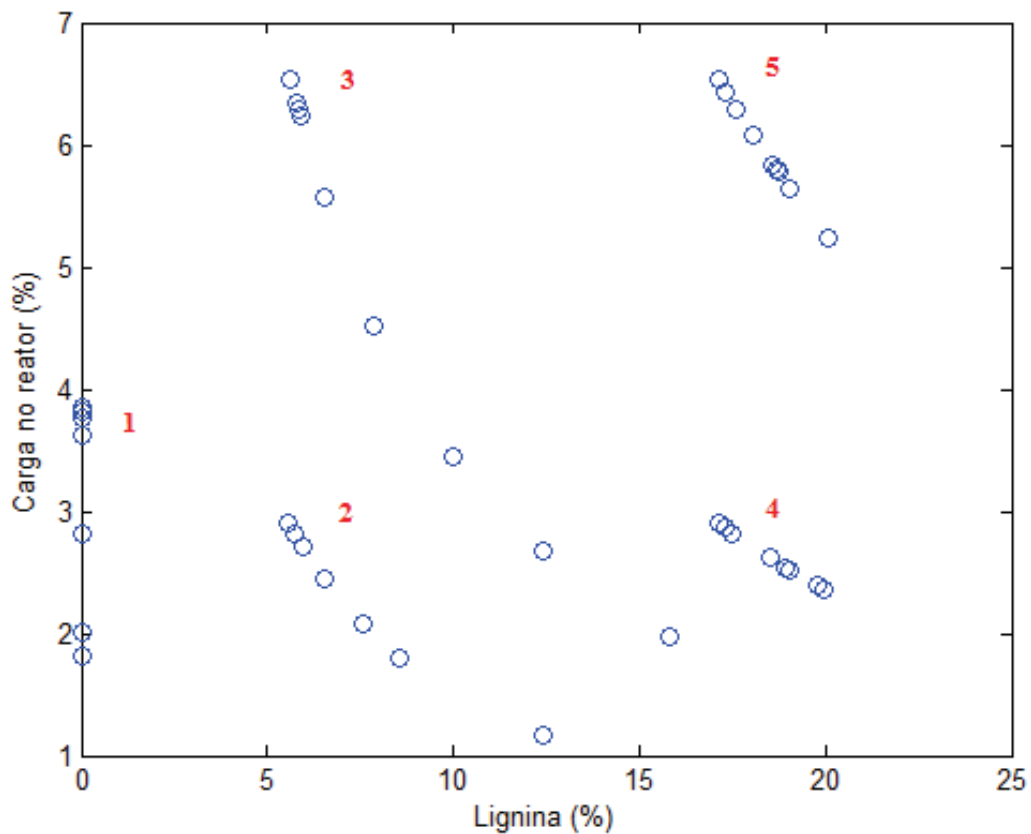


Figura 16: Função Objetivo

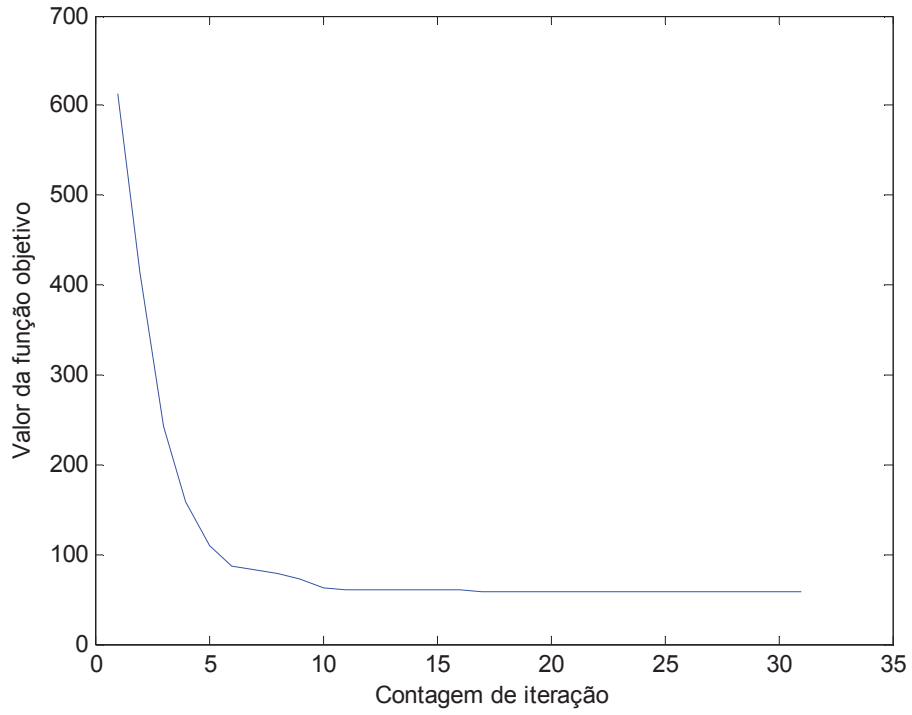
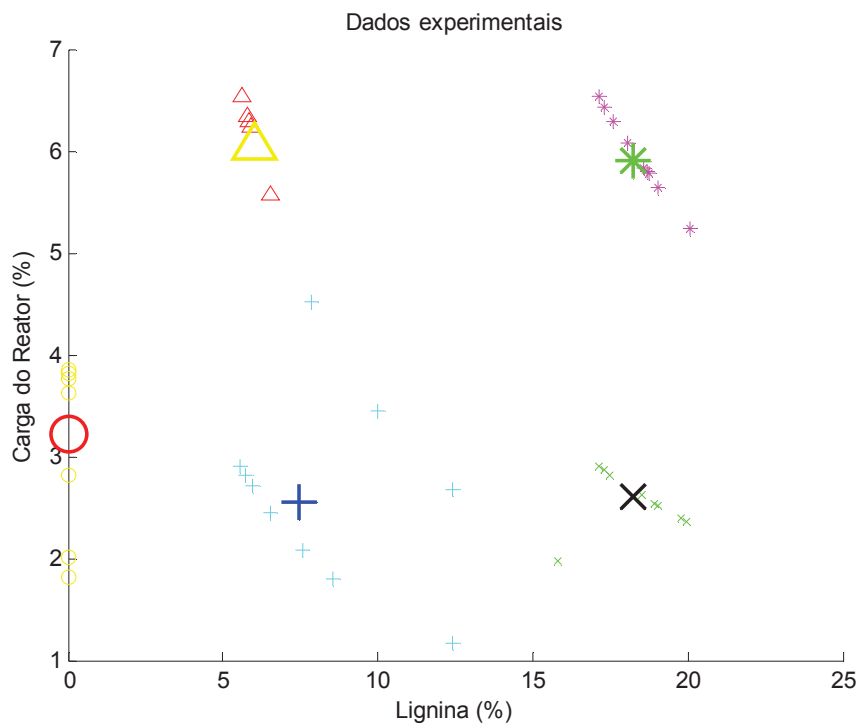


Figura 17: Dados agrupados pela proximidade com o centro do "cluster"



3.4.CONCLUSÃO

Através da análise FCM, pode-se efetuar a classificação em “clusters” dos dados experimentais. Esta análise consistiu em agrupar entidades cujos valores dos principais atributos são próximos entre si (no presente caso, porcentagem de lignina no substrato sólido e carga de sólidos no reator). As informações retornadas pelo FCM podem ser usadas para gerar um sistema de inferência (nebuloso) que represente o comportamento do sistema (velocidade de hidrólise em função das condições de hidrólise).

**4. CAPÍTULO 04: ESTIMATIVA DE
PARÂMETROS CINÉTICOS DE
FERMENTAÇÃO POR S.
CEREVISIAE VISANDO À
PRODUÇÃO INDUSTRIAL DE
ETANOL 1G**

DETERMINATION OF THE KINETIC PARAMETERS OF ALCOHOLIC FERMENTATION FOR INDUSTRIAL ETHANOL PRODUCTION

Aquino, P. L. M.^{1,2}, Fonseca F. S.², Giordano R. C.¹, Sousa Jr. R.^{1,*}

¹ Postgraduate Program in Chemical Engineering, Federal University of Sao Carlos (UFSCar), Rodovia Washington Luís, km 235, Monjolinho, São Carlos, 13565-905, São Paulo, Brazil

* E-mail: ruy@ufscar.br

² Vallée S/A, Av. Comendador Antônio Loureiro Ramos 1500, Distrito Industrial, Montes Claros, 39404-620, Minas Gerais, Brazil

ABSTRACT

Ethanol is a green fuel that, in addition to its economic importance, has major social, political, and environmental impacts. The aim of this work was to evaluate the effectiveness of global minimization using the simulated annealing (SA) technique for determining the kinetic parameters of an alcoholic fermentation model. The minimization was performed using experimental data reported in the literature, considering first a fed-batch cultivation model and then a simple batch process. Minimization was performed of the function describing the difference between the experimental and simulated curves, and the results obtained demonstrated the effectiveness of the SA technique. Compared with classical local search methods, for large parameter spaces the use of the SA technique appears to offer a useful alternative for solving the problem addressed here and others of a similar nature.

Keywords: Fermentation, Ethanol, Simulated annealing, Modeling, Process Simulation.

4.1.INTRODUCTION

4.1.1. Production of ethanol in Brazil

There are two types of ethanol: anhydrous and hydrated. The difference between them is due to the proportion of water present. Anhydrous alcohol has a water content of 0.5%, while the hydrated form contains 5% water (Luchi and Trindade 1982). The conventional industrial process produces hydrated ethanol, which is sold at gas stations for use in motor vehicles. The production of anhydrous ethanol requires an additional procedure to remove excess water (Cerqueira Leite, Verde Leal et al. 2009).

Amongst other applications, anhydrous ethanol is used to produce type C gasoline, which is a mixture of type A gasoline (pure gasoline) and ethanol. The amount of anhydrous alcohol in this mixture ranges from 20% to 25%.

Various plant species can be used to produce ethanol. In Brazil, the most widely used is sugarcane, which also provides an alternative way of generating electricity using the waste bagasse. Each ton of sugarcane produces 140 kg of bagasse (Kiatkittipong, Wongsuchoto et al. 2009). According to data provided by the Brazilian Confederation of Agriculture and Livestock (CNA), Brazil is the world's second largest ethanol producer, after the United States. Brazilian ethanol production is based on sugarcane, while in the US ethanol is produced from maize (Nunez, Oenal et al. 2013).

There are several advantages of ethanol, compared to fossil fuels, including the fact that it leads to less pollution (Furtado, Gaya Scandiffio et al. 2011). However, other stages of the ethanol production chain need to be considered, such as the formation of vinasse as a byproduct of the process (Tsao, Campbell et al. 2012).

4.1.2. Mathematical modeling of bioprocesses

4.1.2.1. Basics

Fermentation systems are highly sensitive to changes in the operating conditions. The attainment of maximum yields requires maintaining the operating conditions inside the fermenter close to preestablished values (Munack 1989). Furthermore, the complexity, efficiency, and reliability of current industrial systems continually increase, requiring the development of new tools for process monitoring and control. In this respect, mathematical models are important for optimizing, monitoring, and controlling bioprocesses (Cosenza, Mannina et al. 2009).

Independent of the area of interest (such as chemistry, biology, physico-chemistry, biochemistry, engineering, psychology, or economics), mathematical modeling of processes seeks to identify a set of relationships between variables that is able to satisfactorily describe the process with the required accuracy (Almeida and Palharini 2012).

4.1.2.2. Monod model

Many unstructured and unsegregated models are available to describe fermentations, considering the quantification of cell growth rates, substrate consumption, product formation, and other related parameters (Kurganov, Dorozhko et al. 1976). In most cases, these models are based on determining the specific growth rate of the microorganism and product formation, while also considering the decrease in the maximum specific rate by means of inhibition and restriction terms (Goldbeter 1974).

One of the most widely used models is the Monod model (Golterman 1991, Carstensen, Harremoes et al. 1995), which describes the specific growth rate as a function of the limiting substrate concentration(s). According to this model, the specific cell growth rate is given by (Babloyan.A and Nicolis 1972):

Model 1. Model proposed by Monod.

$$\frac{dX}{dt} = \frac{\mu_{max}(X.S)}{K_S + S} \quad \text{Eq. 1}$$

where:

$\frac{dX}{dt}$: Difference between the final (X) and initial (X_0) concentrations of biomass, according to time ($\text{g.L}^{-1}.\text{h}^{-1}$);

μ_{max} : Maximum specific growth rate (h^{-1});

X: Biomass concentration at time t (g.L^{-1});

S: Substrate concentration at time t (g.L^{-1});

K_S : Substrate saturation constant of the Monod model (g.L^{-1}).

Several studies have found that inhibition by ethanol only affects μ_{max} , with no effect on K_S (Nashed and Girton 1958, Bultmann, Szabo et al. 1993). Levenspiel (1980) reported a generalized mathematical equation to describe cell growth, containing a product inhibition term, which also considered the effect of the limiting substrate (K_S):

Model 2. Model proposed by Levenspiel (1980).

$$\mu = \left(\frac{\mu_{max} \cdot S}{K_S + S} \right) \cdot \left(1 - \frac{P}{P_{max}} \right)^n \quad \text{Eq. 2}$$

In Eq. 2, P_{max} is the product concentration at which microbial growth ceases. Cell growth becomes non-viable at ethanol concentrations higher than indicated by this parameter. For concentrations of P much lower than P_{max} , the Levenspiel (1980) model becomes reduced to the Monod model (Levenspiel 1980).

The reactor model used in this study was described by Daré (2008) in his Master's thesis, considering an ethanol production process (as used in industrial plants) employing the classical batch mode. The process is divided into two steps. In the first step, the cultivation involves a fed-batch operation starting with an inoculum that occupies a volume of approximately one third of the total volume of the fermenter tank, with the must being fed at a constant rate until the vat is full. The second batch stage is then initiated and is continued until complete exhaustion of the fermentable sugars.

The models found in the literature are reasonably accurate in describing the behavior of alcoholic fermentations (considering cell growth, substrate consumption, and product formation) during the final batch step. However, such models do not provide satisfactory fits to the experimental data during the initial step (semi-continuous or fed-batch), because at the beginning of the feeding there is rapid substrate consumption, with consequent production of ethanol without any observed cell growth (which only starts after a certain period of must feeding). It is as if the cells present two different growth kinetics behaviors during the course of the fermentation process. In order to explain this cell behavior observed in industrial processes, Daré (2008) proposed the model described below for the alcoholic fermentation process.

In the Melle-Boinot process, the yeast is recovered from the fermented wine after the end of each batch and, after treatment, is recycled to the process for use in new batches. During the step of treating the yeasts in the presence of oxygen, the cells derived from the centrifugation step of the previous batch, denoted type 2 (X_2), are transformed into type 1 cells (X_1), which do not grow due to the absence of substrate. During the treatment step, the type 1 cells that comprise the inoculum remain for a long time in the absence of substrate, in a condition of starvation that alters their physiological state. Eq. 3 is used to describe the transformation of X_2 to X_1 :



X_1 : Starved cells;

X_2 : Cells derived from centrifugation of the fermented wine, which consume sugar, grow, and produce ethanol.

4.1.2.3. Fed-batch ethanol production step

When the feeding of the must into the vat is started in the initial fed-batch step, in the presence of substrate the transformation of the type 1 cells (X_1) to type 2 cells (X_2) starts again, according to the stoichiometry shown in Eq. 4.



At the beginning of this step, there is rapid consumption of substrate and simultaneous production of ethanol, without any observed cell growth. Hence, the hypothesis is that the type 1 cells consume the substrate rapidly to generate energy (ATP), producing ethanol due to the oxidative metabolization of hexose sugars during glycolysis. Meanwhile, the energy produced (ATP) in this process is not used for cell growth. It is exclusively used in restructuring the cellular machinery, or in a new physiological change, considered in the model as a transformation of the type 1 cells to type 2 cells, which once again have the ability to grow and produce ethanol, while at the same time consuming the substrate.

The rate of transformation of the type 1 cells to type 2 cells, $r_{t_{X_2}}$, is given by Eq. 5:

$$r_{t_{X_2}} = \frac{k_1 + C_{X_1}}{k_2 + C_{X_1}} = r_{t_{X_1}} \quad \text{Eq. 5}$$

Cell death can occur during the production step, as described by Eq. 6:



where X_{2NV} represents non-viable type 2 cells.

The stoichiometry and kinetics of ethanol production can be described by Eqs. 7 and 8 for the type 1 cells, and by Eqs. 9 and 10 for the type 2 cells.



$$\frac{r_{p1}}{p_1} = -r_s = k_p \cdot C_{X1} \quad \text{Eq. 8}$$



$$\frac{r_{p2}}{p_2} = \frac{r_{X2}}{y_1} \rightarrow r_{p2} = \frac{p_2}{y_1} \cdot r_{X2} \rightarrow r_{p2} = Y_{p/X2} \quad \text{Eq. 10}$$

According to the proposed model, unlike the X_1 cells, the X_2 cells are able to grow, consuming the substrate and producing ethanol (P) and more X_2 cells, according to Eq. 9.

The rate of growth of the X_2 cells is given by Eq. 11.

$$r_{X2} = \mu \cdot C_{X2} \quad \text{Eq. 11}$$

Considering that ethanol inhibition of the growth of the type 2 cells follows the Levenspiel (1976) model, the following expression is valid:

$$r_{X2} = \mu_{max} \cdot \frac{C_s}{K_s + C_s} \cdot \left(1 - \frac{C_p}{C_{p*}}\right)^n \cdot C_{X2} \quad \text{Eq. 12}$$

Having presented the hypotheses of the reaction model and the reaction equations, there now follows a presentation of the systems of mass balance equations that constitute the modeling of the process.

4.1.2.4. Fed-batch step

Model 3. Mass balance for the type 1 cells.

$$\frac{dC_{X1}}{dt} = -\frac{k_1 \cdot C_{X1}^m}{k_2 + C_{X1}} - \left(k_d + \frac{F}{V}\right) \cdot C_{X1} \quad \text{Eq. 13}$$

Model 4. Mass balance for the type 2 cells.

$$\frac{dC_{X2}}{dt} = -\frac{k_1 \cdot C_{X1}^m}{k_2 + C_{X1}} + \mu \cdot C_{X2} - \frac{F}{V} \cdot C_{X2} \quad \text{Eq. 14}$$

Model 5. Mass balance for ethanol.

$$\frac{dC_P}{dt} = \rho_1 \cdot k_p \cdot C_{X1} + Y_{P/X2} \cdot \mu \cdot C_{X2} - \frac{F}{V} \cdot C_P \quad \text{Eq. 15}$$

Model 6. Mass balance for the substrate.

$$\frac{dC_S}{dt} = \frac{F}{V} \cdot (C_{Se} - C_S) - k_p \cdot C_{X1} - \frac{\mu \cdot C_{X2}}{Y_{X2/S}} \quad \text{Eq. 16}$$

Model 7. Overall mass balance.

$$\frac{dV}{dt} = \frac{\rho_M}{\rho_C} (F_0 + k \cdot t) \quad \text{Eq. 17}$$

4.1.2.5. Batch step

Model 8. Mass balance for the type 1 cells.

$$\frac{dC_{X1}}{dt} = -\frac{k_1 \cdot C_{X1}^m}{k_2 + C_{X1}} - k_d \cdot C_{X1} \quad \text{Eq. 18}$$

Model 9. Mass balance for the type 2 cells.

$$\frac{dC_{X2}}{dt} = -\frac{k_1 \cdot C_{X1}^m}{k_2 + C_{X1}} + \mu \cdot C_{X2} \quad \text{Eq. 19}$$

Model 10. Mass balance for ethanol.

$$\frac{dC_P}{dt} = \rho_1 \cdot k_p \cdot C_{X1} + Y_{P/X2} \cdot \mu \cdot C_{X2} \quad \text{Eq. 20}$$

Model 11. Mass balance for the substrate.

$$\frac{dC_S}{dt} = -k_p \cdot C_{X1} - \frac{\mu \cdot C_{X2}}{Y_{X2/S}} \quad \text{Eq. 21}$$

Daré (2008) fitted this model to experimental data obtained for pilot scale cultivations using different vat feeding conditions. The equations of the proposed mathematical model (or resulting from mass balance) were integrated using the fourth order Runge-Kutta-Gill method with variable steps (Cheng and Lin 2012). The model parameters were estimated using the Marquardt (1963) algorithm. However, the results presented by Daré (2008) for parameter estimation were not completely conclusive.

4.1.3. Estimation of model kinetic parameters

The procedures used for parameter estimation are critical for the application of mathematical models. The design, optimization, and advanced control of chemical, physical, and biological processes depend on the model parameter values obtained by fitting of the experimental data (Dette, Melas et al. 2005). It is shown here that many difficulties in estimating parameters can be overcome using heuristic optimization techniques, such as the simulated annealing global optimization method.

During the development of the model, variables that could not be measured (or for which measurement was difficult) could be defined using the model equations. In order to evaluate these variables (model parameters) and improve the reliability of the model, the values of the parameters could be calculated from the available experimental data by minimization of an objective function that weights the distances between the model predictions and the experimental results.

Iterative procedures should normally be used to minimize the objective function. However, care must be taken concerning: (a) the size of the parameter search space; (b) the existence of local minima; (c) the continuity of the objective function; and (d) the sensitivity of the objective function to each of the model parameters (Parthasarathy and Balaji 2008).

In the case of linear models, the confidence region of the parameter estimates is described by a hyper-ellipsoid in the parameter space, with the estimated values of the model parameters located in the center of the hyper-ellipsoid. The assumption of an elliptical shape of the confidence region often provides a poor approximation (Murino, Trucco et al. 1998). Nonetheless, elliptical approximations of confidence regions are widely used, due to the simplicity of the method (Marin and Salmeron 1996).

4.1.3.1. Simulated annealing

Optimization procedures are of great interest in contemporary industry, given the importance of avoiding losses or the need for reworking, and the possibility of higher yields at lower costs. In the case of fermentation processes, the use of kinetic parameters is extremely important to ensure consistent modeling.

In a carefully constructed model of a system, optimization algorithms are used to achieve a higher yield. Simulated annealing is a probabilistic tool widely used for global optimization when the search space is very large. The algorithm, based on processes for annealing of solids (Metropolis, Rosenbluth et al. 1953), tries to find an optimal solution in combined optimization problems by iteratively exploring new areas in the solution space of the problem (Musharavati and Hamouda, 2012). The performance of this technique can be assessed by the value found for the objective function, observing the number of iterations required to achieve a global optimum result (Pinto et al., 2009).

Table 1 shows the general form of the pseudocode of the simulated annealing method, where the main parameters are the model variables, X_j , the objective function, f , the instantaneous and initial temperatures of the system, T and T_0 , the maximum number of iterations of the model, N_{max} , the probability generation function, and the neighbors, $R(x_i)$. The responses of the model are the functions f_b and x_b . The initial temperature and the maximum number of iterations are arbitrary values defined before the start of the iterations. The temperature decay factor (α) is a constant value or a function $\alpha = f(T)$ belonging to the interval (0; 1).

Table 2. Simulated annealing pseudocode.

```

fb=+∞;
x0→Initial solution;
fm=f(x0);
I=0 → Number of iterations;
T=T0 →Initial temperature of the system;
  if i ≤ Nmax then:
    xj=R(x0);
    I=I+1
    If f(xj)<f(x0) then:
      xm=x0;fm=f(x0);
    end if;
    if not:
      Δf=f(xj)-f(x0);P(f)→unif.[0:1];
      if P(f)≤e- Δf/T then:
        xm=x0;fm=f(x0);
      end if;
    end if;
    T=T.α;
    If fb<fm then:
      fb=fm;xb=xm;
    end if;
  end when;

```

In their work, Eftaxias, Font et al. (2002) evaluated the performance of simulated annealing for the determination of nonlinear kinetic parameters, compared to the classical Levenberg-Marquardt algorithm. Both methods were tested in determination of kinetic parameters using a set of three kinetic models of progressively greater complexity, and both models achieved convergence.

Another approach based on simulated annealing was proposed by El-Naggar, AlRashidi et al. (2012) to obtain the parameters of an ideal solar cell model. Comparative studies of different techniques demonstrated the effectiveness of the approach adopted.

In the present work, the use of simulated annealing involved minimization of the error function, employing the experimental data of Daré (2008), with simulation using the probabilistic heuristic model also described by Daré (2008).

4.2. MATERIALS AND METHODS

4.2.1. Model

The alcoholic fermentation simulations were performed using the model proposed by (Daré 2008), described above.

4.2.2. Simulated annealing algorithm

Minimization of the ERROR function (Table 2) was performed using code for the simulated annealing algorithm implemented in Matlab. A new point was generated randomly for each iteration of the simulated annealing algorithm. The distance of the new point to the current point (or the extent of the search) was based on a probability distribution with a scale proportional to the temperature. The algorithm accepted all new points that reduced the objective function, as well as (according to a certain probability) points that increased the objective function. By accepting points that increased the objective function, the algorithm avoided becoming trapped in local minima in the initial iterations and was able to explore globally for better solutions. The simulated annealing pseudocode is given in Table 2. The experimental data used were from the cultivations performed by Daré (2008).

Table 2. Error function to be minimized.

$$\text{ERROR} = [\text{Curve generated by the model simulation} - \text{Curve for the experimental data}]^2$$

4.3. RESULTS AND DISCUSSION

The results of the simulations are presented in Figures 18-25.

Figure 18 Culture P15.

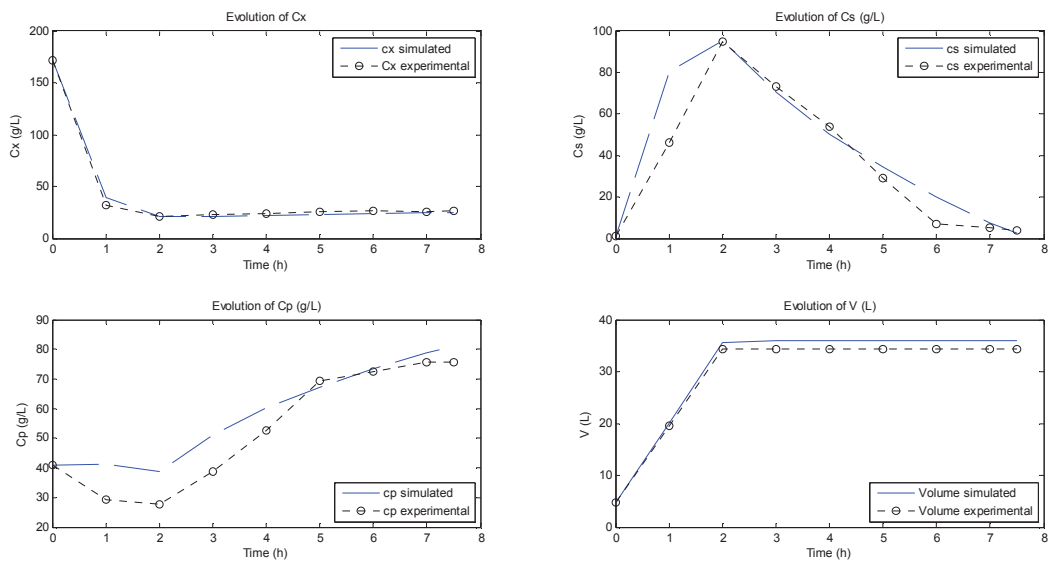


Figure 19. Linear confidence region for culture P15.

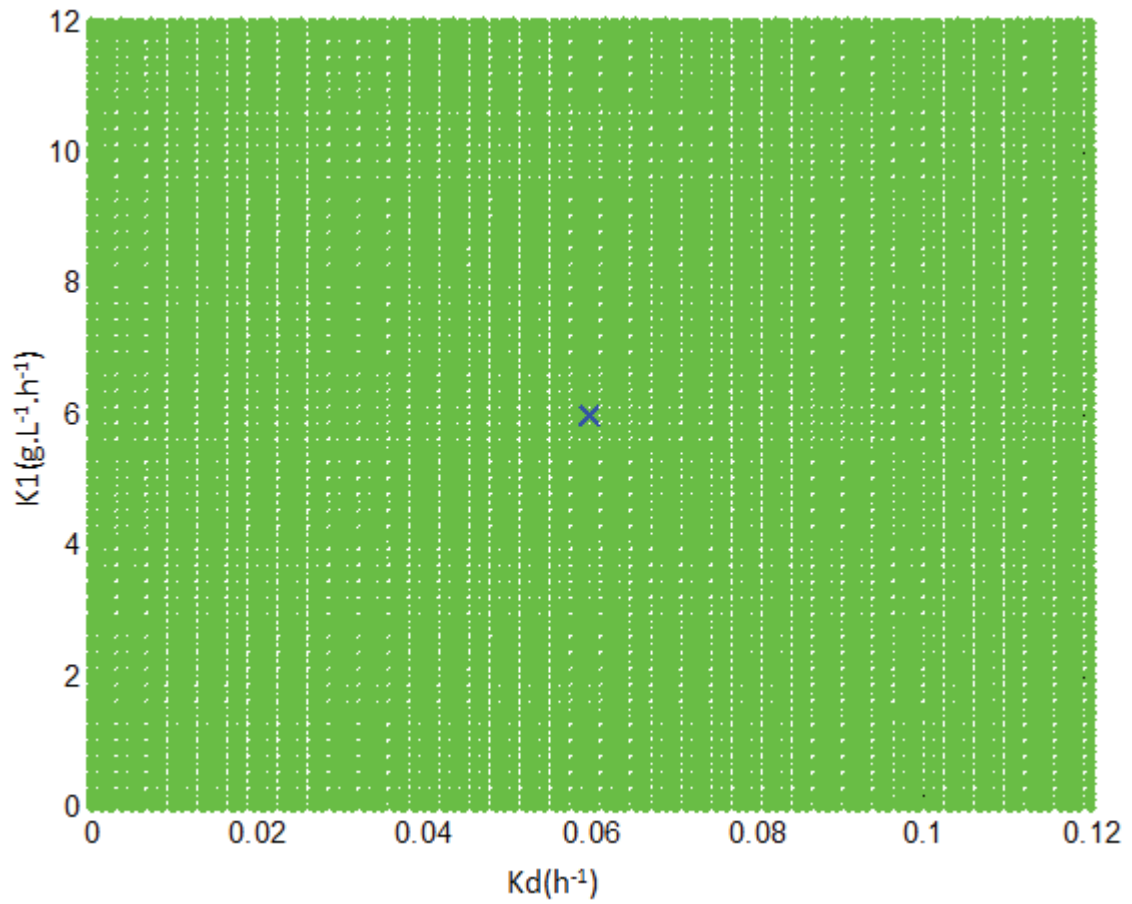


Figure 20. Culture P55A.

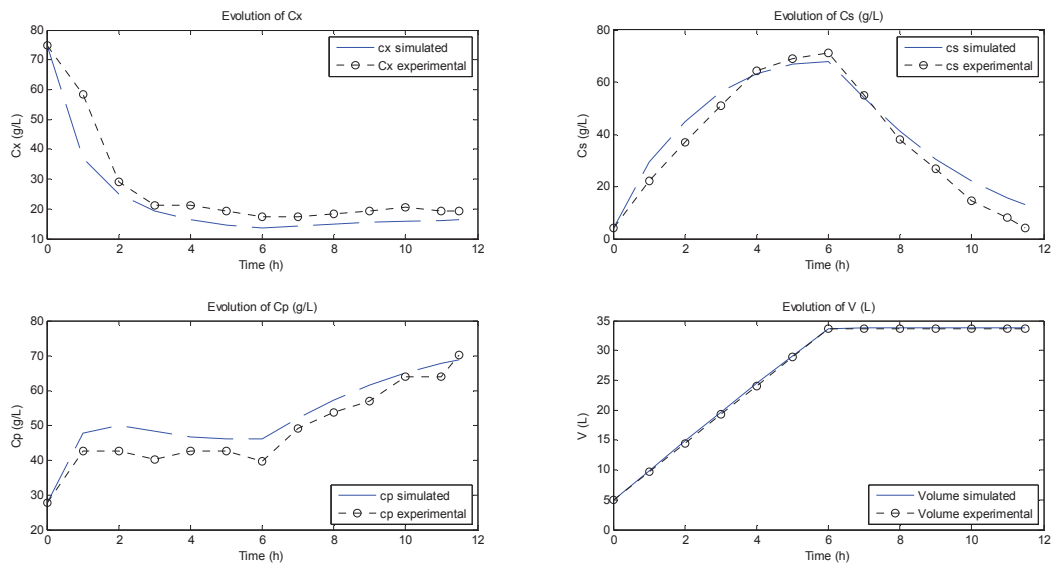


Figure 21. Linear confidence region for culture P55A.

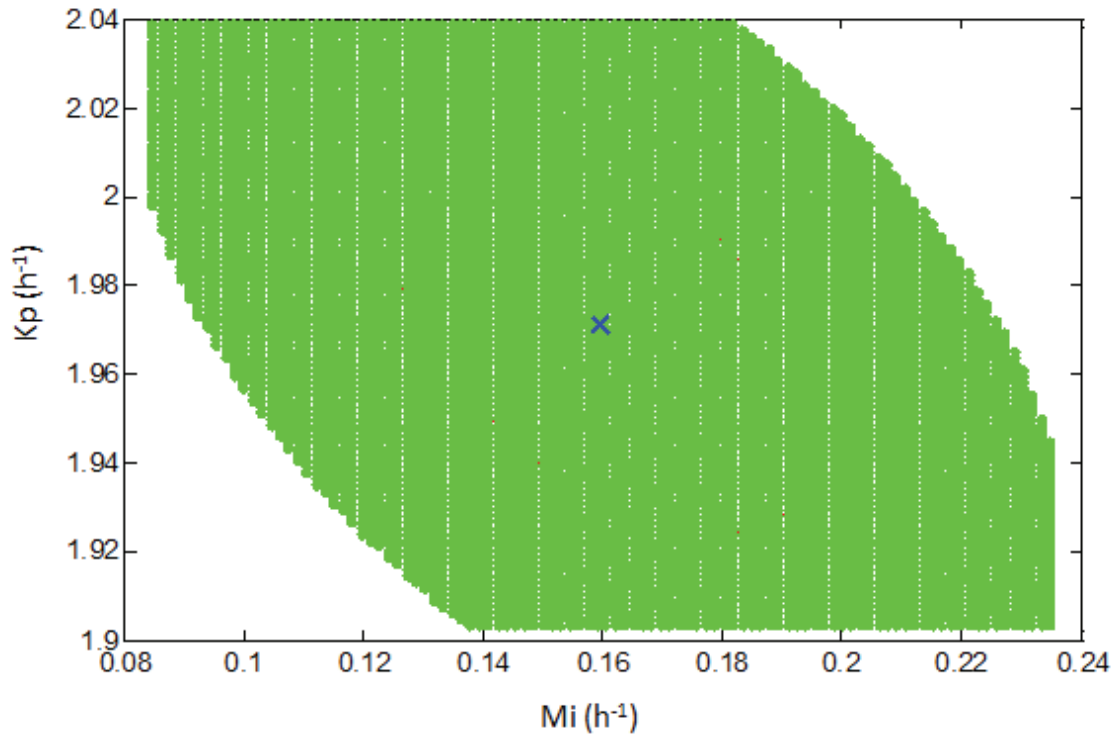


Figure 22. Culture P.7.5-7.5A.

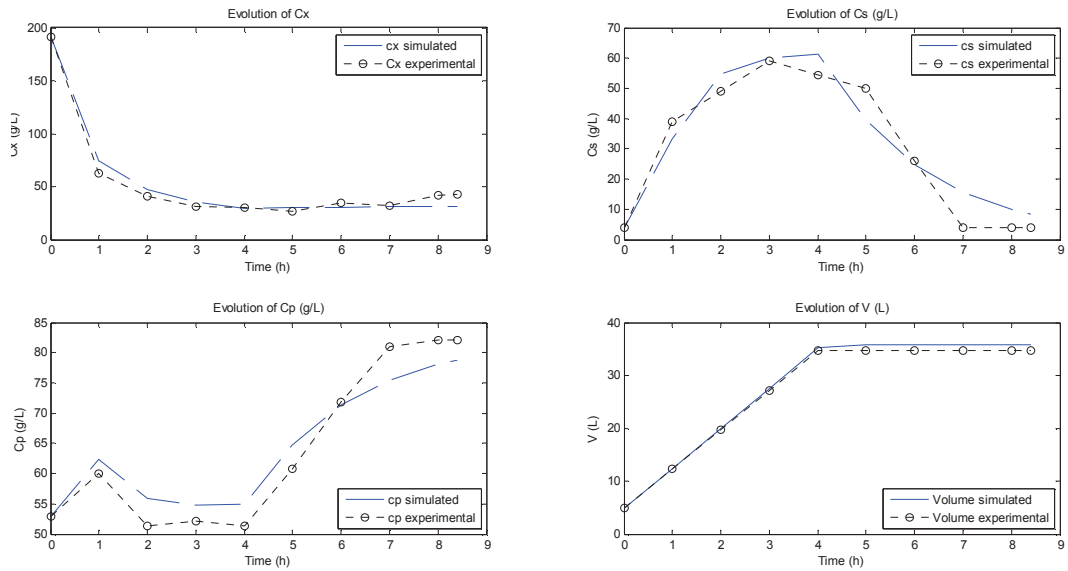


Figure 23. Linear confidence region for culture P.7.5-7.5A.

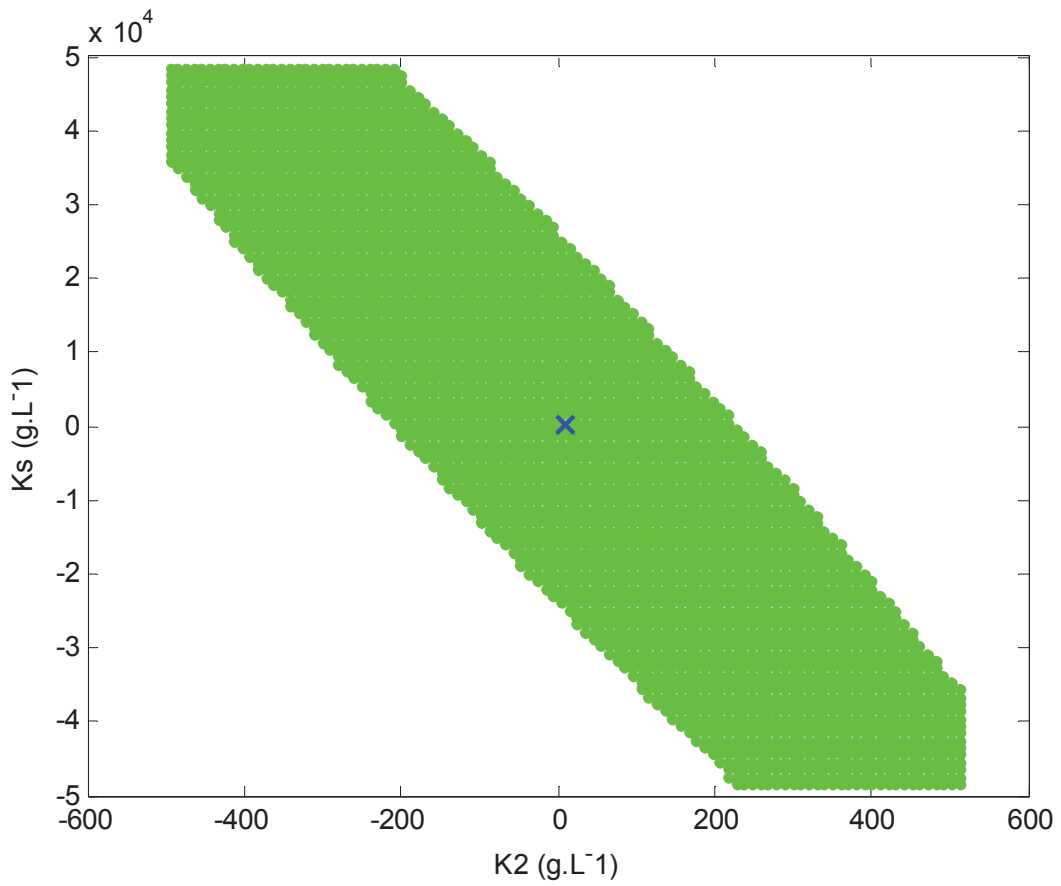


Figure 24. Culture P.7.5-7.5B.

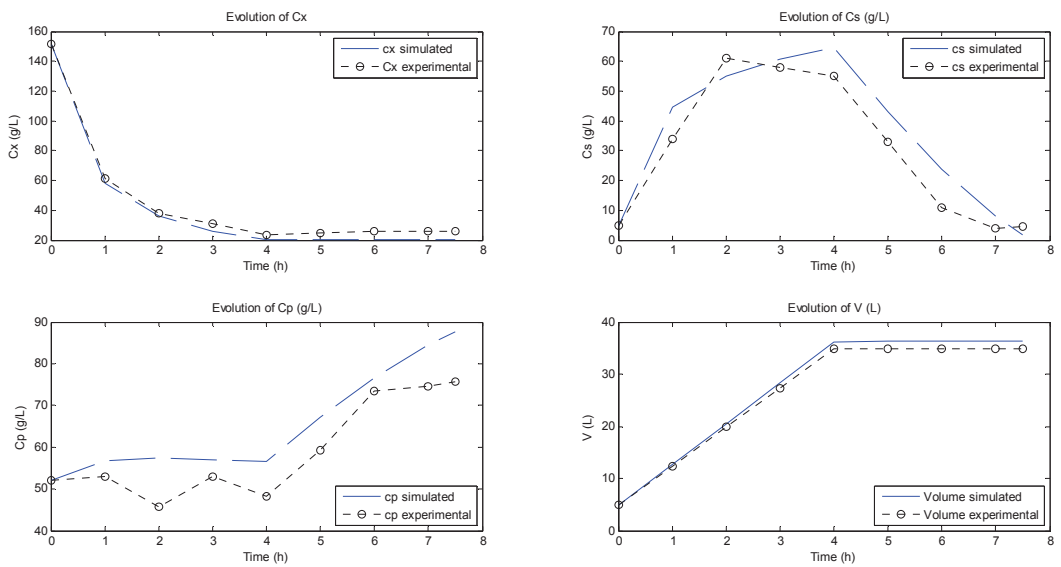


Figure 25. Linear confidence region for culture P.7.5-7.5B.

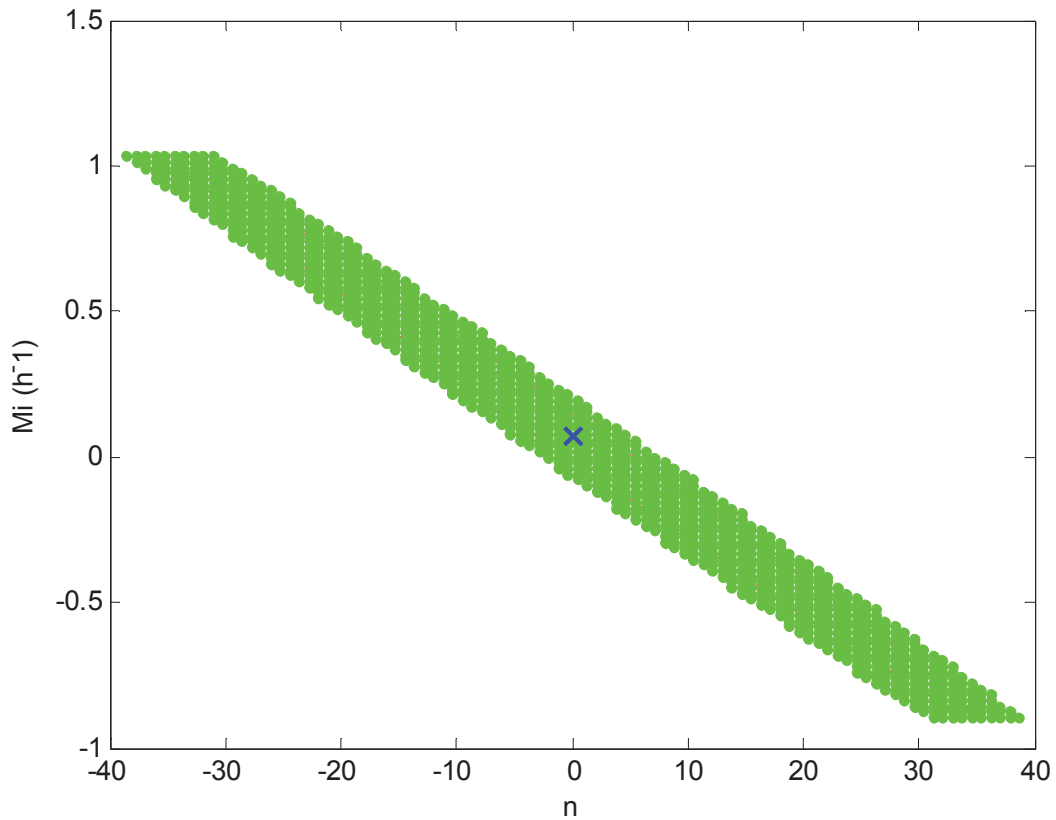
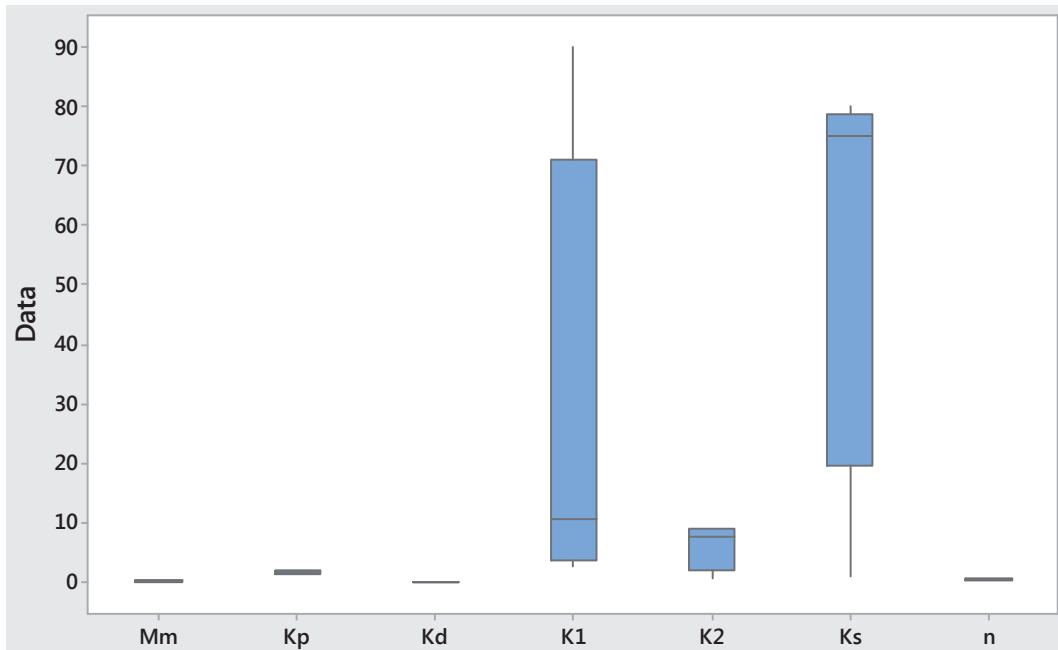


Table 3 shows the parameter values obtained by error minimization using the simulated annealing technique.

Table 3. Kinetic parameter values for the pilot plant cultures.

Culture	μ_m h^{-1}	K_p h^{-1}	K_d h^{-1}	K_1 $g.L^{-1}.h^{-1}$	K_2 $g.L^{-1}$	K_s $g.L^{-1}$	n
P-15-15	0.08	1.57	0.06	6.00	0.60	1.00	0.37
P-5-5 A	0.16	1.97	0.02	14.92	8.89	74.98	0.32
P-7.5-7.5 A	0.13	1.42	0.07	89.98	8.85	74.98	0.47
P-7.5-7.5 B	0.07	1.23	0.01	2.58	6.06	80.00	0.04

Figure 26. Boxplot of Mm, Kp, Kd, K1, K2, Ks; and n.



It can be seen from the graphs shown in Figures 18, 20, 22, and 24, produced using simulated annealing error minimization, that the simulation curves closely followed the experimental points. Figures 19, 21, 23, and 25, constructed using the tool implemented in Matlab, provide information on the confidence intervals for the kinetic parameters (in pairs).

Figures 18, 20, 22, and 24 show that all the pilot plant cultures presented similar concentration profiles, for the range of experimental feed conditions studied. Good agreement was obtained between the simulated and experimental results for the cell profiles. The substrate concentration profiles presented differences that were more significant, which could be attributed, amongst other factors, to the lack of homogeneity of the medium. Greater differences were found for the ethanol concentration profiles, which could also be attributed to the difficulty in collecting homogeneous samples.

Table 3 presents the kinetic parameters of the proposed model determined from fitting of the model to the experimental values obtained for the pilot plant cultures. The graph in Figure 26 shows the variation of the estimated kinetic coefficients, from which

it can be seen that there was acceptable variability between the values obtained for the different cultures, with greater variation only for K1 and Ks.

4.4. CONCLUSIONS

The meta-heuristic simulated annealing technique was used to search for the optimum parameters of the Monod-Levenspiel model, considering a single output. All the parameters were determined simultaneously, avoiding the need to use previously established values, as frequently required in the case of parameter spaces with large dimensions when classical local search algorithms are used (such as the Marquardt algorithm). Satisfactory results were obtained, although it is possible that the global search method could be refined using local search procedures.

4.5. REFERENCES

Luchi, N. R. and S. C. Trindade "ETHANOL VERSUS NAPHTHA UNDER BRAZIL ECONOMY." Hydrocarbon Processing **61**(5): 179-183 (1982).

Cerqueira Leite, R. C. d., et al. "Can Brazil replace 5% of the 2025 gasoline world demand with ethanol?" Energy **34**(5): 655-661 (2009).

Kiatkittipong, W., et al. "Life cycle assessment of bagasse waste management options." Waste Management **29**(5): 1628-1633 (2009).

Nunez, H. M., et al. "Land use and economic effects of alternative biofuel policies in Brazil and the United States." Agricultural Economics **44**(4-5): 487-499 (2013).

Furtado, A. T., et al. "The Brazilian sugarcane innovation system." Energy Policy **39**(1): 156-166 (2011).

Tsao, C. C., et al. "Increased estimates of air-pollution emissions from Brazilian sugarcane ethanol." Nature Climate Change **2**(1): 53-57 (2012).

Munack, A. (1989). OPTIMAL FEEDING STRATEGY FOR IDENTIFICATION OF MONOD-TYPE MODELS BY FED-BATCH EXPERIMENTS.

Cosenza, A., et al. "Parameter estimation and sensitivity analysis of a nitrogen and phosphorus biological removal model." 18th World Imacs Congress and Modsim09 International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences: 3151-3157 (2009).

Almeida, L. M. W. d. and B. N. Palharini "Os "Mundos da Matemática" em Atividades de Modelagem Matemática." Bolema: Boletim de Educação Matemática **26**(43): 907-934 (2012).

Kurganov, B. I., et al. "THEORETICAL-ANALYSIS OF KINETIC-BEHAVIOR OF HYSTERETIC ALLOSTERIC ENZYMES .1. KINETIC MANIFESTATIONS OF SLOW CONFORMATIONAL CHANGE OF AN OLIGOMERIC ENZYME IN MONOD, WYMAN AND CHANGEUX MODEL." Journal of Theoretical Biology **60**(2): 247-269 (1976).

Goldbeter, A. "KINETIC NEGATIVE COOPERATIVITY IN ALLOSTERIC MODEL OF MONOD, WYMAN AND CHANGEUX." Journal of Molecular Biology **90**(1): 185-190 (1974).

Golterman, H. L. "REFLECTIONS ON POST-OECD EUTROPHICATION MODELS." Hydrobiologia **218**(2): 167-176 (1991).

Carstensen, J., et al. "STATISTICAL IDENTIFICATION OF MONOD-KINETIC PARAMETERS FROM ONLINE MEASUREMENTS." Water Science and Technology **31**(2): 125-133 (1995).

Babloyan.A and G. Nicolis "CHEMICAL INSTABILITIES AND MULTIPLE STEADY-STATE TRANSITIONS IN MONOD-JACOB TYPE MODELS." Journal of Theoretical Biology **34**(1): 185-& (1972).

Nashed, R. B. and R. E. Girton "INHIBITION BY ETHANOL OF THE GROWTH AND RESPIRATION OF MAIZE ROOTS AND COLEOPTILES." American Journal of Botany **45**(3): 190-193 (1958).

Bultmann, R., et al. "INHIBITION BY ETHANOL OF CONTRACTIONS OF RAT VAS-DEFERENS - NO EVIDENCE FOR SELECTIVE BLOCKADE OF P(2X)-PURINOCEPTORS." Naunyn-Schmiedebergs Archives of Pharmacology **347**(5): 527-533 (1993).

Levenspiel, O. "THE MONOD EQUATION - A REVISIT AND A GENERALIZATION TO PRODUCT INHIBITION SITUATIONS." Biotechnology and Bioengineering **22**(8): 1671-1687 (1980).

Cheng, W. T. and C. H. Lin "NUMERICAL ANALYSIS OF MASS TRANSFER ON THE MELTING PHENOMENON FROM A VERTICAL PLATE IN A LIQUID-SATURATED POROUS MEDIUM." Journal of Porous Media **15**(5): 485-493 (2012).

Detle, H., et al. "Robust and efficient design of experiments for the Monod model." Journal of Theoretical Biology **234**(4): 537-550 (2005).

Parthasarathy, S. and C. Balaji "Estimation of parameters in multi-mode heat transfer problems using Bayesian inference - Effect of noise and a priori." International Journal of Heat and Mass Transfer **51**(9-10): 2313-2334 (2008).

Murino, V., et al. Edge/region-based segmentation and reconstruction of underwater acoustic images by Markov Random Fields. 1998 Ieee Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings: 408-413.(1998).

Marin, A. and J. Salmeron "Tactical design of rail freight networks .2. Local search methods with statistical analysis." European Journal of Operational Research **94**(1): 43-53 (1996).

Metropolis, N., et al. "EQUATION OF STATE CALCULATIONS BY FAST COMPUTING MACHINES." Journal of Chemical Physics **21**(6): 1087-1092 (1953).

Eftaxias, A., et al. "Nonlinear kinetic parameter estimation using simulated annealing." Computers & Chemical Engineering **26**(12): 1725-1733 (2002).

El-Naggar, K. M., et al. "Simulated Annealing algorithm for photovoltaic parameters identification." Solar Energy **86**(1): 266-274 (2012).

Daré, R. M. AVALIAÇÃO DOS COEFICIENTES DE RENDIMENTO E MODELAGEM DO PROCESSO FERMENTATIVO DE PRODUÇÃO DE ETANOL. Programa de pós graduação em Engenharia Química. São Carlos, Universidade Federal de São Carlos. Mestrado: 67 (2008).

**5. CAPÍTULO 05: OPTIMIZATION OF
THE PRODUCTION OF
INACTIVATED *CLOSTRIDIUM NOVYI*
TYPE B VACCINE USING
COMPUTATIONAL INTELLIGENCE
TECHNIQUES**

Optimization of the production of inactivated *Clostridium novyi* type B vaccine using computational intelligence techniques

Aquino, P. L. M.^{1,2} Fonseca F. S.² Mozzer O. D.² Giordano R. C.¹

Sousa Jr. R.^{1,*}

1- Federal University of Sao Carlos – Programa de Pós-Graduação em Engenharia Química - UFSCar. Rodovia Washington Luís, km 235. Zip code 13565-905, Monjolinho – São Carlos – São Paulo, Brazil

*e-mail: ruy@ufscar.br

2- Vallée S/A - Av. Comendador Antônio Loureiro Ramos, 1500, Zip Code 39404-620, Distrito Industrial, Montes Claros, Minas Gerais, Brazil

Optimization of the Production of Inactivated *Clostridium novyi* Type B Vaccine Using Computational Intelligence Techniques

P. L. M. Aquino^{1,2} · F. S. Fonseca² · O. D. Mozzer² ·
R. C. Giordano¹ · R. Sousa Jr.¹

Received: 19 December 2015 / Accepted: 1 March 2016
© Springer Science+Business Media New York 2016

Abstract *Clostridium novyi* causes necrotic hepatitis in sheep and cattle, as well as gas gangrene. The microorganism is strictly anaerobic, fastidious, and difficult to cultivate in industrial scale. *C. novyi* type B produces alpha and beta toxins, with the alpha toxin being linked to the presence of specific bacteriophages. The main strategy to combat diseases caused by *C. novyi* is vaccination, employing vaccines produced with toxoids or with toxoids and bacterins. In order to identify culture medium components and concentrations that maximized cell density and alpha toxin production, a neuro-fuzzy algorithm was applied to predict the yields of the fermentation process for production of *C. novyi* type B, within a global search procedure using the simulated annealing technique. Maximizing cell density and toxin production is a multi-objective optimization problem and could be treated by a Pareto approach. Nevertheless, the approach chosen here was a step-by-step one. The optimum values obtained with this approach were validated in laboratory scale, and the results were used to reload the data matrix for re-parameterization of the neuro-fuzzy model, which was implemented for a final optimization step with regards to the alpha toxin productivity. With this methodology, a threefold increase of alpha toxin could be achieved.

Keywords *Clostridium novyi* · Neuro-fuzzy · Simulated annealing · Mathematical modeling of bioprocesses · Culture medium optimization

✉ R. Sousa, Jr.
ruy@ufscar.br

¹ Programa de Pós-Graduação em Engenharia Química, UFSCar, Federal University of Sao Carlos, Rodovia Washington Luís, km 235, 13565-905 Monjolinho – São Carlos, São Paulo, Brazil

² Vallée S/A - Av. Comendador Antônio Loureiro Ramos, 1500, Distrito Industrial, 39404-620 Montes Claros, Minas Gerais, Brazil

Introduction

Clostridium novyi causes necrotic hepatitis of sheep and cattle, known as “black disease.” It also causes gas gangrene following wounds [1]. The main characteristics of this microorganism are the following: large rods (measuring 5 to 10 μm in length and 1 to 1.5 μm in diameter) (Fig. 1); central or subterminal oval spores; non-capsulated; peritrichously flagellated (mobile under anaerobic conditions); gram-positive when young, but becoming gram-variable after subculturing several times or after long culture periods; and global distribution, being predominantly present in the soil [2, 3].

Clostridium novyi is strictly anaerobic, fastidious, and difficult to cultivate. Growth rates can be increased by the addition of glucose, blood, or brain to the culture medium [4]. The colonies are flat with irregular borders, spreading over the agar surface in the form of a film and showing transparent growth [5], as in Fig. 2. When grown in solid media with blood, hemolysis occurs after around 48–72 h of incubation [6]. In media containing cooked meat, growth is slow, the meat becomes rose-colored, the gelatin is liquefied, and the serum is coagulated but not digested. Colonies grown in media containing egg yolk develop an opalescent precipitation zone (the Nagler reaction), due to the production of lecithinase (gamma toxin) by *C. novyi* type A or beta toxin by *C. novyi* types B and D. The fermentation of carbohydrates is variable, with acid and gas produced from glucose, while lactose is not fermented [7].

C. novyi is divided into four types (A, B, C, and D), with type D recognized as *Clostridium haemolyticum*, according to the production of toxins (alpha, beta, gamma, and delta) [8]. *C. novyi* type B (the organism studied in the present work) produces alpha and (to a lesser extent) beta toxins. In this microorganism, the production of alpha toxin is linked to the presence of specific bacteriophages. In the absence of these bacteriophages, only beta toxin is produced, in a similar way as *C. novyi* type D (*C. haemolyticum*) [9].

The main method for combating diseases caused by *C. novyi* is vaccination. These vaccines normally contain toxoids (or toxoids and bacterins) in their composition [1].

The cultivation of *C. novyi* is difficult, as reported in the literature and observed in laboratory scale tests. The latter showed that the fermentation time could be about 100 h or more, and that toxin might or might not be produced in sufficient amounts for vaccine production. Very large variations were found in fermentation time and toxin production.

Fig. 1 Rods of *Clostridium novyi*

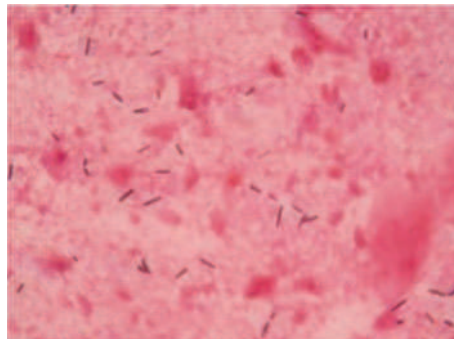


Fig. 2 Colonies of *Clostridium novyi* on agar medium



Objectives

General Objective

The aim of this work was to test different culture medium components and their concentrations in order to maximize productivity of the alpha toxin of *C. novyi* type B.

Specific Objective

The specific objective of the study is to identify the best culture medium for growth of *C. novyi* by maximizing cell density, toxin production and (ultimately) toxin productivity, utilizing computational techniques and laboratory scale tests to confirm the results.

Neuro-Fuzzy Systems

Neuro-fuzzy networks are systems that endeavor to combine the advantages of fuzzy logic and artificial neural networks (ANNs), hence reducing deficiencies and errors associated with the individual use of either technique [10]. The methods can be complementary, because while fuzzy logic assumes previous knowledge of the problem [11], the basic characteristics of artificial neural networks are their ability to learn and adapt starting from a set of input and output data, without any prior knowledge of the system [12].

Fuzzy Logic and Neural Networks

Fuzzy logic was proposed as a way of addressing the constant emergence of uncertainties in a system that it is desired to optimize, because vague and uncertain expressions, common to human thinking and communication, are not accepted by traditional logic [13].

Fuzzy theory emerged in 1965, when Lofti A. Zadeh published the article entitled *Fuzzy Sets* in the journal *Information and Control* [14]. One of the great advantages of fuzzy logic is the ability to use the properties of fuzzy sets to translate linguistic terms employed in natural language into mathematical expressions [11].

Zadeh [15] proposed the flexible association of elements with the various sets comprising a system, creating the notion of degree of membership [16]. In this way, an element could belong partially or wholly to a given set, and belong to more than one fuzzy set, with different degrees of membership [17, 18]. Fuzzy modeling (or control) uses human experience to generate a set

of norms able to translate logical reasoning, representing the mathematical model of a system by means of linguistic variables and a series of rules of the type IF...THEN [19].

Artificial neural networks are systems structured in a computational approach based on connections, where the nodes of the neural network are interlinked to form a mesh. The original inspiration for this technique emerged from examination of the structures of the brain, notably the neurons [20, 21]. The most important property of neural networks is the ability to learn from their environment in order to improve performance. This is achieved by means of an iterative process for tuning the weights of the connections, known as training [22, 23]. The network must respond in the most correct way, by means of the generalization effect, to any other stimuli presented to it in a phase subsequent to the training, denoted testing [24, 25].

Neuro-Fuzzy Hybrid Systems

The neuro-fuzzy hybrid system offers an especially interesting alternative for the modeling of systems, because it can extract the learned content of neural networks, with the advantages of the interpretation capability of fuzzy logic [26].

There are three basic classes of hybridization: embedded, auxiliary, and sequential. Neuro-fuzzy applications usually employ embedded hybrid systems, since this class of hybrid system makes the best use of the advantages of the two component techniques [27, 28]. It should be pointed out that the only hybrid system in which it is not possible to separate the two subsystems is the embedded system, reflecting its greater degree of hybridization [29].

Neuro-fuzzy systems that combine the learning ability of ANNs with the linguistic interpretation capacity of fuzzy systems are becoming more widely adopted. These systems possess several important characteristics, such as the applicability of various learning algorithms used by neural networks and the improved ability to interpret and understand knowledge, mainly because the extraction of knowledge is achieved by means of fuzzy rules [30].

The basic form of implementation of neuro-fuzzy networks is that the learning algorithms of the neural networks are used to adjust the parameters of the fuzzy inference system (FIS) [31]. The structure of a neuro-fuzzy system is illustrated in Fig. 3.

The hybrid method employs fundamental characteristics of each separate technique, combining and adapting them to the class of problem to be solved. The contributions of the fuzzy function of the hybrid system are partitioning of the input and output spaces, fuzzy modeling, and defuzzification. The artificial neural network function, in turn, concerns the type of learning, identification of the structure, and identification of the parameters.

There are various possible configurations for each of the characteristics listed above (fuzzy and neural network). The set of methods employed creates a specific type of neuro-fuzzy system. The most widely used are the following: ANFIS (Adaptive-Neuro-based Fuzzy Inference System) [32], FSOM (Fuzzy Self-Organized Map) [33], NEFCLASS (Neuro Fuzzy Classification) [34], NEFCON (Neuro-Fuzzy Control) [35], and NEFPROX (Neuro-Fuzzy Function Approximation) [36].

As pointed out by [37], it is important to evaluate the advantages and disadvantages of the use of neuro-fuzzy systems.

Simulated Annealing

Process optimization is one of the most important issues in modern industries, due to the need to maximize productivity, while avoiding losses and minimizing costs. The simulated

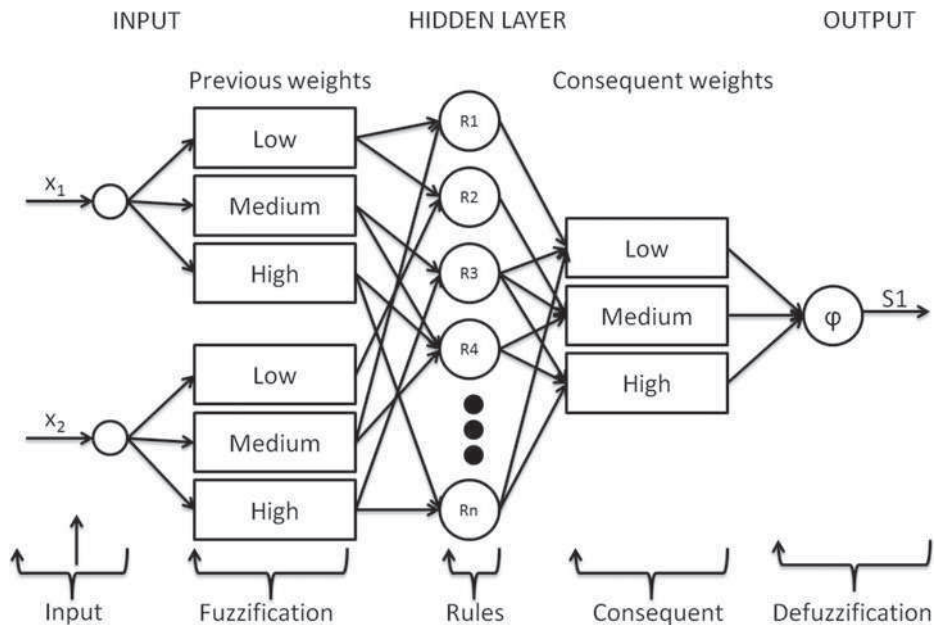


Fig. 3 Basic architecture of a neuro-fuzzy system

annealing technique is a probabilistic tool that is widely used for global optimization when the search space is very extensive. The original inspiration for the algorithm were processes of reheating solids [38]. The analogy with combinatorial optimization was introduced by Kirkpatrick et al. [39] and refined by Cerny [40].

The algorithm tries to find the best solution to problems involving combinatorial optimization, iteratively exploring new areas of the solution space [41]. The stop criterion for the algorithm is the value found for the objective function, observing the restriction of a maximum number of iterations [42, 43].

Figure 4 shows the general form of the pseudocode of the simulated annealing method, designed for general applications, in which the main parameters of the method are the variables of the model, x_j , the objective function, f , the instantaneous and initial temperatures of the system, T and T^0 , the maximum number of iterations of the model, N_{max} , the probability generation function, and the neighbors, $R(x_i)$. The responses of the model are the functions f_b and x_b . The initial temperature and the maximum number of iterations are arbitrary values defined prior to initiation of the iterations. The temperature decay factor (α) is a constant value or a function $\alpha = f(T)$ belonging to the interval (0;1).

In this work, the data matrix with the formulation of the culture media used in the cultivations of *C. novyi* type B was the input for a neuro-fuzzy system, while the output were the respective cell concentrations and productivities with respect to the alpha toxin. A global search was then performed in order to find components of the culture medium, and their concentrations, that could maximize microbial growth and the production of alpha toxin.

Fig. 4 Simulated annealing pseudocode

<pre> $f_b = +\infty$; $x_0 \rightarrow$ Initial solution; $f_m = f(x_0)$; $x_m = x_0$ $i = 0 \rightarrow$ Number of iterations; $T = T^0 \rightarrow$ Initial temperature of the system; While $i \leq N_{\max}$ then: $x_i = x_m$ $x_j = R(x_i)$; If $f(x_j) < f(x_i)$ then: $x_m = x_j$; $f_m = f(x_j)$; end; If not: $\Delta f = f(x_j) - f(x_i)$; $P(f) \rightarrow \text{unif.}[0:1]$; If $P(f) \leq e^{(-\Delta f/T)}$ then: $x_m = x_j$; $f_m = f(x_j)$; end; end; $T = T \cdot \alpha$; $i = i + 1$ If $f_b > f_m$ then: $f_b = f_m$; $x_b = x_m$; end; end. </pre>	<div style="border: 1px solid black; padding: 2px; width: fit-content; margin: auto;"> general minimizing problem </div>
--	--

It is important to comment here on the motivation for using neuro-fuzzy techniques in the context of the present work. Previous works did not use these techniques with similar goals. Usual techniques as optimal design of experiments are the most commonly used [44, 45]. In this sense, the presented approach is innovative.

Methodology

Microbial Cultures

Strain Used

The microorganism used in the fermentations was *C. novyi* FD470, provided by the Technology and Vaccines Laboratory (LTV) of Vallée S/A.

Selection of Culture Medium

Selection of the components of the culture media used for *C. novyi* was based on cultures already tested in the R&D laboratory at Vallée S/A. Documented historical data for the culture

medium formulations provided values for toxin production (quantified in terms of log L+/mL), in different times (t); selection of the media was based on experience with cultures of other *Clostridium* species produced by Vallée S/A.

Quantification of Toxin

Quantification of the product (toxin) can be achieved by means of toxin production tests employing the in vitro cytotoxicity method developed and standardized by Lima et al. [46].

Medium Optimization Procedure

Firstly, a culture medium composition data matrix was constructed using fermentations with 8 L of medium in a Bioflow® benchtop fermenter using the following parameters for all fermentations:

- Agitation speed 50 rpm
- Aeration with nitrogen 1 vvm for 30 min before initiation of fermentation
- NaOH Added automatically to maintain pH 7.5
- HCl Added automatically to maintain pH 7.5

The culture media used as inputs for the neuro-fuzzy network are shown in Table 1.

Table 1 Culture medium formulation input data matrix

Quantity (g/L)					
Component	Chauvoei	LTV29	Thioglycolate fluid	LTV05	LTV103
Component 01	+1	+1	-1	+1	+1
Component 02	-1	+1	-1	+1	-1
Component 03	+1	+1	-1	+1	0
Component 04	+1	+1	-1	+1	-1
Component 05	+1	+1	-1	+1	+1
Component 06	+1	+1	0	+1	+1
Component 07	-1	-1	+1	-1	-1
Component 08*	-1	+1	-1	-1	-1
Component 09*	-1	+1	-1	+1	-1
Component 10	-1	-1	+1	-1	-1
Component 11	-1	-1	0	-1	+1
Component 12	-1	-1	+1	-1	-1
Component 13	-1	-1	+1	-1	-1
Component 14	-1	-1	+1	-1	-1
Component 15	-1	-1	-1	0	+1
Component 16	-1	-1	-1	-1	+1
Component 17	0	0	0	0	0

* Volume fraction (mL/L)

The presentation of quantities for the substrates (components) is done comparatively. An average concentration value (in regard to all the different culture media in Table 1) was considered as reference for each component (C_i^{avg}). Besides, their maximum (p_i^{max}) and minimal (p_i^{min}) concentration values were taken into account (again, for each component). So, the quantity of each component in a specific culture medium is presented as being +1, 0, or -1 considering the following:

$$C_i^{avg} + 0.17 \times (p_i^{max} - p_i^{min}) < C_i \leq p_i^{max}, C_i \text{ is presented as } +1$$

$$C_i^{avg} - 0.17 \times (p_i^{max} - p_i^{min}) \leq C_i \leq C_i^{avg} + 0.17 \times (p_i^{max} - p_i^{min}), C_i \text{ is presented as } 0$$

$$p_i^{min} \leq C_i < C_i^{avg} - 0.17 \times (p_i^{max} - p_i^{min}), C_i \text{ is presented as } -1$$

In summary, for presenting this part of the industrial data from Vallée, the components of the culture medium were classified in three levels:

- +1: for high concentration of substrate (component) i ,
- 0: for an average composition of said substrate, and
- -1: when the said substrate is in minimal amounts in the culture medium.

It is important to stress that this classification is only for presenting the quantities for the substrates (components) in each culture medium (the levels +1, 0, and -1 are not weights of the neuro-fuzzy network).

Chauvoei media is used industrially by Vallée S/A. Thioglycolate fluid is a commercial medium. The LTV29 medium was formulated using design of experiments (DOE) methodology, considering maximization of cellular density, according to the scheme shown in Fig. 5, for each individual LTV-NX experiment.

The LTV05 and LTV103 media were produced using simulated annealing optimization with the output of the neuro-fuzzy model. For initialization of the procedure, the data matrix of the three other media previously described (Chauvoei, Thioglycolate fluid and LTV29) was employed. For generating the LTV05 medium, the objective function was the optical density of the system. For generating LTV103, on the other hand, production of alpha toxin was considered.

As the final output data for neuro-fuzzy network, a data matrix was constructed with all values of toxin production (log L+/mL) (Table 2) and time (Table 3), in order to generate a table for productivity (Table 4), which was to be ultimately maximized. Maximizing cell density, toxin production, and productivity is in fact a multi-objective optimization problem

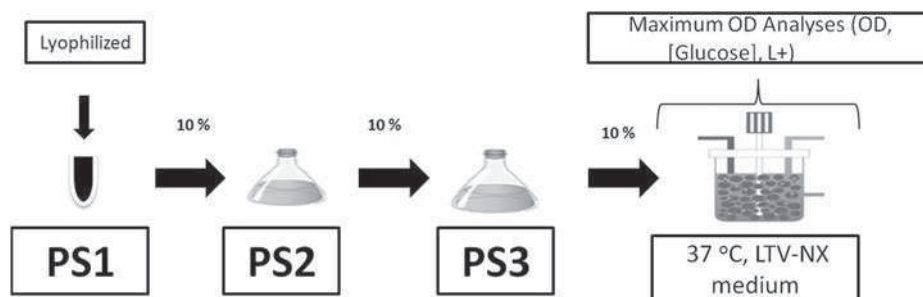


Fig. 5 Laboratory process for production of *Clostridium novyi*

Table 2 Output data matrix for toxin production (log L+/mL)

Log L+/mL				
Chauvoei media	LTV29	Thioglycolate fluid	LTV05	LTV103
2.045	1.491	1.47	1.037	1.62

and could be treated by a Pareto approach [47]. Nevertheless, the approach chosen here was a step-by-step one, i.e., using a global mono-objective algorithm (simulated annealing).

The GENFIS2 (*Generates a Sugeno-type FIS using Subtractive Clustering*) function, implemented in Matlab[®], was used to generate the fuzzy inference system (FIS) by means of diffuse subtractive clustering. As described in the introduction, the fuzziness lie on representing the mathematical model of the system by means of linguistic variables and a series of rules of the type IF...THEN. For our particular scientific question, the rules are of the type:

IF (Concentration of component X1) is (Low, Medium or High) and (Concentration of component X2) is (Low, Medium or High) and ... (Concentration of component X17) is (Low, Medium or High) THEN (Objective function) is (Low, Medium or High).

The neuro-fuzzy system was subsequently generated using the ANFIS (*Adaptive Neuro-Fuzzy Training of Sugeno-type FIS*) function, also implemented in Matlab[®]. Nodes and links of our neuro-fuzzy network are similar to those presented in Fig. 3. As we described in the introduction, the basic form of implementation of neuro-fuzzy networks is that the learning algorithms of neural networks are used to adjust the parameters of the fuzzy inference system (previous weights and consequent weights of the neuro-fuzzy system).

The function ANFIS uses a hybrid learning algorithm to tune the parameters of a Sugeno-type fuzzy inference system (FIS). The algorithm uses a combination of the least-squares and back-propagation gradient descent methods [48] to model a training data set. The main identification parameters used are described in Table 5.

The defuzzification method utilized by function anfis is the centroid, for which the output is the center of gravity of the distribution function for the output variable.

After data modeling, the simulated annealing global search algorithm was applied, using a function implemented in Matlab[®].

Validation of the Optimization

The culture medium composition generated from the optimization procedure was tested in a 10 L bioreactor in order to validate the results.

Table 3 Output data matrix for time (h)

Time (h)				
Chauvoei media	LTV29	Thioglycolate fluid	LTV05	LTV103
99.83	46.5	30	20	30

Table 4 Output data matrix for productivity ((log L+/mL)/h)

Productivity ((Log L+/mL)/h)				
Chauvoei media	LTV29	Thioglycolate fluid	LTV05	LTV103
0.02	0.03	0.05	0.05	0.05

Results

Constructing the Data Matrix

Firstly, tests employing DOE methodology were performed, which generated the LTV29 medium (see Figs. 6 and 7).

The next step involved investigation of new media for inoculation and cultivation of *C. Novyi*, aiming to maximize cell density, but this time modeling the system using neuro-fuzzy and optimizing with simulated annealing. Figure 8 shows the cellular growth results for cultures performed with four media generated by the computational optimization procedure.

The results of the experiments using the media obtained from the optimization were employed to generate (again using computational optimization) another three culture media (medium 5, medium 6 and medium 7), of which medium 5 (LTV05) showed the best performance (as shown in Fig. 9, also including new results obtained with media 3 and LTV29 for comparison purposes).

It should be noticed that the lag phase was considerably reduced as a result of the computational intelligence procedure employed. These results were used to re-model the system, with the output in this case being the production of alpha toxin, hence generating the LTV103 medium from the simulated annealing global search.

Optimizing Toxin Productivity

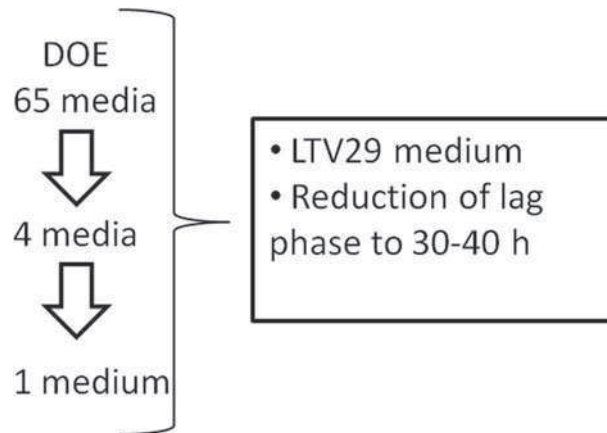
The final (complete) data were modeled using a neuro-fuzzy system with the data matrices shown in Tables 1 and 4 as the input and output, respectively. Optimization was performed using the simulated annealing algorithm to obtain the formulation variables that maximized toxin productivity, hence generating the culture medium formulations described in Table 6.

Three fermentations were performed for each of the proposed culture media, using a Bioflo 110 fermenter containing 10 L of optimized medium. The productivity results (Table 7) showed that the LTVC01 culture medium resulted in greatest productivities

Table 5 Main identification parameters

Number of membership functions for each input:	3
Membership functions type:	Gbellmf (generalized bell function)
Training method:	Hybrid
Defuzzification method:	Centroid

Fig. 6 Optimization route using DOE methodology to obtain the LTV29 medium



of alpha toxin by *C. novyi* and would therefore be most suitable for use on an industrial scale.

Discussion

Compared to our initial process conditions (Chauvoei media), which resulted in a long culture time (about 100 h) and low alpha toxin productivity (0.02 (log L+/mL)/h), the optimized system (LTVC01) provided a significant improvement, with a culture time of 12 h and an alpha toxin productivity of 0.06 (log L+/mL)/h. Considering the Thioglycolate fluid commercial medium, the time of 30 h was necessary to get 0.05 (log L+/mL)/h of alpha toxin productivity.

Besides, the computational tools presented in this work proved to be highly effective for the modeling and optimization of the biotechnological system tested, obtaining more effective results than those obtained by conventional techniques such

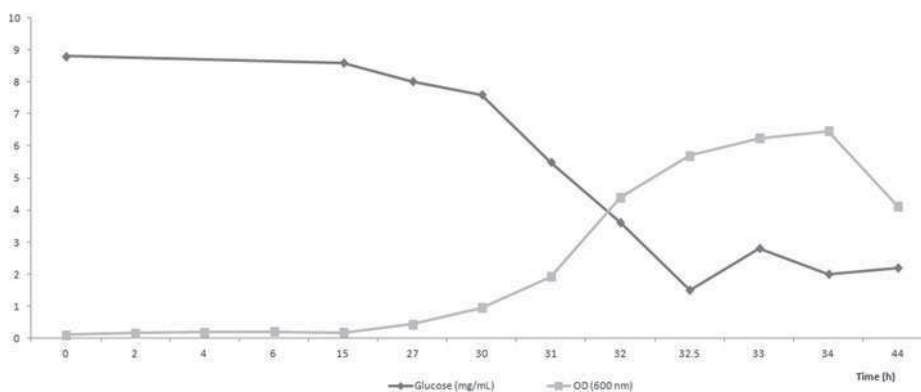


Fig. 7 Cellular growth (OD₆₀₀) and glucose consumption (mg/mL) curves for the culture performed with LTV29 medium

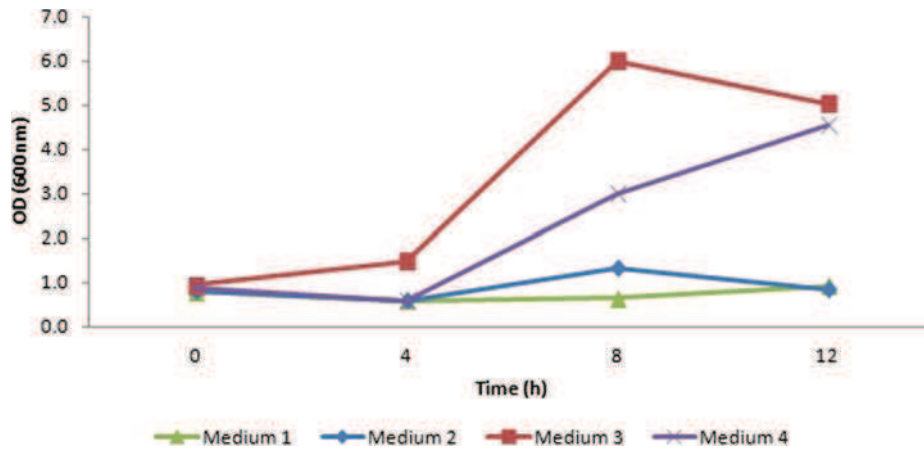


Fig. 8 Growth curves using the media obtained by computational optimization

as optimal design of experiments (LTV29, with time of 46.5 h necessary to get 0.03 (log L+/mL)/h). And the computational time does not matter for this problem. Neuro-fuzzy modeling and simulated annealing optimization need little time (minutes) to be performed.

It is also important to say here that other optimization techniques, such as genetic algorithm (GA), can also be used for optimizing fermentation media. Pal et al. [49], for example, presented a study on the production of a biosurfactant from *Rhodococcus erythropolis* where ANNs coupled with GA were used for media optimization. ANN-GA-optimized media gave about a 3.5-fold enhancement in biosurfactant yield. So, result obtained here (threefold

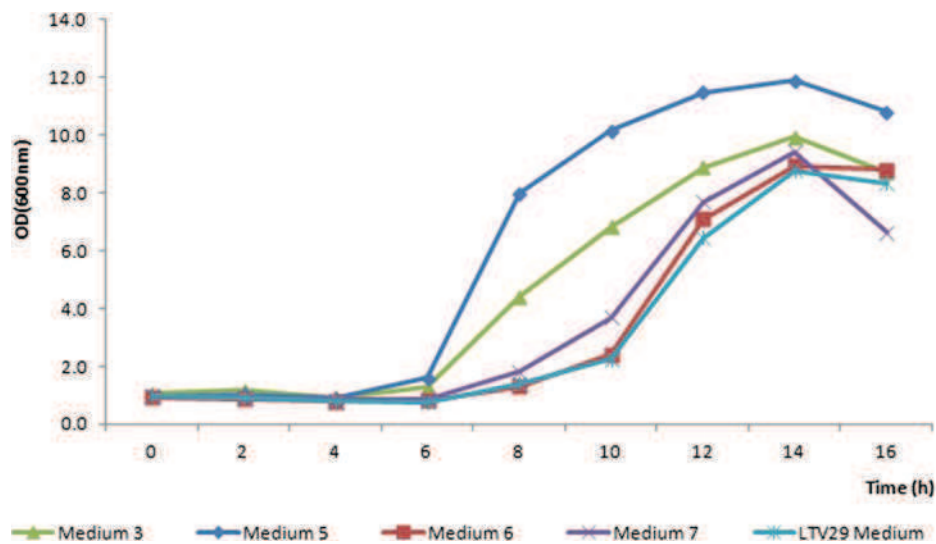


Fig. 9 Microbial growth curves obtained using the media generated by computational optimization

Table 6 Media generated using the proposed optimization method

Quantity (g/L)			
Component	LTVC01	LTVC02	LTVC03
Component 01	+1	+1	+1
Component 02	+1	+1	+1
Component 03	+1	+1	+1
Component 04	+1	+1	+1
Component 05	+1	+1	+1
Component 06	+1	+1	+1
Component 07	+1	+1	+1
Component 08*	0	+1	+1
Component 09*	-1	-1	-1
Component 10	-1	-1	-1
Component 11	0	-1	+1
Component 12	-1	-1	-1
Component 13	-1	-1	-1
Component 14	-1	-1	-1
Component 15	-1	-1	-1
Component 16	-1	-1	-1
Component 17	0	0	0

* Volume fraction (mL/L)

enhancement in alpha toxin productivity) is as good as this previous result using ANNs/GA. In this sense, we infer that our methodology can be used to solve other type of microorganism-production problems.

Conclusions

The computational tools presented in this work proved to be highly effective for the modeling and optimization of the biotechnological system tested. Compared to the initial process conditions, which resulted in a long culture time and low alpha toxin productivity, the optimized system provided a significant improvement. Often, adjustments costs are made through pragmatic cuts and without many criteria in major company expense rubrics. However, these cost optimizations can occur in a more rigorous and effective manner. Thus, in situations that require results with a high level of precision and minimum investment to achieve this goal, the use of computational

Table 7 Experimental productivities obtained using the optimized media

(Log L+/mL)/h		
LTVC01	LTVC02	LTVC03
0.06	0.03	0.02

tools can be very attractive. In the present case, this methodology succeeded in substantially reduce the number of experiments to be performed, achieving better productivity than originally obtained using design of experiments.

References

1. Amimoto, K., Noro, T., Oishi, E., & Shimizu, M. (2007). A novel toxin homologous to large clostridial cytotoxins found in culture supernatant of *Clostridium perfringens* type C. *Microbiology-Sgm*, *153*, 1198–1206.
2. Belyi, Y., & Aktories, K. (2010). Bacterial toxin and effector glycosyltransferases. *Biochimica Et Biophysica Acta-General Subjects*, *1800*, 134–143.
3. Skarin, H., Hafstrom, T., Westerberg, J., & Segerman, B. (2011). *Clostridium botulinum* group III: a group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genomics*, *12*, 185.
4. Ziegler, M. O. P., Jank, T., Aktories, K., & Schulz, G. E. (2008). Conformational changes and reaction of clostridial glycosylating toxins. *Journal of Molecular Biology*, *377*, 1346–1356.
5. Wei, M. Q., Mengesha, A., Good, D., & Anne, J. (2008). Bacterial targeted tumour therapy—dawn of a new era. *Cancer Letters*, *259*, 16–27.
6. Diaz, L. A., Cheong, I., Foss, C. A., Zhang, X. S., Peters, B. A., Agrawal, N., Bettgowda, C., Karim, B., Liu, G. S., Khan, K., Huang, X., Kohli, M., Dang, L. H., Hwang, P., Vogelstein, A., Garrett-Mayer, E., Kobrin, B., Pomper, M., Zhou, S. B., Kinzler, K. W., Vogelstein, B., & Huso, D. L. (2005). Pharmacologic and toxicologic evaluation of C-novyi-NT spores. *Toxicological Sciences*, *88*, 562–575.
7. Brett, M. M., Hood, J., Brazier, J. S., Duerden, B. I., & Hahne, S. J. M. (2005). Soft tissue infections caused by spore-forming bacteria in injecting drug users in the United Kingdom. *Epidemiology and Infection*, *133*, 575–582.
8. Batty, I., & Walker, P. D. (1964). Identification of clostridium novyi (clostridium oedematis) + clostridium tetani by use of fluorescent labelled antibodies. *Journal of Pathology and Bacteriology*, *88*, 327.
9. Eklund, M. W., Poysky, F. T., Meyers, J. A., & Pelroy, G. A. (1974). Interspecies conversion of clostridium-botulinum type-c to clostridium-novyi type-a by bacteriophage. *Science*, *186*, 456–458.
10. Jang, J. S. R., & Sun, C. T. (1995). Neuro-fuzzy modeling and control. *Proceedings of the Ieee*, *83*, 378–406.
11. Mamdani, E. H. (1977). Application of fuzzy logic to approximate reasoning using linguistic-synthesis. *Ieee Transactions on Computers*, *26*, 1182–1191.
12. Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, *12*, 993–1001.
13. Zadeh, L. A. (1983). The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy Sets and Systems*, *11*, 197–198.
14. Moreno, G., & Pascual, V. (2009). A hybrid programming scheme combining fuzzy-logic and functional-logic resources. *Fuzzy Sets and Systems*, *160*, 1402–1419.
15. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.
16. Sousa, R., Jr., & Almeida, P. I. F. (2001). Design of a fuzzy system for the control of a biochemical reactor in fed-batch culture. *Process Biochemistry*, *37*, 461–469.
17. Baldwin, J. F., & Guild, N. C. F. (1980). Feasible algorithms for approximate reasoning using fuzzy logic. *Fuzzy Sets and Systems*, *3*, 225–251.
18. Yager, R. R. (1991). Connectives and quantifiers in fuzzy sets. *Fuzzy Sets and Systems*, *40*, 39–75.
19. Lopez-Cruz, I. L., & Hernandez-Larragoiti, L. (2010). Neuro-fuzzy models for air temperature and humidity of arched and venlo type greenhouses in central mexico. *Agrociencia*, *44*, 791–805.
20. Lin, C. T., & Lee, C. S. G. (1991). Neural-network-based fuzzy-logic control and decision system. *Ieee Transactions on Computers*, *40*, 1320–1336.
21. Wang, L. X., & Mendel, J. M. (1992). Generating fuzzy rules by learning from examples. *Ieee Transactions on Systems Man and Cybernetics*, *22*, 1414–1427.
22. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: structure and dynamics. *Physics Reports-Review Section of Physics Letters*, *424*, 175–308.
23. Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17.

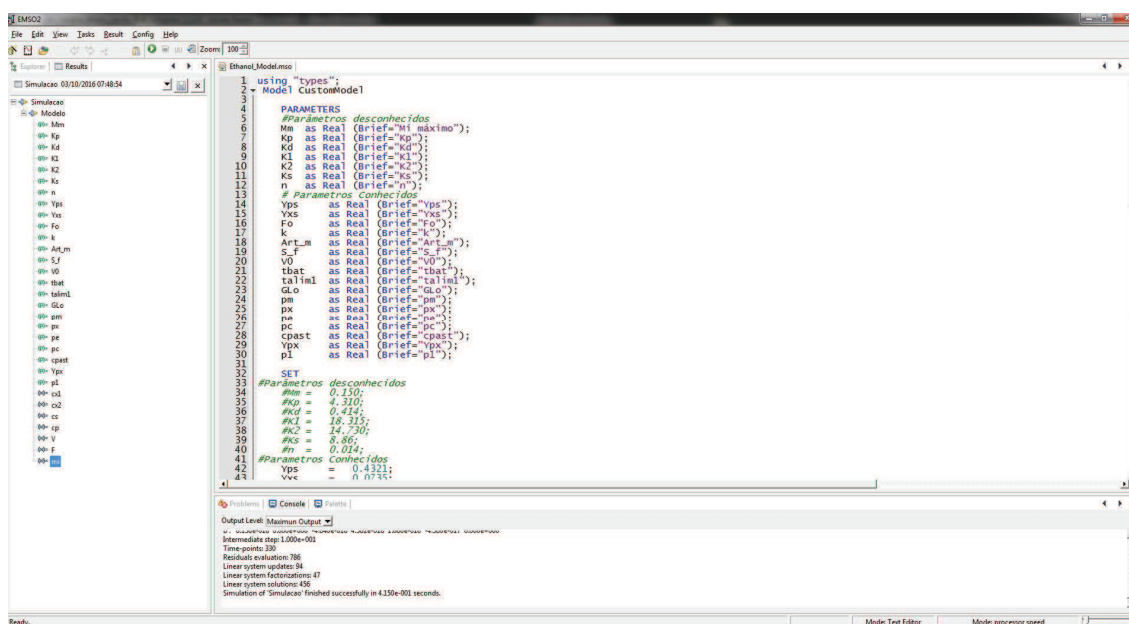
24. Adeli, H., & Jiang, X. M. (2003). Neuro-fuzzy logic model for freeway work zone capacity estimation. *Journal of Transportation Engineering-Asce*, 129, 484–493.
25. Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157–1162.
26. Jiang, X., & Adel, H. (2008). Neuro-genetic algorithm for non-linear active control of structures. *International Journal for Numerical Methods in Engineering*, 75, 770–786.
27. Lei, Y., He, Z., Zi, Y., & Hu, Q. (2007). Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAS. *Mechanical Systems and Signal Processing*, 21, 2280–2294.
28. Mitra, S., & Hayashi, Y. (2000). Neuro-fuzzy rule generation: survey in soft computing framework. *Ieee Transactions on Neural Networks*, 11, 748–768.
29. Kofman, G. I., Lyssenkov, S. A. and Lobachev, R. V. (2005) Automatic creation of neuro-fuzzy expert system from online analytical processing (OLAP) tools. Google Patents
30. Nava, P. A., & Ieee. (1998). Implementation of neuro-fuzzy systems through interval mathematics. *International Symposium on Intelligent Control - Proceedings*, 1, 365–369.
31. Baharani, M., Noori, H., Aliasgari, M., & Navabi, Z. (2014). High-level design space exploration of locally linear neuro-fuzzy models for embedded systems. *Fuzzy Sets and Systems*, 253, 44–63.
32. Jang, J. S. R. (1993). ANFIS—Adaptive-network-based fuzzy inference system. *Ieee Transactions on Systems Man and Cybernetics*, 23, 665–685.
33. Vuorimaa, P. (1994). Fuzzy self-organizing map. *Fuzzy Sets and Systems*, 66, 223–231.
34. Ichihashi, H., Shirai, T., Nagasaka, K., & Miyoshi, T. (1996). Neuro-fuzzy ID3: a method of inducing fuzzy decision trees with linear programming for maximizing entropy and an algebraic method for incremental learning. *Fuzzy Sets and Systems*, 81, 157–167.
35. Rao, D. H., & Kamat, H. V. (1996). Neuro-fuzzy system for robotics applications. *Journal of the Institution of Electronics and Telecommunication Engineers*, 42, 325–333.
36. Nauck, D., & Kruse, R. (1999). Neuro-fuzzy systems for function approximation. *Fuzzy Sets and Systems*, 101, 261–271.
37. Shi, Z. P., & Shimizu, K. (1992). Neurofuzzy control of bioreactor systems with pattern-recognition. *Journal of Fermentation and Bioengineering*, 74, 39–45.
38. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087–1092.
39. Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220, 671–680.
40. Cerny, V. (1985). Thermodynamical approach to the traveling salesman problem—an efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45, 41–51.
41. Musharavati, F., & Hamouda, A. S. M. (2012). Enhanced simulated-annealing-based algorithms and their applications to process planning in reconfigurable manufacturing systems. *Advances in Engineering Software*, 45, 80–90.
42. Pinto, G. A., & Giordano, R. C. (2009). Bioprocess systems engineering applied to the production of protein hydrolysates in a multipurpose plant. *Computer Aided Chemical Engineering*, 27, 1887–1892.
43. Zheng, S., Shu, W., Gao, L. and Ieee. (2006) Task scheduling using parallel genetic simulated annealing algorithm. *Service Operations and Logistics, and Informatics*, 46–50.
44. Yele, V. U., & Desai, K. (2015). A new thermostable and organic solvent-tolerant lipase from *Staphylococcus warneri*; optimization of media and production conditions using statistical methods. *Applied Biochemistry and Biotechnology*, 175, 855–869.
45. Zhang, M., Xiao, G., Thring, R. W., Chen, W., Zhou, H., & Yang, H. (2015). Production and characterization of melanin by submerged culture of culinary and medicinal fungi *Auricularia auricula*. *Applied Biochemistry and Biotechnology*, 176, 253–266.
46. Lima, C. G. R. D., Lobato, Z. I. P., Pires, P. S., Silva, R. O. S., Salvarani, F. M., Assis, R. A., & Lobato, F. C. F. (2011). Padronização de teste de potência in vitro para vacinas que contenham toxoide alfa de *Clostridium novyi* tipo B. *Arquivos do Instituto Biológico*, 78, 507–512.
47. Gonçalves, L. C. G., Furlan, F. F., Soares, R. P., Secchi, A. R., Giordano, R. C., Costa, C. B. B. (2012) Implementation of Pareto multiobjective Particle Swarm Optimization algorithm in EMSO. In: EngOpt 2012 - 3rd International Conference on Engineering Optimization, 2012, Rio de Janeiro.
48. Nelles, O. (2001). *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Berlin: Springer-Verlag.
49. Pal, M. P., Vaidya, B. K., & Desai, K. M. (2009). Media optimization for biosurfactant production by *Rhodococcus erythropolis* MTCC 2794: artificial intelligence versus a statistical approach. *Journal of Industrial Microbiology and Biotechnology*, 36, 747–756.

6. APÊNDICE 1: INTEGRAÇÃO COM O SOFTWARE EMSO

6.1.INTRODUÇÃO

O software EMSO (Environment for Modeling, Simulation, and Optimization, interface demonstrada na Figura 57), foi criado a partir de um projeto chamado Ambiente Livre para Simulação, Otimização e Controle de Processos (ALSOC), coordenado pela Universidade Federal do Rio Grande do Sul. O projeto ALSOC é um trabalho para aproximação da academia com a indústria através da padronização e distribuição, gratuitamente, de especificações e ferramentas de software entre universidades e empresas consorciadas (Soares e Secchi, 2003).

Figura 36: Interface do simulador EMSO



```
1 using "types";
2 Model CustomModel
3
4 #Parâmetros desconhecidos
5
6 Mm as Real (Brief="M máximo");
7 Kd as Real (Brief="kg");
8 Kd as Real (Brief="kd");
9 K1 as Real (Brief="k1");
10 K2 as Real (Brief="k2");
11 Ks as Real (Brief="ks");
12 n as Real (Brief="n");
13
14 # Parâmetros conhecidos
15 Yps as Real (Brief="yps");
16 Yxs as Real (Brief="yxs");
17 Fo as Real (Brief="fo");
18 k as Real (Brief="k");
19 Art_m as Real (Brief="art_m");
20 S_f as Real (Brief="s_f");
21 V0 as Real (Brief="v0");
22 tbat as Real (Brief="tbat");
23 taliml as Real (Brief="taliml");
24 Glo as Real (Brief="glo");
25 pm as Real (Brief="pm");
26 px as Real (Brief="px");
27 pa as Real (Brief="pa");
28 pc as Real (Brief="pc");
29 cpast as Real (Brief="cpast");
30 Ypx as Real (Brief="ypx");
31 pi as Real (Brief="pi");
32
33 SET
34 #Parâmetros conhecidos
35 #M = 0.150;
36 #kp = 4.310;
37 #kd = 0.414;
38 #k1 = 18.215;
39 #k2 = 14.730;
40 #ks = 8.86;
41 #n = 0.014;
42 #yps = 0.4321;
43 #vcc = 0.0751;
44
```

Output Level: Maximum Output
Intermediate step: 1000=00
Time points: 330
Results evaluation: 786
Linear system updates: 81
Linear system factorizations: 87
Linear system solutions: 456
Simulation of 'Simulacao' finished successfully in 4.150e-001 seconds.

De acordo com Soares e Secchi (2003), a programação no EMSO é composta por três entidades principais:

- *Models;*
- *Devices;*
- *Flowsheets.*

Segundo Soares e Secchi (2003) na linguagem do EMSO, um *Model* consiste na abstração matemática de um equipamento real e/ou de uma parte de um processo. Cada modelo pode conter parâmetros, variáveis, equações, condições iniciais e até outros submodelos. *Models* podem ser baseados em outros modelos já existentes e funcionalidades extras podem ser adicionadas. Assim, os conceitos de programação modular e herança (compartilhamento de atributos e métodos pelas classes) são suportados na criação de modelos.

Um *Device* é uma instância de um modelo, ou seja, a realização específica de um tipo de modelo, e representa algum dispositivo real do processo em análise. Assim, um único modelo (*Model*) pode ser utilizado para representar vários *Devices* diferentes, que possuem a mesma estrutura, porém parâmetros e especificações diferentes. *Devices* podem ser conectados entre si para formar um *Flowsheet*, que é uma abstração do processo em análise. (Soares e Secchi, 2003).

6.2.OBJETIVOS

O objetivo deste trabalho foi simular o modelo proposto por Daré (2008) utilizando o simulador de processos orientado a equações EMSO comunicando com o MATLAB através de função específica para realização de busca global dos parâmetros do modelo.

6.3.MATERIAIS E MÉTODOS

6.3.1. Simulador

O simulador EMSO possui sua estrutura orientada a equações. Esse simulador tem uma linguagem à modelagem orientada a objeto o que permite a inserção de novos modelos para a sua biblioteca interna. (Cunha *et al.*, 2009).

6.3.2. Interface Emso MATLAB

Transferiram-se ao MATLAB todos os parâmetros (Tabela 12) e variáveis e dados experimentais do modelo de fermentação implementado no EMSO através da função do EMSO implementada em MATLAB emso2ml.m. A função emso2ml.m importa todos os parâmetros e variáveis do modelo implementado no EMSO, possibilitando o ajuste dos parâmetros no MATLAB.

Tabela 10: Parâmetros ajustados no MATLAB

Parâmetros	Significado
μ_{\max}	Velocidade específica máxima de crescimento celular
Kp	Constante de formação de produto
Kd	Constante de morte celular
K1	Constante do modelo transformação das células do tipo 1 em células do tipo 2
K2	Segunda constante do modelo de transformação das células do tipo 1 em células do tipo 2.
Ks	Constante de saturação
n	Parâmetro do modelo de Levenspiel (1976)

A função utilizada para os ajustes dos parâmetros calcula o erro com os dados experimentais, assim como descrito no capítulo 3 desta Tese de doutorado. Esse valor é otimizado por minimização, realizando, assim, a busca global dos parâmetros pelo algoritmo Simulated Annealing. Os parâmetros ajustados no MATLAB foram retornados ao EMSO para realização da simulação.

6.4.RESULTADOS

Os resultados das simulações são demonstrados nas figuras 2 a 4.

Figura 37: Resultado da simulação no EMSO para a produção de células CX1 e CX2, com parâmetros do modelo ajustados no MATLAB.

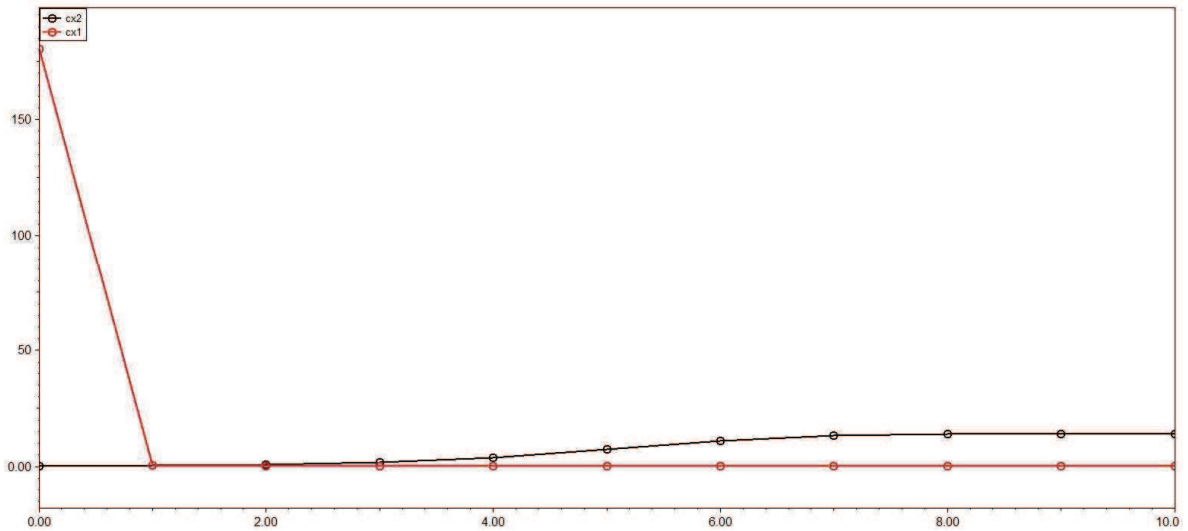


Figura 38: Resultado da simulação no EMSO para a produção de etanol, com parâmetros do modelo ajustados no MATLAB.

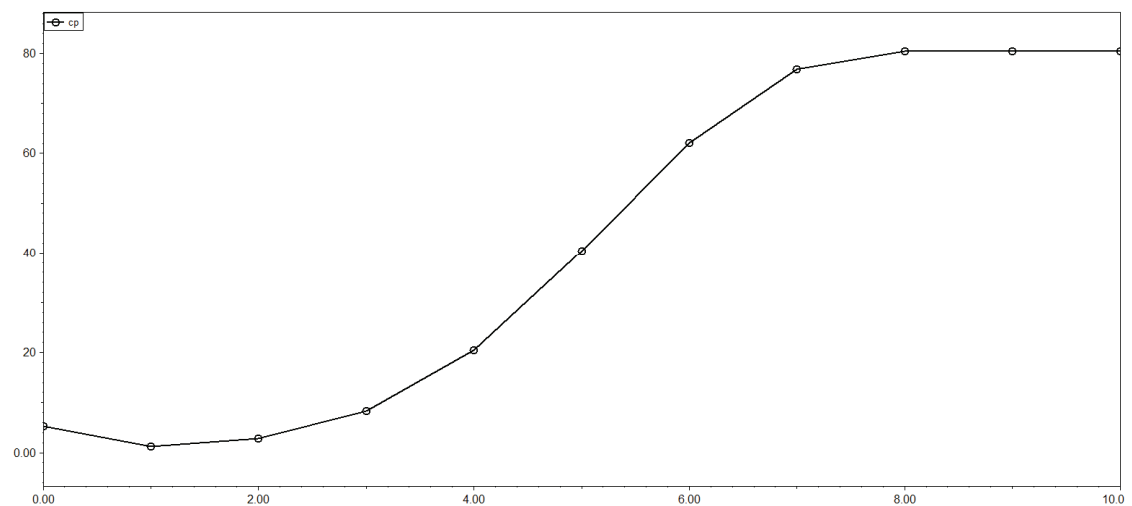
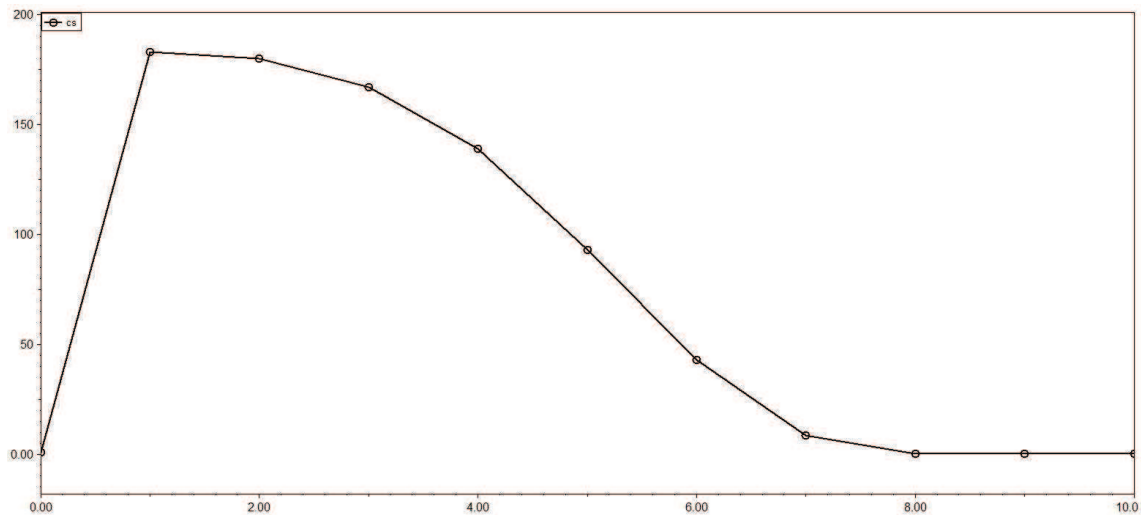


Figura 39: Resultado da simulação no EMSO para o consumo de substrato, com parâmetros do modelo ajustados no MATLAB.



7. APÊNDICE 2: Modeling the Kinetics of Complex Systems: Enzymatic Hydrolysis of Lignocellulosic Substrates

Carlos Alberto Galeano Suarez

Inti Doraci Cavalcanti-Montaño

Renan Gustavo da Costa Marques

Felipe Fernando Furlan

Pedro Luiz da Mota e Aquino

Roberto de Campos Giordano

Ruy de Sousa Jr

Modeling the Kinetics of Complex Systems: Enzymatic Hydrolysis of Lignocellulosic Substrates

Carlos Alberto Galeano Suarez ·
Inti Doraci Cavalcanti-Montaña ·
Renan Gustavo da Costa Marques ·
Felipe Fernando Furlan ·
Pedro Luiz da Mota e Aquino ·
Roberto de Campos Giordano ·
Ruy de Sousa Jr.

Received: 7 February 2014 / Accepted: 7 April 2014
© Springer Science+Business Media New York 2014

Abstract Lignocellulosic biomass is mainly composed of cellulose, hemicellulose, and lignin. Fuzzy logic, in turn, is a branch of many-valued logic based on the paradigm of inference under vagueness. This paper presents a methodology, based on computational intelligence, for modeling the kinetics of a complex reactional system. The design of a fuzzy interpolator to model cellulose hydrolysis is reported, within the perspective of applying kinetic models in bioreactor engineering. Experimental data for various types of lignocellulosic materials were used to develop the interpolator. New experimental data from the enzymatic hydrolysis of a synthetic substrate, on the other hand, were used to validate the methodology. The accuracy of the results indicates that this is a promising approach to extend the application of models fitted for specific situations to different cases, thus enhancing their generality.

Keywords Cellulose · Clustering · Enzymatic hydrolysis · Enzyme kinetics · Fuzzy interpolator · Lignin

Introduction

Recently, lignocellulosic materials such as sugarcane bagasse and other agricultural wastes have been used as raw material for various products. The worldwide search for renewable energy sources makes the use of lignocellulosic biomass waste a promising alternative to produce bioethanol fuel on a large scale and at a competitive cost.

One of the most popular alternatives for production of bioethanol in industrial scale is the biochemical route, based on the fermentation of the sugar monomers of cellulose and

C. A. G. Suarez · I. D. Cavalcanti-Montaña · R. G. da Costa Marques · F. F. Furlan ·
P. L. da Mota e Aquino · R. de Campos Giordano · R. de Sousa Jr. (✉)
Department of Chemical Engineering, Federal University of São Carlos, Rod. Washington Luís,
Km 235, CP 676, 13565-905 São Carlos, SP, Brazil
e-mail: ruy@ufscar.br

hemicellulose. Two techniques may be used for obtaining fermentable sugars from lignocellulosic materials—acid and enzymatic hydrolyses.

Acid chemical hydrolysis has well-known drawbacks—generation of toxic products and corrosiveness of the medium, requiring the use of more costly reactors. Moreover, it is necessary to recover the acid for making the process economically viable. In general, when concentrated acids are used, the monosaccharides released by hemicellulose are degraded due to the fact that the hemicellulose fraction is hydrolyzed more rapidly than the fraction of cellulose. The sugars present in hemicellulose are exposed for a longer time in the reaction medium, which results in their degradation (for instance, to furfural). On the other hand, hydrolysis with dilute acid requires the use of high temperatures. For this reason, a relatively high amount of soluble sugars and lignin is degraded, leading to inhibition during the fermentation process [1].

In the enzymatic hydrolysis, cellulose is broken down by a pool of cellulolytic enzymes, and there is no formation of by-products due to the specificity of enzymes (resulting in a high yield of fermentable sugars). However, it is necessary to pretreat the biomass, to improve its digestibility, and to use high concentrations of enzymes for achieving high conversions of the cellulose, which makes the process more costly [2].

Enzymatic hydrolysis of cellulose yields glucose, which can be further fermented to provide ethanol. It can be operated under mild conditions, avoiding the formation of byproducts which can inhibit the fermentation, such as hydroxymethylfurfural [3], consequently leading to a reduction of the cost of separation of the products [4].

The presence of lignin and hemicellulose hinders the access of cellulases to their substrate; this can reduce significantly the efficiency of the hydrolysis. Therefore, there is a need for pretreatments that weaken the interaction between the main components of the lignocellulosic material. Pretreatments aim at partially separating lignin and hemicellulose from the cellulose matrix and disrupting the crystalline structure of cellulose. The result is increased accessibility to cellulose particles, as a consequence of the solubilization and/or partial degradation of hemicellulose and lignin [5]. To be considered effective, a pretreatment must meet the following requirements: avoiding the need of reducing the particle size of the biomass, limiting the formation of degradation products that inhibit the microbial fermentation, and minimizing the energy demand. Various types of processes have been used for the pretreatment of lignocellulosic materials [1]. The processes described in the literature include physical processes, such as steam explosion and thermal hydrolysis; chemical processes, such as weak acid and alkaline hydrolysis; biological processes using fungi to solubilize the lignin; finally, a combination of processes, as steam explosion catalyzed by adding H_2SO_4 or CO_2 [6, 7]. However, the most common methods are based on the use of alkali, acid hydrolysis, steam explosion, and hot water, all with the purpose of breaking the associative structure of lignocellulose [8–10].

For the enzymatic hydrolysis to be efficient, it is necessary to optimize the reaction conditions since the yield of hydrolysis is governed by several factors such as the type of pretreatment of the substrate, inhibition of enzymatic activity by end products of the biodegradation, ineffective adsorption of enzymes (on lignin, for instance) thermal stability of enzymes, duration of the hydrolysis, pH, concentration of substrate in the medium, and agitation speed [11].

In enzymatic hydrolysis processes, there are some important phenomena that finally control the overall rates of reaction. For hydrolyses in batch-stirred reactors, the overall reaction rate is determined by the rates of three events occurring in sequence: (a) the rate

of (external) mass transfer of the enzyme, (b) the rate of diffusion/adsorption of the enzyme on the substrate surface, and (c) the rate of the cellulase catalytic action. After a first initial stage, the overall reaction rate is subject to (internal) diffusion and adsorption of the enzyme on the solid substrate. Most authors do not consider the external mass transfer resistances. The adsorption of enzymes and the formation of enzyme-substrate complexes are considered crucial steps in this process. Adsorption of cellulase on the insoluble cellulose has been described as reversible, irreversible, and semireversible (thus, literature is inconclusive on this subject [12]).

Modeling the enzymatic hydrolysis of lignocellulosic materials is probably one of the most challenging subjects in bioreactor engineering science. The challenges of this problem can be grouped into three classes—the complexity of the substrate, the action of the enzymes, and the enzyme-substrate interactions. There are different approaches for this modeling. Zhang and Lynd [13] grouped kinetic models of the enzymatic hydrolysis of biomass according to the level of detail of their description of the substrate and/or of the activities of the different enzymes that are acting. According to these authors, models can be classified as nonmechanistic, semimechanistic, functionally based, and structurally based.

Engineering practice for design and optimization of the hydrolysis bioreactor has been based mainly on different empirical and semi-empirical approaches (nonmechanistic and semimechanistic models), including simple kinetic equations [14]. The most common semimechanistic approach is based on pseudohomogeneous Michaelis-Menten models, i.e., despite being actually a solid, the substrate is treated as a soluble reactant, characterized by its concentration. Bezerra and Dias [15] investigated the kinetics of one exoglucanase in the presence of cellobiose using different enzyme/substrate (Avicel) ratios. It was found that the cellulose hydrolysis rate followed a pseudohomogeneous Michaelis-Menten model that takes into account competitive inhibition by cellobiose. A second class of semimechanistic rate equations pictures a system closer to the real one, considering that the substrate is in solid form and that the soluble enzyme has to adsorb to (and desorb from) it. Carrillo et al. [16] studied the kinetics of the hydrolysis of pretreated (with sodium hydroxide) wheat straw using different concentrations of a commercial cellulase (Novozymes A/S). Initial rates of an equation derived from a Michaelis-Menten mechanism were measured, assuming solid substrate and soluble enzyme. The initial rate of hydrolysis can be expressed as a function of the initial enzyme concentration in a modified Michaelis-Menten model.

Finally, in Chrastil's semimechanistic model [17, 18], all time constants of the rate of product formation are ranked, taking into account that in a heterogeneous system the time curves depend strongly on rate-limiting phenomena such as enzyme (internal) diffusion and adsorption.

Many kinetic rate equations were developed for model substrates such as Avicel, CMC, or Solca Floc. However, real substrates do not show the same characteristics as pure cellulose. These substrates may not represent well the hydrolysis of biomass, where cellulose is combined with other components (hemicellulose and lignin). Carvalho et al. [19, 20] investigated the kinetics of the enzymatic hydrolysis of three cellulosic substrates (added to the bioreactor in low and high loads)—filter paper (FP), used as a delignified substrate model, steam-exploded sugarcane bagasse (SB), and acid treated SB. The latter two were also treated with 4 % NaOH. Kinetic models were fit for these different lignocellulosic substrates, and different functional forms of the rate equations (Michaelis-Menten-like and Chrastil equations) were tested for different types of substrates.

However, due to the complexity of the reactional system, it becomes difficult to use a single semimechanistic kinetic model for the different hydrolysis conditions that appear while the reaction proceeds. An interesting idea would be to combine the attributes of the different “local” models adjusted by Carvalho et al. [19, 20] in a “global” system, calculating the rate of hydrolysis (incorporating models fitted for different hydrolysis conditions). A first step towards this development is the classification of the experimental data from Carvalho [19, 20] into an adequate number of clusters. This classification is important for determining kinetically similar regions.

Clustering of numerical data forms the basis of many classification algorithms for preprocessing experimental data. The purpose here is to identify natural groupings in a set of data in order to improve the representation of the system behavior. A popular technique for grouping data is associated with fuzzy logic, and is known as fuzzy C-means (FCM), as described in the “Theory” section of this paper.

The aim of this paper is to develop a fuzzy interpolator for kinetic modeling based on simple models to obtain a “global” model (combining these simple models), covering a broad range of hydrolysis conditions. The present paper intends to assess how the interpolation of simple semimechanistic models adheres to experimental data of enzymatic hydrolysis of cellulosic substrates. Three kinetic models were considered—pseudohomogeneous Michaelis-Menten, modified Michaelis-Menten, and Chrastil models, all previously fitted by Carvalho et al. [19, 20].

Materials and Methods

Fuzzy Interpolator Design

Experimental data from Carvalho [19, 20], of hydrolysis with different lignocellulosic substrates (filter paper; exploded bagasse treated with 4 % NaOH, low and high load; bagasse treated with 1 % H₂SO₄ and 4 % NaOH, low and high load), were grouped using the FCM clustering technique. The data were clustered into different classes, so that each different class matches a different kinetic model. For clustering, the variables used were percentage of cellulose load in the reactor and lignin content (percentage) with regard to the substrate composition.

Initially, Carvalho [19, 20] carried out filter paper studies in 250-mL Erlenmeyer flasks, in a total reaction volume of 20 mL, with 0.5 FPU mL⁻¹_{solution}, at pH 4.8, in 50 mM sodium citrate buffer. The experiments were run using a refrigerated incubator (Marconi MA-832) with agitation set at 250 rpm at 50 °C. The initial substrate concentration of 3.85 % ($w_{\text{cellulose}}/w_{\text{total}}$) was considered. For the exploded sugarcane bagasse treated with 4 % NaOH, hydrolysis experiments were also performed in 250-mL Erlenmeyer flasks, with a liquid volume of 30 mL, by adding 0.85 FPU mL⁻¹_{solution}, at pH 4.8, in 50 mM sodium citrate buffer. The initial percentage of 2.9 % ($w_{\text{cellulose}}/w_{\text{total}}$), corresponding to a potential glucose concentration of 33.33 g_{Glucose} L⁻¹_{solution}, was taken into account. Also, an assay was performed by setting the initial substrate at 6.54 % ($w_{\text{cellulose}}/w_{\text{total}}$), corresponding to a potential glucose concentration of 77.77 g_{Glucose} L⁻¹_{solution}. For the bagasse treated with 1 % H₂SO₄ and delignified with NaOH 4 %, similar procedures were carried out. For the determination of glucose concentration from samples taken during the hydrolysis assays, an enzymatic kit was used by Carvalho [19, 20]. To calculate the cellulose load in the reactor and the lignin content with regard to the substrate composition along time, taking into account the initial chemical characterization of the substrates (for cellulose, lignin, hemicellulose, and ashes) and the

consumption of cellulose along time in each experiment of Carvalho [19, 20], the following calculations were performed:

$$\% \text{ of cellulose in the reactor} = \frac{100 \times \text{remaining cellulose (g)}}{w_{\text{total}} \text{ (g)}} \quad (1)$$

$$\text{lignin content} = \frac{100 \times \text{lignin (g)}}{\text{lignin (g)} + \text{remaining cellulose (g)} + \text{hemicellulose (g)} + \text{ashes (g)}} \quad (2)$$

Based on the clustering, a fuzzy interpolator was generated for the interpolation of different simple semimechanistic models (those presented in Table 1). The study was conducted using the software MATLAB.

The command line function *fcm* of MATLAB starts with an initial estimate for the cluster centers, which will be the average position of each cluster. The initial estimate for these cluster centers is probably incorrect. Furthermore, the function *fcm* calculates, for all data points, 1 degree of membership with regard to each cluster. By iterative updating the position of these cluster centers and the degree of membership of each data point, the function *fcm* iteratively moves the cluster centers within a dataset. These iterations are based on minimizing an objective function that represents the distance from any given data point to a cluster center, weighted by the degree of membership of that point.

The information returned by FCM was used to build a fuzzy inference system by creating membership functions to represent the fuzzy qualities of each cluster. The classification was used to generate an inference algorithm to represent the system behavior (rate of hydrolysis according to the conditions of hydrolysis) using a minimum number of *fuzzy* rules. The predictions of the “global” model proposed here were compared with experimental data of enzymatic hydrolysis of a synthetic (cellulosic) substrate.

Table 1 Semimechanistic kinetic models and kinetic parameters used to develop the fuzzy interpolator for the enzymatic hydrolysis of lignocellulosic materials [19, 20]

Model	Equation	Parameters
Model 1 Pseudohomogeneous Michaelis-Menten (MM) model, with inhibition—studies originated with filter paper	$v = \frac{V_{\max}S}{K_m(1+(P/K_I))+S}$	V_{\max} : 0.09 g L ⁻¹ min ⁻¹ K_m : 45.11 g L ⁻¹ K_I : 4.13 g L ⁻¹
Model 2 Pseudo-homogeneous MM with competitive inhibition	$v = \frac{V_{\max}S}{K_m(1+(P/K_I))+S}$	V_{\max} : 0.105 g L ⁻¹ min ⁻¹ K_m : 8.78 g L ⁻¹ K_I : 5.91 g L ⁻¹
Model 3 Modified MM Model with inhibition	$v = \frac{K_E S}{K_m(1+(P/K_I))+E}$	K : 0.0033 min ⁻¹ K_m : 22.06 g L ⁻¹ K_I : 20.18 g L ⁻¹
Model 4 Chrastil	$P = P_{\infty}[1 - \exp(-k'E_0t)]^n$	P_{∞} : 33.33 g L ⁻¹ k' : 2.49E-7 L g ⁻¹ min ⁻¹ E_0 : 7.83 g L ⁻¹ n : 0.33
Model 5 Chrastil b		P_{∞} : 77.77 g L ⁻¹ k' : 4.96E-8 L g ⁻¹ min ⁻¹ E_0 : 13.05 g L ⁻¹ n : 0.30

Validation Test

Enzyme

The commercial complex of cellulases Accellerase® 1500, from *Trichoderma reesei*, donated by Genencor® (Palo Alto, CA, USA), was used. According to the manufacturer, this complex contains multiple enzymatic activities, different exoglucanases and endoglucanases, beta-glucosidase, and hemicellulase. This complex operates at 50 °C and pH 4.8. In long-term experiments, 50 °C is a good choice to minimize temperature inactivation. To determine the complex overall activity, the method of total reducing sugars (TRS) was applied, making use of 3,5-dinitrosalicylic acid (DNS) and Whatman filter paper no. 1 [21–23].

Enzymatic Hydrolysis of a Synthetic Substrate

Two cellulosic materials were used to compose a synthetic substrate—qualitative filter paper (Satelit, Brazil) and *in natura* sugarcane bagasse (*Saccharum officinarum*), donated by the Center for Sugarcane Technology (CTC), Piracicaba, SP, Brazil.

The synthetic substrate was prepared from the *in natura* sugarcane bagasse and the filter paper with a ratio of 1:1 (*w/w*). This was done in order to obtain a substrate with a lignin percentage between 0 and 20 % (12 %) in order to validate the model interpolator (since this is a new condition, different from Carvalho's [19, 20]).

Experimental Conditions

The hydrolysis experiment was carried out in Erlenmeyer flasks with 50 mL of 50 mM sodium citrate buffer, at pH 4.8 and 20 FPU g⁻¹ cellulose of Accellerase® 1500, corresponding to 1.2 FPU mL⁻¹ solution. The experiment began with the addition of 4.32 g of synthetic substrate (50 % of paper filter–50 % of sugarcane bagasse) in order to obtain an initial concentration of potential glucose of 66.66 g L⁻¹. The synthetic substrate was hydrolyzed at 50 °C and 250 rpm for a total time of 72 h. Experiments were run in triplicate.

Sugar Quantification

Glucose and cellobiose were determined by high performance liquid chromatography (HPLC) using a Shimadzu SCL-10A chromatograph with a refractive index detector Shimadzu RID-10A, column Aminex HPX-87H (300×7.8 mm, Bio-Rad) with 0.005 mol L⁻¹ H₂SO₄ as mobile phase at a flow rate of 0.6 mL min⁻¹ and oven temperature of 45 °C.

Theory

FCM is a technique for grouping data where each data point belongs to all clusters, but in a different intensity—which is specified by the degree of membership of this point to each cluster. Such a technique was originally introduced by Bezdek [24] as an improvement in prior methods [25]. Here, the number of cluster is predefined. There are other techniques that can be used when the number of clusters is not previously known [26].

The FCM algorithm performs the steps listed in the following [25, 27, 28]:

1. Centers of the clusters are chosen randomly, and for each point in the dataset the distance to each center is calculated by Eq. (3).

$$D_{ij}^2 = (u(i) - c_j)^T \sum (u(i) - c_j) \quad (3)$$

where $u(i)$ is a data point, j is the number of clusters, and c is the chosen center.

2. After that, the degree of membership of each data point (for each cluster) is calculated, using Eq. (4).

$$\mu_{ij} = \frac{1}{\sum_{l=1}^C \left(D_{ij}^2 / D_{il}^2 \right)^{\frac{1}{\nu-1}}} \quad (4)$$

where μ is the degree of membership, ν is a weighting exponent, between 1 and ∞ (when there is no prior knowledge, 2 is a good estimate [25, 27, 28]), and C is the number of centers.

3. After this step, the centroid is computed for each cluster, which can be obtained from Eq. (5).

$$c_j = \frac{\sum_{i=1}^N \mu_{ij}^\nu u_i}{\sum_{i=1}^N \mu_{ij}^\nu} \quad (5)$$

N represents the number of existing data. The centers are modified until an objective function, represented by I in Eq. (6) is minimized.

$$I = \sum_{j=1}^C \sum_{i=1}^N \mu_{ij}^\nu (u(i) - c_j)^T \sum (u(i) - c_j) \quad (6)$$

Results and Discussion

Clustering of numerical data forms the basis of many classification and system modeling algorithms. Initially, natural groupings (clusters) from experimental data of Carvalho [19, 20] were identified. Figure 1 shows the experimental dataset, percentage of cellulose in the reactor *versus* percentage of lignin with regard to the substrate composition, during the hydrolysis processes, i.e., along time, to determine the centers of the clusters. It is worth noticing that although we refer to the variable time, the time itself is not an input variable for grouping the dataset (only percentage of cellulose in the reactor and percentage of lignin). The data are for the five different lignocellulosic substrates previously mentioned (filter paper (1); exploded bagasse treated with 4 % NaOH, low (2), and high load (3); bagasse treated with 1 % H₂SO₄ and 4 % NaOH, low (4) and high load (5)) (Fig. 1).

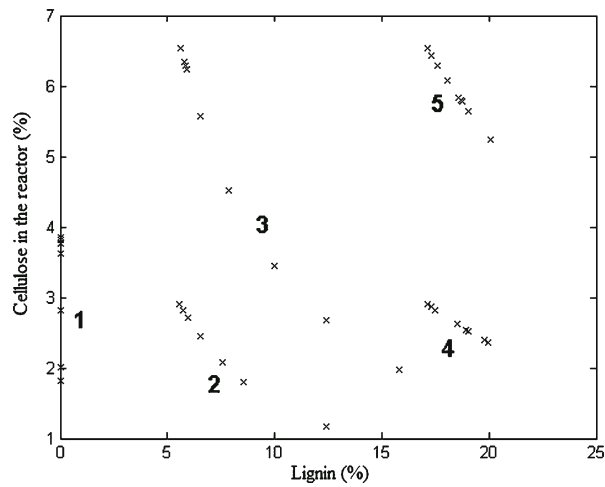


Fig. 1 Percentage of cellulose in the reactor versus percentage of lignin with regard to the substrate composition, during the hydrolysis processes, for the experimental data of Carvalho [19, 20]. The data are for five different lignocellulosic substrates (filter paper (1); exploded bagasse treated with 4 % NaOH, low (2) and high load (3); bagasse treated with 1 % H₂SO₄ and 4 % NaOH, low (4) and high load (5))

We use MATLAB to show the experimental data graphically, where the *Y*-axis represents the percentage of cellulose load in the reactor (%) and the *X*-axis the percentage of lignin with regard to the substrate composition (%). In MATLAB, invoking the command-line function *fcm*, it was possible to find five clusters in this dataset (until the objective function, *objFcn*, was no longer decreased). Function *fcm* returned a vector containing the coordinates of the five cluster centers, other vector

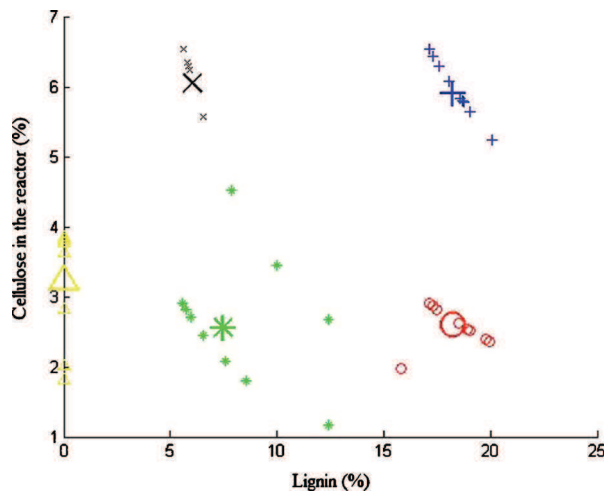


Fig. 2 First output for the clustering process. *Empty triangle* Center of the cluster for filter paper; *asterisk* center of the cluster for exploded bagasse treated with 4 % NaOH (low percentage of cellulose); *multiplication symbol* center of the cluster for exploded bagasse treated with 4 % NaOH (high percentage of cellulose); *empty circle* center of the cluster for bagasse treated with 1 % H₂SO₄ and 4 % NaOH (low percentage of cellulose); *positive symbol* center of the cluster for bagasse treated with 1 % H₂SO₄ and 4 % NaOH (high percentage of cellulose)

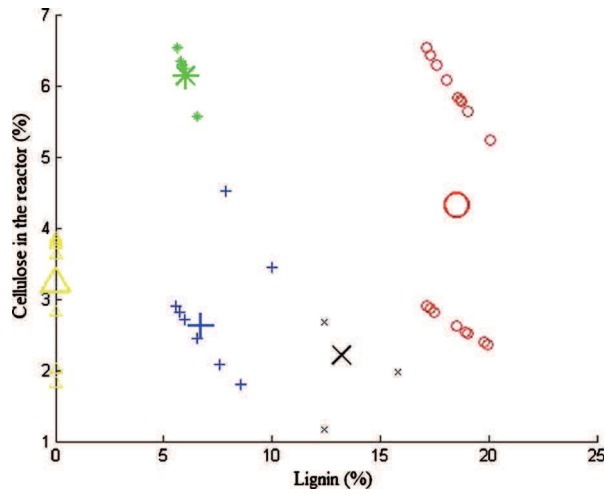


Fig. 3 Second output for the clustering process

containing the membership degrees for each one of the data points, and a history of the objective function along the iterations.

After running *fcm* function, two results were obtained due to the initial random choice of the cluster centers. The two possible outputs are shown in Figs. 2 and 3.

The first output was selected (Fig. 2) since each different cluster center is associated with a different hydrolysis substrate (filter paper (Δ); exploded bagasse treated with 4 % NaOH, low (*) and high (X) load; bagasse treated with 1 % H₂SO₄ and 4 % NaOH, low (O) and high (+) load), providing a more intuitive response.

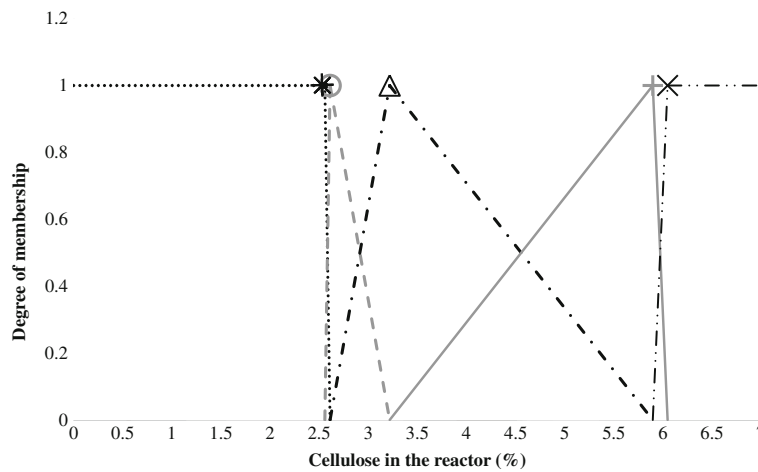


Fig. 4 Membership functions for determining the degree of membership (for cellulose load). Each cluster center corresponds to the maximum degree of membership (=1) of a particular linguistic value (for the linguistic variable cellulose load). The *straight lines* connecting the data points (cluster centers) characterize intermediate degrees of membership within two adjacent linguistic values

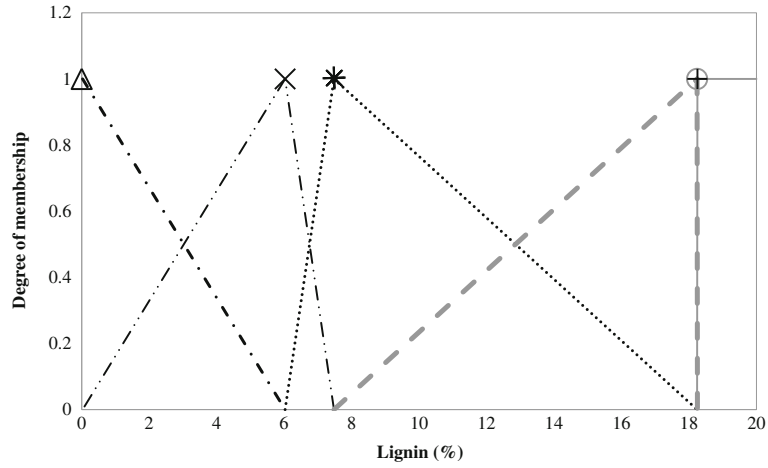


Fig. 5 Membership functions for determining the degree of membership (for lignin content). Each cluster center corresponds to the maximum degree of membership (=1) of a particular linguistic value (for the linguistic variable lignin content). The *straight lines* connecting the data points (cluster centers) characterize intermediate degrees of membership within two adjacent linguistic values

In addition to the figures, the program also provides the coordinates of the cluster centers, which were then used in the creation of functions for determining the degree of membership for the two inputs, cellulose load in the reactor and lignin content, as shown in Figs. 4 and 5.

Each cluster center corresponds to the maximum degree of membership (=1) of a particular linguistic value (for the linguistic variables cellulose load and lignin content).

As one can see in Figs. 4 and 5, some maximum degree of memberships (=1) were very close to each other, making it suitable to fuse some linguistic values. The final membership functions are shown in Figs. 6 and 7.

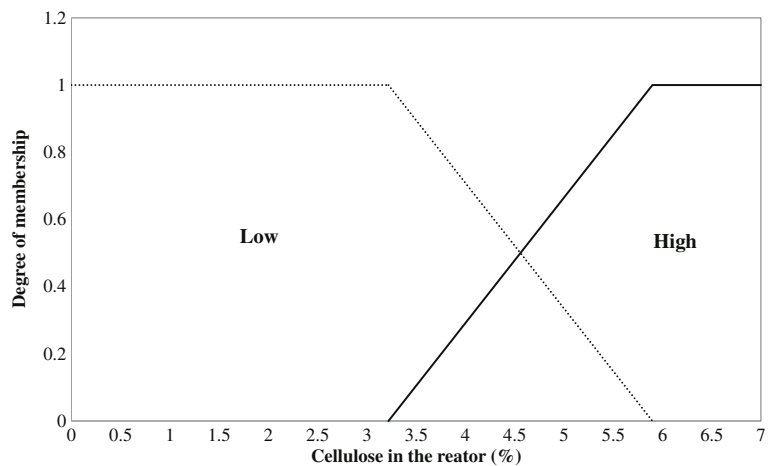


Fig. 6 Final membership functions to determine the degree of membership (for cellulose load)

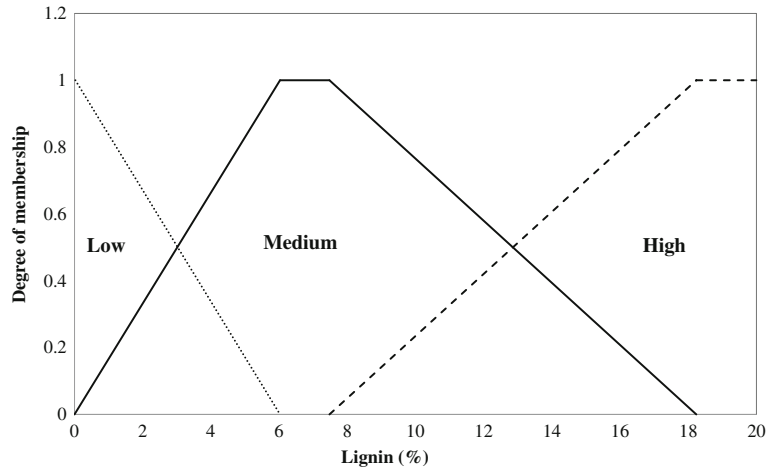


Fig. 7 Final membership functions to determine the degree of membership (for lignin content)

For input variable percentage of cellulose in the reactor, the left trapezoidal membership function in Fig. 6 was assigned the linguistic value “low load”, and the function on the right was assigned the value “high load”. For the input variable percentage of lignin (Fig. 7), in turn, the triangular function on the left was assigned the value “low lignin,” the central trapezoidal region, “medium lignin”, and the right, “high lignin.”

With the equations of the straight lines (linear equations) associated with the linguistic values of the linguistic variables, a fuzzy inference system was generated representing the system behavior (rate of hydrolysis depending on the conditions of hydrolysis) using a minimum number of rules. These rules are:

Rule (1) If the percentage of cellulose is low or high and the percentage of lignin is low, the hydrolysis rate follows kinetic model 1.

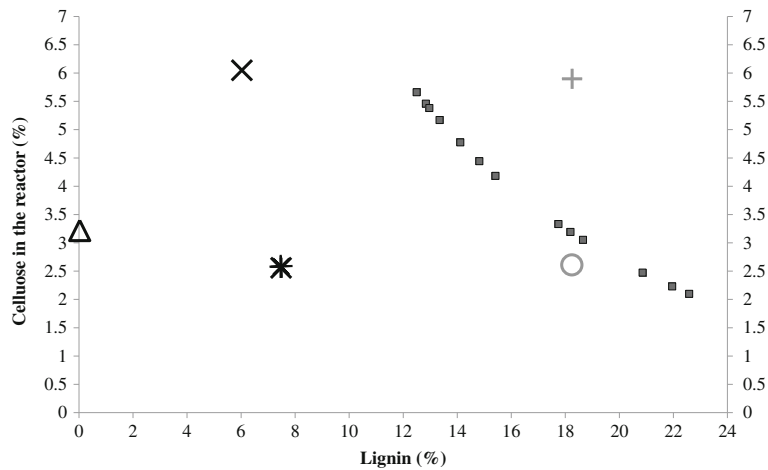


Fig. 8 Experimental data (percentage of cellulose in the reactor and percentage of lignin with regard to the substrate composition) for the validation experiment, in relation to cluster centers

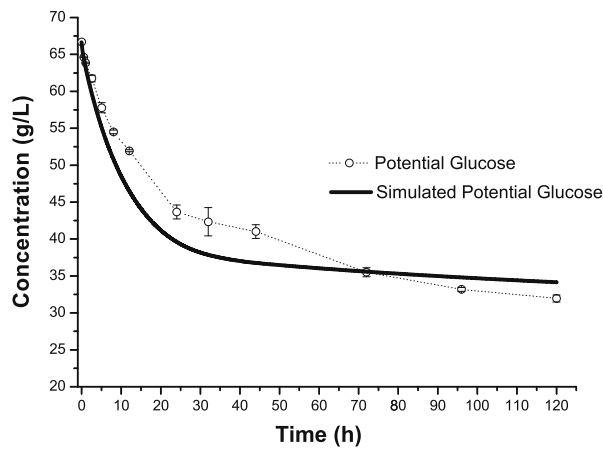


Fig. 9 Potential glucose concentration profiles. Comparison of experimental and simulated data for synthetic substrate. *Error bars* are standard errors of independent triplicate experiments. Operating conditions are 50 °C, pH 4.8, enzyme concentration of 1,200 FPU L⁻¹ solution, and initial potential glucose concentration of 66.66 g_{potential_glucose} L⁻¹ solution

- Rule (2) If the percentage of cellulose is low and the percentage of lignin is medium, the hydrolysis rate follows kinetic model 2.
- Rule (3) If the percentage of cellulose is high and the percentage of lignin is medium, the hydrolysis rate follows kinetic model 3.
- Rule (4) If the percentage of cellulose is low and the percentage of lignin is high, the hydrolysis rate follows kinetic model 4.

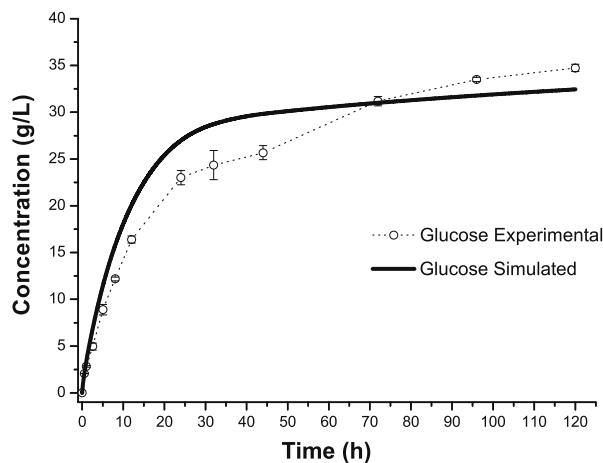


Fig. 10 Product concentration profiles. Comparison of experimental and simulated data for synthetic substrate. *Error bars* are standard errors of independent triplicate experiments. Operating conditions are 50 °C, pH 4.8, enzyme concentration of 1,200 FPU L⁻¹ solution and initial potential glucose concentration of 66.66 g_{potential_glucose} L⁻¹ solution

Rule (5) If the percentage of cellulose is high and the percentage of lignin is high, the hydrolysis rate follows kinetic model 5.

In these rules, the kinetic models 1–5 are:

- Kinetic model 1 Pseudohomogeneous Michaelis-Menten (MM) model, with inhibition studies originated with filter paper.
- Kinetic model 2 Pseudo-homogeneous MM with competitive inhibition.
- Kinetic model 3 Modified MM Model with inhibition.
- Kinetic model 4 Chrastil model.
- Kinetic model 5 Chrastil model b.

Validation Test

Figure 8 shows the experimental data for percentage of cellulose in the reactor and percentage of lignin with regard to the substrate composition for the validation experiment (in relation to cluster centers).

Figures 9 and 10 show the comparison of the “global” model predictions compared to experimental data of enzymatic hydrolysis of the synthetic (validation) substrate, which was not used in the processes of clustering and model building. The simulated responses shown in these figures, provided by the integrator, were built initially considering the relevance of models 2, 3, 4, and 5. With the progress of the experiment, the interpolator gently tended to consider mainly Chrastil model 4.

Figure 10 compares the glucose simulated by the fuzzy interpolator (“global” model) and glucose obtained experimentally for a total time of 120 h of hydrolysis. It can be seen that between 0 and 15 h, the interpolator represents the experimental data very well. However, within the interval between 15 and 45 h, the interpolator response does not adhere perfectly to the experimental data due to the higher standard error associated with the experimental measurements of glucose concentration within this interval. Finally, after 45 h, the experimental and simulated data increased gradually up to $\sim 33 \text{ g L}^{-1}$, until hydrolysis is completed. A major advantage in using this interpolator is that retuning of the parameters of the simple “local” models is not necessary.

Conclusion

The presented results support the idea that a fuzzy interpolator for enzymatic hydrolysis of lignocellulosic materials can be an important modeling tool and a feasible methodology for representing the process in a robust and reliable way. Other important factor to be noticed is that this methodology can be applied for enzymatic hydrolysis of lignocellulosic materials different from sugar cane bagasse, such as soy hulls and sugar cane straw. After “local” kinetic models are adjusted and characterized for the lignocellulosic materials, it is possible to develop a fuzzy interpolator for kinetic modeling based on simple models for obtaining a “global” model covering a broad range of hydrolysis conditions.

Acknowledgments The authors would like to thank the support of Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP-BIOEN), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

- Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology*, *83*, 1–11.
- Eklund, R., Galbe, M., & Zacchi, G. (1990). Optimization of temperature and enzyme concentration in the enzyme saccharification of steam-pretreated willow. *Enzyme and Microbial Technology*, *12*, 225–228.
- Granda, C. B. (2007). Enzymatic hydrolysis of lime-pretreated corn stover and investigation of the HCH-1 Model: Inhibition pattern, degree of inhibition, validity of simplified HCH-1 Model. *Bioresource Technology*, *98*, 2969–2977.
- Hahn-Hagerdal, B., Galbe, M., Gorwa-Grauslund, M. F., Lidén, G., & Zacchi, G. (2006). Bio-ethanol—the fuel of tomorrow from the residues of today. *Trends in Biotechnology*, *24*, 549–556.
- Pandey, A., Soccol, C. R., Nigam, P., Brand, D., Mohan, R., & Roussos, S. (2000). Biotechnology potential of agro-industrial residues. I: sugarcane bagasse. *Bioresource Technology*, *74*, 69–80.
- Hamelinck, C. N., Hooijdonk, G. V., & Faaij, A. P. C. (2005). Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. *Biomass and Bioenergy*, *28*, 384–410.
- Rabelo, S. C. (2007). Avaliação do desempenho do pré-tratamento com peróxido de hidrogênio alcalino para a hidrólise enzimática de cana-de-açúcar. Dissertation. State University of Campinas.
- Cunha, C. M., Silva, F. T. (2001). Characterization of carbohydrates present in hydrolyzate obtained from sugar cane bagasse pretreated by explosion. In: 6th Brazilian Symposium Chemistry Lignins and other Wood Components, VII, pp. 221–226.
- Gámez, S., González-Cabriales, J. J., Ramírez, J. A., Garrote, G., & Vázquez, M. (2006). Study of hydrolysis of sugar cane bagasse using phosphoric acid. *Journal of Food Engineering*, *74*, 78–88.
- Laser, M., Schulman, D., Allen, S. G., Lichwa, J., Antal, M. J., Jr., & Lynd, L. R. (2002). A comparison of liquid hot water and steam pretreatments of sugar cane bagasse for bioconversion to ethanol. *Bioresource Technology*, *81*, 33–44.
- Vallander, L., & Eriksson, K. E. L. (1985). Enzymatic saccharification of pretreated wheat straw. *Biotechnology and Bioengineering*, *27*, 650–659.
- Sousa, R., Jr., Carvalho, M. L., Giordano, R. L. C., & Giordano, R. C. (2011). Recent trends in the modeling of cellulose hydrolysis. *Brazilian Journal of Chemical Engineering*, *28*, 545–564.
- Zhang, Y. H. P., & Lynd, L. R. (2004). Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulose systems. *Biotechnology and Bioengineering*, *88*, 797–824.
- Cavalcanti-Montañó, I. D., Suarez, C. A. G., Rodríguez-Zúñiga, U. F., Giordano, R. L. C., & Sousa, R., Jr. (2013). Optimal bioreactor operational policies for the enzymatic hydrolysis of sugarcane bagasse. *Bioenergy Research*, *6*, 776–785.
- Bezerra, R. M. F., & Dias, A. A. (2004). Discrimination among eight modified Michaelis-Menten kinetics models of cellulose hydrolysis with a large range of substrate/enzyme ratios. *Applied Biochemistry and Biotechnology*, *112*, 173–184.
- Carrillo, F., LIS, M. J., Colom, X., López-Mesas, M., & Valldeperas, J. (2005). Effect of alkali pretreatment on cellulose hydrolysis of wheat straw: Kinetic study. *Process Biochemistry*, *40*, 3360–3364.
- Chrastil, J. (1988). Determination of the first order consecutive reaction rate constants from final products. *Computers and Chemistry*, *12*, 289–292.
- Chrastil, J. (1988). Enzymic product formation curves with the normal or diffusion limited reaction mechanism and in the presence of substrate receptors. *International Journal of Biochemistry*, *20*, 683–693.
- Carvalho, M. L., Sousa, R., Jr., Rodríguez-Zúñiga, U. F., Suarez, C. A. G., Rodrigues, D. S., Giordano, R. C., et al. (2013). Kinetic study of the enzymatic hydrolysis of sugarcane bagasse. *Brazilian Journal of Chemical Engineering*, *30*, 437–447.
- Carvalho, M. (2011). Estudo cinético da hidrólise enzimática de celulose de bagaço de cana-de-açúcar. Dissertation, Federal University of Sao Carlos.
- Adney, B., Baker, J. Chemical analysis and testing task: <https://engineering.purdue.edu/LORRE/research/LAP-006.pdf>. Revised 2013.
- Ghose, T. K. (1987). Measurement of cellulase activity. *Pure and Applied Chemistry*, *59*, 257–268.
- Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry*, *31*, 426–428.
- Bezdek, J. (1981). *Fuzzy models and algorithms for pattern recognition and image processing*. New York: Wiley.
- Nelles, O. (2001). *Nonlinear system identification*. Berlin: Springer.
- Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, *2*, 267–278.
- Bezdek, J. C., Ehrlich, R., et al. (1984). FCM: the fuzzy c-means clustering algorithm. *Computers and Geosciences*, *10*, 191–203.
- Pedrycz, W. (1990). Processing in relation structures: fuzzy relational equations. *Fuzzy Sets and Systems*, *40*, 87–100.

7.1.REFERÊNCIAS

CUNHA, F. C. et al. Dynamic Simulation of a Compressor Located in a Natural Gas Processing Unit Using EMSO Simulator. In: RITA MARIA DE BRITO ALVES, C. A. O. D. N. e EVARISTO CHALBAUD, B. (Ed.). **Computer Aided Chemical Engineering**: Elsevier, v. Volume 27, 2009. p.603-608. ISBN 1570-7946.

DARÉ, R. M. **AVALIAÇÃO DOS COEFICIENTES DE RENDIMENTO E MODELAGEM DO PROCESSO FERMENTATIVO DE PRODUÇÃO DE ETANOL**. 2008. 67 (Mestrado). Programa de pós graduação em Engenharia Química, Universidade Federal de São Carlos, São Carlos.

FURLAN, F. F. et al. A simple approach to improve the robustness of equation-oriented simulators: Multilinear look-up table interpolators. **Computers & Chemical Engineering**, v. 86, p. 1-4, 3/4/ 2016. ISSN 0098-1354. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0098135415003786> >.

SOARES, R. D. P.; SECCHI, A. R. EMSO: A new environment for modelling, simulation and optimisation. In: ANDRZEJ, K. e ILKKA, T. (Ed.). **Computer Aided Chemical Engineering**: Elsevier, v. Volume 14, 2003. p.947-952. ISBN 1570-7946.

ABBOD, M. F. et al. Survey of utilisation of fuzzy technology in Medicine and Healthcare. **Fuzzy Sets and Systems**, v. 120, n. 2, p. 331-349, Jun 1 2001. ISSN 0165-0114. Disponível em: < <Go to ISI>://WOS:000168316500012 >.

ADACHI, E. et al. Modification of metabolic pathways of *Saccharomyces cerevisiae* by the expression of lactate dehydrogenase and deletion of pyruvate decarboxylase genes for the lactic acid fermentation at low pH value. **Journal of Fermentation and Bioengineering**, v. 86, n. 3, p. 284-289, 1998. ISSN 0922-338X. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0922338X98801311> >.

AHMED, M. N. et al. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. **Ieee Transactions on Medical Imaging**, v. 21, n. 3, p. 193-199, Mar 2002. ISSN 0278-0062. Disponível em: < <Go to ISI>://WOS:000175063900001 >.

ALMEIDA, L. M. W. D.; PALHARINI, B. N. Os "Mundos da Matemática" em Atividades de Modelagem Matemática. **Bolema: Boletim de Educação Matemática**, v. 26, n. 43, p. 907-934, 2012-08 2012. ISSN 0103-636X. Disponível em: < <Go to ISI>://SCIELO:S0103-636X2012000300008 >.

AMIMOTO, K. et al. A novel toxin homologous to large clostridial cytotoxins found in culture supernatant of *Clostridium perfringens* type C. **Microbiology-Sgm**, v. 153, p. 1198-1206, Apr 2007. ISSN 1350-0872. Disponível em: <<Go to ISI>://WOS:000245959200028 >.

ATANASSOV, E.; DIMOV, I. T. What Monte Carlo models can do and cannot do efficiently? **Applied Mathematical Modelling**, v. 32, n. 8, p. 1477-1500, 8// 2008. ISSN 0307-904X. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0307904X07001564> >.

BABLOYAN.A; NICOLIS, G. CHEMICAL INSTABILITIES AND MULTIPLE STEADY-STATE TRANSITIONS IN MONOD-JACOB TYPE MODELS. **Journal of Theoretical Biology**, v. 34, n. 1, p. 185-&, 1972 1972. ISSN 0022-5193. Disponível em: <<Go to ISI>://WOS:A1972L305500014 >.

BAHARANI, M. et al. High-level design space exploration of locally linear neuro-fuzzy models for embedded systems. **Fuzzy Sets and Systems**, v. 253, p. 44-63, Oct 16 2014. ISSN 0165-0114. Disponível em: <<Go to ISI>://WOS:000341480700003 >.

BALDWIN, J. F.; GUILD, N. C. F. Feasible algorithms for approximate reasoning using fuzzy logic. **Fuzzy Sets and Systems**, v. 3, n. 3, p. 225-251, 1980. ISSN 0165-0114. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/0165011480900226> >.

BANNAI, H.; OE, T. A New Phase in New Generation Cooperatives in Ethanol Production-Case Study of the State of Minnesota, US. **Tohoku Journal of Agricultural Research**, v. 60, n. 3-4, p. 91-105, Mar 2010. ISSN 0040-8719. Disponível em: <<Go to ISI>://BIOABS:BACD201000226586 >.

BATTY, I.; WALKER, P. D. IDENTIFICATION OF CLOSTRIDIUM NOVYI (CLOSTRIDIUM OEDEMATIENS) + CLOSTRIDIUM TETANI BY USE OF FLOURESCENT LABELLED ANTIBODIES. **Journal of Pathology and Bacteriology**, v. 88, n. 1, p. 327-&, 1964 1964. Disponível em: <<Go to ISI>://WOS:A19646343B00004 >.

BELIYI, Y.; AKTORIES, K. Bacterial toxin and effector glycosyltransferases. **Biochimica Et Biophysica Acta-General Subjects**, v. 1800, n. 2, p. 134-143, Feb 2010. ISSN 0304-4165. Disponível em: <<Go to ISI>://WOS:000274415500009 >.

BEZDEC, J. C. **Pattern Recognition with Fuzzy Objective Function Algorithms**. New York: 1981. 256.

BEZDEK, J. C. **Pattern Recognition with Fuzzy Objective Function Algorithms**. New York: Plenum Press 1981.

BORG, D.; SUETAKE, M.; BRANDÃO, D. A neural network developed in a Foundation Fieldbus environment to calculate flow rates for compressible fluid. **Flow Measurement and Instrumentation**, v. 40, n. 0, p. 142-148, 2014. ISSN 0955-5986. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0955598614001198> >.

BOYKOV, Y.; VEKSLER, O.; ZABIH, R. Fast approximate energy minimization via graph cuts. **Ieee Transactions on Pattern Analysis and Machine Intelligence**, v. 23, n. 11, p. 1222-1239, Nov 2001. ISSN 0162-8828. Disponível em: < <Go to ISI>://WOS:000172108300002 >.

BRETT, M. M. et al. Soft tissue infections caused by spore-forming bacteria in injecting drug users in the United Kingdom. **Epidemiology and Infection**, v. 133, n. 4, p. 575-582, Aug 2005. ISSN 0950-2688. Disponível em: < <Go to ISI>://WOS:000231202000001 >.

BULTMANN, R.; SZABO, B.; STARKE, K. INHIBITION BY ETHANOL OF CONTRACTIONS OF RAT VAS-DEFERENS - NO EVIDENCE FOR SELECTIVE BLOCKADE OF P(2X)-PURINOCEPTORS. **Naunyn-Schmiedebergs Archives of Pharmacology**, v. 347, n. 5, p. 527-533, May 1993. ISSN 0028-1298. Disponível em: < <Go to ISI>://WOS:A1993LB43500012 >.

CARSTENSEN, J.; HARREMOES, F.; MADSEN, H. STATISTICAL IDENTIFICATION OF MONOD-KINETIC PARAMETERS FROM ONLINE MEASUREMENTS. **Water Science and Technology**, v. 31, n. 2, p. 125-133, 1995 1995. ISSN 0273-1223. Disponível em: < <Go to ISI>://WOS:A1995RE55100013 >.

CARVALHO, M. L. D. **ESTUDO CINÉTICO DA HIDRÓLISE ENZIMÁTICA DECELULOSE DE BAGAÇO DE CANA-DE-AÇÚCAR**. 2011. 112 (Dissertação de Mestrado). Departamento de Engenharia Química, UFSCar, São Carlos - SP.

CERNY, V. THERMODYNAMICAL APPROACH TO THE TRAVELING SALESMAN PROBLEM - AN EFFICIENT SIMULATION ALGORITHM. **Journal of Optimization Theory and Applications**, v. 45, n. 1, p. 41-51, 1985 1985. ISSN 0022-3239. Disponível em: < <Go to ISI>://WOS:A1985ACY7400004 >.

CERQUEIRA LEITE, R. C. D. et al. Can Brazil replace 5% of the 2025 gasoline world demand with ethanol? **Energy**, v. 34, n. 5, p. 655-661, 2009. ISSN 0360-5442. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0360544208002958> >.

CHEN, C.-H. et al. Optimization of teacher volunteer transferring problems using greedy genetic algorithms. **Expert Systems with Applications**, v. 42, n. 1, p. 668-678,

1// 2015. ISSN 0957-4174. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0957417414004965>>.

CHENG, W. T.; LIN, C. H. NUMERICAL ANALYSIS OF MASS TRANSFER ON THE MELTING PHENOMENON FROM A VERTICAL PLATE IN A LIQUID-SATURATED POROUS MEDIUM. **Journal of Porous Media**, v. 15, n. 5, p. 485-493, 2012 2012. ISSN 1091-028X. Disponível em: <<Go to ISI>://WOS:000303944000007 >.

CHOI, Y. A Study of Brazilian Military Government (1964-1985): Focusing on the Pre-Democracy Period. **Korean Journal of Latin American and Caribbean Studies**, v. 28, n. 2, p. 445-478, 2010 2010. ISSN 1229-7046. Disponível em: <<Go to ISI>://KJD:ART001423303 >.

CINQUE, L.; FORESTI, G.; LOMBARDI, L. A clustering fuzzy approach for image segmentation. **Pattern Recognition**, v. 37, n. 9, p. 1797-1807, Sep 2004. ISSN 0031-3203. Disponível em: <<Go to ISI>://WOS:000222701700004 >.

COSENZA, A.; MANNINA, G.; VIVIANI, G. Parameter estimation and sensitivity analysis of a nitrogen and phosphorus biological removal model. **18th World Imacs Congress and Modsim09 International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences**, p. 3151-3157, 2009 2009. Disponível em: <<Go to ISI>://WOS:000290045003029 >.

CUNHA, F. C. et al. Dynamic Simulation of a Compressor Located in a Natural Gas Processing Unit Using EMSO Simulator. In: RITA MARIA DE BRITO ALVES, C. A. O. D. N. e EVARISTO CHALBAUD, B. (Ed.). **Computer Aided Chemical Engineering**: Elsevier, v. Volume 27, 2009. p.603-608. ISBN 1570-7946.

DA COSTA, A. C. A.; JUNIOR, N. P.; ARANDA, D. A. G. The situation of biofuels in Brazil: New generation technologies. **Renewable and Sustainable Energy Reviews**, v. 14, n. 9, p. 3041-3049, 2010. ISSN 1364-0321. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S1364032110002388>>.

DA SILVA VELLASCO, P. C. G. et al. Capítulo 3 - Modelagem Numérica com o Método dos Elementos Finitos. In: (Ed.). **Modelagem de Estrutura de Aço e Mistas**. Rio de Janeiro: Elsevier Editora Ltda., 2014. p.165-284. ISBN 978-85-352-7197-3.

DARÉ, R. M. **AVALIAÇÃO DOS COEFICIENTES DE RENDIMENTO E MODELAGEM DO PROCESSO FERMENTATIVO DE PRODUÇÃO DE ETANOL**. 2008. 67 (Mestre). Programa de pós graduação em Engenharia Química, Universidade Federal de São Carlos, 2008.

DE OLIVEIRA, A. Reassessing the Brazilian alcohol programme. **Energy Policy**, v. 19, n. 1, p. 47-55, 1991/2// 1991. ISSN 0301-4215. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/030142159190076Z>>.

DEMBELE, D.; KASTNER, P. Fuzzy C-means method for clustering microarray data. **Bioinformatics**, v. 19, n. 8, p. 973-980, May 22 2003. ISSN 1367-4803. Disponível em: <<Go to ISI>://WOS:000183095300010 >.

DETTE, H. et al. Robust and efficient design of experiments for the Monod model. **Journal of Theoretical Biology**, v. 234, n. 4, p. 537-550, Jun 21 2005. ISSN 0022-5193. Disponível em: <<Go to ISI>://WOS:000228494500008 >.

DHERE, N. G. The development of solar photovoltaic energy in Brazil. **Solar Cells**, v. 26, n. 1-2, p. 13-23, 1989. ISSN 0379-6787. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/037967878990063X>>.

DIAZ, L. A. et al. Pharmacologic and toxicologic evaluation of C-novy-NT spores. **Toxicological Sciences**, v. 88, n. 2, p. 562-575, Dec 2005. ISSN 1096-6080. Disponível em: <<Go to ISI>://WOS:000233215700030 >.

DOLAN, W. B.; CUMMINGS, P. T.; LEVAN, M. D. ALGORITHMIC EFFICIENCY OF SIMULATED ANNEALING FOR HEAT-EXCHANGER NETWORK DESIGN. **Computers & Chemical Engineering**, v. 14, n. 10, p. 1039-1050, Oct 1990. ISSN 0098-1354. Disponível em: <<Go to ISI>://WOS:A1990EQ44000001 >.

EFTAXIAS, A. et al. Nonlinear kinetic parameter estimation using simulated annealing. **Computers & Chemical Engineering**, v. 26, n. 12, p. 1725-1733, 2002. ISSN 0098-1354. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0098135402001564>>.

EKLUND, M. W. et al. INTERSPECIES CONVERSION OF CLOSTRIDIUM-BOTULINUM TYPE-C TO CLOSTRIDIUM-NOVYI TYPE-A BY BACTERIOPHAGE. **Science**, v. 186, n. 4162, p. 456-458, 1974 1974. ISSN 0036-8075. Disponível em: <<Go to ISI>://WOS:A1974U453000026 >.

EL-NAGGAR, K. M. et al. Simulated Annealing algorithm for photovoltaic parameters identification. **Solar Energy**, v. 86, n. 1, p. 266-274, 2012. ISSN 0038-092X. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0038092X11003586>>.

FERREIRA FILHO, J. B. D. S.; HORRIDGE, M. Ethanol expansion and indirect land use change in Brazil. **Land Use Policy**, v. 36, n. 0, p. 595-604, 2014. ISSN 0264-8377. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S026483771300210X>>.

FORBES, N. S.; CLARK, D. S.; BLANCH, H. W. Using isotopomer path tracing to quantify metabolic fluxes in pathway models containing reversible reactions. **Biotechnology and Bioengineering**, v. 74, n. 3, p. 196-211, Aug 5 2001. ISSN 0006-3592. Disponível em: <<Go to ISI>://WOS:000169684000002 >.

FURTADO, A. T.; GAYA SCANDIFFIO, M. I.; BARBOSA CORTEZ, L. A. The Brazilian sugarcane innovation system. **Energy Policy**, v. 39, n. 1, p. 156-166, Jan 2011. ISSN 0301-4215. Disponível em: <<Go to ISI>://WOS:000285492000014 >.

GARIAZZO, C. et al. Assessment of population exposure to Polycyclic Aromatic Hydrocarbons (PAHs) using integrated models and evaluation of uncertainties. **Atmospheric Environment**, v. 101, n. 0, p. 235-245, 1// 2015. ISSN 1352-2310. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1352231014009005> >.

GÍRIO, F. M. et al. Hemicelluloses for fuel ethanol: A review. **Bioresource Technology**, v. 101, n. 13, p. 4775-4800, 2010. ISSN 0960-8524. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0960852410001744> >.

GOLDBETER, A. KINETIC NEGATIVE COOPERATIVITY IN ALLOSTERIC MODEL OF MONOD, WYMAN AND CHANGEUX. **Journal of Molecular Biology**, v. 90, n. 1, p. 185-190, 1974 1974. ISSN 0022-2836. Disponível em: <<Go to ISI>://WOS:A1974V131000016 >.

GOLTERMAN, H. L. REFLECTIONS ON POST-OECD EUTROPHICATION MODELS. **Hydrobiologia**, v. 218, n. 2, p. 167-176, Jul 12 1991. ISSN 0018-8158. Disponível em: <<Go to ISI>://WOS:A1991GE26500008 >.

HAMELINCK, C. N.; HOOIJDONK, G. V.; FAAIJ, A. P. C. Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. **Biomass and Bioenergy**, v. 28, n. 4, p. 384-410, 2005. ISSN 0961-9534. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0961953404001801> >.

HANSEN, L. K.; SALAMON, P. NEURAL NETWORK ENSEMBLES. **Ieee Transactions on Pattern Analysis and Machine Intelligence**, v. 12, n. 10, p. 993-1001, Oct 1990. ISSN 0162-8828. Disponível em: <<Go to ISI>://WOS:A1990EA04200005 >.

HU, B. G.; MANN, G. K. I.; GOSINE, R. G. A systematic study of fuzzy PID controllers - Function-based evaluation approach. **Ieee Transactions on Fuzzy Systems**, v. 9, n. 5, p. 699-712, Oct 2001. ISSN 1063-6706. Disponível em: <<Go to ISI>://WOS:000172014100003 >.

ICHIHASHI, H. et al. Neuro-fuzzy ID3: A method of inducing fuzzy decision trees with linear programming for maximizing entropy and an algebraic method for incremental learning. **Fuzzy Sets and Systems**, v. 81, n. 1, p. 157-167, Jul 8 1996. ISSN 0165-0114. Disponível em: <<Go to ISI>://WOS:A1996UR78400013 >.

JANG, J. S. R. ANFIS - ADAPTIVE-NETWORK-BASED FUZZY INFERENCE SYSTEM. **Ieee Transactions on Systems Man and Cybernetics**, v. 23, n. 3, p. 665-685, May-Jun 1993. ISSN 0018-9472. Disponível em: <<Go to ISI>://WOS:A1993MF08900003 >.

JANG, J. S. R.; SUN, C. T. NEURO-FUZZY MODELING AND CONTROL. **Proceedings of the Ieee**, v. 83, n. 3, p. 378-406, Mar 1995. ISSN 0018-9219. Disponível em: <<Go to ISI>://WOS:A1995QJ99800003 >.

JIANG, X.; ADEL, H. Neuro-genetic algorithm for non-linear active control of structures. **International Journal for Numerical Methods in Engineering**, v. 75, n. 7, p. 770-786, Aug 13 2008. ISSN 0029-5981. Disponível em: <<Go to ISI>://WOS:000258874000002 >.

KAY, L. E. From logical neurons to poetic embodiments of mind: Warren S. McCulloch's project in neuroscience. **Science in Context**, v. 14, n. 4, p. 591-614, Dec 2001. ISSN 0269-8897. Disponível em: <<Go to ISI>://WOS:000177178900005 >.

KIATKITTIPONG, W.; WONGSUCHOTO, P.; PAVASANT, P. Life cycle assessment of bagasse waste management options. **Waste Management**, v. 29, n. 5, p. 1628-1633, May 2009. ISSN 0956-053X. Disponível em: <<Go to ISI>://WOS:000264977200023 >.

KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. OPTIMIZATION BY SIMULATED ANNEALING. **Science**, v. 220, n. 4598, p. 671-680, 1983 1983. ISSN 0036-8075. Disponível em: <<Go to ISI>://WOS:A1983QN77100006 >.

KOFMAN, G. I.; LYSSENKOV, S. A.; LOBACHEV, R. V. **Automatic neuro-fuzzy expert system generation method for artificial intelligence application, involves receiving location type and structure of e.g. spreadsheet, using which neural network of neuro-fuzzy expert system is trained**: Dintecom Inc.

KORDYLEWSKI, H.; GRAUPE, D.; LIU, K. A novel large-memory neural network as an aid in medical diagnosis applications. **Ieee Transactions on Information Technology in Biomedicine**, v. 5, n. 3, p. 202-209, Sep 2001. ISSN 1089-7771. Disponível em: <<Go to ISI>://WOS:000170769300004 >.

KOUROU, K. et al. Machine learning applications in cancer prognosis and prediction. **Computational and Structural Biotechnology Journal**, v. 13, n. 0, p. 8-17, 12// 2015. ISSN 2001-0370. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S2001037014000464> >.

KREUTZBERG, G. W. Microglia: A sensor for pathological events in the CNS. **Trends in Neurosciences**, v. 19, n. 8, p. 312-318, Aug 1996. ISSN 0166-2236. Disponível em: <<Go to ISI>://WOS:A1996VA75000003 >.

KURGANOV, B. I. et al. THEORETICAL-ANALYSIS OF KINETIC-BEHAVIOR OF HYSTERETIC ALLOSTERIC ENZYMES .1. KINETIC MANIFESTATIONS OF SLOW CONFORMATIONAL CHANGE OF AN OLIGOMERIC ENZYME IN MONOD, WYMAN AND CHANGEUX MODEL. **Journal of Theoretical Biology**, v. 60, n. 2, p. 247-269, 1976 1976. ISSN 0022-5193. Disponível em: <<Go to ISI>://WOS:A1976CC60200001 >.

LEI, Y. et al. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAS. **Mechanical Systems and Signal Processing**, v. 21, n. 5, p. 2280-2294, Jul 2007. ISSN 0888-3270. Disponível em: <<Go to ISI>://WOS:000246266700021 >.

LEVENSPIEL, O. THE MONOD EQUATION - A REVISIT AND A GENERALIZATION TO PRODUCT INHIBITION SITUATIONS. **Biotechnology and Bioengineering**, v. 22, n. 8, p. 1671-1687, 1980 1980. ISSN 0006-3592. Disponível em: <<Go to ISI>://WOS:A1980KD08700009 >.

LIMA, C. G. R. D. et al. Padronização de teste de potência in vitro para vacinas que contenham toxoide alfa de Clostridium novyi tipo B. **Arq. Inst. Biol**, v. 78, p. 507-512, 12/2011 2011.

LIN, C. T.; LEE, C. S. G. NEURAL-NETWORK-BASED FUZZY-LOGIC CONTROL AND DECISION SYSTEM. **Ieee Transactions on Computers**, v. 40, n. 12, p. 1320-1336, Dec 1991. ISSN 0018-9340. Disponível em: <<Go to ISI>://WOS:A1991GU89000002 >.

LINDEGREN, C. C.; SPIEGELMAN, S.; LINDEGREN, G. SURVEY OF GROWTH AND GAS PRODUCTION OF GENETIC VARIANTS OF SACCHAROMYCES CEREVISIAE ON DIFFERENT SUGARS. **Archives of Biochemistry**, v. 6, n. 2, p. 185-198, 1945 1945. ISSN 0096-9621. Disponível em: <<Go to ISI>://WOS:A1945UG19600002 >.

LOPEZ-CRUZ, I. L.; HERNANDEZ-LARRAGOITI, L. NEURO-FUZZY MODELS FOR AIR TEMPERATURE AND HUMIDITY OF ARCHED AND VENLO TYPE GREENHOUSES IN CENTRAL MEXICO. **Agrociencia**, v. 44, n. 7, p. 791-805, Oct-

Nov 2010. ISSN 1405-3195. Disponível em: <<Go to ISI>://WOS:000284202900006 >.

LUCHI, N. R.; TRINDADE, S. C. ETHANOL VERSUS NAPHTHA UNDER BRAZIL ECONOMY. **Hydrocarbon Processing**, v. 61, n. 5, p. 179-183, 1982 1982. ISSN 0018-8190. Disponível em: <<Go to ISI>://WOS:A1982NS44400025 >.

MAMDANI, E. H. APPLICATION OF FUZZY LOGIC TO APPROXIMATE REASONING USING LINGUISTIC-SYNTHESIS. **Ieee Transactions on Computers**, v. 26, n. 12, p. 1182-1191, 1977 1977. ISSN 0018-9340. Disponível em: <<Go to ISI>://WOS:A1977EF28400002 >.

MAMDANI, E. H.; ASSILIAN, S. EXPERIMENT IN LINGUISTIC SYNTHESIS WITH A FUZZY LOGIC CONTROLLER. **International Journal of Man-Machine Studies**, v. 7, n. 1, p. 1-13, 1975 1975. ISSN 0020-7373. Disponível em: <<Go to ISI>://WOS:A1975V780800001 >.

MARIN, A.; SALMERON, J. Tactical design of rail freight networks .2. Local search methods with statistical analysis. **European Journal of Operational Research**, v. 94, n. 1, p. 43-53, Oct 11 1996. ISSN 0377-2217. Disponível em: <<Go to ISI>://WOS:A1996VK77300004 >.

MARJOTTA-MAISTRO, M. C.; BARROS, G. S. A. D. C. Relações comerciais e de preços no mercado nacional de combustíveis. **Revista de Economia e Sociologia Rural**, v. 41, n. 4, p. 829-858, 2003 2003. ISSN 0103-2003. Disponível em: <<Go to ISI>://SCIELO:S0103-20032003000400006 >.

MARQUARDT, D. W. AN ALGORITHM FOR LEAST-SQUARES ESTIMATION OF NONLINEAR PARAMETERS. **Journal of the Society for Industrial and Applied Mathematics**, v. 11, n. 2, p. 431-441, 1963 1963. ISSN 0368-4245. Disponível em: <<Go to ISI>://WOS:A1963WW43200015 >.

MARTINEZ, S. H. et al. Analysis of socio-economic impacts of sustainable sugarcane-ethanol production by means of inter-regional Input-Output analysis: Demonstrated for Northeast Brazil. **Renewable & Sustainable Energy Reviews**, v. 28, p. 290-316, Dec 2013. ISSN 1364-0321. Disponível em: <<Go to ISI>://WOS:000328868700023 >.

MCCULLOCH, W. S.; PITTS, W. A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY (REPRINTED FROM BULLETIN OF MATHEMATICAL BIOPHYSICS, VOL 5, PG 115-133, 1943). **Bulletin of Mathematical Biology**, v. 52, n. 1-2, p. 99-115, 1990 1990. ISSN 0092-8240. Disponível em: <<Go to ISI>://WOS:A1990DC78200005 >.

METROPOLIS, N. et al. EQUATION OF STATE CALCULATIONS BY FAST COMPUTING MACHINES. **Journal of Chemical Physics**, v. 21, n. 6, p. 1087-1092, 1953. ISSN 0021-9606. Disponível em: <<Go to ISI>://WOS:A1953UC11200020 >.

MIRA, J. M. Symbols versus connections: 50 years of artificial intelligence. **Neurocomputing**, v. 71, n. 4-6, p. 671-680, Jan 2008. ISSN 0925-2312. Disponível em: <<Go to ISI>://WOS:000253663800027 >.

MITRA, S.; HAYASHI, Y. Neuro-fuzzy rule generation: Survey in soft computing framework. **Ieee Transactions on Neural Networks**, v. 11, n. 3, p. 748-768, May 2000. ISSN 1045-9227. Disponível em: <<Go to ISI>://WOS:000087732100020 >.

MORAN, A. L. INDUSTRIALIZATION OF THE SPRAY FORMING PROCESS FOR MILITARY APPLICATIONS. **Naval Engineers Journal**, v. 104, n. 3, p. 286-290, May 1992. ISSN 0028-1425. Disponível em: <<Go to ISI>://WOS:A1992HP32900027 >.

MOREIRA DA SILVA, R.; CARDOSO, P. J. A. A. P. D. S. 13 - Energia e sustentabilidade. In: (Ed.). **Gestão Ambiental De Unidades Produtivas**. Rio de Janeiro: Elsevier Editora Ltda., 2013. p.385-423. ISBN 978-85-352-5159-3.

MORENO, G.; PASCUAL, V. A hybrid programming scheme combining fuzzy-logic and functional-logic resources. **Fuzzy Sets and Systems**, v. 160, n. 10, p. 1402-1419, 2009. ISSN 0165-0114. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0165011408005307> >.

MORIMURA, S.; LING, Z. Y.; KIDA, K. Ethanol production by repeated-batch fermentation at high temperature in a molasses medium containing a high concentration of total sugar by a thermotolerant flocculating yeast with improved salt-tolerance. **Journal of Fermentation and Bioengineering**, v. 83, n. 3, p. 271-274, 1997. ISSN 0922-338X. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0922338X97809919> >.

MOTIZUKI LOPES, O. M.; VASCONCELOS MARTINS CARRILHO, E. N.; RIBEIRO CASIMIRO LOPES-ASSAD, M. L. EFFECT OF ROCK POWDER AND VINASSE ON TWO TYPES OF SOILS. **Revista Brasileira De Ciencia Do Solo**, v. 38, n. 5, p. 1547-1557, Sep-Oct 2014. ISSN 0100-0683. Disponível em: <<Go to ISI>://WOS:000344966700020 >.

MUNACK, A. OPTIMAL FEEDING STRATEGY FOR IDENTIFICATION OF MONOD-TYPE MODELS BY FED-BATCH EXPERIMENTS. 1989. 195-204 ISBN 1-85166-397-5. Disponível em: <<Go to ISI>://WOS:A1989BQ52K00020 >.

MURAI, T. et al. Evaluation of the function of arming yeast displaying glucoamylase on its cell surface by direct fermentation of corn to ethanol. **Journal of Fermentation and Bioengineering**, v. 86, n. 6, p. 569-572, 1998. ISSN 0922-338X. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0922338X99800087> >.

MURINO, V. et al. Edge/region-based segmentation and reconstruction of underwater acoustic images by Markov Random Fields. In: (Ed.). **1998 Ieee Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings**, 1998. p.408-413. (Proceedings - Ieee Computer Society Conference on Computer Vision and Pattern Recognition). ISBN 1063-6919

0-8186-8497-6.

MUSHARAVATI, F.; HAMOUDA, A. S. M. Enhanced simulated-annealing-based algorithms and their applications to process planning in reconfigurable manufacturing systems. **Advances in Engineering Software**, v. 45, n. 1, p. 80-90, Mar 2012. ISSN 0965-9978. Disponível em: < <Go to ISI>://WOS:000299803100009 >.

NASHED, R. B.; GIRTON, R. E. INHIBITION BY ETHANOL OF THE GROWTH AND RESPIRATION OF MAIZE ROOTS AND COLEOPTILES. **American Journal of Botany**, v. 45, n. 3, p. 190-193, 1958 1958. ISSN 0002-9122. Disponível em: < <Go to ISI>://WOS:A1958WN38400006 >.

NASS, R. FUZZY-LOGIC FINALLY GAINS ACCEPTANCE IN THE UNITED-STATES. **Electronic Design**, v. 40, n. 13, p. 37-40, Jun 25 1992. ISSN 0013-4872. Disponível em: < <Go to ISI>://WOS:A1992JT59800006 >.

NAUCK, D.; KRUSE, R. Neuro-fuzzy systems for function approximation. **Fuzzy Sets and Systems**, v. 101, n. 2, p. 261-271, Jan 16 1999. ISSN 0165-0114. Disponível em: < <Go to ISI>://WOS:000077405600007 >.

NAVA, P. A.; IEEE. **Implementation of neuro-fuzzy systems through interval mathematics**. 1998. 365-369 ISBN 0-7803-4423-5. Disponível em: < <Go to ISI>://WOS:000075973800065 >.

NELLES, O. Structure optimization of Takagi-Sugeno fuzzy models. **International Journal of Uncertainty Fuzziness and Knowledge-Based Systems**, v. 6, n. 2, p. 161-170, Apr 1998. ISSN 0218-4885. Disponível em: < <Go to ISI>://WOS:000073412100007 >.

NETO, R. M.; PUREZA, V.; PUREZA, P. A. C. M. F. H. P. M. N. N. P. D. L. B. T. L. H. M. S. E. G. D. C. Capítulo 8 - Modelagem e Simulação. In: (Ed.). **Metodologia De Pesquisa Em Engenharia De Produção (Segunda edição)**: Elsevier Editora Ltda., 2012. p.169-198. ISBN 978-85-352-4891-3.

NUNEZ, H. M.; OENAL, H.; KHANNA, M. Land use and economic effects of alternative biofuel policies in Brazil and the United States. **Agricultural Economics**, v. 44, n. 4-5, p. 487-499, Jul 2013. ISSN 0169-5150. Disponível em: <<Go to ISI>://WOS:000321442000012 >.

OIH YU, A. S. et al. **Supplier Involvement in Flex-Fuel Technology Development: The General Motors and Volkswagen Brazilian Cases**. 2009. 1573-1582 ISBN 978-1-8908-4319-9. Disponível em: <<Go to ISI>://WOS:000277502500162 >.

OIH YU, A. S. et al. **The Evolution of Flex-Fuel Technology in Brazil: The Bosch Case**. 2010. ISBN 978-1-890843-22-9. Disponível em: <<Go to ISI>://WOS:000287527800129 >.

OMAR, K.; FÁVERO, P. B. P. Capítulo 2 - Introdução à programação linear: formulação geral e modelagem de problemas reais. In: (Ed.). **Pesquisa operacional Para Cursos De Administração**: Elsevier Editora Ltda., 2012. p.17-63. ISBN 978-85-352-3421-3.

PANDEY, A. et al. Biotechnological potential of agro-industrial residues. I: sugarcane bagasse. **Bioresource Technology**, v. 74, n. 1, p. 69-80, 2000. ISSN 0960-8524. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S096085249900142X> >.

PANDEY, A. et al. Biotechnological potential of agro-industrial residues. II: cassava bagasse. **Bioresource Technology**, v. 74, n. 1, p. 81-87, 2000. ISSN 0960-8524. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0960852499001431> >.

PARTHASARATHY, S.; BALAJI, C. Estimation of parameters in multi-mode heat transfer problems using Bayesian inference - Effect of noise and a priori. **International Journal of Heat and Mass Transfer**, v. 51, n. 9-10, p. 2313-2334, May 2008. ISSN 0017-9310. Disponível em: <<Go to ISI>://WOS:000255811700016 >.

PASSOS, R. G. Anexo II - Noções Introdutórias em Redes Neurais. In: (Ed.). **Data Mining: Um Guia Prático**. Rio de Janeiro: Elsevier Editora Ltda., 2005. p.175-181. ISBN 978-85-352-1877-0.

PEARSON, R. THE RESURRECTION OF LUKASIEWICZ, JAN. **Data Processing**, v. 4, n. 11, p. 29-31, 1962 1962. ISSN 0011-684X. Disponível em: <<Go to ISI>://WOS:A1962CDB2300005 >.

PESSÔA, M.; SPINOLA, M. 4 - Modelagem de sistemas dinâmicos com Transformadas de Laplace. In: (Ed.). **Introdução à Automação**. Rio de Janeiro: Elsevier Editora Ltda., 2014. p.77-108. ISBN 978-85-352-4889-0.

PINTO, G. A.; GIORDANO, R. C.; RITA MARIA DE BRITO ALVES, C. A. O. D. N. A. E. C. B. Bioprocess Systems Engineering Applied to the Production of Protein Hydrolysates in a Multipurpose Plant. In: (Ed.). **Computer Aided Chemical Engineering**; Elsevier, v. Volume 27, 2009. p.1887-1892. ISBN 1570-7946.

RAI, A.; WARRIER, M.; SCHNEIDER, R. A hierarchical multi-scale method to simulate reactive–diffusive transport in porous media. **Computational Materials Science**, v. 46, n. 2, p. 469-478, 2009. ISSN 0927-0256. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0927025609001578> >.

RAO, D. H.; KAMAT, H. V. Neuro-fuzzy system for robotics applications. **Journal of the Institution of Electronics and Telecommunication Engineers**, v. 42, n. 4-5, p. 325-333, Jul-Oct 1996. ISSN 0377-2063. Disponível em: < <Go to ISI>://WOS:A1996WH31300014 >.

REZVANI, N.; BOLDUC, D. L. Monte Carlo Analysis. In: WEXLER, P. (Ed.). **Encyclopedia of Toxicology (Third Edition)**. Oxford: Academic Press, 2014. p.393-396. ISBN 978-0-12-386455-0.

RIESENBERG, D.; GUTHKE, R. High-cell-density cultivation of microorganisms. **Applied Microbiology and Biotechnology**, v. 51, n. 4, p. 422-430, Apr 1999. ISSN 0175-7598. Disponível em: < <Go to ISI>://WOS:000080179400002 >.

ROSILLO-CALLE, F. The Brazilian ethanolchemistry industry (a review). **Biomass**, v. 11, n. 1, p. 19-38, 1986. ISSN 0144-4565. Disponível em: < <http://www.sciencedirect.com/science/article/pii/0144456586900181> >.

ROSILLO-CALLE, F.; HALL, D. O. Brazilian alcohol: Food versus fuel? **Biomass**, v. 12, n. 2, p. 97-128, 1987. ISSN 0144-4565. Disponível em: < <http://www.sciencedirect.com/science/article/pii/0144456587900503> >.

ROWLEY, A. 1973, 1979, 1985 - THE 3 OIL CRISES. **Histoire**, n. 122, p. 18-26, May 1989. ISSN 0182-2411. Disponível em: < <Go to ISI>://WOS:A1989U585900003 >.

SAMANEZ, C. P.; FERREIRA, L. D. R.; NASCIMENTO, C. C. D. Avaliação da opção de troca de combustível no -carro brasileiro flex: um estudo por região geográfica -usando teoria de opções reais e simulação estocástica. **Production**, v. 24, n. 3, p. 628-643, 2014-09 2014. ISSN 0103-6513. Disponível em: < <Go to ISI>://SCIELO:S0103-65132014000300012 >.

SCHADD, M. P. D. et al. Single-player Monte-Carlo tree search for SameGame. **Knowledge-Based Systems**, v. 34, n. 0, p. 3-11, 10// 2012. ISSN 0950-7051. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0950705111001869> >.

SHI, Z. P.; SHIMIZU, K. NEUROFUZZY CONTROL OF BIOREACTOR SYSTEMS WITH PATTERN-RECOGNITION. **Journal of Fermentation and Bioengineering**, v. 74, n. 1, p. 39-45, 1992 1992. ISSN 0922-338X. Disponível em: < <Go to ISI>://WOS:A1992JG67400008 >.

SINGER, T. et al. Empathy for pain involves the affective but not sensory components of pain. **Science**, v. 303, n. 5661, p. 1157-1162, Feb 20 2004. ISSN 0036-8075. Disponível em: < <Go to ISI>://WOS:000189074700036 >.

SKARIN, H. et al. Clostridium botulinum group III: a group with dual identity shaped by plasmids, phages and mobile elements. **Bmc Genomics**, v. 12, Apr 12 2011. ISSN 1471-2164. Disponível em: < <Go to ISI>://WOS:000290794800001 >.

SMINCHISESCU, C.; WELLING, M. Generalized darting Monte Carlo. **Pattern Recognition**, v. 44, n. 10-11, p. 2738-2748, 10// 2011. ISSN 0031-3203. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0031320311000537> >. Acesso em: 2011/11//.

SMITH, L. S. INTRODUCTION TO THE THEORY OF NEURAL COMPUTATION - HERTZ,J, KROGH,A, PALMER,RG. **British Journal of Psychology**, v. 83, p. 425-426, Aug 1992. ISSN 0007-1269. Disponível em: < <Go to ISI>://WOS:A1992JL45300009 >.

SOARES, R. D. P.; SECCHI, A. R. EMSO: A new environment for modelling, simulation and optimisation. In: ANDRZEJ, K. e ILKKA, T. (Ed.). **Computer Aided Chemical Engineering**: Elsevier, v. Volume 14, 2003. p.947-952. ISBN 1570-7946.

SOUSA JR, R.; ALMEIDA, P. I. F. Design of a fuzzy system for the control of a biochemical reactor in fed-batch culture. **Process Biochemistry**, v. 37, n. 5, p. 461-469, 2001. ISSN 1359-5113. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0032959201002394> >.

SUBRAHMANYAM, S. et al. 'Bite-and-Switch' approach using computationally designed molecularly imprinted polymers for sensing of creatinine. **Biosensors & Bioelectronics**, v. 16, n. 9-12, p. 631-637, Dec 2001. ISSN 0956-5663. Disponível em: < <Go to ISI>://WOS:000172169500004 >.

SWAIN, J. E. et al. PROGRESS ON THE PATERNAL BRAIN: THEORY, ANIMAL MODELS, HUMAN BRAIN RESEARCH, AND MENTAL HEALTH IMPLICATIONS. **Infant Mental Health Journal**, v. 35, n. 5, p. 394-408, Sep-Oct 2014. ISSN 0163-9641. Disponível em: <<Go to ISI>://WOS:000342071200002 >.

SWANSON, W. H.; CLIFTON, C. E. GROWTH AND ASSIMILATION IN CULTURES OF SACCHAROMYCES-CEREVISIAE. **Journal of Bacteriology**, v. 56, n. 1, p. 115-124, 1948 1948. ISSN 0021-9193. Disponível em: <<Go to ISI>://WOS:A1948UK03900016 >.

TEMPO, R.; ISHII, H. Monte Carlo and Las Vegas Randomized Algorithms for Systems and Control*: An Introduction. **European Journal of Control**, v. 13, n. 2-3, p. 189-203, // 2007. ISSN 0947-3580. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0947358007708191> >.

TSAO, C. C. et al. Increased estimates of air-pollution emissions from Brazilian sugarcane ethanol. **Nature Climate Change**, v. 2, n. 1, p. 53-57, Jan 2012. ISSN 1758-678X. Disponível em: <<Go to ISI>://WOS:000299495500021 >.

VELOSO, F. A.; VILLELA, A.; GIAMBIAGI, F. Determinantes do "milagre" econômico brasileiro (1968-1973): uma análise empírica. **Revista Brasileira de Economia**, v. 62, n. 2, p. 221-246, 2008-06 2008. ISSN 0034-7140. Disponível em: <<Go to ISI>://SCIELO:S0034-71402008000200006 >.

VUORIMAA, P. FUZZY SELF-ORGANIZING MAP. **Fuzzy Sets and Systems**, v. 66, n. 2, p. 223-231, Sep 9 1994. ISSN 0165-0114. Disponível em: <<Go to ISI>://WOS:A1994PH84300007 >.

WANG, L. X.; MENDEL, J. M. GENERATING FUZZY RULES BY LEARNING FROM EXAMPLES. **Ieee Transactions on Systems Man and Cybernetics**, v. 22, n. 6, p. 1414-1427, Nov-Dec 1992. ISSN 0018-9472. Disponível em: <<Go to ISI>://WOS:A1992KN76300013 >.

WEI, M. Q. et al. Bacterial targeted tumour therapy-dawn of a new era. **Cancer Letters**, v. 259, n. 1, p. 16-27, Jan 18 2008. ISSN 0304-3835. Disponível em: <<Go to ISI>://WOS:000253023000002 >.

YAO, X.; LIU, Y.; LIN, G. M. Evolutionary programming made faster. **Ieee Transactions on Evolutionary Computation**, v. 3, n. 2, p. 82-102, Jul 1999. ISSN 1089-778X. Disponível em: <<Go to ISI>://WOS:000082713300002 >.

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, n. 3, p. 338-353, 1965. ISSN 0019-9958. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S001999586590241X> >.

_____. The role of fuzzy logic in the management of uncertainty in expert systems. **Fuzzy Sets and Systems**, v. 11, n. 1–3, p. 197-198, 1983. ISSN 0165-0114. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0165011483800815> >.

ZANETTI, S. S. et al. Estimação da evapotranspiração de referência no estado do Rio de Janeiro usando redes neurais artificiais. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 12, n. 2, p. 174-180, 2008-04 2008. ISSN 1807-1929. Disponível em: <Go to ISI>://SCIELO:S1415-43662008000200010 >.

ZHENG, S. et al. **Task scheduling using parallel genetic simulated annealing algorithm**. 2006. 46-50 ISBN 1-4244-0317-0. Disponível em: <<Go to ISI>://WOS:000241425100009 >.

ZIEGLER, M. O. P. et al. Conformational changes and reaction of clostridial glycosylating toxins. **Journal of Molecular Biology**, v. 377, n. 5, p. 1346-1356, Apr 11 2008. ISSN 0022-2836. Disponível em: <<Go to ISI>://WOS:000254979500006 >.

ZIMMERMANN, H. J. Applications of fuzzy set theory to mathematical programming. **Information Sciences**, v. 36, n. 1–2, p. 29-58, 1985/8// 1985. ISSN 0020-0255. Disponível em: < <http://www.sciencedirect.com/science/article/pii/0020025585900258> >.