UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA QUÍMICA
DEPARTAMENTO DE PROCESSOS QUÍMICOS

OTIMIZAÇÃO DE BIOPROCESSOS: AVALIAÇÃO DE DESEMPENHO DAS ABORDAGENS DETERMINÍSTICA E POR ALGORITMOS GENÉTICOS

Elmer Alberto Ccopa Rivera

Mestre em Engenharia de Alimentos - UNICAMP

Prof. Dr. Rubens Maciel Filho

Orientador

Prof[a]. Dr[a]. Aline Carvalho da Costa

Co-orientadora

Tese apresentada à Faculdade de Engenharia Química como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Química

Campinas – São Paulo

Dezembro de 2006

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

Ccopa Rivera, Elmer Alberto
C319o Otimização de bioprocessos: avaliação de desempenho das abordagens determinística e por algoritmos genéticos / Elmer Alberto Ccopa Rivera.-- Campinas, SP: [s.n.], 2006.

Orientador: Rubens Maciel Filho; Aline Carvalho da Costa.
Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Álcool. 2. Biotecnologia. 3. Modelamento matemático. 4. Otimização. I. Maciel Filho, Rubens. II. Costa, Aline Carvalho da. III. Universidade Estadual de Campinas. Faculdade de Engenharia Química. IV. Título.

Título em Inglês: Bioprocess optimization: evaluation of performance using deterministic and genetic algorithm approaches.
Palavras-chave em Inglês: Bioprocess, Bioethanol, Mathematical modeling, Optimization techniques.
Área de concentração: Desenvolvimento de Processos Químicos.
Titulação: Doutor em Engenharia Química
Banca examinadora: Beatriz Vahan Kilikian, Francisco Maugeri Filho, Carlos Eduardo Vaz Rossell e Daniel Ibraim Pires Atala.
Data da defesa: 13/12/2006
Programa de Pós-Graduação: Engenharia Química

Tese de Doutorado defendida por Elmer Alberto Ccopa Rivera e aprovada em 13 de Dezembro de 2006 pela banca examinadora constituída pelos doutores:

Prof. Dr. Rubens Maciel Filho – Orientador

Profa. Dra. Beatriz Vahan Kilikian

Prof. Dr. Francisco Maugeri Filho

Dr. Carlos Eduardo Vaz Rossell

Dr. Daniel Ibraim Pires Atala

Este exemplar corresponde a versão final da Tese de Doutorado em Engenharia Química.

Prof. Dr. Rubens Maciel Filho – Orientador

Dedicatória

A mis padres Santiago y María, y a mis hermanas Luz y Rosa, quienes siempre me han brindado su amor y apoyo incondicional.

Agradecimentos

À Deus pela vida.

Ao Prof. Rubens Maciel Filho, pela orientação amiga e entusiástica. Agradeço seu profissionalismo, sua confiança e apoio.

À Profª. Aline Carvalho da Costa, minha co-orientadora, por compartilhar comigo sua experiência como pesquisadora e pela inestimável colaboração na qualidade final desta tese.

Aos membros da banca examinadora pela participação e sugestões.

À UNICAMP pela infra-estrutura que proporcionou e pelo ambiente agradável do campus.

À FAPESP, que financiou este projeto de pesquisa e possibilitou a realização do estágio na Universidade Politécnica de Catalunya, Espanha.

Aos amigos e colegas do LOPCA, pelo ambiente agradável de trabalho e pela disposição em ajudar no surgimento de dúvidas e problemas.

À minha família, a quem este trabalho é dedicado, como um pequeno gesto de agradecimento.

Minha gratidão a todos os amigos que estiveram comigo neste tempo, sempre presentes nos momentos alegres ou difíceis.

Sumário

Lista de Tabelas

CAPÍTULO 8

APÊNDICE A

APÊNDICE B

Lista de Figuras

CAPÍTULO 2

CAPÍTULO 3

CAPÍTULO 4

CAPÍTULO 5

CAPÍTULO 6

CAPÍTULO 7

CAPÍTULO 8

APÊNDICE A

RESUMO

Esta tese aborda aplicações de modelagem, metodologias sistemáticas e confiáveis de otimização, além de ferramentas computacionais. Estas áreas podem capturar características quantitativas essenciais dos bioprocessos e possuem grande impacto nas práticas modernas da melhoria de desempenho dos processos petroquímicos e mecânicos. É evidente que o mesmo deve ocorrer com os bioprocessos, uma área em que ainda poucos estudos e desenvolvimentos foram feitos. Porém, espera-se que os bioprocessos construídos em cima dos métodos que consideram a análise do problema, codificados como uma ferramenta matemática apropriada para o uso das ferramentas da engenharia de processos assistida por computador, convertam resultados numéricos em soluções de engenharia úteis para produzir, finalmente, uma melhor compreensão de tais sistemas. Diferentes tipos de metodologias, apropriadas para aplicações em computador, foram estudadas com esta finalidade. Os métodos propostos foram aplicados ao desenvolvimento e otimização da produção de bioetanol. A parte central desta tese consiste de artigos científicos (Capítulos 3 - 8). O Capítulo 3 concentra-se nas técnicas de otimização para estimar os parâmetros do modelo cinético do processo de fermentação alcoólica em batelada para a produção de etanol usando *Saccharomyces cerevisiae*. O Algoritmo Quasi-Newton (QN) e o Algoritmo Genético de código real (RGA) foram usados para resolver o problema de estimação e encontrar uma solução ótima. O Capítulo 4 fornece um modelo híbrido adaptativo que considera o efeito da temperatura na cinética da fermentação alcoólica, representada por redes neurais. O potencial de um RGA e de um algoritmo genético de código binário (BGA), para re-estimar os parâmetros da rede neural, foram avaliados. O desempenho destes algoritmos foi comparado com o algoritmo Quasi-Newton (QN). No Capítulo 5 foi avaliado um procedimento para o desenvolvimento de um modelo matemático robusto para um processo de fermentação alcoólica industrial. O modelo proposto é um modelo híbrido neural, que combina equações de balanço de massa e energia com redes do tipo Functional Link Network para descrever a cinética. Estas redes apresentaram uma boa capacidade de aproximação não-linear, embora a estimação dos seus pesos

seja linear. O Capítulo 6 propõe uma nova metodologia para estimação de parâmetros cinéticos. A primeira etapa consiste na obtenção de valores iniciais para todos os parâmetros do modelo. Conseqüentemente, um RGA foi usado para a estimação simultânea dos parâmetros. Na terceira etapa, os parâmetros mais significativos foram identificados usando o planejamento Placket-Burman (PB). Finalmente, os parâmetros mais significativos foram otimizados usando o algoritmo QN, que converge para o ótimo muito mais rápido que o RGA. No Capítulo 7 a modelagem de processos biotecnológicos foi estudada com foco no desenvolvimento das metodologias que podem ser usadas sempre que uma re-estimação dos parâmetros seja necessária. O desempenho de um modelo híbrido neural e de um modelo fenomenológico, ambos considerando o efeito da temperatura na cinética, foi avaliado não somente em termos de sua exatidão em descrever os dados experimentais, mas principalmente pelas dificuldades envolvidas na adaptação de seus parâmetros. No Capítulo 8, o modelo não-linear de um processo de fermentação alcoólica extrativa, representado por Redes Neurais Multilayer Perceptron foi otimizado, usando RGA e BGA, para encontrar condições ótimas de operação. A fim de verificar a validade do modelo neural, os resultados foram comparados à otimização de um modelo determinístico, onde os parâmetros cinéticos foram determinados experimentalmente como funções da temperatura. Adicionalmente, o Apêndice A apresenta informações suplementares a respeito da verificação do comportamento não-linear das variáveis do processo extrativo de fermentação alcoólica. Com a finalidade de comparar os resultados de otimização deste processo empregando algoritmos genéticos, o Apêndice B apresenta os resultados de otimização usando Programação Quadrática Sucessiva.

Palavras-chave

Bioprocessos, bioetanol, modelagem matemática, técnicas de otimização

ABSTRACT

This thesis addresses the issues of modeling, systematic and reliable optimization methodologies, analysis and computational tools. These fields can capture essential quantitative features of bioprocesses and have a great impact on the modern practices of petrochemical and mechanical process performance improvement. It is evident that the same should occur with bioprocesses, a field in which very little studies and developments has been made. However, it is expected that the bioprocesses built upon methods that consider the problem analysis, codified into a mathematical tool appropriate for the use of computer aided process engineering tools, will transform numerical results into useful engineering solution to finally produce a better understanding of the bioprocess. Different types of models, suitable for computer-aided applications, have been studied for this purpose. The proposed methods are applied to the development and optimization of bioethanol production. The main part of the thesis consists of journal papers (Chapters 3 - 8). Chapter 3 focuses on the optimization techniques to estimate the kinetic model parameters of batch fermentation process for ethanol production using *Saccharomyces cerevisiae*. The potential of Quasi-Newton (QN) and Real-Coded Genetic Algorithm (RGA) to solve the estimation problem is considered to find out the optimal solution. Chapter 4 provides an adaptive hybrid model that considers the effect of temperature on the kinetics of alcoholic fermentation, represented by neural networks. The potential of a RGA and a binary-coded genetic algorithm (BGA) to re-estimate the parameters of the neural network is evaluated. These algorithms were compared to the quasi-newton algorithm (QN) in terms of performance of the hybrid model. In the Chapter 5 a procedure for the development of a robust mathematical model for an industrial alcoholic fermentation process was evaluated. The proposed model is a hybrid neural model, which combines mass and energy balance equations with Functional Link Networks to describe the kinetics. These networks have been shown to have a good non linear approximation capability, although the estimation of its weights is linear. Chapter 6 proposes a new methodology to estimation of kinetic parameters. The first step is to obtain initial values for all parameters in the model and then a

RGA is used to calculate the parameters in the model. The third step is to identify the most significant of the parameters using Plackett and Burman design (PB) and finally the most significant are optimized using a QN algorithm, which converges much more quickly than RGA to the optimal. In the Chapter 7 the modeling of biotechnological processes is studied with focus on developing methodologies that can be used always that a re-estimation of parameters is necessary. The performance of a hybrid neural model and a first-principles model, both considering the effect of temperature on the kinetics, are evaluated not only by their accuracy in describing experimental data, but mainly by the difficulties involved in the adaptation of their parameters. In the Chapter 8 the non-linear model of an extractive alcoholic fermentation process, represented by neural networks, is optimized using RGA and BGA to determine the optimal operational conditions. In order to check the validity of the computational modeling, the results were compared to the optimization of a deterministic model, whose kinetic parameters were experimentally determined as functions of the temperature. Furthermore, Appendix A provides additional information regarding the verification of the non-linear behavior of the variables of the extractive alcoholic fermentation process. With the purpose to compare the optimization performances of this process by genetic algorithms, Appendix B presents the results of optimization using successive quadratic programming.

Keywords

Bioprocess, bioethanol, mathematical modeling, optimization techniques.

CAPÍTULO 1. INTRODUÇÃO

Prefácio

Este capítulo apresenta as motivações, objetivos e contribuições vinculadas ao desenvolvimento desta tese. Inicialmente, busca-se ressaltar a importância dos bioprocessos e as vantagens do uso de fontes renováveis de energia. Em seguida, parte-se para a apresentação de diversas abordagens sobre diferentes aspectos da modelagem e otimização de processos de fermentação alcoólica. O capítulo é, posteriormente, finalizado com a descrição da tese, onde o conteúdo de cada capítulo é apresentado resumidamente.

1.1 Introdução Geral e Objetivos

Nos últimos anos a utilização de bioprocessos vem crescendo de maneira significativa com a obtenção de produtos já tradicionalmente obtidos por via biotecnológica de forma mais eficiente à procura de rotas de outros produtos requerendo tecnologias cada vez mais sofisticadas e grande número de pesquisas (**Komives and Parker, 2003**). O assunto é considerado como uma das áreas de ponta de pesquisa em Engenharia Química. Portanto, é uma nova etapa para o desenvolvimento e implementação de técnicas de otimização, controle avançado e operação assistida por computador.

O interesse em fontes renováveis tende a crescer no futuro conforme as fontes não renováveis, como os combustíveis fósseis vão tornando-se escassos ou economicamente inviáveis. As constantes altas no preço do petróleo e as perspectivas de extração cada vez mais complexas e caras têm dado um incentivo a mais para pesquisas em fontes alternativas, buscando a obtenção de produtos químicos. É necessário, então, procurar substitutos para o petróleo como matéria-prima na indústria química e também como fonte de energia, principalmente quando se leva em conta que o Brasil importa petróleo e, tendo dívida externa, gasta não só na importação, mas com pagamento de juros. Para o Brasil, um substituto em potencial para o petróleo é a biomassa, já que o país apresenta

custos internacionalmente competitivos desta matéria-prima (**Macedo, 2003**), destacando-se a cana-de-açúcar.

No Brasil, a produção de álcool foi iniciada em larga escala em 1975 a partir dos incentivos do Programa Nacional do Álcool (Pró-Álcool) do governo federal. Entre 1975 e 1985 a produção de cana-de-açúcar quadruplicou e o álcool se tornou um combustível muito importante no país. Posteriormente, houve, no Brasil, uma mudança de prioridade e a produção da frota automotiva movida a álcool passou de cerca de 95% na década de 1980 (auge do programa Pró-Álcool) para menos de 5%. Vários fatores contribuíram para que a conveniência do modelo adotado no Proálcool fosse posta em discussão, entre eles a política de subsídios mesmo aos produtores mais ineficientes. Ainda assim o Pro-álcool tem sido citado como o mais importante programa de energia a partir de biomassa do mundo **(Goldemberg et al., 2004)**. As vantagens do uso de etanol como combustível, seja na forma hidratada em substituição à gasolina, seja misturado à gasolina como etanol anidro, são inúmeras. Entre elas estão a redução da poluição em centros urbanos, a eliminação total dos aditivos à base de chumbo, redução de SO_x e particulados, redução de 40 a 70% no CO, emissões de VOCs (Volatile Organic Compounds) com menor toxicidade e reatividade.

O crescente interesse pelo álcool no mercado externo, além do retorno da utilização do álcool em veículos bi-combustíveis, que possibilitam a queima de álcool e gasolina em qualquer proporção, mostram que as perspectivas para o mercado doméstico e internacional do álcool são muito boas. O consumo de álcool nos carros bi-combustíveis avança a cada mês e em maio de 2004 o álcool foi lançado como commodity na bolsa de valores de Nova Iorque. Com isso o álcool (etanol, referenciado no exterior como bioetanol) passa a ter uma larga aceitação no mercado internacional, principalmente quando se leva em conta que leis ambientais obrigam dezenas de países a adicionar índices que variam de 10% a 30% de álcool à gasolina. Por outro lado, as várias reuniões mundiais têm deixado claro que políticas para uso de energia renovável são essenciais para o desenvolvimento sustentável. Com a adesão da Rússia ao protocolo de Quioto em novembro de 2004, surge com força total o comércio de créditos de carbono, que

está previsto no mecanismo de flexibilização do protocolo e permite que cada tonelada de gás carbônico que deixa de ser emitida por um país que não precisa reduzir suas emissões seja negociada no mercado mundial. O Brasil pode vir a ter uma grande oferta de créditos de carbono com diversos projetos, entre eles projetos relacionados à produção e uso de biocombustíveis **(Macedo, 2003)**.

Atualmente, há muitos problemas menores relacionados ao processo industrial de fermentação alcoólica, porém com impacto no valor final do produto. Entre eles a falta de robustez da fermentação na presença das flutuações nas condições de operação, levando a mudanças no comportamento cinético com impacto no rendimento, produtividade e conversão **(Macedo, 2003)**. Estas mudanças são bastante comuns em plantas de fermentação alcoólica; elas ocorrem não só devido a variações na qualidade da matéria-prima, mas também devido a variações de linhagens de leveduras dominantes no processo. Uma das fontes de variação da matéria-prima é a operação das fábricas de açúcar, que altera as quantidades de melaço e caldo de cana tratado utilizados na composição do meio de alimentação aos fermentadores. Além disso, o melaço sofre variações de composição nas diferentes safras. As alterações nas populações de levedura são comuns devido à alta carga microbiana contida na matéria-prima **(Andrietta et al., 1999)**.

Para tratar com esta falta de robustez é importante realizar uma monitoração eficiente do processo de fermentação alcoólica, se possível, em tempo real, de algumas variáveis e inferênciar em outras que são difíceis ou caras de serem medidas, fazendo-se ajustes nas condições de operação e controle para manter o processo operando nas condições ideais.

Neste contexto, um dos maiores desafios na otimização e controle de bioprocessos é a determinação de modelos confiáveis do sistema que apresentam soluções compatíveis com as suas aplicações. Este é um ponto crítico nas aplicações em tempo real e esta dificuldade está associada à natureza complexa do metabolismo microbiano e à natureza altamente não-linear de sua cinética. A dificuldade é maior quando se trata de processos contínuos ou em batelada, já que, sendo estes processos transientes, suas variáveis apresentam, normalmente,

grandes mudanças com o tempo. Embora modelos advindos de leis físico-químicas proporcionem bastante conhecimento sobre o processo, a experiência prática mostra que eles somente são válidos para as condições específicas nas quais foram determinados. Quando há variações nas condições de operação, devem-se reestimar os parâmetros do modelo. A reestimação freqüente, no entanto, é usualmente difícil e demorada e, se tiver que ser feita em tempo real, o problema de robustez dos procedimentos deve ser considerado. Dificuldades surgem não só na complexidade matemática dos modelos e conseqüente otimização, como também em se reproduzir os fenômenos.

Dentro do contexto exposto acima, este trabalho de tese aborda, em geral, o desenvolvimento e comparação de desempenho de metodologias de otimização desenvolvidas, tanto para o estudo dos processos de fermentação alcoólica em batelada como contínuos em regime dinâmico. Além disso, foram desenvolvidos procedimentos para determinação de modelos matemáticos robustos frente a alterações nas condições operacionais, principalmente flutuações na qualidade da matéria-prima. O objetivo foi obter modelos que possam ser usados como simuladores, tornando mais simples não só o desenvolvimento e implementação de novos controladores e otimizadores, mas também a re-sintonização de controladores existentes e a determinação de novas condições ótimas de operação quando ocorrerem mudanças operacionais.

1.2 Organização da Tese

Esta tese consta de dez capítulos, sendo que os **Capítulos 1 e 2**, referentes à **Introdução** e **Revisão Bibliográfica**, respectivamente, seguiram as normas para redação de tese da Faculdade de Engenharia Química da Universidade Estadual de Campinas e destinou-se a inserir ao leitor o tema central desta tese, dando um embasamento teórico aos estudos desenvolvidos, os quais são apresentados nos Capítulos 3 a 8 na forma de artigos científicos, que individualmente, abrangem cada um dos objetivos estabelecidos, podendo ser lidos independentemente dos demais.

A fim de manter a característica original dos artigos (redigidos conforme as normas da revista), estes capítulos não foram traduzidos quando inseridos na tese.

O artigo relatado no **Capítulo 3**, intitulado "Evaluation of optimization techniques for parameter estimation: Application to ethanol fermentation considering the effect of temperature" (publicado na revista *Process Biochemistry*, 2006), avalia técnicas de otimização para estimar os parâmetros do modelo cinético do processo de fermentação alcoólica em batelada para a produção de etanol usando *Saccharomyces cerevisiae*. Observações experimentais em cinco temperaturas diferentes (28°C, 31°C, 34°C, 37°C e 40°C) foram usadas para formular o problema de estimação de parâmetros. O Algoritmo Quasi-Newton (QN) e o Algoritmo Genético de código real (RGA) foram usados para resolver o problema de estimação e encontrar uma solução ótima. Posteriormente, os parâmetros otimizados (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px}) foram caracterizados por funções de correlação que adotam dependência da temperatura.

No capítulo 3, devem ser observadas as correções conforme o Apêndice C.

O **Capítulo 4**, referente ao artigo intitulado "An adaptive hybrid model for bioethanol production by continuous extractive fermentation" (publicado no livro *European Symposium on Computer Aided Process Engineering*, Vol. 21, 2006), traz os resultados de um modelo híbrido adaptativo para um processo de fermentação alcoólica extrativa. No modelo híbrido, equações diferenciais baseadas em balanço de massa e energia foram combinadas com redes neurais do tipo Multilayer Perceptron (MLPNN) que descrevem a cinética do processo. Inicialmente, realiza-se um treinamento off-line para escolher a arquitetura da rede neural. Posteriormente, os pesos da rede foram re-estimados por uma aproximação integral usando algoritmos evolucionários na forma de um algoritmo genético de código real e um algoritmo genético de código binário. O desempenho destes algoritmos foi comparado com o algoritmo Quasi-Newton.

Continuando com a mesma abordagem de modelos híbridos descrita no Capítulo 4, o artigo exposto no **Capítulo 5**, intitulado "Hybrid neural network model of an industrial ethanol fermentation process considering the effect of temperature" (aceito para publicação na revista *Applied Biochemistry and*

Biotechnology, 2007), propõe um procedimento para o desenvolvimento de um modelo matemático robusto para um processo de fermentação alcoólica industrial. O modelo proposto é um modelo híbrido neural, que combina equações de balanço de massa e energia com redes do tipo Functional Link Network para descrever a cinética. Estas redes apresentaram uma boa capacidade de aproximação não-linear, embora a estimação dos seus pesos seja linear. O modelo proposto considera o efeito da temperatura na cinética e tem os pesos da rede neural re-estimados sempre que ocorram mudanças em circunstâncias operacionais.

O artigo descrito no **Capítulo 6**, intitulado "Development o a methodology for the estimation of kinetic parameters in a structured model for ethanol production" (submetido à revista *Brazilian Journal of Chemical Engineering*, 2006), apresenta uma metodologia de estimação de parâmetros cinéticos usando diferentes técnicas de otimização. Este trabalho está ligado ao escopo desta tese pela importância do aperfeiçoamento dos métodos de estimação de parâmetros, devido principalmente às não-linearidades, ao grande número de parâmetros e às interações entre eles. O procedimento consiste em quatro etapas. Primeiro, os valores iniciais foram encontrados da literatura. Subseqüentemente, o potencial de busca global do algoritmo genético de codificação real (RGA) foi aplicado para a estimação simultânea dos parâmetros. Na terceira etapa, os parâmetros mais significativos foram identificados usando o planejamento Placket-Burman (PB). Finalmente, o algoritmo quasi-Newton (QN) foi usado para a otimização dos parâmetros mais significativos, perto da região ótima global, já que os valores iniciais já tinham sido determinados pelo algoritmo de busca global RGA.

O artigo apresentado no **Capítulo 7**, intitulado "Development of adaptive modeling techniques to describe the temperature dependent kinetics of biotechnological process" (aceito para publicação na revista *Biochemical Engineering Journal*, 2007), aperfeiçoa a abordagem híbrida apresentada nos capítulos 4 e 5, para situações com resultados experimentais. Assim, neste capítulo foi estudada a modelagem de processos biotecnológicos com foco no desenvolvimento das metodologias que podem ser usadas sempre que uma re-estimação dos parâmetros seja necessária. O processo de fermentação alcoólica foi

usado como um estudo de caso. Dois modelos foram propostos considerando o efeito da temperatura na cinética: um modelo determinístico e um modelo híbrido, o qual combina balanços de massa com redes neurais do tipo Multilayer Perceptron. As dificuldades e os desempenhos envolvidos na adaptação dos parâmetros para os dois modelos foram avaliados e comparados.

O **Capítulo 8**, referente ao artigo intitulado "Ethyl alcohol production optimization by coupling genetic algorithm and multilayer perceptron neural network" (publicado na revista *Applied Biochemistry and Biotechnology*, 2006), de uma maneira geral, encerra os objetivos previstos para este trabalho de tese, mostrando que a inteligência artificial é uma abordagem muito promissora nas áreas de modelagem e otimização de bioprocessos em aplicações de engenharia, uma vez que apresenta diversas vantagens sobre as topologias comumente usadas na literatura. Algoritmos genéticos (GA) e Redes Neurais do tipo Multilayer Perceptron (Multilayer Perceptron Neural Network - MLPNN) foram combinados de forma eficiente para reduzir a complexidade do problema de otimização. Destaca-se neste trabalho, um método de identificação baseado em MLPNN e planejamento de experimentos. O modelo não-linear de um processo de fermentação alcoólica extrativa, representado por Redes Neurais Multilayer Perceptron foi otimizado, usando um Algoritmo Genético de Código Real e um Algoritmo Genético de Código Binário, para encontrar condições ótimas de operação de modo que a conversão seja maximizada aos limites definidos. A fim de verificar a validade do modelo MLPNN, os resultados foram comparados à otimização de um modelo determinístico, onde os parâmetros cinéticos foram determinados experimentalmente como funções da temperatura.

O **Capítulo 9**, referente às **Considerações Finais,** apresena os principais resultados obtidos nos artigos apresentados nesta tese. Em seguida, articulam-se propostas de perspectivas futuras para a área de modelagem e otimização de bioprocessos.

O **Apêndice A** apresenta informações adicionais a respeito da verificação do comportamento não-linear das variáveis do processo extrativo de fermentação alcoólica contínua. O mapeamento da dinâmica do processo é feito utilizando

técnicas de planejamento fatorial completo de dois níveis e a metodologia de Plackett-Burman.

Com o intuito de comparar os resultados de otimização do processo de fermentação alcoólica extrativa utilizando algoritmos genéticos, expostos no capítulo 8, o **Apêndice B** apresenta os resultados de otimização dos modelos não-lineares do processo (um modelo neural e um modelo determinístico) otimizados utilizando um algoritmo de programação quadrática sucessiva (SQP). Finalmente o **Apêndice C** apresenta as correções observadas no artigo que corresponde ao capítulo 3.

Referências Bibliográficas

1. Andrietta, S.R., Migliari, P.C. e Andrietta, M.G.S., Classificação das cepas de leveduras de processos industriais de fermentação alcoólica utilizando capacidade fermentativa. *STAB – Açúcar Álcool e Subprodutos*, 1999, **17**(5) 54-59.
2. Goldemberg, J., Coelho, S.T., Nastari, P. M. e Lucon, O., Ethanol learning curve-the Brazilian experience. *Biomass and Bioenergy*, 2004, **26** 301-304.
3. Komives, C. e Parker, R.S., Bioreactor state estimation and control. *Current Opinion in Biotechnology*, 2003, **14** 1-7.
4. Macedo, Isaias (coordenador). Estado da arte e tendências das tecnologias para energia. Brasil - Ministério da Ciência e da Tecnologia. CTEnerg Secretaria Técnica do Fundo Setorial de Energia. CGEE Centro de gestão e Estudos estratégicos. Brasília, enero 2003.

CAPÍTULO 2. REVISÃO BIBLIOGRÁFICA

Prefácio

Nesta seção, é apresentada uma visão geral dos sistemas de fermentação alcoólica e a importância do potencial de exploração da cana-de-açúcar como uma matéria-prima. Além disso, é destacada a abordagem de tópicos específicos da teoria de sistemas computacionais bio-inspirados, mais especificamente a Redes Neurais Artificiais (RNA) e Algoritmos Genéticos (AG), para otimização de bioprocessos. Adicionalmente, são apresentadas referências bibliográficas para um maior aprofundamento.

2.1 Fermentação alcoólica

2.1.1 Matéria-prima

A matéria-prima utilizada na produção de bioetanol, do ponto de vista de fermentação, pode ser qualquer material contendo açúcares, em especial hexoses (glicose ou frutose) ou sacarídeos feitos a partir de estes açúcares.

Para o Brasil as matérias-primas, de importância econômica imediata para a produção de bioetanol, são os melaços e a cana-de-açúcar. Atualmente, cerca de 50% da cana cultivada é usada na produção de bioetanol (**Macedo e Cortez, 2005**).

A energia é um assunto extremamente complicado, porém, o Brasil está analisando com cuidado o programa do bioetanol como combustível. Uma dos assuntos cruciais é que, para que uma economia baseada em bioetanol seja viável, a energia do combustível total utilizado em produzir o bioetanol não deve exceder o índice de energia do produto. É, entretanto, muito difícil estabelecer se há tal equilíbrio de energia. Muitos processos são envolvidos, incluindo fertilização, cultivo, colheita, transporte, fermentação, destilação e distribuição. É necessário considerar, também, o combustível usado na construção de fazendas e equipamentos para plantas de bioetanol. Os cálculos dependem do que é incluído ou excluído, e das metodologias empregados na estimação dos valores energéticos

dos co-produtos, além da consideração de usos alternativos das matérias-primas (**Grad, 2006**).

Hoje em dia, um dos principais objetivos é o aproveitamento integral da biomassa produzindo açúcares e álcool, como intermediários; produzindo, subseqüentemente, por rota química ou bioquímica inovadora, derivados do bioetanol de alto valor agregado (algunos deles listados na Figura 2.1).

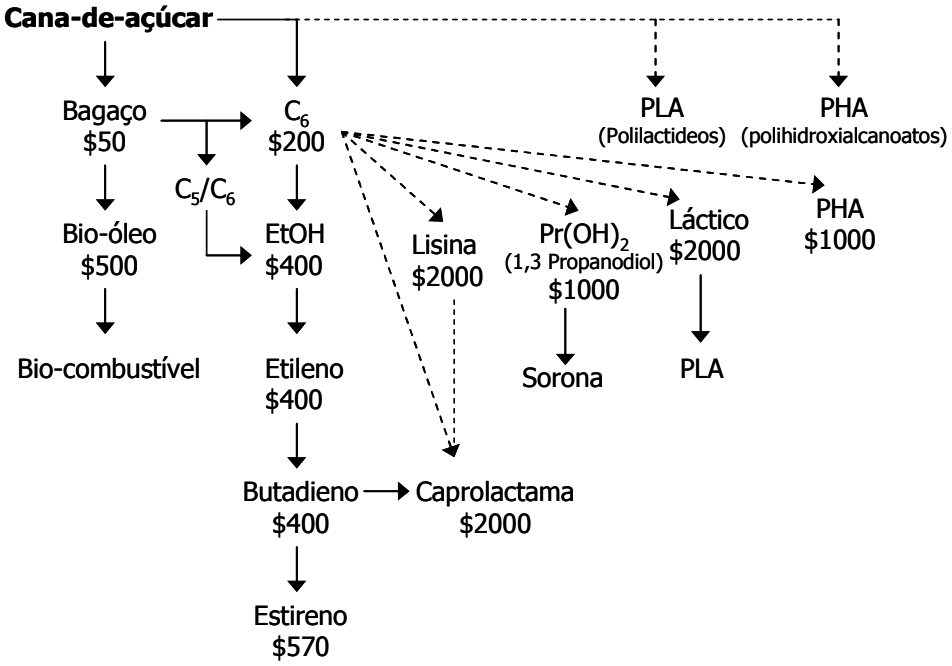


Figura 2.1. Rotas de obtenção de diversos produtos químicos usando cana-de-açúcar como matéria-prima. Os preços estão em US$ por tonelada de composto

Cabe destacar que a síntese fermentativa de produtos de alto valor agregado a partir de biomassa renovável torna-se viável somente quando os substratos usados são baratos e facilmente disponíveis, como é o caso do bagaço de cana no Brasil, que já se encontra nas usinas, as quais tem facilidades de administração, de logística para distribuição interna e externa, de emissão de efluentes, de energia, de processos e que em outros países ainda é preciso desenvolver (**Rossell, 2006**).

2.1.2 Sistemas de fermentação alcoólica

Existem várias maneiras de se conduzir uma fermentação alcoólica. Distingui-se-ão aqui os processos descontínuos e contínuos. O conhecimento do processo

descontínuo simples significa o conhecimento básico da cinética do processo, sendo, portanto, de extremo interesse. Por outro lado, o processo contínuo possui evidentes vantagens sobre o processo intermitente, entre as quais: menor tempo de fermentação, o que permite para uma determinada instalação uma produção maior do que a fermentação descontínua; possibilidade de inserção de controle automático e, portanto, menor emprego de mão-de-obra; e uniformidade do produto final, o que se deve ao regime estacionário.

No Brasil, houve uma resistência inicial para introduzir os processos contínuos nas destilarias. As dificuldades na época foram atribuídas às dificuldades técnicas para manter o funcionamento normal assepticamente e sem interrupções por um longo período de tempo, e ao desconhecimento parcial e às vezes total do comportamento cinético dos processos fermentativos contínuos. Somente mais tarde, ocorreu uma retomada de esforços, devido principalmente a que o processo de fermentação passou por uma transformação muito grande, guiada pela necessidade de maior produtividade e, mais tarde, pelo aumento da eficiência. Os principais avanços tecnológicos foram:

1980 – melhorias na engenharia: centrifugação, resfriamento;

1982 – controle microbiológico;

1985 – primeiro processo de fermentação contínua em larga escala;

1989 – dinâmica de população de leveduras $\rightarrow$ identificação de DNA;

1992 – leveduras selecionadas;

1994 – segunda geração da fermentação contínua (10^6/ l/dia).

Nos últimos anos, no Brasil, estão sendo estudadas novas tecnologias em escala de bancada, com o objetivo de otimizar a operação da fermentação industrial existente e ainda desenvolver novos projetos para que o custo de produção do álcool seja sempre mínimo. Entre as muitas abordagens, pode-se mencionar:

2003 – Sistema de fermentação alcoólica contínua utilizando reatores tipo torre e leveduras auto-imobilizáveis (**Ferreira, 2003**).

2004 – Processo fermentativo extrativo de produção de etanol (**Atala, 2004**).

A Figura 2.2 apresenta alguns resultados dos ganhos de produtividade obtidos na fermentação (figura gerada a partir de dados de **Macedo e Cortez, 2005**).

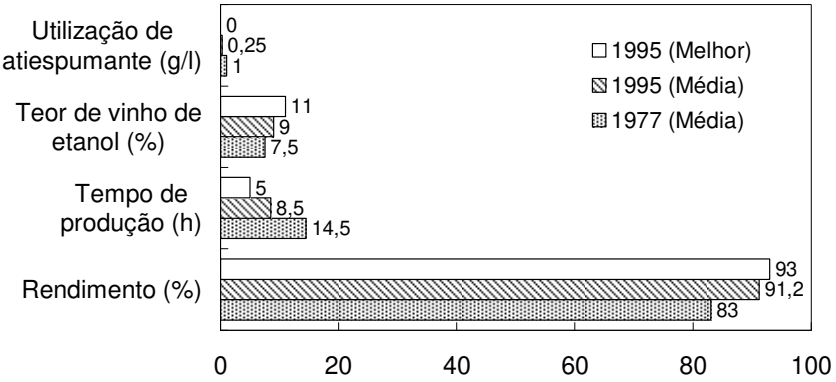


Figura 2.2. Ganhos de produtividade na fermentação

Atualmente, a corrida pelo desenvolvimento da tecnologia de produção de bietanol é mundial. Considerando que apenas um terço da biomassa contida na planta da cana é, atualmente, aproveitada para a produção de açúcar e, eventualmente, bioetanol. O primeiro grande desafio a ser enfrentado nesse assunto é como aproveitar os dois terços restantes da planta. A mudança mais importante no futuro próximo é transformar em açúcar, ou em bioetanol, a celulose e a hemicelulose contidas no bagaço da cana e, também, nas pontas e na palha que, hoje, são queimadas ou deixadas no campo. Assim, é muito importante que haja um planejamento apropriado da interface entre a produção agrícola (taxas de corte e a seleção de espécies adequadas) e a indústria (processamento rápido e eficiente) para evitar a deterioração e as perdas. A composição media da cana-de-açúcar para o processamento, quando entra na indústria é de 8% a 14% de fibras, 12% a 23% de sólidos solúveis e de 65% a 75% de água.

Em termos de tecnologia, a indústria de cana-de-açúcar brasileira ocupa uma posição privilegiada, uma vez que o Brasil poderia muito bem ser considerado líder mundial na tecnologia de produção de álcool, tanto em pequena como em larga escala industrial (**Macedo e Cortez, 2005**).

O fluxograma básico do processo de produção de açúcar e bioetanol é apresentado na Figura 2.3.

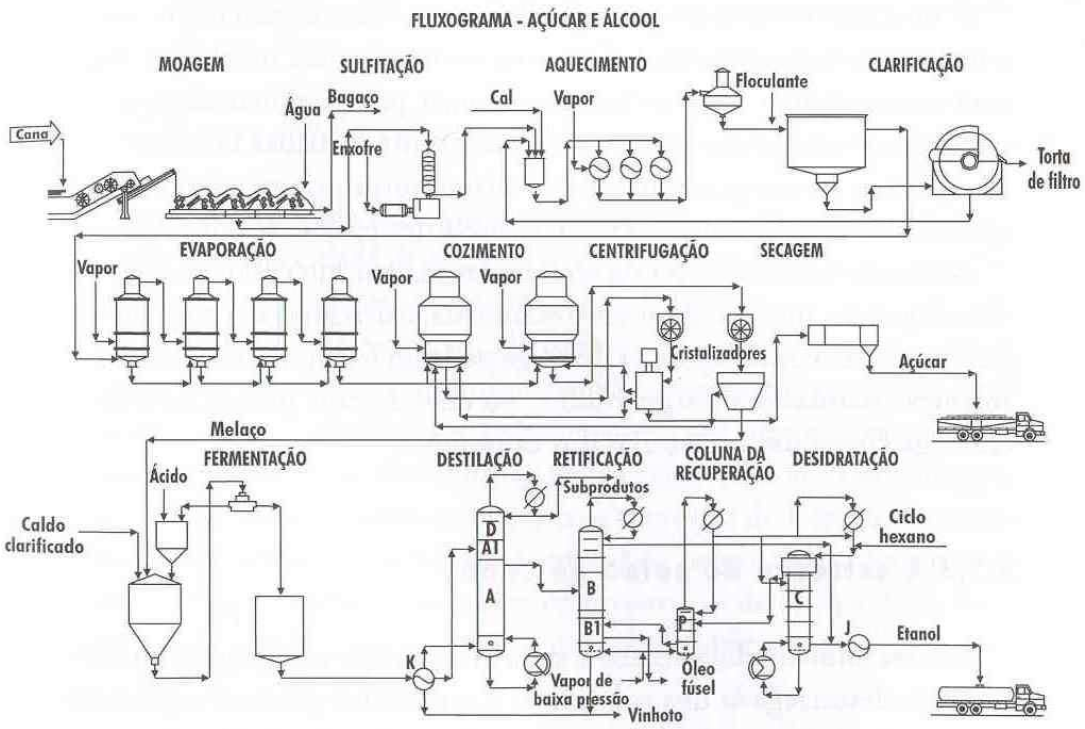


Figura 2.3. Diagrama do fluxo de processamento do açúcar e bioetanol (Macedo e Cortez, 2005)

2.2 Teoria de sistemas computacionais bio-inspirados para modelagem e otimização de bioprocessos

Idéias extraídas de sistemas naturais já vêm sendo utilizadas com muito sucesso para o desenvolvimento de ferramentas tecnológicas capazes de resolver problemas de complexidade elevada. Atualmente, pode-se afirmar que a inteligência computacional é uma área de atuação derivada da ciência da computação, que procura inspiração no comportamento do ser humano considerando-se tanto os aspectos cognitivos quanto os aspectos evolutivos para criar técnicas computacionais que apresentem um grau de inteligência em seu processamento de informação.

Técnicas como as redes neurais artificiais e os algoritmos genéticos são, freqüentemente, considerados como pertencentes ao campo da inteligência computacional. É importante mencionar que, diversos pesquisadores têm constatado que as técnicas acima são, em muitos aspectos, complementares. Assim, muitas pesquisas têm sido realizadas no sentido de investigar possíveis

formas de interação entre estes métodos, e mesmo entre métodos de inteligência computacional e métodos tradicionais de engenharia, tais como, técnicas estatísticas e técnicas baseadas em gradientes.

Técnicas de inteligência computacional já foram aplicadas com sucesso na otimização e controle dos bioprocessos **(Stephanopoulos e Han, 1996; Shioya et al., 1999; Rowland, 2003; Oliveira, 2004** e **Zuo et al., 2006)**. Avanços neste sentido foram descritos por **Schügerl (2001)** em uma revisão do desenvolvimento da engenharia de bioprocessos nas últimas duas décadas.

A seguir, uma abordagem de tópicos específicos discorre sobre as técnicas de inteligência computacional: redes neurais artificiais (seção 2.2.1) e algoritmos genéticos (seção 2.2.2), as quais irão compor sistemas complexos apresentados nos capítulos 3 a 8.

2.2.1 Redes Neurais Artificiais

Devido ao seu grau de complexidade e capacidade de processamento de informação, aquele dentre os sistemas naturais que tem recebido maior atenção, é o cérebro humano. A inteligência artificial se desenvolveu em um estágio inicial de investigação acerca da inteligência humana, visando compreender como um cérebro de dimensões físicas reduzidas e consumo de energia limitada, seja ele biológico ou eletrônico, poderia perceber, compreender, prever e manipular um mundo extremamente diversificado (**Russel e Norvig, 1995**).

Haykin (1999) define a rede neural artificial (RNA) como sendo um sistema de processamento massivamente paralelo, composto por unidades simples com capacidade natural de armazenar conhecimento e disponibilizá-lo para uso futuro. Existem muitas razões para que sejam empregadas as RNA. A razão mais importante é sua capacidade de aproximação universal e sua flexibilidade para formar soluções de qualidade para uma ampla classe de problemas.

As RNA têm sido desenvolvidas como generalizações de modelos matemáticos de cognição humana ou neurobiologia, assumindo que o processamento de informação ocorre com o auxilio de várias unidades de processamento chamados neurônios. Os sinais são propagados de um elemento a

outro através de conexões; cada conexão possui um peso associado, que em uma rede neural típica, pondera o sinal transmitido; e cada neurônio (ou unidade) possui uma função de ativação, que tem como argumento a soma ponderada dos sinais de entrada, para determinar sua saída. Uma rede neural pode ser caracterizada por três aspectos principais: (i) padrões de conexões entre as unidades (arquitetura), (ii) método de determinação dos pesos das conexões (algoritmo de treinamento ou aprendizagem) e (iii) função de ativação.

McCulloch e Pitts (1943) projetaram a estrutura que é conhecida como a unidade básica de uma rede neural. Estes pesquisadores propuseram um modelo de neurônio como uma unidade de processamento binária. Matematicamente, o neurônio da Figura 2.4 pode ser expresso pela seguinte equação:

$$y_k = f(v_k) = f(\sum_{j=1}^{p} w_{kj} x_j) = f(w^T x) \quad (1)$$

onde y_k é a saída do neurônio, v_k é a ativação do neurônio, $f(\cdot)$ sua função de ativação, $\mathbf{x} = [x_0\ x_1\ ...\ x_n]^T$ é o vetor de entradas do neurônio k e $\mathbf{w} = [w_{k0}\ w_{k1}\ ...\ w_{kn}]^T$ é o vetor de pesos do neurônio k.

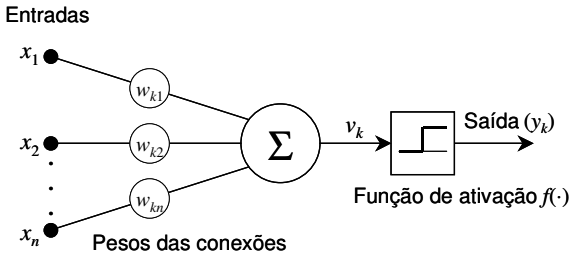


Figura 2.4. Representação funcional de um neurônio artificial

Rosenblatt (1957) introduziu uma nova abordagem para o problema de reconhecimento de padrões como o desenvolvimento do perceptron, definida por:

$$y_k = f(\sum_{j=1}^{p} w_{kj} x_j - w_{k0}) = f(w^T x) \quad (2)$$

uma entrada de polarização x_0 juntamente com o peso w_{k0} a ela associada, tem o efeito de transladar a função de ativação em torno da origem, fazendo com que a ativação interna v_k do neurônio não seja nula quando todas as demais entradas $\{x_1, x_2, ..., x_n\}$ forem nulas.

2.2.1.1 Redes neurais *feedforward* do tipo perceptron de múltiplas camadas (MLP)

A forma pela qual os neurônios de uma RNA estão estruturados está relacionada ao algoritmo de aprendizagem a ser utilizado para treiná-la.

As redes *feedforward* do tipo perceptron de múltiplas camadas (MLP) consiste tipicamente de um conjunto de unidades sensoriais (neurônios) que formam uma camada de entrada, uma ou mais camadas intermediárias ou escondidas e uma camada de saída. As redes MLP são geralmente treinadas usando o algoritmo de retro-propagação do erro (*error backpropagation*). Este algoritmo requer a propagação direta (*feedforward*) do sinal de entrada através da rede, e a retro-propagação do sinal do erro (**Haykin, 1999**).

Função de ativação

A função de ativação f(·) representa a ativação de saída do neurônio em termos do seu nível de ativação interna. Esta função geralmente é monotonicamente não-decrescente e apresenta um tipo de não-linearidade associada ao efeito. A Figura 2.5 apresenta alguns tipos de função de ativação utilizados na literatura.

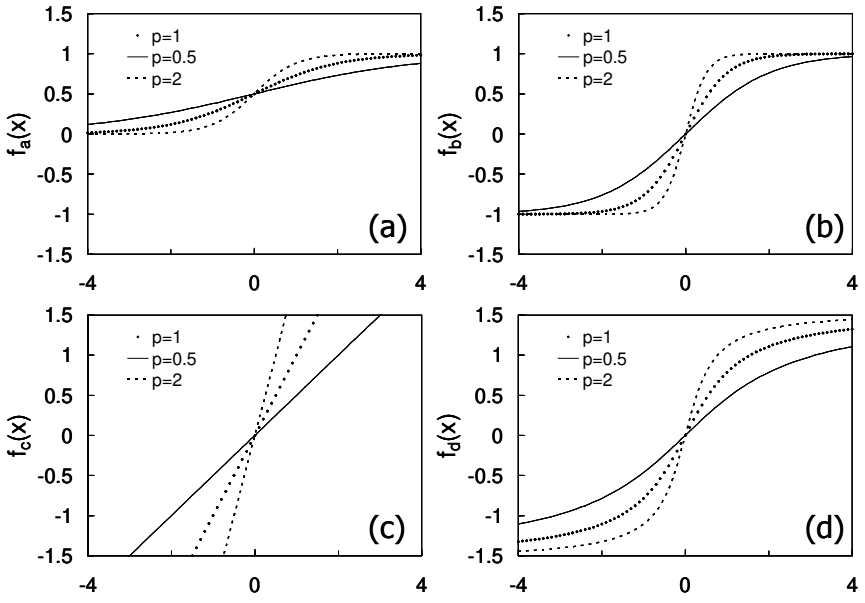


Figura 2.5. Funções de ativação. a) sigmoidal: $f_a(x) = 1/(1+\exp(-px))$; b) tangente hiperbólica $f_b(x) = \tanh(px)$; linear: $f_c(x) = px$; arco-tangente: $f_d(x) = \text{atan}(px)$

A origem da função sigmoidal (Figura 2.5.a) está vinculada à preocupação em limitar o intervalo de variação da derivada da função, pela inclusão de um efeito de saturação. A função sigmoidal é contínua, não-constante, limitada e monotonicamente crescente. Sendo monotônica, fornece também um treinamento mais eficiente da rede (**Baughman, 1995**). A função tangente hiperbólica (Figura 2.5.b) assume valores no intervalo (–1, +1), e pode ser transformada para o intervalo (0, 1), assim como a função logística pode ser transladada para o intervalo (-1, 1), uma vez que $f_b(x) = (f_a(x) + 1)/2$. A função de ativação linear (Figura 2.5.c) é muito utilizada nas unidades que compõem a camada de saída das arquiteturas MLP. Outra função de ativação utilizada é a função arco-tangente (Figura 2.5.d) que possui valores de ativação no intervalo $(-\pi/2, \pi/2)$, e pode ser apresentada como uma alternativa à função tangente hiperbólica para a implementação computacional, pois requer menos cálculos para sua elaboração.

Algoritmo de otimização para treinamento supervisionado

O treinamento do tipo MLP é um problema de otimização não-linear irrestrito, onde uma função de erro global é minimizada a partir do ajuste de parâmetros da rede neural (pesos). Este fato traz grandes vantagens devido à vasta literatura existente sobre métodos de otimização não-linear.

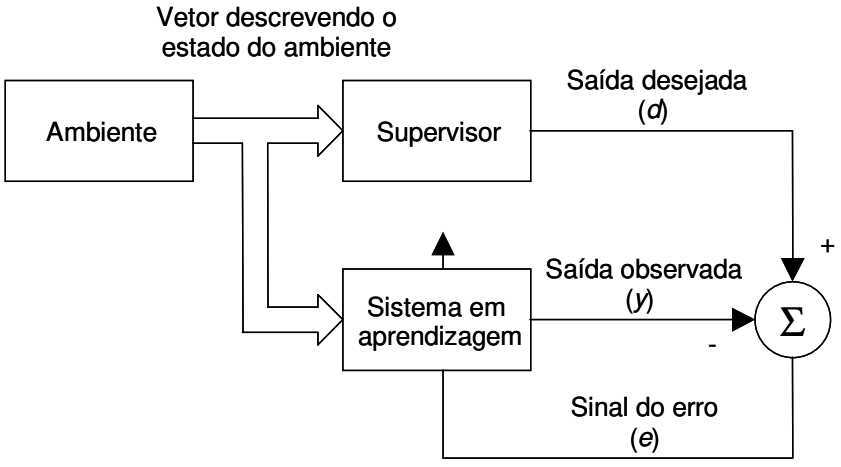


Figura 2.6. Diagrama de blocos do processo de aprendizagem supervisionada

A retro-propagação (*backpropagation*) do gradiente do erro provou sua utilidade no treinamento supervisionado de redes multicamadas para aplicação a

muitos problemas de classificação e mapeamento estático de funções não-lineares. A Figura 2.6 ilustra a propagação do erro em uma rede MLP no qual um supervisor possui conhecimento sobre o ambiente em que a rede está inserida. Este conhecimento está representado sob a forma de um conjunto de amostras de entrada-saída. O ambiente, por sua vez, é desconhecido. Esta abordagem é denominada aprendizagem supervisionada, onde os parâmetros da rede são ajustados pela combinação do sinal de entrada com um sinal de erro, que é a diferença entre a saída desejada e a fornecida pela rede.

Uma forma de implementar uma estratégia de treinamento supervisionado em redes neurais é através de procedimentos iterativos de correção do erro. Seja t o índice que denota o intervalo de tempo do processo iterativo responsável pelo ajuste de pesos do neurônio k. O único sinal de saída $y_k(t)$ do neurônio k é comparado com uma saída desejada, denominada $d_k(t)$. Conseqüentemente, um sinal de erro $e_k(t)$ é produzido:

$$e_k = d_k(t) - y_k(t) \tag{3}$$

Esta perspectiva do processo de treinamento supervisionado permite desenvolver algoritmos de treinamento baseados em resultados fundamentados da teoria de análise numérica convencional.

Métodos passíveis de implementação numérica que utilizam apenas o gradiente local da função (métodos de primeira ordem), ou então, métodos que utilizam também as derivadas de segunda ordem. Os métodos de primeira ordem são conhecidos por serem ineficientes no tratamento de problemas de larga escala, pois apresentam taxas de convergência muito pobres, especialmente em regiões próximas a mínimos locais (**Bazaara et al., 1991**). Em vista disso, e do fato de que os dados para treinamento geralmente apresentam grande dimensionalidade dos espaços de busca, é justificável a utilização de um método de otimização não-linear de segunda ordem.

Há casos em que a velocidade de aprendizagem é um fator limitante para possibilitar a implementação prática deste tipo de ferramenta computacional no

processo de solução de problemas que requerem otimalidade, robustez e rapidez na convergência do processo de ajuste de parâmetros.

Como o treinamento de redes MLP é entendido como um problema geral de aproximação de funções (**Cybenko, 1989**), será apresentado a seguir uma breve descrição sobre a teoria de aproximação universal de funções.

Aproximação universal de funções

As redes MLP com uma única camada de neurônios escondidos, como a representada na Figura 2.7, são consideradas estruturas poderosas para realizar uma aproximação uniforme, dado um conjunto de treinamento suficientemente significativo para representar a função.

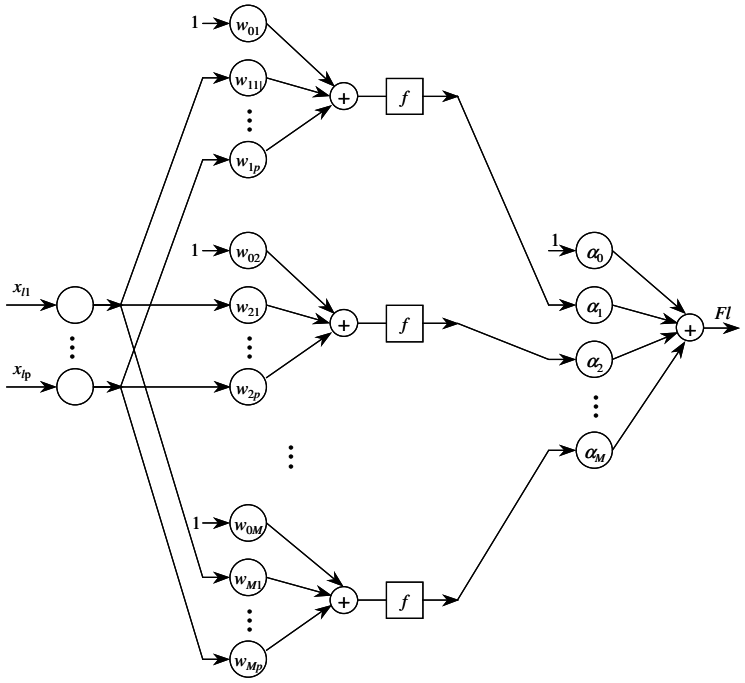


Figura 2.7. Rede MLP com uma única camada escondida. Os parâmetros da rede compõem a equação 4.

Cybenko (1989) demonstrou rigorosamente que uma rede MLP com uma única camada intermediária é suficiente para aproximar uniformemente qualquer função contínua que encaixe em um hipercubo unitário. O teorema de aproximação universal aplicável à arquitetura do tipo MLP é descrito a seguir.

Teorema: Seja f(·) uma função contínua não-constante, limitada, e monotonicamente crescente. Seja I_p um hipercubo unitário p-dimensional $(0,1)^p$. O espaço das funções contínuas em I_p é denominado $C(I_p)$. Então, dada qualquer função $g \in C(I_p)$ e $\varepsilon > 0$, existe um inteiro M e conjuntos de constantes reais α_i e w_{ij}, onde i=1,...,M e j=1,...,p, tais que pode-se definir:

$$F(x_1, x_2, ..., x_p) = \sum_{i=1}^{M} \alpha_i f\left(\sum_{j=1}^{p} w_{ij} x_j - w_{0i} \right) \tag{4}$$

como uma aproximação da função g(·) tal que,

$$\left| F(x_1, x_2, ..., x_p) - g(x_1, x_2, ..., x_p) \right| < \varepsilon \tag{5}$$

Para todo $\{x_1,...,x_p\} \in I_p$

Prova: veja Cybenko (1989).

O Teorema afirma que existe uma rede neural com uma única camada escondida e com um número adequado de neurônios nesta camada, que seja capaz de aproximar com precisão ε, uma função contínua desconhecida, dado um conjunto de amostras para treinamento. Notamos que a função logística ou tangente hiperbólica, utilizada como não-linearidade de um neurônio é contínua, não-constante, limitada, e monotonicamente crescente; satisfazendo assim, as condições impostas à função f(·). Verificamos que a Equação 4, representa uma rede do tipo MLP. A rede possui p nós de entrada e uma única camada intermediária consistindo de M unidades; o conjunto de entradas é $\{x_1,...,x_p\}$. O neurônio intermediário i possui pesos $w_{i1},...,w_{ip}$ e bias w_{01}. A saída da rede é uma combinação linear das saídas das unidades intermediárias, com $\alpha_1,..., \alpha_M$ definindo os coeficientes dessa combinação. Entretanto, o teorema não afirma que uma arquitetura do tipo MLP com uma única camada escondida é ótima no sentido de tempo de treinamento, facilidade de implementação e também, não indica como obter M, α e w, sendo apenas uma representação preestabelecida.

Como o treinamento de redes MLP é entendido como um problema geral de aproximação de funções (**Cybenko, 1989**), apresenta-se a seguir uma breve

descrição do algoritmo Levenberg-Marquardt que foi utilizado para a otimização dos sistemas de redes neurais resultantes nos capítulos 4, 5, 7 e 8.

Levenberg-Marquardt (LM)

O método de Levenberg-Marquardt é uma variante do método de Newton, que procura manter o cálculo da matriz Hessiana sempre positiva definida.

O método de Newton é baseado na função S que se minimiza e se aproxima localmente por uma função quadrática. Assim, perto de x_k, pode-se aproximar S pela serie de Taylor truncada:

$$S(x_{k+1}) = S(x_k + \Delta x_k) \approx S(x_k) + g_k^T \Delta x_k + \tfrac{1}{2} \Delta x_k^T H_k \Delta x_k \tag{6}$$

onde g_k é o gradiente avaliado em x_k, isto é, $g_k \equiv \nabla S(x)\big|_{x=x_k}$, $H_k \equiv \nabla^2 S(x)\big|_{x=x_k}$ é a matriz Hessiana (matriz direção) e $\Delta x_k = x_{k+1} - x_k$.

O lado direito da equação (6) minimiza-se com respeito à direção Δx_k, ou seja, $\frac{\partial S}{\partial \Delta x_k} = 0$, assim:

$$g_k + H_k \Delta x_k = 0 \tag{7}$$

Resolvendo para Δx_k, teremos:

$$\Delta x_k = -H_k^{-1} g_k \tag{8}$$

Então o método de Newton é definido por:

$$x_{k+1} = x_k - H_k^{-1} g_k \tag{9}$$

Suponha que S(x) é a soma dos quadrados:

$$S(x) = \sum_{i=1}^{N} v_i^2(x) = v^T(x) v(x) \tag{10}$$

então, o j-ésimo elemento do gradiente será:

$$[\nabla S(x)]_j = \frac{\partial S(x)}{\partial x_j} = 2 \sum_{i=1}^{N} v_i(x) \frac{\partial v_i(x)}{\partial x_j} \tag{11}$$

O gradiente pode ser escrito na forma matricial:

$$\nabla S(x) = 2J^T(x)v(x) \tag{12}$$

onde:

$$J^T(x) = \begin{bmatrix} \frac{\partial v_1(x)}{\partial x_1} & \frac{\partial v_1(x)}{\partial x_2} & \cdots & \frac{\partial v_1(x)}{\partial x_n} \\ \frac{\partial v_2(x)}{\partial x_1} & \frac{\partial v_2(x)}{\partial x_2} & \cdots & \frac{\partial v_2(x)}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial v_N(x)}{\partial x_1} & \frac{\partial v_N(x)}{\partial x_2} & \cdots & \frac{\partial v_N(x)}{\partial x_n} \end{bmatrix} \tag{13}$$

é a matriz Jacobiana.

Para o cálculo da matriz Hessiana, os elementos k, j da matriz são:

$$[\nabla^2 S(x)]_{k,j} = \frac{\partial^2 S(x)}{\partial x_k x_j} = 2\sum_{i=1}^{N}\left\{\frac{\partial v_i(x)}{\partial x_k}\frac{\partial v_i(x)}{\partial x_j} + v_i(x)\frac{\partial^2 v_i(x)}{\partial x_k \partial x_j}\right\}. \tag{14}$$

A matriz Hessiana também pode ser representada da seguinte forma:

$$\nabla^2 S(x) = 2J^T(x)J(x) + 2F(x) \tag{15}$$

onde

$$F(x) = \sum_{i=1}^{N} v_i(x)\nabla^2 v_i(x) \tag{16}$$

Suponha que F(x) é muito pequeno, então, podemos aproximar a matriz Hessiana como:

$$\nabla^2 S(x) \cong 2J^T(x)J(x) \tag{17}$$

Substituindo a Equação 17 e a Equação 12 na Equação 9, obtemos o método de Gauss-Newton:

$$x_{k+1} = x_k - [2J^T(x_k)J(x_k)]^{-1}2J^T(x_k)v(x_k) = x_k - [J^T(x_k)J(x_k)]^{-1}J^T(x_k)v(x_k) \tag{18}$$

Tendo como base o método de Gauss-Newton, a modificação de Levenberg-Marquardt na Equação (18) serve para resolver o problema causado quando a

matriz $J^T(x_k)J(x_k)$ do sistema do método de Gauss-Newton é singular (não tem inversa). Esta modificação transforma o algoritmo resultante global em:

$$x_{k+1} = x_k - [J^T(x_k)J(x_k) + \mu_k I]^{-1} J^T(x_k)v(x_k) \quad (19)$$

O objetivo da matriz $\mu_k I$ é adicionar μ a cada autovalor de $J^T(x_k)J(x_k)$. Uma vez que a matriz $J^T(x_k)J(x_k)$ é semi-definida positiva e, portanto, o autovalor mínimo possível é zero, qualquer valor positivo mesmo pequeno, mas numericamente significativo de μ, será suficiente para restaurar a matriz aumentada e produzir uma direção descendente de busca. As propriedades de convergência sugerem que o algoritmo LM pode ter convergência local lenta em problemas de grandes resíduos ou acentuadamente não-lineares. No entanto, as variadas implementações que têm surgido deste algoritmo provaram que é bem sucedido na maior parte dos problemas não-lineares (**Wolfe, 1978**; **Dennis e Schnabel, 1983**).

2.2.2 Aplicações de redes neurais artificiais em bioprocessos

Diversos estudos encontrados na literatura comprovam a eficácia das RNA aplicadas aos bioprocessos. Alguns trabalhos propuseram uma visão geral (**Baughman e Liu, 1995**; **Shioya et al. (1999)**; **Basheer e Hajmeer, 2000**). Entretanto, diversas aplicações bem sucedidas são citadas abaixo:

- **Modelagem de processos.** Muitos trabalhos têm surgido nos últimos anos aplicando RNA à modelagem de bioprocessos. Alguns exemplos são citados a seguir: **Zhu et al. (1996)** usaram RNA para a produção em grande escala de lisina e mostraram com sucesso a formação do produto baseada em medidas *on-line* e simulação recursiva do consumo de açúcar. **Becker et al. (2002)** usaram RNA para modelar a fermentação da cerveja e otimizar os perfis de temperatura para reduzir o tempo do processo. **Horiuchi et al. (2004)** usaram RNA para predizer o comportamento do processo de produção de queijo e concluíram que é uma técnica eficiente para predizer o comportamento de vários processos na indústria com características complexas e não-lineares. **Wouwer et al. (2004)**

propuseram uma metodologia sistemática para descrever com eficácia a cinética de reação de um processo biotecnológico utilizando RNA de base radial.

➢ **Processamento de dados**. As RNA podem ser aplicadas à compressão e filtragem de dados (**Baughman e Liu, 1995**). É possível conseguir isto através de redes neurais auto-associativas onde, as entradas são mapeadas através de uma camada oculta com menos neurônios do que o número de entradas. As RNA de arquitetura recorrente podem ser aplicadas para reconstrução de dados (*on-line*) pela remoção de ruído e de erros grosseiros (grandes perturbações em tempos curtos). As redes recorrentes distinguem-se das redes *feedforward* pela existência de pelo menos um laço (*loop*) de recorrência (*feedback*). Por exemplo, uma rede recorrente pode consistir de uma única camada de neurônios alimentando seu sinal de saída de volta para a entrada de todos os outros neurônios. Trabalhos recentes apontam para um interesse da combinação das RNA com outras técnicas como os *Wavelets*, para gerar uma função de reconstrução que é usada para o pré-processamento de dados de treinamento escassos (**Bing e Woodley, 2002**).

➢ ***Soft sensors*** **(Sensores vituais).** As RNA são utilizadas também para desenvolver *soft sensors* capazes de inferir variáveis de difícil medida, a partir de medidas simples e de baixo custo (variáveis secundárias) (**Assis e Maciel, 2000**). Medidas de variáveis secundárias, geralmente *on-line*, são relacionadas às medidas *off-line* do processo (**Karim et al., 1997**; **Linko et al., 1999**). Por exemplo, as medidas *on-line* de pH, concentração de amônia e as entradas para o controle da agitação, foram aplicadas para predizer as concentrações de biomassa e produto (**Linko et al., 1997**). **James et al. (2002)** estudaram o uso das RNA para a predição (*on-line)* da concentração de biomassa na produção de biopolímeros. *Soft sensors* podem também ser aplicados para substituir procedimentos analíticos difíceis ou caros. Por exemplo, **Kang et al. (1998)** desenvolveram um *soft sensor* para medir a concentração do ácido clavulânico baseado na espectrometria de massa de amostras de cultivo.

➢ **Otimização do meio de cultura**. **Kennedy e Spooner (1996)** compararam as RNA com um modelo clássico polinomial de segunda ordem para predizer os efeitos das concentrações dos componentes na produção de glutation em sua forma reduzida por *S. cerivisiae*, obtendo melhores resultados utilizando as RNA. **Baishan et al. (2003)** empregaram com sucesso a combinação de RNA e algoritmos genéticos na otimização do meio de fermentação de xilitol por *Candida mogii*.

➢ **Controle baseado em modelo**. RNA são utilizadas como modelos entrada-saída para aplicações em estratégias de controle preditivo não-linear (MPC- *Model Predictive Control*). O desenvolvimento de modelos e a falta de técnicas de medição *on-line* são os principais limitantes na aplicação de MPC nos bioprocessos. Recentes desenvolvimentos nas técnicas de medição *on-line*, estimação de parâmetros, modelagem qualitativa, assim como, a procura de um controle de qualidade mais rígido e produção de melhores produtos, motivaram o desenvolvimento de aplicações de controle MPC. Por exemplo, **Preuss et al. (2000)** aplicaram o MPC para controle de glicose na produção de *S. cerevisiae*, onde a concentração de etanol foi controlada com *set-point* baixo e fixo. **Hodge e Karim (2002)** usaram o controle preditivo para a otimização da produção de etanol em uma fermentação em batelada alimentada usando *Zymomonas mobilis* recombinantes. O controle MPC foi empregado para manter uma concentração constante do produto durante a fase de alimentação intermediária, para deslocar os efeitos inibitórios sobre o produto.

➢ **Reconhecimento de padrões**. RNA foram aplicadas para predizer o desempenho *on-line* e para detecção de falhas que ocorrem nos sensores dos bioprocessos (**Glassey et al., 1997**). Outro trabalho aborda a aplicação de redes de Kohonen (RNA auto-organizadas) para relacionar dados de qualidade para predizer o desempenho no estágio final da fermentação (**Ignova et al., 1999**). As redes de Kohonen possuem aprendizado competitivo ou auto-organizado, onde os neurônios de saída competem entre si para estarem ativos ou não. **Huang et al. (2002)** empregaram com sucesso redes neurais auto-associativas para detecção *on-line* de falhas na produção de virginiamicina por *Streptomyces*

virginiae. As redes auto-associativas possuem uma propriedade emergente que ajuda a recuperar informações e lidar com ruídos.

2.2.3 Algoritmos Genéticos

Os Algoritmos genéticos (AG) são procedimentos computacionais estocásticos, cujos métodos de busca modelam fenômenos biológicos de herança genética e seleção natural (**Michalewicz, 1996**). Os AG foram propostos por **John Holland (1975)**, com o objetivo de formalizar matematicamente e explicar rigorosamente processos de adaptação em sistemas naturais e desenvolver sistemas simulados em computador que retenham os mecanismos originais encontrados em sistemas naturais. O processo de evolução executado por um algoritmo genético corresponde a um processo de busca sobre uma população de pontos, e não sobre um único ponto. Outras características importantes são a utilização de funções de custo (funções objetivo), ao invés de derivadas ou outro tipo de conhecimento auxiliar e a utilização de regras de transição probabilística, e não determinística.

Algoritmos genéticos são uma alternativa para tentar superar as limitações apresentadas por métodos tradicionais, embora não garantam a obtenção da solução exata. Existem duas desvantagens principais ao tratar problemas complexos (**Yang et al., 1999**). O primeiro é o custo computacional elevado devido à convergência lenta e estratégias de busca paralela. O segundo está relacionado à complexidade do espaço de busca; otimizar todas as variáveis ao mesmo tempo é difícil devido à dimensionalidade elevada do espaço de busca e variações na sensibilidade dos parâmetros.

Os algoritmos genéticos considerados como pertencentes ao campo dos sistemas computacionais bio-inspirados consideram os seguintes princípios básicos (**Atmar, 1994; Bäck et al., 2000**):

i) Utilizam o processo de aprendizagem coletivo de uma população inicial de indivíduos que é escolhida aleatoriamente. O tamanho N da população é geralmente grande. Na maioria das vezes, cada indivíduo representa (ou codifica)

um ponto de busca no espaço de possíveis soluções potenciais para um determinado problema.

ii) Os descendentes dos indivíduos são gerados como erro, de forma que cada pai gere n descendentes. O mecanismo de introdução de erro durante a reprodução deve ser estocástico e variável.

iii) A qualidade do comportamento dos indivíduos em seu ambiente (comportamento do sistema) é avaliada para todos os indivíduos da população, pais e filhos, baseada em uma medida de adaptabilidade ou *fitness* de cada individuo. Uma das duas condições é geralmente implementada: 1) os N melhores indivíduos são selecionados para reproduzir e compor a próxima geração; 2) N dos melhores indivíduos são selecionados probabilisticamente. O tamanho da população é geralmente restrito.

iv) Repete-se o processo, voltando ao passo 2. A convergência é assumida quando uma solução pré-definida é atingida, ou um número fixo de gerações foi executado.

Algoritmo genético clássico e modificado

Para um problema particular, o algoritmo genético básico possui as seguintes características: (i) codificação binária; (ii) reprodução e seleção natural via o método estocástico da roleta (*Roullete Wheel*); iii) crossover simples; e iv) mutação. Este algoritmo pode ser descrito pela Figura 2.8 (**De Jong, 1994**).

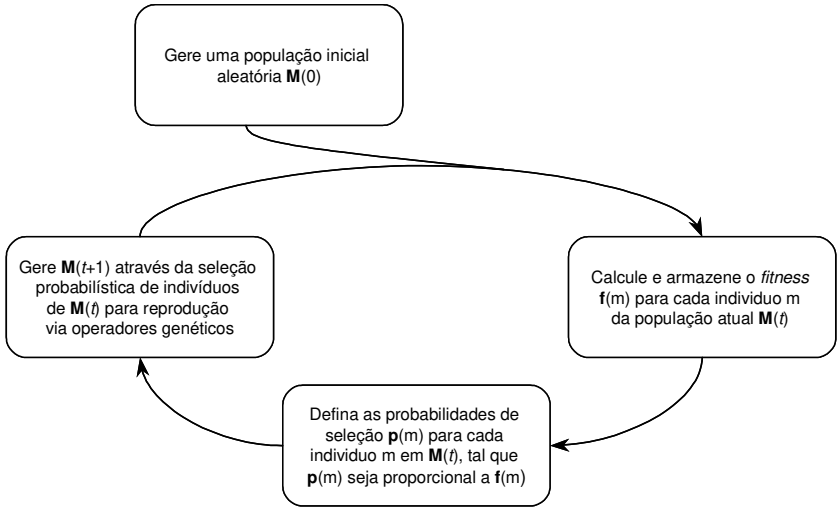
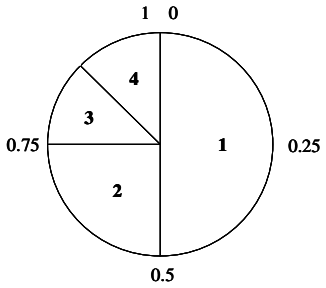


Figura 2.8. Algoritmo genético clássico

O algoritmo genético clássico utiliza um esquema de seleção de indivíduos para a próxima geração chamado *Roulette Wheel* (RW) (**Michalewicz, 1996**). Nesta abordagem a probabilidade de seleção é proporcional ao *fitness* do individuo. A Figura 2.9 ilustra o RW para uma população composta por quatro indivíduos. Girar a roleta é equivalente a gerar aleatoriamente um número com distribuição uniforme no intervalo [0, 1]. O valor obtido vai definir o cromossomo escolhido de acordo com as marcações na Figura 2.9 (a).



(a)

N	Cromossomo	Fitness	Graus	% do Total
1	0001110101	6.0	180	50.0
2	0101110011	3.0	90	25.0
3	1101010001	1.5	45	12.5
4	1101010011	1.5	45	12.5

(b)

Figura 2.9. Exemplo de *Roullete Wheel* para uma população com quatro indivíduos

O cruzamento (*crossover*) simples (Figura 2.10) é realizado da seguinte forma: dados dois cromossomos (pais), escolhe-se uma posição aleatória destas cadeias e efetua-se a troca de material genético.

No algoritmo genético clássico, cada indivíduo tem uma probabilidade de crossover fixa aos indivíduos da população.

No processo de mutação, uma posição do cromossomo é escolhida aleatoriamente e seu gene (bit) trocado de $0 \rightarrow 1$ ou $1 \rightarrow 0$, conforme a Figura 2.11. A probabilidade de mutação define a taxa com que cada posição pode ser mutado.

Dentre os principais problemas com o algoritmo genético clássico destacam-se: o esquema de seleção de indivíduos por RW pode fazer como que o melhor indivíduo não passe para a próxima geração; a posição do gene no cromossomo influi na probabilidade de realizar crossover; e a dificuldade de codificação quando os parâmetros são número reais. Para solucionar estes problemas, algumas estratégias foram propostas: a utilização de mecanismos alternativos de seleção; crossover uniforme; e codificação em *strings* (cadeias) de número reais.

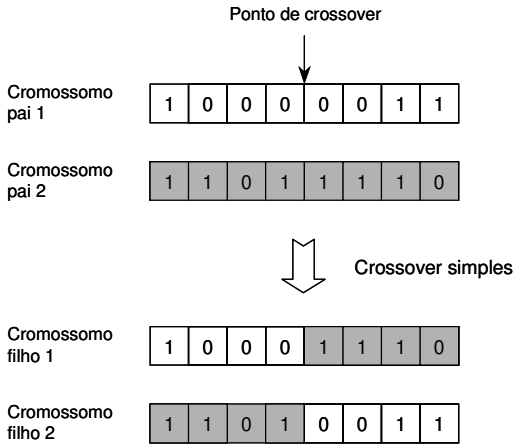


Figura 2.10. Crossover simples

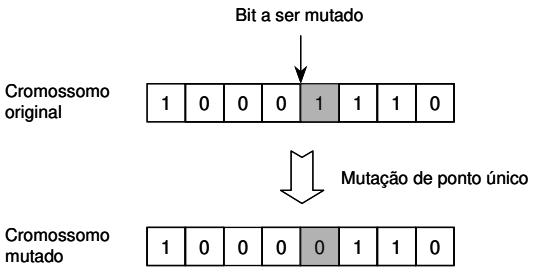


Figura 2.11. Mutação de ponto único

Outros operadores de recombinação

i) Crossover simples entre indivíduo aleatório e o melhor indivíduo. O objetivo deste operador é preservar seqüências genéticas (partes da cadeia) do melhor indivíduo na população durante todo o processo evolutivo.

ii) Crossover uniforme entre indivíduos aleatórios. Para cada gene (bit) no primeiro filho é decidido (com alguma probabilidade fixa p), qual pai vai contribuir com seu valor para aquela posição (Figura 2.12). Como o crossover uniforme troca genes ao invés de segmentos de genes, ele pode combinar características independentemente da sua posição relativa no cromossomo.

iii) Crossover uniforme entre indivíduos aleatórios e o melhor indivíduo.

iv) Mutação indutiva (somente para codificação real).

v) Mutação do melhor indivíduo. Pode ser aleatória ou indutiva.

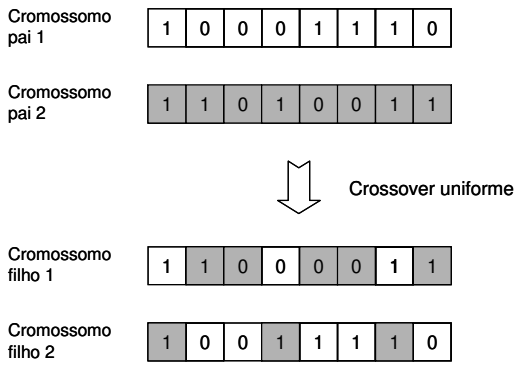


Figura 2.12. Crossover uniforme

Os operadores de recombinação descritos, também podem ser utilizados em cromossomos com codificação de ponto flutuante. Existem operadores de crossover especialmente desenvolvidos para uso com codificação em ponto flutuante. Um exemplo é o chamado crossover aritmético (**Michalewicz, 1996**). Este operador é definido como uma combinação linear de dois níveis (cromossomos): sejam x_1 e x_2 dois indivíduos selecionados para crossover; então os dois filhos resultantes serão $x'_1 = ax_1 + (1-a)x_2$ e $x'_2 = (1-a)x_1 + ax_2$ onde a é um número aleatório no intervalo [0, 1]. Quando a região viável é convexa, este operador é particularmente apropriado em problemas de otimização, devido a que combinações convexas de x_1 e x_2 serão também viáveis. Deste modo, garante-se que o crossover não gera indivíduos inválidos para o problema em questão.

Outros operadores crossover utilizados particularmente em problemas de otimização numérica restritos e codificação de ponto flutuante são o crossover geométrico e o crossover esférico, apresentados em **Michalewicz e Schoenauer (1996)**.

O operador de mutação, no caso de problemas com codificação em ponto flutuante, mais utilizado é a mutação gaussiana (**Michalewicz e Schoenauer, 1996**), que modifica todos os componentes de um cromossomo $x = [x_1 ... x_n]$ na forma: $x' = x + N(0, \sigma)$, onde $N(0, \sigma)$ é um vetor de variáveis aleatórias gaussianas independentes, com média zero e desvio padrão σ. Outro operador que utiliza codificação em ponto flutuante é a mutação uniforme (**Michalewicz, 1996**). Este operador seleciona aleatoriamente um componente $k \in \{1,2,...,n\}$ do cromossomo $x = [x_1 ... x_k ... x_n]$ e gera um indivíduo $x' = [x_1 ... x'_k ... x_n]$, onde x'_k é um número aleatório (com distribuição de probabilidade uniforme) amostrado no intervalo [LB, UB] e LB e UB são, respectivamente, os limites inferior e superior da variável x_k.

Outras estratégias de seleção

A seleção de indivíduos por *Roullete Wheel* pode fazer com que o melhor indivíduo da população seja perdido, ou seja, não passe para a próxima geração. A seguir são apresentadas algumas outras possíveis estratégias de seleção:

i) Elitista: manter os N melhores indivíduos da geração atual na geração seguinte.

ii) Por diversidade: são selecionados os indivíduos mais diversos da população.

iii) Bi-classista: são selecionados os P% melhores indivíduos e os (100-P)% piores indivíduos.

iv) Aleatória: são selecionados aleatoriamente N indivíduos da população intermediária. Podemos subdividir este mecanismo de seleção em:

Salvacionista: seleciona-se o melhor indivíduo e os outros aleatoriamente.

Não-salvacionista: todos os indivíduos são selecionados aleatoriamente.

v) Por torneio binário: dois indivíduos são escolhidos aleatoriamente. Um número aleatório $r \in [0, 1]$ é gerado. Caso $r < k$ (k é um parâmetro, como por exemplo, 0.7), o indivíduo com maior fitness é selecionado, senão o outro indivíduo é escolhido.

Técnicas para melhorar algoritmos genéticos

Algoritmos Genéticos são apropriados para problemas complexos, mas algumas melhorias devem ser feitas, permitindo trabalhar com problemas multimodal e multi-objetivo. Pode-se considerar, por exemplo, integração de AG com algoritmos de busca baseado em gradiente para melhorar a convergência combinando os AG com métodos de otimização tradicionais, por exemplo, o método de Levenberg-Marquardt (**Park e Froment, 1998**) e simplex (**Kasprzyk e Jaskula, 2004**).

O compartilhamento de *fitness* (*fitness sharing*), que é o método de nichos (*niching*) mais empregado na literatura, é utilizado para melhorar a convergência para um ou mais ótimos da função. O método de nichos é aquele que faz com que os algoritmos genéticos consigam manter uma população diversa de indivíduos, tornando-os capazes de localizar múltiplas soluções ótimas (sub-populações) dentro de uma única população (**Goldberg, 1989**). O estudo de nichos em algoritmos genéticos está diretamente relacionado ao problema da diversidade populacional. Uma das principais motivações para o desenvolvimento dos métodos

de nichos foi a busca pela diversidade, que possui duas funções dentro dos AG: retardar a convergência, objetivando aumentar a exploração da superfície de adaptação, e permitir a determinação de múltiplas soluções. **Golberg e Richardson (1987)** propuseram um esquema prático que utiliza uma metáfora de compartilhamento de *fitness* para solucionar o problema de otimização de funções multimodais. Esta técnica limita o crescimento descontrolado de espécies particulares dentro de uma população. Similaridades de espécies pode ser medida tanto no espaço genotípico (atributos iguais entre os indivíduos), como no espaço fenotípico (exemplos de treinamento cobertos pela regra). Neste esquema, uma função de compartilhamento é definida para determinar a vizinhança e grau de compartilhamento para cada indivíduo da população.

2.2.4 Aplicações de algoritmos genéticos em bioprocessos

Os métodos evolutivos são aplicados na indústria e pesquisa química e bioquímica, por exemplo, projeto do processo, identificação de parâmetros, modelagem de processos, entre outros.

- **Controle de processos.** Algumas aplicações em controle de bioprocessos foram desenvolvidas para o cálculo *off-line* de condições ótimas de temperatura (**Moriyama e Shimizu, 1996**), condições ótimas de alimentação para um processo de fermentação em batelada alimentada (**Sarkar e Modak, 2004**) e estratégias de controle preditivo, baseada na linearizacão dos modelos, de uma planta para produção de etanol (**Campello et al., 2003**).

- **Estimação de parâmetros.** AG foram aplicados para a estimação de parâmetros em bioprocessos contínuos, em batelada e batelada alimentada (**Park e Froment, 1998**; **Nougués et al., 2002**). **Park e Froment (1998)** estimaram os parâmetros da equação que descreve a taxa de uma reação catalítica heterogênea, utilizando um método híbrido AG-Levenberg-Marquardt. Os resultados mostraram que o modelo híbrido foi exato e eficiente, mantendo um equilíbrio apropriado entre a convergência e a diversidade durante o funcionamento do AG.

➢ **Otimização do meio de cultura.** AG são usados para otimizar meios de cultura de forma direta em procedimentos experimentais iterativos, isto é, experimentos reais são usados para otimizar a função objetivo para cada individuo (conjunto de variáveis) na população de soluções, em vez de simulações. Por exemplo, **Weuster-Botz e Wandrey (1995)** otimizaram com sucesso 14 componentes do meio para um processo de fermentação contínua. **Weuster-Botz et al. (1997)** aplicaram um procedimento similar para otimizar simultaneamente 13 componentes do meio de cultura para a produção de lisina por batelada alimentada. **Etschmann et al. (2004)** utilizaram AG, 13 meios de cultura e variações de temperatura para melhorar a bio-conversão de L-fenilalanina em 2-feniletanol por *Kluyveromyces marxianus* CBS 600, obtendo um aumento de 87% quando comparado ao meio não otimizado.

➢ **Otimização de parâmetros do processo.** De forma similar aos métodos para a otimização do meio, pode-se aplicar também métodos baseados em AG para otimizar condições operacionais. Isto foi feito por **Dutta et al. (2004)** para produção de protease por *Pseudomonas sp*. O algoritmo genético foi utilizado para minimização do tempo de reação e maximização da concentração do produto.

➢ ***Scheduling* (programação) e configuração do processo.** Os AG foram combinados com simulação e representação de processos (fluxograma ou *flowsheet*) para viabilizar o planejamento de produção proposta, por exemplo, **Wang et al. (2000)** usaram os AG para manter a viabilidade de soluções dos problemas na indústria de polímeros que permitiram a obtenção de soluções ótimas de produção.

➢ **Seleção do modelo de entrada. Abdollahi e Bagheri (2004)** usaram AG para a seleção do comprimento de onda para a determinação espectrofotométrica simultânea de Vitamina K_3 e 1,4–naftoquinona, com um modelo de calibração construído com a utilização da regressão por mínimos quadrados parciais.

➢ **Identificação da estrutura do modelo.** A Programação Genética (PG) constitui um algoritmo evolutivo, no qual as estruturas de dados que sofrem

adaptações são programas executáveis. A PG envolve uma busca baseada na evolução, junto aos candidatos pertencentes ao espaço de possíveis programas computacionais (**Bäck et al., 2000**). Uma comparação entre a PG, redes neurais artificiais (RNA) e modelos híbridos para produção de α-amilase por *B. Bacillus* foi feita por **Marenbach (1998)**. Todos os métodos resultaram em modelos exatos, entretanto, uma desvantagem principal da PG, quando comparada as RNA e técnicas de modelagem híbrida, foi a carga computacional elevada devido a que se utilizou um conjunto de computadores em paralelo para identificar modelos de bioprocessos que foram relativamente simples. Porém, a PG deu lugar a estruturas de modelo de melhor interpretação que as RNA.

Referências Bibliográficas

1. Abdollahi, H. e Bagheri, L. (2004). Simultaneous spectrophotometric determination of Vitamin K_3 and 1,4-naphthoquinone after cloud point extraction by using genetic algorithm based wavelength selection-partial least squares regression. *Analytica Chimica Acta*. 514:211-218.
2. Assis, A.J. e Maciel Filho, R. (2000). Soft sensors development for on-line bioreactor state estimation. *Comput. Chem.Eng.* 24:1099-1103.
3. Atala, D. I. P. (2004) Montagem, Instrumentação, Controle e Desenvolvimento Experimental de um Processo de Fermentação Alcoólica Extrativo. Campinas - SP. Tese (Doutor em Engenharia de Alimentos). Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas.
4. Atmar, W. (1994). Notes on the simulation of evolution. *IEEE Transactions on Neural Networks*. 5(1):130-147.
5. Bäck, T., Fogel, D.B. e Michalewicz, Z. (2000). *Evolutionary Computation 1, Basic algorithms and operators*. Institute of Physiscs Publishing (IOP), Bristol and Philadelphia.
6. Baishan, F., Hongwen, C., Xiaolan, X., Ning, W. e Zongding, H. (2003). Using genetic algorithms coupling neural networks in a study of xylitol production: medium optimisation. *Process Biochemistry*. 38:979-985.

7. Basheer, I.A. e Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *J. Microbiol. Meth.* 43:3–31.
8. Baughman, D.R. e Liu, Y.A. (1995). *Neural networks in bioprocessing and chemical engineering*. New York: Academic Press.
9. Bazaara, M.S., Sherali, C.M. e Shetty, C.M. (1991). *Nonlinear Programming: Theory and Algorithms*. Wiley, New York.
10. Becker, T., Enders, T. e Delgado, A. (2002). Dynamic neural networks as a tool for the online optimization of industrial fermentation. *Bioproc. Biosyst. Eng.* 24:347–354.
11. Bing, H. C. e Woodley, J.M. (2002). Wavelet shrinkage data processing for neural networks in bioprocess modeling. *Comput. Chem.Eng.* 26:1611-1620.
12. Campello, R.J.G.B., Von Zuben, F.J., Amaral, W.C., Meleiro, L.A.C. e Maciel Filho, R. (2003). Hierarchical fuzzy models within the framework of orthonormal basis functions and their application to bioprocess control. *Chem. Engng. Sci.* 58(18):4259-4270.
13. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and System*. 2:303-314.
14. De Jong, K. A. (1994), Genetic Algorithms: A 25 Year Perspective. *Computational Intelligence: Imitating Life.* IEEE Press. pp. 125-134.
15. Dennis, J. E. e Schnabel, R. B. (1983). *Numerical methods for unconstrained optimization and nonlinear equations*. Prentice-Hall, Inc.
16. Dutta, J.R., Dutta, P.K. e Banerjee, R. (2004). Modeling on optimization of protease production by a newly isolated *Pseudomonas* sp. using a genetic algorithm. *Process Biochem.* in press.
17. Etschmann, D., Shell, D. e Schrader, J. (2004). Medium optimization for the production of the aroma compound 2-phenylethanol using a genetic algorithm. *J. Molecular Catalysis B: Enzymatic*. 29:187-193.
18. Ferreira, R. M. (2003) Modelagem e simulão de biorreatores tipo torre operando com leveduras auto-imobilizáveis para produção de etanol. Campinas - SP. Tese (Doutor em Engenharia Química). Faculdade de Engenharia Química, Universidade Estadual de Campinas.

19. Glassey, J., Ignova, M., Ward, A.C., Montague, G. e Morris, A.J. (1997). Bioprocess supervision: neural networks and knowledge based systems. *J. Biotechnol.* 52:201–205.
20. Goldberg, D. E. & Richardson, J. (1987), Genetic algorithms with sharing for multimodal function optimization, *genetic algorithms and their applications: Proc. of the II ICGA*, pp. 41-49.
21. Goldberg, D. E. (1989), *Genetic algorithms in search, optimization and machine learning.* Addisson-Wesley Reading, Massachussets.
22. Grad, P. (2006). Biofuelling Brazil: An overview of the bioethanol success story in Brazil. *Refocus* 7:56-59.
23. Haykin, S. (1999). *Neural networks; A comprehensive foundation*. 2nd Ed. Prentice Hall.
24. Hodge, D.B. e Karim, M.N. (2002). Modeling and advanced control of recombinant *Zymomonas mobilus* fed-batch fermentation. *Biotechnol. Progr. 18*(3); 572-579.
25. Holland, J. H. (1998). *Adaptation in Natural and Artificial Systems.* 5a Ed. MIT Press.
26. Horiuchi, J., Shimada, T., Funahashi, H., Tada, K., Kobayashi, M. e Kanno, T. (2004). Artificial neural network model with a culture database for prediction of acidification step in cheese production. *J. Food Eng.* 63:459-465.
27. Huang, J., Shimizu, H. e Shioya, S. (2002). Data preprocessing and output evaluation of an autoassociative neural network model of online fault detection in virginiamycin production. *J. Biosc. Bioeng.* 87:261–266.
28. Ignova, M., Montague, G.A., Ward, A.C. e Glassey, J. (1999). Fermentation seed quality analysis with self-organising neural networks. *Biotechnol. Bioeng.* 64:82–91.
29. James, S., Legge, R. e Budman, H. (2002). Comparative study of black-box and hybrid estimation methods in fed-batch fermentation. *J. Process Control.* 12:113-121.
30. Kang, S.G., Lee D.H., Ward, A.C. e Lee, K.J. (1998). Rapid and quantitative analysis of clavulanico acid production by the combination of pyrolysis mass

spectrometry and artificial neural network. *J. Microbiol. Biotechnol.* 8:523–530.

31. Karim, M.N., Yoshida, T., Rivera, S.L., Saucedo, V.M., Eikens, B. e Oh, G.S. (1997). Global and local neural network models in biotechnology: application to different cultivation processes. *J. Ferment. Bioeng.* 83:1–11.
32. Kasprzyk, G. P., Jaskula, M. (2004). Application of the hybrid genetic-simplex algorithm for deconvolution of electrochemical responses in SDLSV method. *Journal of Electroanalytical Chemistry*. 567:39-66.
33. Kennedy, M.J. e Spooner, N.R. (1996). Using fuzzy logic to design fermentation media: A comparison to neural networks and factorial design. *Biotechnol. Techn.* 10:47–52.
34. Linko, S., Luopa, J. e Zhu, Y.H. (1997). Neural networks as 'software sensors' in enzyme production. *J. Biotechnol.* 52:257–266.
35. Linko, S., Zhu, Y.H. e Linko, P. (1999). Applying neural networks as software sensors for enzyme engineering. *Trends Biotechnol.* 17:155–162.
36. Macedo, I.C. e Cortez, L.A.B. (2005). O processamento industrial da cana-de-açúcar no Brasil, em Rosillo-Calle, F., Bajay, S.V. e Rothman, H. (Editores): Uso da biomassa para produção na indústria brasileira. Unicamp, SP, Brasil.
37. Marenbach, P. (1998). Using prior knowledge and obtaining process insight in data based modeling of bioproceses. *Syst. Anal. Model. Sim.* 31:39–59.
38. McCulloch W. e Pitts W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 5, pp. 115-133.
39. Michalewiz, Z. e Schoenauer, M. (1996), Evolutionary algorithms for constrained parameter optimization problems. *Evolutionary Computation*. 4(1): 1-32.
40. Moriyama, H. e Shimizu, K. (1996). On-line optimisation of culture temperature for ethanol fermentation using a genetic algorithm. *J. Chem. Technol. Biot.* 66:217–222.
41. Nougués, J.M., Grau, M.D. e Puigjaner, L. (2002). Parameter estimation with genetic algorithm in control of fed-batch reactors. *Chem. Eng. Process.* 41:303-309.

42. Oliveira, R. (2004). Combining first principles modeling and artificial neural networks: a general framework. *Comput. Chem.Eng.* 28:755-766.
43. Park, T.Y., Froment, G.F. (1998). A hybrid genetic algorithm for the estimation of parameters in detailed kinetic models. *Computers chem. Engng.* 22(suppl.): 103–110.
44. Preuss, K., Le Lann M.V., Proth, J. e Pingaud, H. (2000). Modelling and predictive control of fed-batch yeast growth on industrial plant scale. *Chem. Eng. J.* 78:53–59.
45. Rosenblatt, F. (1957). The Perceptron: A perceiving and recognizing automation. *Technical Report 85-460-1*. Cornell Aeronautical Laboratory, 1957.
46. Rossell, C. (2006). Principal especialista em obtenção de etanol a partir de resíduos de cana diz quais são as vantagens do Brasil e como mantê-las. *Inovação Unicamp*. Campinas, 24 de julho de 2006. Entrevista concedida a Janaína Simões.
47. Rowland, J.J. (2003). Model selection methodology in supervised learning with evolutionary computation. *Biosystem*. 72:187-196.
48. Russel, S. e Norvig, P. (1995). *Artificial Intelligence A Modern Approach*. Prentice Hall.
49. Sarkar, D. e Modak, J.M. (2004). Optimization of fed-batch bioreactors using genetic algorithm: multiple control variables. *Comput. Chem.Eng.* 28:789-798.
50. Schügerl, K. (2001). Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *J. Biotechnol.* 85:149–173.
51. Shioya, S., Shimizu, K. e Yoshida, T. (1999). Knowledge-based design and operation of bioprocess systems. *J. Biosc. Bioeng.* 87:261–266.
52. Stephanopoulos, G. e Han, C. (1996). Intelligent systems in process engineering: a review. *Comput.Chem. Eng.* 20:743–791.
53. Wang, K., Löhl, T., Stobbe, M. e Engell, S. (2000). A genetic algorithm for online-scheduling of a multiproduct polymer batch plant. *Comput. Chem.Eng.* 24:393-400.

54. Weuster-Botz, D. e Wandrey, C. (1995). Medium optimization by genetic algorithm for continuous production of formate dehydrogenase. *Process Biochem.* 30:563–571.

55. Weuster-Botz, D., Kelle, R., Franten M. e Wandrey C. (1997). Substrate controlled fed-batch production of L-Lysine with *Corynebacterium glutamicum*. *Biotechnol. Progr.* 13:387–393.

56. Wolfe, M. A. (1978). *Numerical methods for unconstrained optimization*. Van Nostrand Reinhold Company Ltd.

57. Wouwer, A.V., Renotte, C. e Bogaerts, Ph. (2004). Biological reaction modeling using radial basis function networks. *Comput. Chem.Eng.* in press.

58. Yang, L., Yen, J., Rajesh. A. e Kihm, K.D. (1999). A supervisory architecture and hybrid GA for the identification of complex systems. *Proc. Congress on Evolutionary Computing*, Volume 2, USA, pp. 862–869.

59. Zhu, Y. H., Rajalahti, T. e Linko, S. (1996). Application of neural networks to lysine production. *Chem. Eng. J.* 62:207–214.

60. Zuo, K., Cheng H.P., Wu, S.C. e Wu, W.T. (2006). A hybrid model combining hydrodynamic and biological effects for production of bacterial cellulose with a pilot scale airlift reactor. *Biochem. Eng. J.* 29:81-90.

CAPÍTULO 3. EVALUATION OF OPTIMIZATION TECHNIQUES FOR PARAMETER ESTIMATION: APPLICATION TO ETHANOL FERMENTATION CONSIDERING THE EFFECT OF TEMPERATURE

Trabalho publicado na revista **Process Biochemistry**, v. 41, p. 1682-1687, 2006.
Elmer Ccopa Rivera, Aline C. Costa, Daniel I. P. Atala, Francisco Maugeri, Maria R. Wolf Maciel, Rubens Maciel Filho.

Abstract

Optimization techniques are evaluated to estimate the kinetic model parameters of batch fermentation process for ethanol production using *Saccharomyces cerevisiae*. Batch experimental observations at five temperatures (28°C, 31°C, 34°C, 37°C and 40°C) are used to formulate the parameter estimation problem. The potential of Quasi-Newton (QN) and Real-Coded Genetic Algorithm (RGA) to solve the estimation problem is considered to find out the optimal solution. Subsequently, the optimized parameters (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px}) were characterized by correlation functions assuming temperature dependence. The kinetic models optimized by QN and RGA describe satisfactorily the batch fermentation process as demonstrated by the experimental results.

Keywords

Ethanol fermentation, batch fermentation, parameter estimation, temperature effect, genetic algorithm, quasi-newton algorithm.

3.1 Introduction

As a large tropical country, Brazil has a high potential for the use of biomass. Sugarcane products are today the most economically important biomass source for energy cogeneration as well as a suitable alternative feedstock for chemicals. Impelled by the creation of the "Proálcool" project in the 70's and 80's, the country has become in the last years, the world largest ethanol producer via fermentation, developing and improving many fermentative processes **[1]**. Bioethanol (ethanol from biomass) is an attractive, sustainable energy source to fuel. Based on the premise that fuel bioethanol can contribute to a cleaner environment and with the implementation of environmental protection laws in many countries, demand for this fuel is increasing **[2]**. A very common argument against ethanol is its economic competitiveness against fossil fuels. Nevertheless, Goldemberg **[1]** demonstrated, through the Brazilian experience with ethanol, that economy of scale and technological advances can lead to increased competitiveness of this renewable alternative, reducing the gap with conventional fossil fuels. In

consequence, there is an intensified interest in the study of all the steps involved in ethanol production in order to reduce costs. Modeling potentially reduces the cost of alcoholic fermentation process development by eliminating unnecessary experimental work, because it allows the study of the various process parameters interactions through simulation. Besides, it provides understanding of the process, which is helpful for operational policy definitions and can be applied for posterior optimization and control.

The lack of robustness of the fermentation in the presence of fluctuations in the quality of the raw material, which leads to changes in the kinetic behavior with impact on yield, productivity and conversion, is among the main current problems related to the alcoholic fermentation process. Changes in the operational conditions are quite common in plants of alcoholic fermentation; they occur not only due to the variations in the quality of the raw material but also due to variations of dominant yeast in the process. Also, the alcoholic fermentation process is exothermic and small deviations in temperature can dislocate the process from optimal operational conditions. Thus, it is clear that the main difficulty in model-based techniques for definition of operational strategies, control and optimization, is the problem of obtaining an accurate model.

It is known that even small changes in temperature affect the kinetics, and, thus, productivity and conversion. The temperature in a typical industrial fermentor for ethanol production varies from 33.5ºC (during the night) to 35 ºC (at the end of the day) due to fluctuations in the cooling water temperature. In plants with poor temperature control the temperature in the fermentor goes up to 40ºC. Moreover, terms of temperature influence on ethanol fermentation kinetics can be useful in strategies for process optimization **[3].** Still, there are very few works in the literature on the mathematical modeling of the fuel ethanol fermentation considering the effect of temperature on the kinetic parameters. Among them are the works of Atala et al. **[4]** and Aldiguier et al. **[5]**. Aldiguier et al. **[5]**, however, determine different values for kinetic parameters in different temperatures, but do not determine a function to describe the kinetic parameters as functions of temperature.

Estimation of kinetic parameters of differential models is usually complex, mainly due to non-linearities, great number of parameters and interactions among them. In biochemical engineering, the most classical method involves the mathematical estimation of model parameters based on the minimization of some cost function built-up with the parameters to be estimated. Several kinetic models have been proposed for the alcoholic fermentation process **[6-9]**. Many techniques are available for minimizing the error of estimation, generally methods based on gradient search, such as Quasi-Newton Algorithm (QN).

Artificial intelligence (AI), such as Genetic Algorithms (GA), covers a wide range of techniques and tools that facilitate decision making and have often been found to be as powerful and effective as gradient search methods in many engineering applications. These methods have already been successfully applied in the optimization and control of bioprocesses for more than 20 years **[10]**. GA have been successfully utilized for kinetic parameters estimation in biotechnological process **[11-13].**

This work focuses on the process operation aspects using model-based kinetic parameters optimization of an alcoholic fermentation process. Kinetics was determined for batch fermentations at temperatures between 28 and 40^{o}C. Based on experimental data, a differential model consisting of rate expressions for cell growth, substrate consumption and product formation was proposed. A comparison is made between the performance of the model when the kinetic parameters are optimized using a Quasi-Newton Algorithm and a Real-Coded Genetic Algorithm. The primary objective is to assess the performance of the model and no consideration is given to the speed of convergence. In general, Quasi-Newton, which is a local search algorithm, will converge to a solution much more quickly than a global search algorithm such as the Genetic Algorithm. Here, the focus is on the accuracy of the achieved solution using optimization algorithms.

3.2 Process description

3.2.1 Experimental

The microorganism used was *Saccharomyces cerevisiae* cultivated in the Bioprocess Engineering Laboratory in the Faculty of Food Engineering/State

University of Campinas and obtained from an industrial fermentation plant. The growth medium for the inoculum contained 50 kg/m^3 of glucose, 5 kg/m^3 of KH_2PO_4, 1.5 kg/m^3 of NH_4Cl, 0.7 kg/m^3 of $MgSO_4 \cdot 7H_2O$, 1.2 kg/m^3 of KCl, and 5 kg/m^3 of yeast extract. The production medium was diluted sugarcane molasses to which 1 kg/m^3 of yeast extract and 2.4 kg/m^3 of $(NH_4)_2SO_4$ were added. Sterilization was done at 121^{o}C for 30 min.

A bioreactor (Bioflow III System; New Brunswick Scientific) with a total working volume of 2.5 L was used. The dry cell mass was determined gravimetrically after centrifuging, washing, and drying the cells at 105^{o}C. Viable cells were counted with the methylene blue staining technique **[14]**. In this work the cell viability during batch fermentation was always close to 100%. Total reducing sugar (TRS) and ethanol concentrations were determined by high-performance liquid chromatography (Varian 9010 model). A column SHODEX KS 801 at 70^{o}C was used. Ultrapure Milli-Q water was used as the eluent at a flow rate of 0.5 mL/min. The standards were mixed solutions of sucrose, fructose, and ethanol at concentrations from 0.1 to 40 kg/m^3.

3.2.2 Batch model

For batch fermentation process, the mass balance equations that describe microorganism growth, substrate consumption and ethanol formation are:

$$\frac{dX}{dt} = r_x \tag{1}$$

$$\frac{dS}{dt} = -r_s \tag{2}$$

$$\frac{dP}{dt} = r_p \tag{3}$$

where X is the concentration of cell mass (kg/m^3), S is the concentration of substrate (kg/m^3) and P is the concentration of ethanol (kg/m^3).

In this study the rates of cell growth, r_x (kg/[m^3·h]), substrate consumption, r_s (kg/[m^3·h]), and product formation, r_p (kg/[m^3·h]), were expressed by an unstructured model for batch fermentation.

For fermentation with *Saccharomyces cerevisiae*, experimental data showed that cellular, substrate and product inhibitions are of importance **[4]**. In this study, the cell growth rate, r_x, includes terms for such types of inhibitions.

$$r_x = \mu_{max} \frac{S}{K_s + S} \exp(-K_i S)\left(1 - \frac{X}{X_{max}}\right)^m \left(1 - \frac{P}{P_{max}}\right)^n X \quad (4)$$

where μ_{max} is the maximum specific growth rate (h^{-1}), K_s the substrate saturation parameter (kg/m^3), K_i the substrate inhibition parameter, X_{max} the biomass concentration when cell growth ceases (kg/m^3), P_{max} the product concentration when cell growth ceases (kg/m^3), and m and n are the parameters used to describe cellular and product inhibitions, respectively.

Luedking-Piret expression **[15]** was used to account for the ethanol formation rate, r_p.

$$r_p = Y_{px} r_x + m_p X \quad (5)$$

where Y_{px} is the yield of product based on cell growth (kg/kg) and m_p is the ethanol production associated with growth (kg/[kg·h]).

The substrate consumption rate, r_s, is given by

$$r_s = (r_x / Y_x) + m_x X \quad (6)$$

describing the sugar consumption during fermentation, which leads to cell mass and ethanol formation. Y_x and m_x are the limit cellular yield (kg/kg) and maintenance parameter (kg/[kg·h]), respectively.

3.3 Parameter estimation

According to the above description, there are 11 parameters that have to be estimated from experimental observations and within them there are a number of temperature-dependent parameters in the kinetic model, μ_{max}, X_{max}, P_{max}, Y_x and Y_{px}, and this dependence can be described by the following expression **[16]**:

$$\text{temperature-dependent parameter} = A \exp\left(\frac{B}{T}\right) + C \exp\left(\frac{D}{T}\right) \quad (7)$$

In this equation, *A*, *B*, *C* and *D* are constants, and *T* is the temperature in ^{o}C.

3.3.1 Optimization criteria

Let θ specify the parameters vector, which contains all the temperature-dependent parameters. The objective of optimization is to find out θ by minimizing the objective function, min $E(\theta)$:

$$E(\theta) = \sum_{n=1}^{np}\left[\frac{(X_n - Xe_n)^2}{Xe_{\max}^2} + \frac{(S_n - Se_n)^2}{Se_{\max}^2} + \frac{(P_n - Pe_n)^2}{Pe_{\max}^2}\right] = \sum_{n=1}^{np}\varepsilon_n(\theta) \qquad (8)$$

by a Quasi-Newton Algorithm as well as by a Real-Coded Genetic Algorithm using the experimental data from fermentations performed in the temperature range of 28-40^{o}C. Xe_n, Se_n and Pe_n are the measured concentrations of cell mass, substrate and ethanol at the sampling time *n*. X_n, S_n and P_n are the concentrations computed by the model at the sampling time *n*. Xe_{max}, Se_{max} and Pe_{max} are the maximum measured concentrations and the term *np* is number of sampling points. Here $\varepsilon_n(\theta)$ is the error in the output due to the *n*th sample.

3.4 Results and Discussion

All computation were performed on an AMD Athlon 2600+ 1.25 MHz processor with 512 MB RAM. Eqs. (1-6) were solved using a Fortran program with integration by an algorithm based on the fourth-order Runge-Kutta method **[17]**.

To model the fermentation experiment, the temperature dependent parameters (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px} in Eqs. 4-6) were determined by minimizing Eq. (8) (by a Real-Coded Genetic Algorithm and a Quasi-Newton Algorithm). This procedure was repeated for each temperature value (28, 31, 34, 37 and 40^{o}C). The parameters that are not temperature dependent had their values fixed according to the work of Atala et al. **[4]** as follows: K_s = 4.1, K_i = 0.002, m_p = 0.1, m_x = 0.2, m = 1.0 and n = 1.5. On the other hand, Yalçin and Özbas **[18]** investigate the effect of different substrates (glucose, fructose and sucrose) on growth of *S. cerevisiae* and demonstrated that for the same microbial

species the values of parameters K_s and K_i were found to be different and dependent on all substrates.

3.4.1 Optimization by Quasi-Newton Algorithm

Quasi-Newton Algorithm was used to minimize the cost function given by Eq. (8). The FORTRAN IMSL routine DBCONF was used for this purpose. The straightforward idea is to implement the optimization problem as a nonlinear programming problem (NLP) and can be written as:

Minimize Equation 8

Subjet to $l_p \leq x_p \leq u_p$, p=1,…,5

where x_p is the temperature-dependent parameters. The l_p and u_p are specified lower and upper bounds on the parameters, with $l_p \leq u_p$.

3.4.2 Optimization by Genetic Algorithms

The GA considered in this work is based on the freeware versions written in FORTRAN of a Real-Coded Genetic Algorithm (RGA) developed by Yedder **[19],** adapted and enhanced to the specific needs of the parameter optimization. In the proposed strategy, the chromosome or individual are vectors where each real-value (gene) stands for each one of unknown temperature-dependent parameter in the kinetic model. The stopping criterion is selected when the maximum number of fixed generations is reached. The RGA parameters were adjusted to minimize the number of generations (iterations) required to reach a satisfactory fitness (objective function) value by minimizing Eq. (8). Table 3.1 shows the Real-Coded Genetic Algorithm used.

3.4.3 Influence of temperature on the kinetics

Figure 3.1 depicts the optimized parameters (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px}) by Quasi-Newton Algorithm (triangle data points) and Real-Coded Genetic Algorithm (square data points), estimated from the experimental data in the range 28-40^{o}C. Besides, the mean of the optimized parameters by QN and RGA (and 95% confidence limits) are showed in Table 3.2.

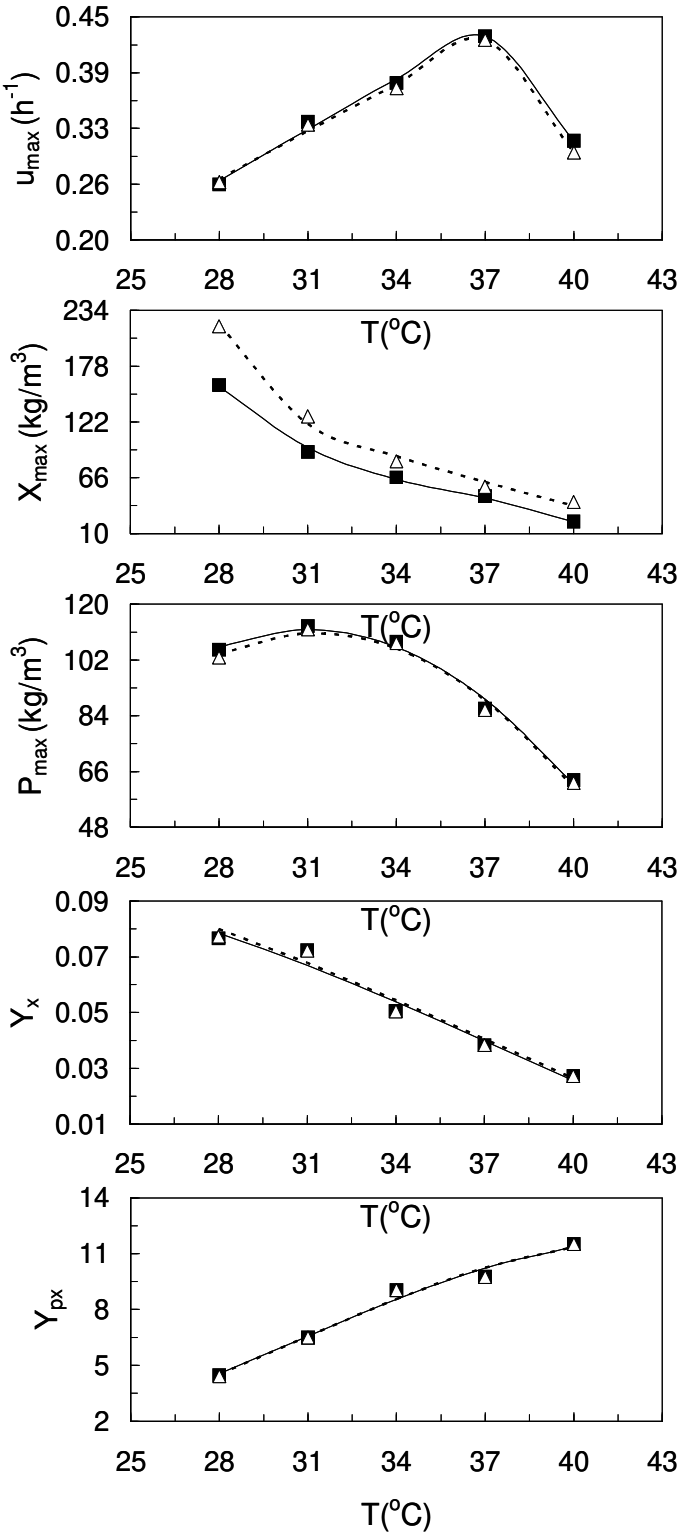


Figure 3.1. Optimized parameters (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px}) by Quasi-Newton Algorithm (triangle data points) and Real-Coded Genetic Algorithm (square data points), estimated from the experimental data in the range 28-40°C. The fitting results, using Eq. (7), are represented by lines (solid lines that fit the optimized parameters by QN and dashed lines that fit the optimized parameters by RGA)

Table 3.1. Real-Coded Genetic Algorithm

Step i: Set the probability of crossover (P_c)=0.9, the probability of mutation (P_m)=0.003, the total number of individuals in the population (N_I)=10 and the desired number of generations (N_g)=500 (these same setting factors were used for parameter optimization in all the tested temperatures). k was set zero and then the initial population $P(0)$ was generated with N_I individuals.

Step ii: Stop if $k>N_g$. Otherwise, go to the next step.

Step iii: Compute the fitness (Eq. 8) $F(I_j)$ of each individual I_j (j=1,…,N_I) of the current population $P(k)$.

Step iv: Niching. Take individuals with the smallest degree of similarity (using fitness sharing **[20]**) with the individual corresponding to the best fitness and include a copy of them into an auxiliary population $Pa(k)$.

Step v: Probability of selection **[21]**. The probability of selection $P_s(I_j)$ of each individual I_j (j=1,…, 10) of the current population is computed as $P_s(I_j)=F(I_j)/\sum_{I=1}^{N_I} F(I_I)$, in such a way that $P_s(I_j)$ to select the individual I_I is proportional to $F(I_j)$.

Step vi: Elitism. Take the individual with the best fitnesses and to compare it with the best individual in the prior generation. The best one of them is kept into the current auxiliary population $Pa(k)$.

Step vii: Selection and Reproduction. Execute the substeps below until N_I new individuals are obtained. (a) Selection. Take two individuals out the current auxiliary population $Pa(k)$. (b) Reproduction. Select a value $\varepsilon\in[0, 1]$ at random. If $\varepsilon>P_c$, then include a copy of the individual into update auxiliary population $Pa(k)$. Otherwise, apply barycentric crossover **[19]** to the individuals (parents) and include the two resulting individuals (off-spring) into $Pa(k)$.

Step viii: Mutation and generation update. Apply mutation to every individual of $Pa(k)$, as follows: Select a value $\varepsilon \in [0, 1]$ at random for each gene of the individual. If $\varepsilon < P_m$, then apply non-uniform mutation **[19]**. If $\varepsilon > P_m$ do nothing.

Finally, set $P(k+1) = Pa(k)$, $k=k+1$ and go back to **Step ii**.

The temperature dependency of the reaction rate was fitted by Eq. (7) using the Levenberg-Marquardt algorithm (solid lines that fit the optimized parameters by QN and dashed lines that fit the optimized parameters by RGA, in Figure 3.1). The estimate constants (A, B, C and D) in Eq. (7) are shown in Table 3.3. This table also shows the coefficient of determination (r^2) of the optimal parameters fitting as functions of temperature. Under such optimal model, the computed profiles for ethanol (P), substrate (S) and biomass (X) are shown in the solid-dotted (QN) and dash-dotted (RGA) curves of Figure 3.2 at 28, 31, 34, 37 and 40°C.

The residual standard deviation (RSD) **[4],** Eq. (9), written as a percentage of the average of the experimental values was the measurement used for characterizing the quality of the prediction of the model.

Table 3.2. Mean of the kinetic parameters optimized by a Quasi-Newton Algorithm and a Real-Coded Genetic Algorithm (and 95% confidence limits) at all temperatures

Param.	T=28°C	T=31°C	T=34°C	T=37°C	T=40°C
μ_{max}	0.264	0.331	0.373	0.426	0.305
	(0.261, 0.266)	(0.328, 0.333)	(0.368, 0.379)	(0.423, 0.430)	(0.292, 0.318)
X_{max}	189.400	105.843	76.866	54.474	29.877
	(129.426, 249.375)	(78.095, 133.590)	(55.993, 97.738)	(40.638, 68.311)	(14.240, 45.515)
P_{max}	104.072	112.392	107.685	86.072	62.752
	(101.676, 106.469)	(111.355, 113.429)	(107.306, 108.064)	(85.637, 86.507)	(61.905, 63.598)
Y_x	4.439	6.488	9.041	9.763	11.514
	(4.396, 4.482)	(6.487, 6.489)	(9.024, 9.058)	(9.754, 9.773)	(11.513, 11.515)
Y_{px}	0.07696	0.07229	0.05055	0.03831	0.02713
	(0.07622, 0.07769)	(0.07224, 0.07233)	(0.05050, 0.05060)	(0.03826, 0.03835)	(0.02712, 0.02714)

Table 3.3. Constant values in the Eq . (7) and coefficient of determination (r^2) of parameters fitting as functions of temperature to the batch model optimized by Quasi-Newton Algorithm (QN) and Real-Coded Genetic Algorithm (RGA)

Param.	(r^2)	QN				(r^2)	RGA			
		A	*B*	*C*	*D*		*A*	*B*	*C*	*D*
μ_{max}	(0.98)	1.772	-52.857	-6.638×10^{28}	-2.724×10^{3}	(0.99)	-1.488×10^{16}	-1.559×10^{3}	2.006	-56.511
X_{max}	(0.99)	139.360	154.734	-138.492	154.734	(0.99)	-1.436×10^{31}	-2.768×10^{3}	0.977	142.289
P_{max}	(0.99)	3.197×10^{5}	-104.778	-3.273×10^{5}	-105.823	(0.99)	3.529×10^{5}	-100.781	-3.596×10^{5}	-101.611
Y_x	(0.99)	5.622×10^{4}	-166.992	-5.794×10^{4}	-168.726	(0.99)	5.756×10^{4}	-165.030	-5.913×10^{4}	-166.600
Y_{px}	(0.99)	5.354×10^{4}	-35.269	-5.354×10^{4}	-35.269	(0.98)	-6.864	-39.383	6.378	-36.047

$$\text{RSD}(\%) = \frac{\sqrt{\text{RSD}}}{\bar{d}_p} \times 100 \tag{9}$$

where $\text{RSD} = \frac{1}{np}\sum_{p=1}^{np}(d_p\text{-}x_p)^2$, in which x_p and d_p are, respectively, the value predicted by the mathematical model and experimental value, $\bar{d}_p$ is the average of the experimental values and np is the number of experimental points.

Table 3.4. Residual standard deviation (RSD) written as a percentage of average of experimental values, used to characterize the prediction quality of the model

Output Variable	RSD(%)									
	(T=28°C)		(T=31°C)		(T=34°C)		(T=37°C)		(T=40°C)	
	QN	RGA	QN	RGA	QN	RGA	QN	RGA	QN	RGA
X	6.7	6.8	4.8	4.9	9.4	9.4	9.0	9.0	6.1	6.3
S	6.6	6.5	6.3	6.9	7.1	6.2	7.1	5.9	3.5	5.3
P	9.2	9.1	6.8	6.8	10.6	10.7	7.6	7.6	10.2	10.2

The RSD(%) for the batch model optimized by QN and RGA are shown in Table 3.4. It can be seen that the two optimization algorithms presented almost the same RSD. The deviations are from 3.5% to 10.7%. In bioprocess engineering, values of RSD(%) below 10% can be considered acceptable **[4]**. Thus, the estimated model was able to fit batch experimental observations very satisfactorily and, therefore, the model can be applied for posterior optimization and controller process design.

3.5. Concluding remarks

It is well known from past studies that models which do not have their parameters re-estimated when changes in operational conditions occur do not describe the process accurately. We focused on developing an estimation methodology that can be used always that a re-estimation of parameters is necessary. The methodology consisted in fixing the values of parameters which are not dependent on temperature and estimating the temperature dependent parameters. The potential of a Quasi-Newton Algorithm and a Real-Coded Genetic Algorithm were considered for estimation of these parameters and the two algorithms have shown similar good performances.

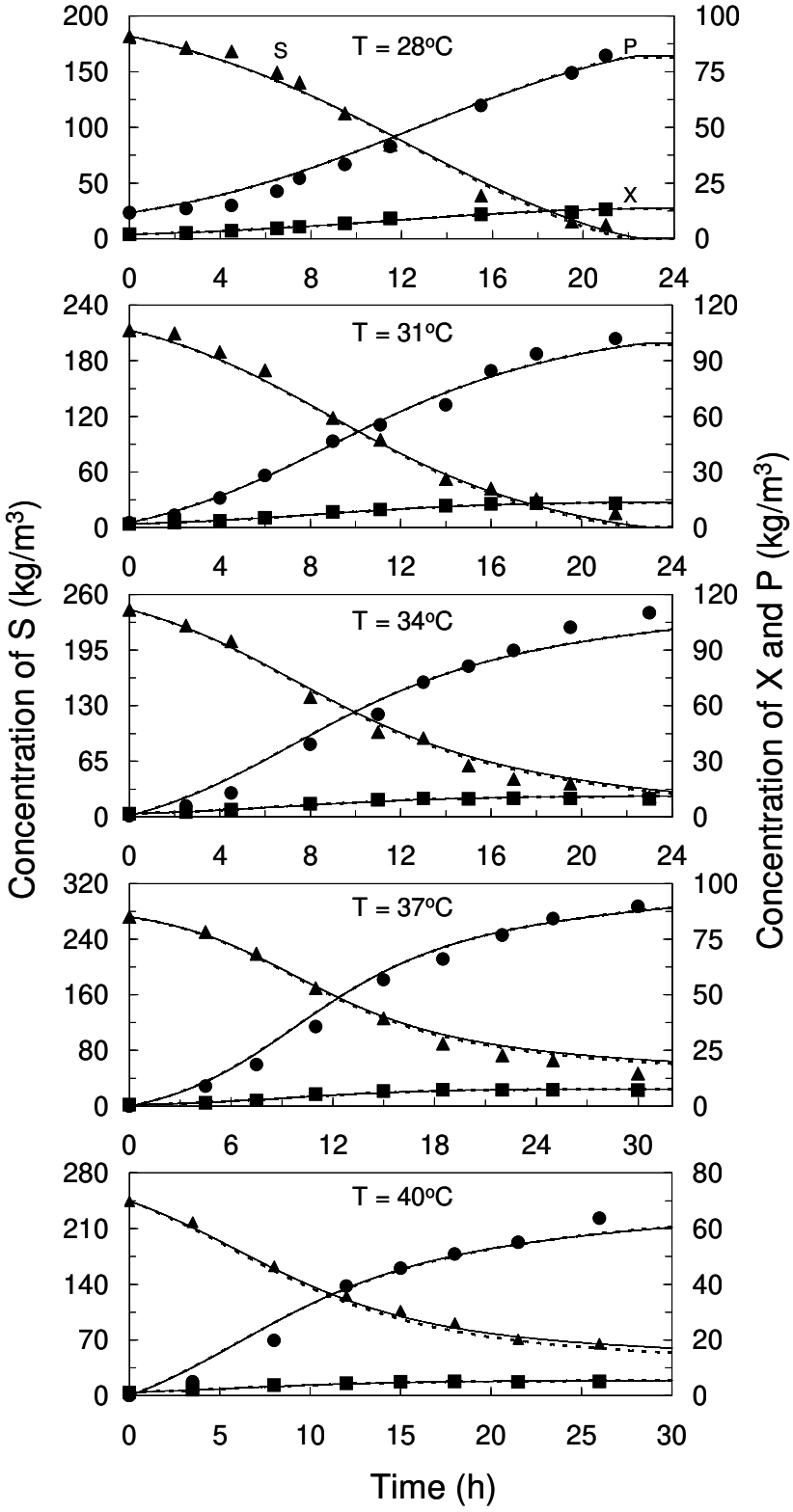


Figure 3.2. Experimental and simulated data for batch fermentation experiments at 28, 31, 34, 37 and 40°C. The experimental data are for concentrations of substrate, S (▲); ethanol, P (●) and cell mass, X (■). Simulated results are represented by lines (with re-estimation —; without re-estimation ---)

Once the parameters were estimated for each temperature, the function described by Eq. (7) was used to fit the parameters as functions of temperature. Curves of optimized parameters versus temperature presented coefficient of determination greater than 98%. In the current methodology, low coefficients of determination would induce to additional error in the model.

The formulation of optimization criteria (objective function) on the basis of process simulation is one of the crucial steps in the application of parameter estimation problem. The determination of the feasible region of the total search space in the multiparameter optimization of the deterministic model is complex. For that reason, the optimization was evaluated using different optimization techniques. Quasi-Newton Algorithm (deterministic technique) was compared to the Real-Coded Genetic Algorithm (stochastic technique) to solve the problem of parameter estimation. From the computational results, it was observed that Real-Coded Genetic Algorithm (stochastic technique) provided an alternative solution. The most significant advantages of Genetic Algorithms are that it avoids the initial guess selection problem, provides several acceptable local solutions and are well suited to large-scale problems since it does not demand the computation of any Hessian matrix. This leads Genetic Algorithms to have relatively small memory and processing requirements. Also, it ensures the convergence of the optimization procedure.

Finally, it should be noted that the QN and RGA can be successfully applied in parameter estimation problem for batch fermentation and can be used as reliable tools for optimizing other types of fermentation processes.

Acknowledgements

The authors acknowledge FAPESP and CNPq for financial support.

Nomenclature

K_i substrate inhibition parameter (m^3/kg)

K_s substrate saturation parameter (kg/m^3)

m parameter used to describe cellular inhibition

m_p ethanol production associated with growth (kg/[kg·h])

m_x maintenance parameter (kg/[kg·h])

n parameters used to describe product inhibitions

P product concentration (kg/m^3)

P_{max} product concentration when cell growth ceases (kg/m^3)

r_p kinetic rate of product formation (kg/[m^3·h])

r_s kinetic rate of substrate consumption (kg/[m^3·h])

r_x kinetic rate of growth (kg/[m^3·h])

S substrate concentration (kg/m^3)

T temperature into the fermentor (^{o}C)

X biomass concentration (kg/m^3)

X_{max} biomass concentration when cell growth ceases (kg/m^3)

Y_{px} yield of product based on cell growth (kg/kg)

Y_x limit cellular yield (kg/kg)

μ_{max} maximum specific growth rate (h^{-1})

References

1. Goldemberg, J., Coelho, S.T., Nastari, P.M. and Lucon, O., Ethanol learning curve – the Brazilian experience. *Biomass and Bioenergy*, 2004, **26** 301-304.
2. Zaldivar, J., Nielsen, J. and Olson, O., Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Applied Biochemistry and Biotechnology*, 2001, **56**(1-2) 17-34.
3. Malherbe, S., Fromion, V., Hilgert, N. and Sablayrolles, J.M., Modeling the effects of assimilable nitrogen and temperature on fermentation kinetics in enological conditions. *Biotechnology and Bioengineering*, 2004, **86**(3) 261-272.
4. Atala, D.I.P., Costa, A.C., Maciel, R. and Maugeri, F., Kinetics of ethanol fermentation with high biomass concentration considering the effect of temperature. *Applied Biochemistry and Biotechnology*, 2001, **91-93**(1-9) 353-366.

5. Aldiguier, A.S., Alfenor, S., Cameleyre, X., Goma, G., Uribelarrea, J.L., Guillouet, S.E. and Molina-Jouve, C., Synergistic temperature and ethanol effect on *Saccharomyces cerevisiae* dynamic behaviour in ethanol bio-fuel production. *Bioprocess and Biosystem Engineering*, 2004, **26** 217-222.
6. Kalil, S.J., Maugeri, F. and Rodrigues, M.I., Response surface analysis and simulation as a tool for bioprocess design and optimization. *Process Biochemistry*, 2000, **35**(6) 539-550.
7. Dourado, A., Goma, G., Albuquerque, U. and Sevely, Y., Modeling and static optimization of the ethanol production in a cascade reactor. I. Modeling. *Biotechnology and Bioengineering*, 1986, **29** 187-194.
8. Andrietta, S.R. and Maugeri, F., Optimum design of a continuous fermentation unit of an industrial plant for alcohol production. *Advances in Bioprocess Engineering*, 1994, **1** 47-52.
9. Pascal, F., Dagot, C., Pingaud, H., Corriou, J.P., Pons, M.N. and Engasser, J.M., Modeling of an industrial alcohol fermentation and simulation of the plant by a process simulator. *Biotechnology and Bioengineering*, 1994, **46** 202-217.
10. Schügerl, K., Progress in monitoring, modeling and control of bioprocess during the last 20 years. *Journal of Biotechnology*, 2001, **85** 149-173.
11. Park, L.J., Park, C.H., Park, C. and Lee, T., Application of genetic algorithms to parameter estimation of bioprocesses. *Medical and Biological Engineering and Computing*, 1997, **35** 47-49.
12. Park, T.Y. and Fromet, G.F., A hybrid genetic algorithm for the estimation of parameters in detailed kinetic models. *Computers and Chemical Engineering*, 1998, **22** S103-S110.
13. Sáiz, J.M.G., Pizarro, C., Vidal, D.G., Evaluation of kinetic models for industrial acetic fermentation: Proposal of a new model optimized by genetic algorithm. *Biotechnology Progress*, 2003, **19** 599-611.
14. Lee, S.S., Robinson, F.M. and Wang, H.Y., Rapid determination of yeast viability. *Biotechnology Bioengineering Symposium*, 1981, **11** 641-649.

15. Luedeking, R. and Piret, E.L., Transient and steady states in continuous fermentation. Theory and experiment. *Journal of Biochemical and Microbiological Technology and Engineering*, 1959, **1** 431-450.
16. Akerberg, C., Hofvendahl, K., Zacchi, G. and Hahn-Hagerdal, B., Modelling the influence of pH, temperature, glucose and lactic aced concentrations on the kinetics of lactic acid production by *Lactococcus lactis* ssp. *lactis* ATCC 19435 in whole-wheat flour. *Applied Microbiology and Biotechnology*, 1998, **49** 682-690.
17. Petzold, L.R., Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. Report SAND80-8230, Sandia National Laboratories, Livermore, California, 1980.
18. Yalçin, K. S. and Özbas Z. Y., Effects of different substrates on growth and glycerol production kinetics of a wine yeast strain *Saccharomyces cerevisiae* Narince 3. *Process Biochemistry*, 2004, **39** 1285-1291.
19. Yedder, R.B., Optimisation numérique et Contrôle optimal: (applications en chimie moléculaire). PhD thesis, Ecole Nationale des Ponts et Chaussées (ENPC), Paris, France, 2002.
20. Goldberg, D.E., *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Massachusetts, 1989.
21. Michalewicz, Z., Genetic algorithms+Data Structures=Evolution Programs. Springer, 3rd ed., New York, 1996.

CAPÍTULO 4. AN ADAPTIVE HYBRID MODEL FOR BIOETHANOL PRODUCTION BY CONTINUOUS EXTRACTIVE FERMENTATION

Capítulo do livro **Computer Aided Process Engineering**, v. 21, 2006, editado pela **ELSEVIER BV** (versão ampliada).
Elmer Ccopa Rivera, Aline C. Costa, Rubens Maciel Filho.

Abstract

In this work, an adaptive hybrid model that considers the effect of temperature on the kinetics was implemented. The hybrid model for bioethanol production by continuous extractive fermentation combines mass and energy balance equations with neural networks that describe the process kinetics. Initially, an off-line training is performed in order to choose appropriate neural network architecture and in a second stage, the neural network parameters are re-estimated in an adaptive scheme that simulated changes in the microorganism kinetic behavior. The potential of a real-coded genetic algorithm and a binary-coded genetic algorithm to re-estimate the parameters of the neural network is evaluated. These algorithms were compared to the quasi-newton algorithm in terms of performance of the hybrid model. The results show that the proper choice of the neural network architecture and the use of an appropriate optimization algorithm results in a robust model, allowing further on-line optimization and control studies.

Keywords

Bioreactors, fermentation, optimization, mathematical modeling, multilayer perceptron neural network, genetic algorithms.

4.1 Introduction

Bioethanol (ethanol from biomass) is an attractive, sustainable energy source to fuel transportation. Based on the premise that fuel bioethanol can contribute to a cleaner environment and with the implementation of environmental protection laws in many countries, demand for this fuel is increasing (Zaldivar et al., 2001).

Impelled by the creation of the "Proálcool" program in the 70's and 80's, Brazil has become the world largest ethanol producer via fermentation, developing and improving many fermentative processes (Goldemberg et al., 2004). Nowadays, Brazilian annual production capacity of ethanol is over 18 billions of liters and with the increase in demand, the objective is to multiply this capacity through the installation of new factories and operation optimization of the existing ones. In

consequence, there is an intensified interest in the study of all the steps involved in ethanol production.

The lack of robustness of fermentation in the presence of fluctuations in the quality of the raw material, which leads to changes in the kinetic behaviour with impact on yield, productivity and conversion, is among the main current aspects related to the alcoholic fermentation process. Changes in the operational conditions are quite common in plants of alcoholic fermentation; they occur not only due to the changes in the quality of the raw material but also due to variations of dominant yeast in the process. One of the sources of raw material quality fluctuations is the alteration in the amounts of treated cane juice and molasses used in the composition of the medium fed to the fermentor. These alterations are made due to the sugar factories operation. Furthermore, molasses undergoes variation of composition in different crops. On the other hand, variations in the yeast population are frequent due to the high microbial load in the raw material.

The lack of robustness can be corrected by adjustments in the operational and control parameters of the process when fluctuations occur. In order to accomplish this, it is important that a mathematical model is available to aid in the decision making, mainly when the difficulties of monitoring the key process variables (concentrations of biomass, substrate and product, and temperature) are taken into account. However, the operational changes described make the prediction of the dynamic behavior of the process with a single model difficult, as they lead to changes in microorganism kinetics. Thus, it would be of great advantage to have a mathematical model that could be easily adapted to changes in operational conditions. This can be accomplished by re-estimating the kinetic parameters when necessary.

The frequent re-estimation of parameters of deterministic models, however, is usually slow and difficult, mainly due to non-linearities, great number of parameters and interactions among them. The use of hybrid neural modeling techniques can be a good option in this case. These models are based on first principles (physical knowledge) associated with artificial neural networks (ANN)

which represent the kinetic relations (Georgieva et al., 2003; Harada et al., 2002; Psichogios and Ungar, 1992; Zorzetto et al., 2000). Thus, the problem of re-estimating the kinetic parameters is transformed into a problem of updating the neural network parameters, which is a simpler procedure.

Several architectures of neural networks are available in the literature and there are no methods to choose the best one for a given application (Tarca et al., 2004). Multilayer perceptron neural network (MLPNN) is the most widely used type of ANN. When used in hybrid neural models, an appropriate topology may fail to produce the best hybrid model, unless a suitable training algorithm is used.

Genetic Algorithms (GAs), have often been found to be as powerful and effective as gradient search methods in many engineering applications (Modak and Sarkar, 2005). They require only the objective function values and are highly likely to locate the global minimum. Genetic algorithms have been successfully utilized for training neural networks (Fogel et al., 1990; Sexton et al., 1998; Whitley et al., 1990).

In this work, an adaptive hybrid model for bioethanol production by extractive continuous fermentation process was implemented. The adaptive scheme was performed using real-coded and binary-coded genetic algorithms. In a previous work, genetic algorithms were employed with success for modeling the effect of temperature on the kinetics of a batch fermentation process for ethanol production (Ccopa Rivera et al., 2006a) and to determine the optimal operational conditions of a fermentation process modeled with neural networks (Ccopa Rivera et al., 2006b).

The present work extends this area of investigations using the potential of the genetic algorithms in a hybrid approach that considers the effect of temperature on the kinetics. Thus, the adaptive hybrid model should be easily updated in the case of changes in operational conditions and fluctuations in the quality of raw material, both very frequent situations in alcoholic fermentation plants.

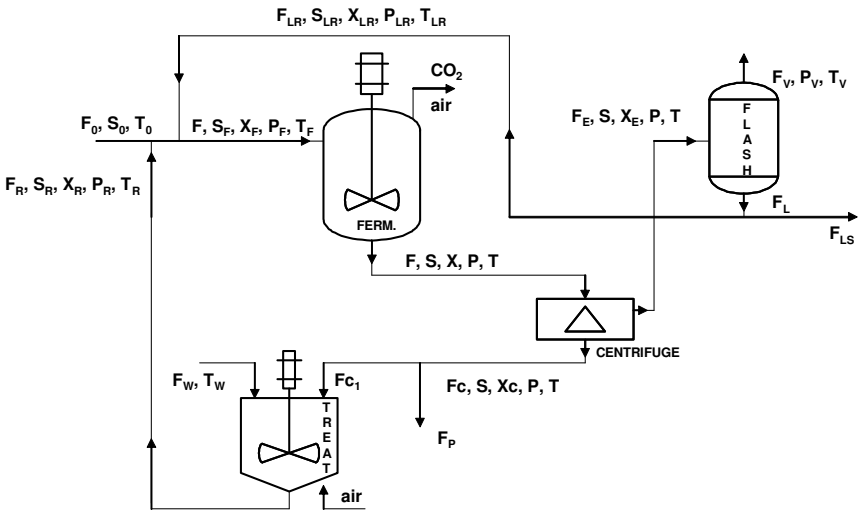


Figure 4.1. Extractive alcoholic fermentation scheme

4.2 Process description and modeling

4.2.1 Extractive continuous alcoholic fermentation process

High concentrations of ethanol inhibit the fermentation process, particularly when a fermentative medium with high concentration of substrate is used, as is the case in the majority of the industrial processes. Taking these into account, Silva et al. (1999) studied a process of fermentation combined with a vacuum flash vessel, which selectively extracts ethanol from the medium as soon as it is produced. These authors have shown that this scheme presents many positive features and better performance than a conventional industrial process (Andrietta and Maugeri, 1994).

The extractive fermentation process for bioethanol production proposed by Silva et al. (1999) is shown in Figure 4.1. The process is composed of four interlinked units: fermentor (ethanol production unit), centrifuge (cell separation unit), cell treatment unit and vacuum flash vessel (ethanol–water separation unit). This scheme attempts to simulate industrial conditions (Andrietta and Maugeri, 1994), with the difference that only one fermentor is used instead of a cascade system, besides the flash is used to extract part of the ethanol.

The mathematical modeling of the process consists of mass and energy balance equations. All equipment, with the exception of the fermentor, are modeled assuming the hypothesis of 'pseudo' steady state.

Assuming constant volume, the mass and energy balance equations for the fermentor using the intrinsic model (Monbouquette, 1987, 1992) are as follows:

Viable cells: $$\frac{dX_v}{dt} = (\mu_1 - \mu_2)X_v - \frac{F}{V}(X_v - X_{vF}) \quad (1)$$

Dead cells: $$\frac{dX_d}{dt} = \mu_2 X_v - \frac{F}{V}(X_d - X_{dF}) \quad (2)$$

substrate: $$\frac{d\left[\left(1 - \frac{X_t}{\rho}\right)SV\right]}{dt} = F(S_F - S) - V\sigma X_v \quad (3)$$

product: $$\frac{d\left[\left(1 - \frac{X_t}{\rho}\right)PV + \frac{X_t}{\rho}\gamma PV\right]}{dt} = V\pi X_v + F(P_F - P) \quad (4)$$

$$\frac{dT}{dt} = D(T_F - T) + \frac{\Delta Hr_s}{\rho_m Cp} \quad (5)$$

In these equations X_v is the concentration of the viable biomass (kg/m^3), X_d, is the concentration of dead biomass (kg/m^3), X_t is the total biomass (kg/m^3), S is the concentration of substrate (kg/m^3), P is the concentration of ethanol (kg/m^3) and T is the fermentor temperature (^{o}C). The specific rates of cell growth, μ_1 (h^{-1}), cell death, μ_2 (h^{-1}), substrate consumption, σ (h^{-1}), and product formation, π (h^{-1}), are given by Eqs. 6-9.

The specific kinetic rates of growth, death, ethanol formation and substrate consumption are as follows:

$$\mu_1 = \mu_{\max}\frac{S}{K_s + S}\exp(-K_i S)\left(1 - \frac{X_t}{X_{\max}}\right)^m\left(1 - \frac{P}{P_{\max}}\right)^n \quad (6)$$

$$\mu_2 = K_{dT}\exp(K_{dP}P) \quad (7)$$

$$\pi = Y_{px}\mu_1 + m_p \quad (8)$$

$$\sigma = \frac{\mu_1}{Y_x} + m_x \quad (9)$$

The parameters were adjusted as functions of the temperature from experimental data and are given in Atala et al. (2001). A detailed description of the process and mathematical model can be found in Costa et al. (Costa et al., 2001).

This deterministic model was used to simulate the extractive continuous alcoholic fermentation process for the development and validation of the hybrid model.

4.2.2 Building up the hybrid model

In this work, the structure of the hybrid model is derived considering the mass and energy balances (Eqs. 1-5) for the fermentor and all the other process units (see Figure 4.1), with a MLPNN describing the specific rate expression for cell growth, μ_1 (Eq. 6). If a neural network describes μ_1 well, the development of neural networks to describe π and σ , Eq. 8 and Eq. 9, respectively, is straightforward, because these specific rates are described by variables multiplied by μ_1.

A MLPNN consists of three types of layers: an input layer, an output layer and one or more hidden layers (see Figure 4.2). Each layer may have a different number of neurons, and even a different transfer function. An appropriate architecture would help to achieve higher model accuracy (Haykin, 1994).

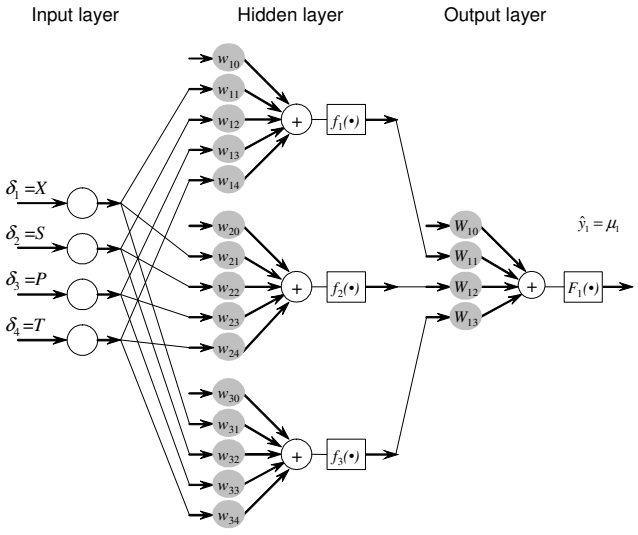


Figure 4.2. Representation of a multilayer perceptron neural network with one hidden layer

For the current work, the specific rate of cell growth, expressed by Eq. 6, was modeled with a multilayer perceptron neural network with four inputs (concentrations of biomass, substrate and product, and temperature) and a single hidden layer, as shown in Figure 4.2 and described mathematically by Eq. 10.

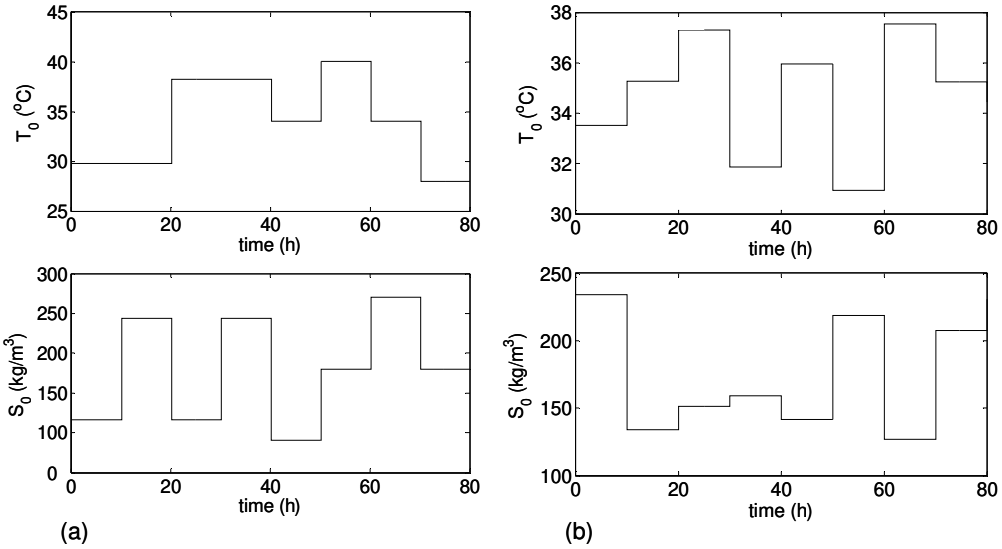


Figure 4.3. Disturbances in T_0 and S_0, used to obtain the optimal neural network architecture using the cross-validation technique: (a) Disturbances using a factorial design to generate the training data set; (b) Random disturbances to generate the validation data set

It follows from Cybenko's theorem (Cybenko, 1989) that all continuous functions can be approximated to any desired accuracy with a network of one hidden layer of sigmoidal ($f(\delta) = 1/(1+\exp^{-\delta})$) hidden units (nodes) and a layer of linear output nodes. Such structure was used in this work.

$$\hat{y}_1 = g_1[\delta, \theta] = F_1[\sum_{j=1}^{3} W_{1j} f_j(\sum_{l=1}^{4} w_{jl}\delta_l + w_{j0}) + W_{10}] \tag{10}$$

In this equation, θ specifies the parameters vector, which contains all the adjustable parameters of the network; i.e., the weights and biases $\{w_{j,l}\ W_{1,j}\}$. The Levenberg–Marquardt algorithm in Matlab's neural network toolbox was used for the parameters adjustment for the current off-line training that is used to choose neural network architecture.

The appropriate number of nodes to include in the hidden layer was addressed with the cross-validation technique that splits the sample data into a

training sample and a validation sample. Then neural networks with different hidden nodes are trained with the estimation sample, and their performance monitored with the validation sample.

The training was performed using a full factorial design 2^2+star configuration (Milton and Arnold, 1990), involving the variables S_0 (kg/m^3) and T_0 (^{o}C). Both variables had their values changed every 10 hours (this period was sufficient to reach the steady state) within the ranges 90 kg/m^3 to 270 kg/m^3 (S_0) and 28°C to 40°C (T_0). Thus, representative data containing the input and output signals of the process were generated. Table 4.1 shows the coded factor levels and the real values for the disturbance inputs. For validation, the disturbances were also a sequence of steps with random changes every 10 hours within the operational interval 120 kg/m^3 to 230 kg/m^3 (S_0) and 30°C to 38°C (T_0). The disturbances used to generate the training data set are shown in Figure 4.3a and those used to generate the validation data set in Figure 4.3b.

Table 4.1. Coded factor levels and real values for factorial design, used to generate a training data set, used to choice of the neural network architecture.

Variable	-α (-1.4142)	-1	0	+1	-α (-1.4142)
T_0 (^{o}C)	28	29.75	34	38.24	40
S_0 (kg/m^3)	90	116.36	180	243.63	270

Using the cross-validation criterion, the network with three hidden nodes was found to be appropriate to avoid model over-fitting. Therefore, the specific rate of growth, μ_1, was modeled by a MLPNN containing 19 scalar parameters.

4.3 Results and Discussion

4.3.1 Performance evaluation of the hybrid model trained with off-line measurements

The prediction quality of the hybrid model can be characterized using the residual standard deviation (RSD), Eq. 11, which provides an indication of the accuracy of the prediction, as suggested by Cleran et al. (1991):

$$\mathrm{RSD} = \frac{\sqrt{\sum_{i=1}^{n}(y_i - y_{Pi})^2}}{n} \tag{11}$$

In Eq. 11, y_i is the "real" value (deterministic model), y_{Pi} is the value predicted by the hybrid model and n is the number of points. As the magnitude of the RSD varies depending on the magnitude of the variable to be predicted, it is easier to analyze the RSD as a percentage of the average of the "real" values $\bar{y}_i$ (Atala et al., 2001), Eq. 12.

$$\text{RSD}(\%) = \frac{\text{RSD}}{\bar{y}_i} \times 100 \tag{12}$$

In addition, the prediction quality of the hybrid model can be also characterized using the correlation coefficient (COR) (Box et al., 1978), Eq. 13.

$$\text{COR} = \left(1 - \frac{\text{SEE}}{\text{S}_{\text{TT}}}\right) \times 100 \tag{13}$$

where $\text{SEE} = \sum_{i=1}^{n} (y_i - y_{pi})^2$ and $\text{S}_{\text{TT}} = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2$.

Figure 4.4 shows that the hybrid model described well the dynamic behavior of the process when the process is subjected to the random disturbances shown in Figure 4.3b.

Table 4.2 shows the prediction results of the hybrid model for concentrations of biomass, substrate and product, and temperature. It can be seen that all the variables present correlations greater than 99% and the RSD(%) in the range of 0.019 to 0.649.

Table 4.2. Comparison between the correlation coefficient, COR, and the residual standard deviations, RSD(%), of the average of "real" values, to characterize the prediction quality of the hybrid model trained with the off-line measurements

Output variable	COR(%)	RSD(%)
Biomass (kg/m^3)	99.976	0.019
Substrate (kg/m^3)	99.983	0.649
Product (kg/m^3)	99.872	0.359
Temperature (^{o}C)	99.985	0.042

4.3.2 Adaptive hybrid scheme

The proposed scheme for adaptation of the hybrid model is shown in Figure 4.5 and should be used always that changes in process kinetic behavior make re-estimation of the kinetic parameters to be necessary. In this procedure, the

architecture of the MLPNN is considered known (defined by off-line training) and the neural network parameters are re-estimated using on-line measured data.

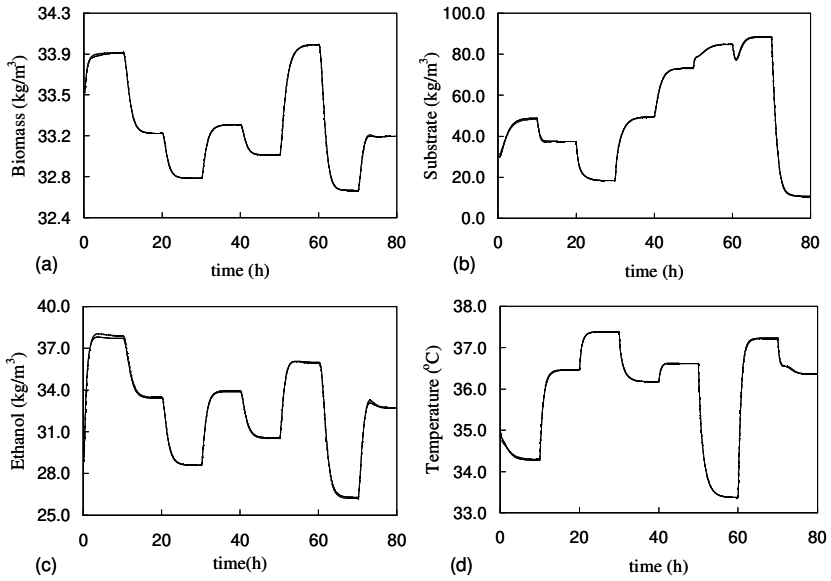


Figure 4.4. (a) Biomass concentration, (b) substrate concentration, (c) ethanol concentration and (d) temperature. Solid curves obtained by using the deterministic model. Dashed curves obtained by using the hybrid model training with off-line measurements

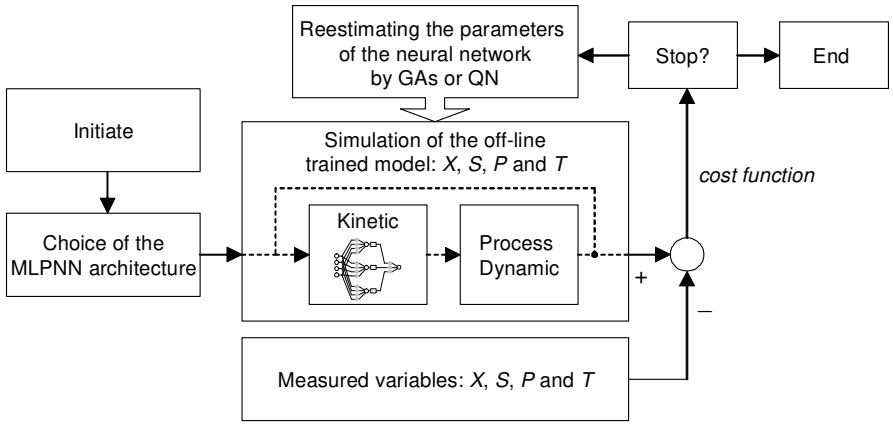


Figure 4.5. Adaptive scheme that re-estimates the parameters of the neural network

The adaptive algorithm minimizes the following cost function, which is dependent of θ, the vector containing the weights of the neural network:

$$E(\theta) = \sum_{i=1}^{n} \left| y_i - y_{Pi} \right| = \sum_{i=1}^{n} \left| e_i(\theta) \right| \qquad (14)$$

In the Eq. 14, y_i are the measured values of concentrations of biomass, substrate and product, and temperature; y_{Pi} are the values of concentrations of biomass, substrate and product, and temperature predicted by the off-line trained hybrid model and n is the number of points. The term $|e_i(\theta)|$ is the absolute value of the error in the output due to the ith sample. For the current work, the weights vector, θ, was re-estimated by a real-coded genetic, a binary-coded genetic algorithm and a quasi–newton algorithm.

The performance of the proposed adaptive scheme to re-estimate the neural network parameters was tested by simulating changes in the microorganism kinetic behavior. In order to accomplish this, a modified deterministic model was used, in which the specific growth rate, Eq. 6, was replaced by another specific growth rate expression, given by Dourado et al. (1987) and shown in Eq. 15. This modification aims to simulate the changes in kinetic behavior expected in real ethanol production plants. So, in the test of the adaptive scheme, the modified deterministic model represents the real plant and the concentrations of biomass, substrate and product and temperature determined using it are considered on-line measured variables.

The off-line model was trained using data of the original model and, without updating of its parameters, will not describe the new dynamic behavior appropriately. The simulation of the modified deterministic model, in these conditions, employed the same random disturbance in T_0 and S_0 used to evaluate the performance of the hybrid model trained with off-line measurements.

$$\mu_1 = \mu_{\max}\left(\frac{S}{K_s + S}\right)\left(\frac{K_p}{K_p + P}\right)\left(1 - \frac{P}{P_{\max}}\right) \tag{15}$$

In this equation, K_p and K_s are kinetic constants equal to 5 kg/m^3 and 4.5 kg/m^3 respectively, $\mu_{\max}$ is 0.3 h^{-1} and $P_{\max}$ is 85 kg/m^3.

4.3.3 Re-estimating the parameters of the neural network by genetic algorithms

The GAs considered in the present paper are based on the freeware versions written in FORTRAN of a real-coded genetic algorithm (RGA) developed by Yedder

(2002) and a binary-coded genetic algorithm (BGA) developed by Carroll (1998), adapted and enhanced to re-estimate the parameters of the neural network. These algorithms were employed with success in a previous work by the authors (Ccopa Rivera et al., 2006a, 2006b).

When re-estimating the neural network parameters using GAs, the different neural networks (population of chromosomes) were vectors where each real-value or binary-value (gene) stands for a parameter of the neural network. The stopping criterion was selected simply when the maximum number of iterations (generations) fixed is reached. The RGA and BGA parameters were adjusted to minimize the number of generations required to reach a satisfactory fitness value (minimization of the cost function, Eq. 14), reflecting good agreement between y_i and y_{Pi}. The main technical features of the genetic algorithms are reported in Table 4.3.

Table 4.3. Main technical features used by the genetic algorithms in the adaptive scheme

Option chosen - RGA	Parameters	Parameter values
Niching, elitism, barycentric crossover, non-uniform mutation	Individual length	19
	Population size	10
	Crossover probability	0.6
	Mutation probability	0.0015
Option chosen - BGA		
Elitism, single-point crossover, jump mutation	Individual length	19
	Population size	10
	Crossover probability	0.5
	Mutation probability	0.02

Table 4.4. Comparison between the correlation coefficient, COR, and the residual standard deviations, RSD(%), of the average of "real" values, to characterize the prediction quality of the hybrid model where the neural network parameters were re-estimated by a real-coded genetic algorithm (RGA), a binary-coded genetic algorithm (BGA) and a quasi-newton algorithm (QN)

Output variable	COR(%)			RSD(%)		
	BGA	RGA	QN	BGA	RGA	QN
Biomass (kg/m^3)	98.2	98.4	97.4	0.03	0.03	0.04
Substrate (kg/m^3)	99.9	99.9	99.9	0.8	0.8	0.9
Product (kg/m^3)	99.1	99.2	97.8	0.8	0.8	0.9
Temperature (^oC)	99.9	99.9	99.9	0.10	0.09	0.11

4.3.4 Re-estimating the parameters of the neural network by quasi-newton algorithm

The quasi-newton algorithm was used to minimize the cost function given by Eq. 14. The FORTRAN IMSL routine DBCONF was used for this purpose. The straightforward idea is to implement the optimization problem as a nonlinear programming problem (NLP) and can be written as:

Minimize Equation 14

Subjet to $l_p \leq w_p \leq u_p$, p=1,…,19

where w_p is the neural network weights. l_p and u_p are specified lower and upper bounds on the parameter, with $l_p \leq u_p$

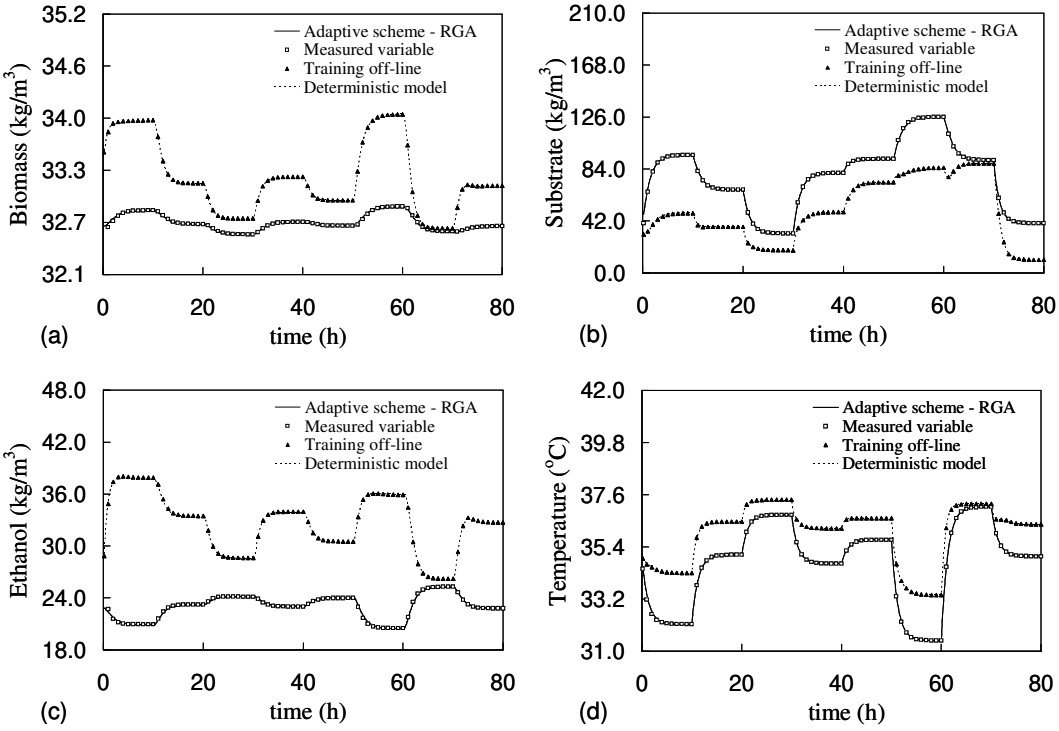


Figure 4.6. (a) Biomass concentration, (b) substrate concentration, (c) ethanol concentration and (d) temperature. Dashed curves obtained by using the deterministic model. Triangle data points obtained by using the hybrid model trained with off-line measurements. Solid curves obtained by using the proposed adaptive scheme using the real-coded genetic algorithm. Square data points for "on-line measured data" (modified deterministic model that used a specific rate of growth cited by Dourado et al. (1987), Eq. 15)

4.3.5 Performance evaluation of the adaptive hybrid model

A comparison was carried out between the hybrid model where the neural network parameters were re-estimated by a real-coded genetic algorithm (RGA), a binary-coded genetic algorithm (BGA) and a quasi-newton algorithm (QN). Table 4.4 shows the prediction quality of the adaptive hybrid model for concentrations of biomass, substrate and product, and temperature. This table includes the residual standard deviation, RSD(%), and the correlation coefficient, COR. For all the algorithms tested, it can be seen that the hybrid model reflected well the dynamic behavior of the modified deterministic model that used a specific rate of growth cited by Dourado et al. (1987), Eq. 15. The results obtained using RGA and BGA to re-estimate the neural network parameters were slightly better than the QN results.

Figure 4.6 shows the performances of the original deterministic model, the off-line trained hybrid model and the adaptive hybrid model in describing the concentrations of biomass, substrate and product, and temperature when the specific growth rate is described by equation 15. It can be seen that in the presence of changes in dynamic behavior, neither the original deterministic model nor the off-line trained hybrid model present good description of the data. The adaptive hybrid model, however, described the dynamic behavior accurately. As the performances of the adaptive algorithms using RGA, BGA and QN were similar, in Figure 4.6 it is shown only the results using RGA.

4.4. Concluding remarks

Fermentation processes are quite difficult to model, since their operation involves microbial growth under constantly changing conditions. Hence, there is a need for the development of simple though realistic mathematical models, which can be used to develop optimal operating strategies. The objective is to avoid the decisions based on infrequent measured data generally used in plants of fermentation due to the lack of a robust mathematical model.

This study was focused on modeling an extractive continuous alcoholic fermentation process using a hybrid strategy in order to obtain a reliable model to

be used in further evaluation of control structures, optimization and design of process controllers.

Neural network modelling is a powerful and flexible tool and its use in the alcoholic fermentation process can be advantageous because updating the neural network parameters is a simpler procedure than re-estimating kinetic parameters of deterministic models. Frequent re-estimation of kinetic parameters (or in the case of this work re-estimating the neural network parameters) is necessary due to changes in the dynamic behaviour caused by fluctuations in the quality of the raw material, changes in microorganism metabolism and variation of the dominant yeast strains present in the fermentation process.

In this work an adaptive scheme was proposed to update the weights of the neural network, part of a hybrid model. The adaptive algorithm is started considering an off-line trained hybrid model and minimizes an objective function defined as the absolute error between the measured variables and the variables calculated by the off-line trained hybrid model. The optimization variables are the weights of the neural network.

The performance of the proposed adaptive hybrid model was tested by simulating changes in the microorganism kinetic behavior. It was observed that the real-coded and binary-coded genetic algorithms (stochastic technique) were successfully applied to re-estimate the neural network parameters and were slightly better, in terms of RSD(%) and COR, when compared to the results of quasi-newton algorithm (deterministic technique). Here, the hybrid model took advantage of the natural capacity of storing knowledge and the massively parallel processing of the MLPNN, and the use of appropriate global optimization algorithms, such as genetic algorithms. In many applications, genetic algorithms perform better than gradient search methods for multi-dimensional and high precision problems; in this work the performances of the two stochastic techniques were similar to that of the deterministic technique.

Acknowledgements

The authors acknowledge FAPESP and CNPq for financial support.

Nomenclature:

Cp	Heat capacity, Kcal/(kg.^{o}C)
$D = F / V$	Dilution rate, h^{-1}
F	Feed stream flow rate, m^3/h
F_L	Liquid outflow from the vacuum flash tank, m^3/h
F_{LR}	Liquid phase recycling flow rate, m^3/h
F_{LS}	Liquid phase flow to rectification column, m^3/h
F_0	Fresh medium flow rate, m^3/h
F_R	Cell recycling flow rate, m^3/h
F_V	Vapor outflow from the vacuum flash tank, m^3/h
K_{dP}	Coefficient of death by ethanol, m^3/kg
K_{dT}	Coefficient of death by temperature, h^{-1}
K_{ei}	Equilibrium constant
K_i	Substrate inhibition constant, m^3/kg
K_S	Substrate saturation constant, kg/m^3
m	Constant in Equation 6
m_p	Ethanol production associated to growth, kg/(kg.h)
m_x	Maintenance coefficient, kg/(kg.h)
n	Constant in Equation 6
P	Product concentration into the fermentor, kg/m^3
P_F	Feed product concentration, kg/m^3
P_{LR}	Product concentration in the light phase from centrifuge, kg/m^3
P_{max}	Product concentration when cell growth ceases, kg/m^3
P_V	Product concentration in the vapor phase from the flash tank, kg/m^3
$r = F_{LR} / F_L$	Flash recycle rate
r_d	Kinetic rate of death, $kg/(m^3.h)$
r_p	Kinetic rate of product formation, $kg/(m^3.h)$

r_s	Kinetic rate of substrate consumption, kg/(m^3.h)
r_x	Kinetic rate of growth, kg/(m^3.h)
$R = F_R / F$	Cells recycle rate
S	Substrate concentration into the fermentor, kg/m^3
S_F	Feed substrate concentration, kg/m^3
S_0	Inlet substrate concentration, kg/m^3
T	Temperature into the fermentor, ^{o}C
T_F	Feed temperature, ^{o}C
t_r	Residence time, h
V	Reactor volume, m^3
X_d	Dead biomass concentration into the fermentor, kg/m^3
X_{dF}	Dead biomass concentration in the feed stream, kg/m^3
X_F	Feed biomass concentration, kg/m^3
$X_{\max}$	Biomass concentration when cell growth ceases, kg/m^3
$X_t = X_v + X_d$	Total biomass concentration into the fermentor, kg/m^3
X_v	Viable biomass concentration into the fermentor, kg/m^3
X_{vF}	Viable biomass concentration in the feed stream, kg/m^3
Y_{px}	Yield of product based on cell growth, kg/kg
Y_x	Limit cellular yield, kg/kg
ΔH	Reaction heat, kcal/(kg TRS)
γ	Ratio of concentration of intracellular to extracellular ethanol, kg/m^3
μ_1	Specific rates of growth, (h^{-1})
μ_2	Specific rates of death, (h^{-1})
$\mu_{\max}$	Maximum specific growth rate, h^{-1}
ρ	Ratio of dry cell weight per wet cell volume, kg/m^3
ρ_m	Density, kg/m^3
π	Specific rates of product formation, (h^{-1})

σ Specific rates of substrate consumption, (h^{-1})

References

1. Andrietta, S.R., Maugeri, F., (1994). Optimum design of a continuous fermentation unit of an industrial plant for alcohol production. *Advances in Bioprocess Engineering*, **1**, 47-52.
2. Atala, D.I.P., Costa, A.C., Maciel Filho, R., Maugeri, F., (2001). Kinetics of ethanol fermentation with high biomass concentration considering the effect of temperature. *Applied Biochemistry and Biotechnology*, **91-93**(1-9), 353-366.
3. Box, G.E.P., Hunter, W.G., Hunter J.S., (1978). Statistics for experimenters: An introduction to design, data analysis, and model building, John Wiley and Sons, New York.
4. Carroll, D.L., (1998). FORTRAN Genetic Algorithm Driver, Last updated 12-11-1998. University of Illinois, USA (carroll@cuaerospace.com).
5. Ccopa Rivera E., Costa, A.C., Atala, D.I.P., Maugeri, F., Wolf, M.R., Maciel, R., (2006a). Evaluation of optimization techniques for parameter estimation problems of ethanol fermentation considering the effect of temperature, *Process Biochemistry*, **41**, 1682-1687.
6. Ccopa Rivera E., Costa, A.C., Wolf, M.R., Maciel Filho, R., (2006b). Ethyl alcohol production optimization by coupling genetic algorithm and multilayer perceptron neural network, *Applied Biochemistry and Biotechnology*, **129-132**, 969-984.
7. Cleran, Y., Thibault, J., Cheruy, A., Corrieu, G., (1991). Comparison of Predition Performances Between Models Obtained by the Group Method of Data Handling and Neural Networks for the Alcoholic Fermentation Rate in Enology. *Journal of Fermentation and Bioengineering*, **71**(5), 356-362.
8. Costa, A.C., Atala, D.I.P., Maugeri, F., Maciel Filho, R., (2001). Factorial design and simulation for the optimization and determination of control structures for an extractive alcoholic fermentation. *Process Biochemistry*, **37**, 125-137.

9. Cybenko, G., (1989). Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals, and Systems*, **2**, 303-314.
10. Dourado, A., Goma, G., Albuquerque, U., Sevely, Y., (1987). Modeling and static optimization of the ethanol production in a cascade reactor. I. Modeling. *Biotechnology and Bioengineering*, **29**, 187-194.
11. Fogel, D. B., Fogel, L. J., Porto, V. W., (1990). Evolving neural networks. *Biological Cybernetic*, **63**, 487-493.
12. Georgieva, P., Meireles, M.J., Feyo de Azevedo, S., (2003). Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena. *Chemical Engineering Science*, **58**, 3699-3713.
13. Goldemberg, J., Coelho, S.T., Nastari, P.M., Lucon, O., (2004). Ethanol learning curve – the Brazilian experience. *Biomass and Bioenergy*, **26**, 301-304.
14. Harada, L.H.P., Costa, A.C., Maciel Filho, R., (2002). Hybrid neural modeling of bioprocess using functional link networks. *Applied Biochemistry and Biotechnology*, **98-100**(1-9), 1009-1024.
15. Haykin, S., (1994). Neural networks: A comprehensive foundation, Macmillan College Publishing Company, 2nd ed., New York.
16. Milton, J.S., Arnold, J.C., (1990). Introduction to probability and statistics, McGraw Hill, 2nd ed., New York.
17. Modak, J.M., Sarkar, D., (2005). Pareto-optimal solutions for multi-objective optimization of fed-batch bioreactors using nondominated sorting genetic algorithm. *Chemical Engineering Science*, **60**, 481-492.
18. Monbouquette, H.G., (1987). Models for high cell density bioreactors must consider biomass volume fraction: Cell recycle example. *Biotechnology and Bioengineering*, **29**, 1075-1080.
19. Monbouquette, H.G., (1992). Modeling high-biomass-density cell recycle fermentors. *Biotechnology and Bioengineering*, **39**,498-503.

20. Psichogios, D.C., Ungar, L.H., (1992). A hybrid neural network-first principles approach to process modeling. *American Institute of Chemical Engineers Journal*, **38**, 1499–1511.
21. Sexton, R.S., Dorsey, R.E., Johnson, J.D., (1998). Toward global optimization of neural networks: a comparison of the genetic algorithm and backpropagation. *Decision Support Systems*, **22**, 171-185.
22. Silva, F.L.H., Rodrigues, M.I., Maugeri, F., (1999). Dynamic modelling, simulation and optimization of an extractive continuous alcoholic fermentation process. *Journal of Chemical Technology and Biotechnology*, **74**, 176-182.
23. Tarca, L.A., Grandjean, B.P.A., Larachi, F., (2004). Designing supervised classifiers for multiphase flow data classification. *Chemical Engineering Science*, **59**, 3303-3313.
24. Whitley, D., Starkweather, T., Bogart, C., (1990). Genetic algorithms and neural networks: Optimizing connections and connectivity. *Parallel Computing*, **14**, 347-361.
25. Yedder, R.B., (2002). Optimisation numérique et Contrôle optimal: (applications en chimie moléculaire). PhD thesis, Ecole Nationale des Ponts et Chaussées (ENPC), Paris, France.
26. Zaldivar, J., Nielsen, J., Olson, O., (2001). Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Applied Biochemistry and Biotechnology*, **56**(1-2), 17-34.
27. Zorzetto, L.F.M., Maciel Filho, R., Wolf-Maciel, M.R., (2000). Process modelling development through artificial neural networks and hybrid models. *Computer and Chemical Engineering*, **24**, 1355–1360.

CAPÍTULO 5. HYBRID NEURAL NETWORK MODEL OF AN INDUSTRIAL ETHANOL FERMENTATION PROCESS CONSIDERING THE EFECT OF TEMPERATURE

Trabalho aceito para publicação na revista **Applied Biochemistry and Biotechnology**, 2007.
Ivana C. C. Mantovaneli, Elmer Ccopa Rivera, Aline C. Costa, Rubens Maciel Filho

Abstract

In this work a procedure for the development of a robust mathematical model for an industrial alcoholic fermentation process was evaluated. The proposed model is a hybrid neural model, which combines mass and energy balance equations with Functional Link Networks to describe the kinetics. These networks have been shown to have a good non linear approximation capability, although the estimation of its weights is linear. The proposed model considers the effect of temperature on the kinetics and has the neural network weights reestimated always that a change in operational conditions occurs. This allows to follow the system behavior when changes in operating conditions occur.

Keywords

Mathematical modeling, alcoholic fermentation, functional link networks, process simulation, kinetic parameters estimation, bioreactors.

5.1 Introduction

A potential substitute for petroleum in Brazil is biomass, particularly sugar cane. The sugar cane industry is the largest alternative commercial energy production program in the world with ethanol and the almost complete use of sugar cane bagasse as fuel. Although the bioethanol production is running for several years it is clear that improvements are required to increase the process performance. The influence of temperature in the kinetics is a very important factor in the alcoholic fermentation process. It is difficult to support a constant temperature during large-scale alcoholic fermentation and variations in temperature affects productivity through changes in kinetics as well as in microorganism lifetime. The temperature in a typical industrial fermentor varies from 33.5ºC (during the night) to 35 ºC (at the end of the day) owing to fluctuations in the cooling water temperature. In plants with poor temperature control the temperature in the fermentor goes up to 40ºC. Thus, an accurate mathematical model must take the influence of temperature into account. Also, changes in operational conditions and fluctuations in raw material composition,

both very common in alcoholic fermentation plants, results in kinetic changes and affects yield and productivity.

Owing to the difficulties described in the paragraph above, the main difficulty in model-based techniques for definition of operational strategies, control and optimization for the ethanol production process is the problem of obtaining an accurate model. Thus, in this work a procedure for the development of robust mathematical models for the alcoholic fermentation process is evaluated. The objective is to obtain a model to be used as a simulator, making simpler not only the development and implementation of new control and optimization techniques, but also retuning of existing controllers as well as finding out new optimal operational conditions when operational changes occur.

In the last years, many studies of the mathematical modeling of alcoholic fermentation process have been performed **(1, 2)**. However, although temperature has an important influence on this process, there are very few works in the literature considering the effect of temperature on the kinetic parameters. Among them are the works of Atala et al. **(3)** and Aldiguier et al. **(4)**. Aldiguier et al., however, determine different values for kinetic parameters in different temperatures, but do not determine a temperature function to describe the kinetic parameters. Although phenomenological mathematical models provide understanding about the process, practical experience shows that they are only valid for the specific conditions in which they were determined. When there are changes in operational conditions, the model kinetic parameters have to be reestimated. The frequent reestimation, mainly when the parameters are described as functions of temperature, is difficult and time-consuming because nonlinearities, great number of parameters and interactions among them.

One way to deal with this problem is to use hybrid neural models **(5, 6)**. As was demonstrated by Psichogious and Ungar **(7)**, it is simple to propose a bioreactor mathematical model through application of mass balances for the process variables. The really difficult part is the mathematical representation of the kinetics. A cell is a complex organism wherein thousands of enzyme catalyzed reactions take place. The kinetic rates are most often poorly understood nonlinear

functions; whereas the corresponding parameters are in general time-varying **(8)**. Literature presents a great number of approximate kinetic models, which take different factors into account, and the proper choice of the mathematical description of these expressions is object of discussion in many works **(2**, **3**, **9**, **10)**. Thus, hybrid neural models are suitable to model biotechnological processes, as they combine mass and energy balance equations with neural networks, which represent the unknown kinetics **(11)**.

In this work a hybrid neural model is proposed to model an industrial alcoholic fermentation process. The specific growth rate is described by a functional link network (FLN) **(12)**. These networks have been shown to have a good nonlinear approximation capability, although the estimation of its weights is linear. Because of this linear estimation, the FLNs training is rapid, requires low computational effort, and convergence is guaranteed, so that it has a large potential for online control and optimization implementations **(6)**.

5.2 Industrial alcoholic fermentation process – Phenomenological model

The alcoholic fermentation process to be modeled is illustrated in Figure 5.1. This is an existing plant, designed using the procedure of Andrietta and Maugeri **(13)** to a maximum production of 320 m^3 of anhydrous ethanol/d. The system is a typical large-scale industrial process made up of five continuous-stirred tank reactors (CSTR) attached in series and operating with cell recycle. Each reactor has an external heat exchanger to keep the temperature constant at an ideal level for the fermentation process. The feedstock, a mixture made up of sugar cane molasses and syrup and sources of nitrogen and mineral salts, is converted into ethanol by a fermentation process carried out using the yeast *Saccharomyces cerevisiae*. A set of centrifuges splits the outlet-fermented product into two phases. The light phase (F_V) is sent to a distillation tower, in which the alcohol is finally obtained. The heavy phase is mixed with acid and diluted with water (flow rate F_a) before being recycled (flow rate F_R) to the first reactor. F_S (m^3/h) is the flow rate of the cell purge stream, used to permit cell renovation and the withdrawal of

secondary products accumulated into the fermentor. These process-operating conditions are real conditions of typical industrial distilleries in Brazil.

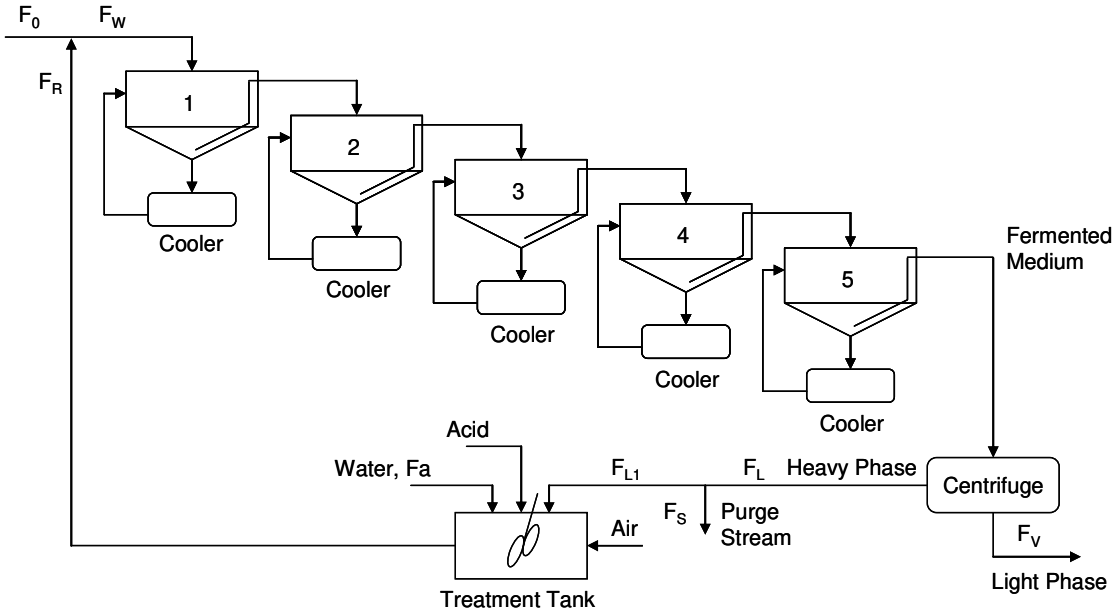


Figure 5.1. Schematic illustration of an industrial plant for ethanol production

The mass and energy balances for the five fermentors are:

Global mass balance in *i* th fermentor:

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

Substrate balance in *i* th fermentor:

$$\frac{d(S_iV_i)}{dt} = F_{i-1}S_{i-1} - F_iS_i - V_iX_i\frac{1}{Y_{X/S}}\mu_i \tag{2}$$

Product balance in *i* th fermentor:

$$\frac{d(P_iV_i)}{dt} = F_{i-1}P_{i-1} - F_iP_i + V_iX_i\frac{Y_{P/S}}{Y_{X/S}}\mu_i \tag{3}$$

Biomass balance in *i* th fermentor:

$$\frac{d(X_iV_i)}{dt} = F_{i-1}X_{i-1} - F_iX_i + V_iX_i\mu_i \tag{4}$$

Energy balance in *i* th fermentor:

$$\frac{d(V_i T_i)}{dt} = F_{i-1} T_{i-1} - F_i T_i + F_{Ci}(T_{Ci} - T_i) + \frac{V_i \Delta HX_i}{\rho_i \, Cp_i Y_{X/S}} \mu_i \quad (5)$$

Energy balances for the reagent fluid in *i* th heat exchanger:

$$\frac{d(T_{Ci})}{dt} = F_{Ci}(T_i - T_{Ci})\rho_i Cp_i - \left(\frac{UA_i}{V_{Ci}\rho_i Cp_i} \right) LMTD_i \quad (6)$$

Energy balances for the cooling fluid in *i* th heat exchanger:

$$\frac{d(T_{Ji})}{dt} = \frac{F_{Ji}}{V_{Ji}}(T_{Je} - T_{Ji}) + \left(\frac{UA_i}{V_{Ji}\rho_j Cp_j} \right) LMTD_i \quad (7)$$

Logarithmic mean temperature difference (LMTD) for *i* th heat exchanger:

$$LMTD_i = \frac{(T_i - T_{Ji}) - (T_{Ci} - T_{Je})}{\ln\left(\frac{(T_i - T_{Ji})}{(T_{Ci} - T_{Je})} \right)} \quad (8)$$

where the subscripts i refer to each stage (fermentor), J_i and J_e are indexes referring to water inlet and outlet of the cooler, C is an index referring to process fluid at the cooler; V (m^3) is the reactor volume; T (^oC) is the temperature; F (m^3/h) is the flow rate; S (kg/m^3), X (g/l) and P (kg/m^3) are the substrate, cell and ethanol concentrations, respectively; ρ (kg/m^3), Cp ($J/kg.^oC$) and ΔH (J/kg) are density, specific heat and reaction heat, respectively; and $Y_{X/S}$ and $Y_{P/S}$ are the kinetic parameters of yield. A kinetic model was validated with typical industrial conditions for this process **(14)**. The specific growth rate used was proposed by Lee et al. **(2)**:

$$\mu = \mu_{MAX} \frac{S}{K_S + S} \left(1 - \frac{P}{P_{MAX}} \right)^n \left(1 - \frac{X}{X_{MAX}} \right)^m \quad (9)$$

In which μ_{MAX} (h^{-1}), P_{MAX} (kg/m^3) and X_{MAX} (kg/m^3) are the maximum specific growth rate, the product concentration when cell growth ceases and the biomass concentration when cell growth ceases, respectively, and n and m are inhibitor

terms coefficients. The maximum specific growth rate, μ_{max}, is influenced by temperature and is calculated by the Arrhenius equation **(15)**:

$$\mu_{MAX} = Ae^{\frac{-E}{RT}} \tag{10}$$

According to Andrietta **(14)**, P_{MAX} has a constant value below a critical temperature (32ºC) and above this temperature its value is described by:

$$P_{MAX} = K_0 e^{aT} \tag{11}$$

The kinetic parameters and constants used with Eqs. 1-11 are given by E=6417 J/mol, $A=4.5\times10^{10}$, R=8.314 J/mol·K, K_0=895.6 kg/m^3, a=-0.0676 °C^{-1}, X_{max}=100 kg/m^3, n=3.0, m=0.9, K_s=1.6 kg/m^3, $Y_{P/S}$=0.445, $Y_{X/S}$=0.033 **(14)**.

Table 5.1 shows the design parameters for the five fermentors.

Table 5.1. Design parameters for the five fermentors

Reactor	A_i (m^2)	V(m^3)	F_c (m^3/h)	F_J (m^3/h)	V_c (m^3)	V_J (m^3)
1	76.361	433	400	400	20	20
2	63.242	370	350	350	20	20
3	31.061	366	180	180	20	20
4	6.809	359	60	60	20	20
5	2.869	333	28	30	20	20

The operations involved in the recycle (centrifuge, purge, dilution and mixing) presents fast dynamics when compared to the fermentation process and can be considered in pseudo-steady-state. The concentrations of ethanol and substrate can be considered to remain constant in the value of the fermentor outlet at the centrifuge exit. Also, variations in density are considered negligible through the process. The equations are as follows:

$$F_W = \frac{F_0}{(1 - RR)} \tag{12}$$

$$F_R = F_W - F_0 \tag{13}$$

Cell concentration is kept constant in the recycle by manipulation of water flow rate, F_a. Mass balance on the dilution tank leads to:

$$F_{L1} = \frac{F_R X_R}{X_L} \tag{14}$$

$$F_a = F_R - F_{L1} \tag{15}$$

$$S_R = \frac{F_{L1} S_5}{F_R} \tag{16}$$

$$P_R = \frac{F_{L1} P_5}{F_R} \tag{17}$$

Mass balance on the centrifuge leads to:

$$F_V = F_W \frac{X_L - X_5}{X_L - X_V} \tag{18}$$

$$F_L = F_W - F_V \tag{19}$$

In the purge point we have:

$$F_S = F_L - F_{L1} \tag{20}$$

The concentrations of the first fermentor feeding are:

$$S_W = \frac{F_R S_R + F_0 S_0}{F_W} \tag{21}$$

$$P_W = \frac{F_R P_R}{F_W} \tag{22}$$

$$X_W = \frac{F_R X_R}{F_W} \tag{23}$$

The mathematical model is composed by Eqs. 1-23 and solved using the fourth order Runge-Kutta method. The must feed flow rate used was F_0 = 100 m^3/h. Other parameters used in the equations are: ΔH=-6.509×10^5 J/kg, ρ= 950 Kg/m^3, Cp=4184.1 J/Kg·°C, U=1.464×10^7 J/h·°C·m^2**,** ρ_j **=** 1000 Kg/m^3, Cp_j= 4184.1 J/Kg·°C **(14)**. The model does not consider cellular death because purge and recycle maintain the average age of cells in the process from 7 to 10 days (young cells). Death rate is equal to the rate of cellular replacement and there is no accumulation of live cells or dead cells.

5.3 Hybrid neural model

The phenomenological model described by Eqs. 1-23 was determined by Andrietta **(14)** considering a real industrial process operating in continuous mode. It was validated with industrial data and shown to describe the process dynamic behavior accurately. However, this model is not able to describe the process in the presence of changes in operational conditions. In 2005, Andrietta evaluated this same model and concluded that, for the current operational conditions in the factory evaluated, the model of Ghose and Thyagi **(1)** with some modifications led to a better description of the process dynamic behavior. He also determined the existence of four groups of kinetic parameters in different harvesting periods in the year, which indicates that, in order to describe the dynamic behavior of the plant for long periods of time, the mathematical model should have its kinetic parameters updated periodically. However, the reestimation of kinetic parameters, mainly if they are considered functions of temperature, is a difficult and time-consuming task.

In order to solve this problem we used a hybrid neural model in which the kinetics is described by FLNs. The great advantage of these neural networks is that the estimation of its weights is a linear problem, and so the reestimation is rapid and convergence is guaranteed. The hybrid model consists of the mass and energy balance equations (Eqs. 1-8 and 12-23) and a FLN, which describes the specific biomass growth rate. The choice of this kinetic rate to be described by the neural network was made based on the results of a sensitivity analysis of the process.

5.4 Functional Link Networks

A neural network typically consists on many simple computational elements or nodes arranged in layers and operating in parallel. The weights, which define the strength of connection between the nodes, are estimated to yield good performance. Usually, in the training of neural networks, the inputs to a node are linearly weighted before the sum is passed through some nonlinear activation function that ultimately gives the network its nonlinear approximation ability. However, the same nonlinearity creates problems in learning the network weights,

as nonlinear learning rules must be used, the learning rate is often unacceptably slow, and local minima may cause problems **(12)**. One way of avoiding nonlinear learning is the use of FLNs. In these networks, a nonlinear functional transformation or expansion of the network inputs is initially performed and the resulting terms are combined linearly. The obtained structure has a good non-linear approximation capability and the estimation of the network weights is linear.

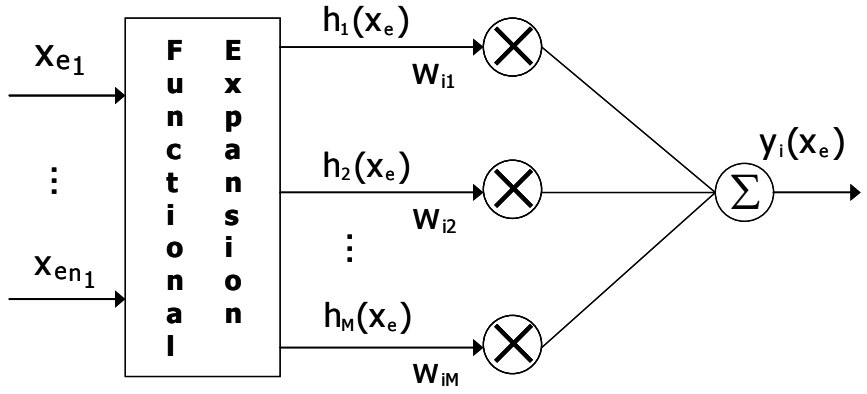


Figure 5.2. General structure of a Functional Link Network

The general structure of an FLN is shown in Fig. 5.2, where $\mathbf{x_e}$ is the input vector and $y_i(\mathbf{x_e})$ is an output. The hidden layer performs a functional expansion on the inputs, which maps the input space, of dimension n_1, onto a new space of increased dimension, M ($M > n_1$). The output layer consists on m nodes, each one, in fact, a linear combiner. The input-output relationship of the FLN is

$$y_i(\mathbf{x_e}) = \sum_{j=1}^{M} w_{ij} h_j(\mathbf{x_e}), \qquad 1 \le i \le m \tag{24}$$

We use a modification in the structure of the FLNs, where the output given by eq. (24) is transformed by an invertible non-linear activation function. The new output is

$$y_i(\mathbf{x_e}) = f_i\left(\sum_{j=1}^{M} w_{ij} h_j(\mathbf{x_e}) \right), \qquad 1 \le i \le m \tag{25}$$

where f_i is an invertible non-linear function such as, for example, the sigmoidal function. Another modification in the FLNs used in this work is that the real inputs ($\mathbf{x_e}$) are transformed into a greater number (n_z) of auxiliary inputs ($\mathbf{z}$) before the

functional expansion is performed. These modifications increase the non-linear approximation ability of the network and yet, the estimation of the parameters remains a linear problem. A polynomial expansion of degree six is then performed on the new inputs. The generated monomials ($h_j[\mathbf{z}]$) are shown in Table 5.2.

Table 5.2. Polynomial expansion of degree six

Degree	Monomials
0	1
1	z_i $i=1,n_z$
2	z_iz_j $i=1,n_z$; $j=i,n_z$
3	$z_iz_jz_k$ $i=1,n_z$; $j=i,n_z$; $k=j,n_z$
4	$z_iz_jz_kz_l$ $i=1,n_z$; $j=i,n_z$; $k=j,n_z$; $l=k,n_z$
5	$z_iz_jz_kz_lz_m$ $i=1,n_z$; $j=i,n_z$; $k=j,n_z$; $l=k,n_z$; $m=l,n_z$
6	$z_iz_jz_kz_lz_mz_n$ $i=1,n_z$; $j=i,n_z$; $k=j,n_z$; $l=k,n_z$; $m=l,n_z$; $n=m,n_z$

Once the monomials are generated, the orthogonal least-squares estimator proposed by Billings et al. **(16)** is used to calculate the network weights (w_{ij}) and to eliminate the monomials which are not significant in explaining the output variance. This reduces the size and complexity of the neural network and avoids overfitting of the data.

5.4.1 Training of the neural network

The network inputs ($\mathbf{x_e}$) are the process state variables: biomass, substrate and product concentrations, and temperature. The output is the specific growth rate. In this work, the state variables are considered accessible by direct measurement, but the output is not measurable. It can be estimated from measured experimental data by the discretization of the biomass balance equation. When dealing with noisy data, some kind of smoothing algorithm has to be applied to produce reliable rate values from biomass concentration data.Training was performed using a full factorial design, considering axial points and one center

point, involving the variables S_0 and T_0. Both variables had their values changed every 10 h within the ranges 150-210 g/L (S_0) and 28-35 °C (T_0). Ten hours is enough time for the process to reach a new steady state. Table 5.3 shows the factorial design.

Table 5.3. Full factorial design with axial points and one center point

Time (h)	S_0 g/L	T_0 °C
10	158.8	29
20	158.8	34
30	201.2	29
40	201.2	34
50	150	31.5
60	210	31.5
70	180	28
80	180	35
90	180	31.5

Table 5.4. Input vectors of the network

Reactor	Vector Z
1	$z = \left[X^{3.02} \quad \frac{1}{S} \quad P^{4.92} \quad T^{0.59} \right]$
2	$z = \left[X^{2.7} \quad \frac{1}{S} \quad \frac{1}{P^{4.0}} \quad T^{0.6} \right]$
3	$z = \left[X^{3.0} \quad \frac{1}{S} \quad P^{4.90} \quad T \right]$
4	$z = \left[X \quad \frac{1}{S} \quad P \quad T^{0.4} \right]$
5	$z = \left[X^{1.25} \quad \frac{1}{S} \quad P \quad T^{0.59} \right]$

During training and validation the performance of the FLN was measured by equation 26, given by Milton and Arnold **(17)**:

$$\text{cor} = \left(1 - \frac{\sum_{k=1}^{N} (y_e(k) - y(k))^2}{\sum_{k=1}^{N} (y_e(k) - \bar{y}_e)^2} \right) 100\% \qquad (26)$$

where $y_e(k)$ is the real output, $y(k)$ is the network output, $\bar{y}_e$ is the mean value of the real outputs and N is the number of training points. For validation, the disturbances encompassed a sequence of steps with random changes every 10 h within the ranges 162-198 g/L (S_0) and 28-32°C (T_0). During training and validation, the activation function that led to the best performance is given by Eq. 27:

$$f\left(\sum_{j=1}^{M} w_{ij} h_j(z)\right) = \frac{1}{\sum_{j=1}^{M} w_{ij} h_j(z)} \tag{27}$$

The auxiliary input vectors were chosen for each reactor after many tests and are given in Table 5.4.

5.5 Results

For the five process fermentors, the best performance of the FLNs was obtained using fifth degree monomials. Using four auxiliary inputs (as shown in Table 5.4) we generated 126 monomials for each of the five FLNs (each one describing the specific growth rate of one fermentor). After using the orthogonal least-squares estimator proposed by Billings et al. **(16)** to eliminate nonsignificant monomials, the number of monomials for each network was drastically reduced, as can be seen in Table 5.5, which shows the number of monomials and the performance of the five FLNs during validation (measured by Eq. 26).

Table 5.5. Performance and number of monomials for the Functional Link Networks

Reactor	Number of monomials	Performance (%)
1	7	99.6870
2	6	99.6868
3	11	99.9895
4	18	99.7873
5	18	96.6931

The number of monomials in the FLNs can be considered equivalent to the number of connections (or weights) in a multilayer perceptron neural network.

The five FLNs are described by Eqs. 28-32:

$$\mu_{FLN}(\text{fermentor 1}) = \frac{1}{-1.4295T^{1.18} + 1.1165\times10^{-23}X^{6.04}\frac{1}{S}P^{9.84} + 6.8689\times10^{-7}X^{6.04}P^{4.92}T^{1.18}} \frac{1}{-5.0705\times10^{-13}X^{3.02}P^{4.92}T^{1.77} + 7.5585x10^{-7}X^{3.02}T^{2.36} - 5.104\times10^{-24}\frac{1}{S}P^{14.76}T^{0.59}} \frac{1}{+9.7211\times10^{-10}P^{4.92}T^{2.36}} \quad (28)$$

$$\mu_{FLN}\ (\text{fermentor 2}) = \frac{1}{-334.11 + 0.0060987X^{8.1}\frac{1}{P^{4.0}} + 1.1656\times10^{-14}X^{10.8}\frac{1}{S} - 7.2676\times10^{-7}X^{10.8}\frac{1}{P^{4.0}}} \frac{1}{-4.5488x10^{-11}X^{8.1}\frac{1}{S^{2.0}} + 8.6769\times10^{-9}X^{5.4}T^{1.8}} \quad (29)$$

$$\mu_{FLN}(\text{fermentor 3}) = \frac{1}{9039.3 - 323.37T - 0.00033194\frac{1}{S}P^{4.9} + 3.0911\times10^{-5}\frac{1}{S}P^{4.9}T - 9.6048\times10^{-7}\frac{1}{S}P^{4.9}T^{2}} \frac{1}{-8.9072\times10^{-11}P^{4.9}T^{3.0} - 5.2582\times10^{-11}X^{6.0}T^{3.0} + 9.2656\times10^{-8}X^{3.0}T^{4.0}} \frac{1}{+2.7862\times10^{-5}\frac{1}{S^{3.0}}T^{2.0} + 9.9624\times10^{-9}\frac{1}{S}P^{4.9}T^{3.0} + 2.8483\times10^{-12}P^{4.9}T^{4.0}} \quad (30)$$

$$\mu_{FLN}(\text{fermentor 4}) = \frac{1}{-2899.6 - 0.53749X\frac{1}{S^{2.0}}P + 0.27431\frac{1}{S^{2.0}}P^{2.0} - 0.017078\frac{1}{S}P^{3.0} + 2.0252PT^{1.2}} \frac{1}{-2.3187\times10^{-6}X\frac{1}{S^{4.0}} - 0.0053256X\frac{1}{S^{2.0}}P^{2.0} + 0.21679X\frac{1}{S^{2.0}}PT^{0.4}} \frac{1}{-0.043293XPT^{1.2} + 1.6239\times10^{-9}\frac{1}{S^{5.0}} + 1.8985\times10^{-7}\frac{1}{S^{4.0}}P + 1.5343\times10^{-5}\frac{1}{S^{4.0}}T^{0.4}} \frac{1}{-3.8448\times10^{-7}\frac{1}{S^{3.0}}P^{2.0} + 0.0010677\frac{1}{S^{2.0}}P^{3.0} - 0.072847\frac{1}{S^{2.0}}P^{2.0}T^{0.4}} \frac{1}{-2.9298\frac{1}{S^{2.0}}T^{1.2} + 2.3907\times10^{-5}\frac{1}{S}P^{4.0} + 0.0039741\frac{1}{S}P^{3.0}T^{0.4}} \quad (31)$$

$$\mu_{FLN}(\text{fermentor 5}) = \frac{1}{-7511.3\frac{1}{S} - 0.0046055\frac{1}{S^{2.0}}P + 10.992\frac{1}{S}P^{2.0} + 13.016T^{1.77} - 9.8779\times10^{-9}\frac{1}{S^{4.0}}} \frac{1}{-0.23645\frac{1}{S}P^{3.0} - 0.00058067X^{3.75}\frac{1}{S^{2.0}} + 0.01669X^{2.5}\frac{1}{S^{2.0}}T^{0.59}} \frac{1}{-0.15954X^{1.25}\frac{1}{S^{2.0}}T^{1.18} - 0.019771X^{1.25}T^{2.36} - 1.1657\times10^{-14}\frac{1}{S^{5.0}}} \frac{1}{+1.3446\times10^{-9}\frac{1}{S^{4.0}}T^{0.59} + 5.2132\times10^{-9}\frac{1}{S^{3.0}}P^{2.0} - 7.4404\times10^{-8}\frac{1}{S^{3.0}}PT^{0.59}} \frac{1}{+1.8348x10^{-7}\frac{1}{S^{3.0}}T^{1.18} + 0.50787\frac{1}{S^{2.0}}T^{1.77} + 0.0013531\frac{1}{S}P^{4.0}} \frac{1}{+0,00091873\frac{1}{S}P^{3.0}T^{0.59}} \quad (32)$$

The quality of prediction of the proposed hybrid neural model was measured by the residual standard deviation (RSD) described as a percentage of average of the 'real' values, as described by Atala et al. **(3)**:

$$RSD(\%) = \frac{RSD}{\overline{y_i}} x100 \tag{33}$$

$$RSD = \frac{\sqrt{\sum_{i=1}^{n}(y_i - y_{Pi})^2}}{n} \tag{34}$$

In this equation y_i is the 'real' value (calculated by the deterministic model), y_{Pi} is the value predicted by the hybrid model and n is the number of points. Figure 5.3 shows the performance of the hybrid model to describe the dynamic behavior of the fifth fermentor when the process is subjected to the disturbances shown in Figure 5.4. In this figure the disturbances performed for training of the neural networks are shown for comparison. Table 5.6 shows the values of RSD% obtained when we use the mathematical model to calculate concentrations of biomass, substrate and product in the five fermentors. If a model without update of its parameters is used it is not possible to obtain good predictions, and errors over 50% might be found depending on the type and level of changes.

Table 5.6. Model quality of prediction described by the Residual Standard Deviation (RSD %)

Reactor	Biomass X (kg/m^3)	Substrate S (kg/m^3)	Product P (kg/m^3)
1	0.0037	0.1506	0.0323
2	0.0021	0.3097	0.0147
3	0.0020	0.7540	0.0126
4	0.0024	1.4294	0.0149
5	0.0023	1.9731	0.0143

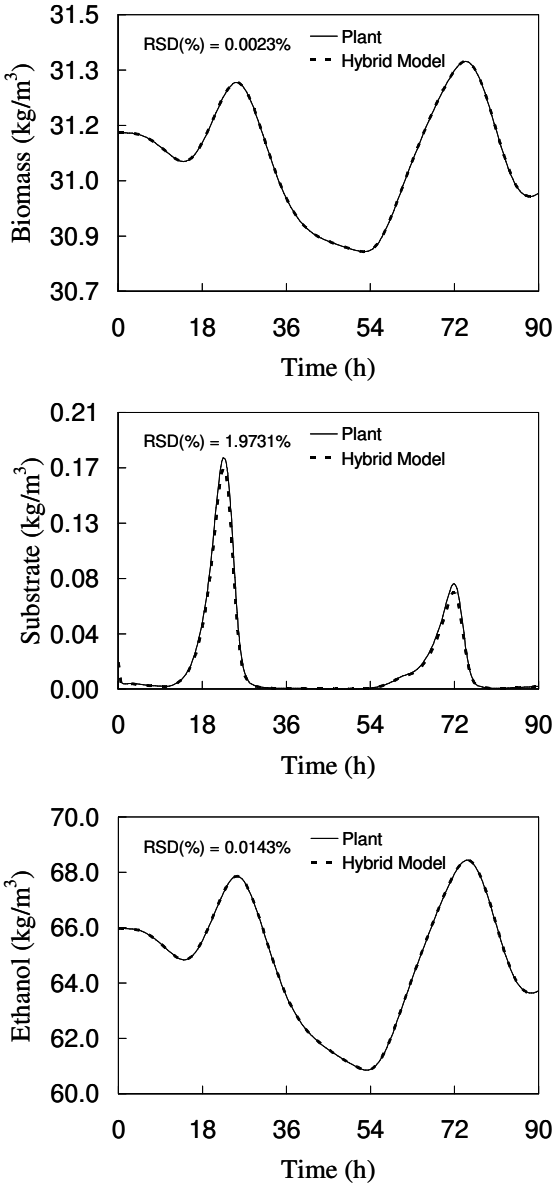


Figure 5.3. Concentrations of biomass, substrate and product in the fifth fermentor (Deterministic Model —; Hybrid Model ---)

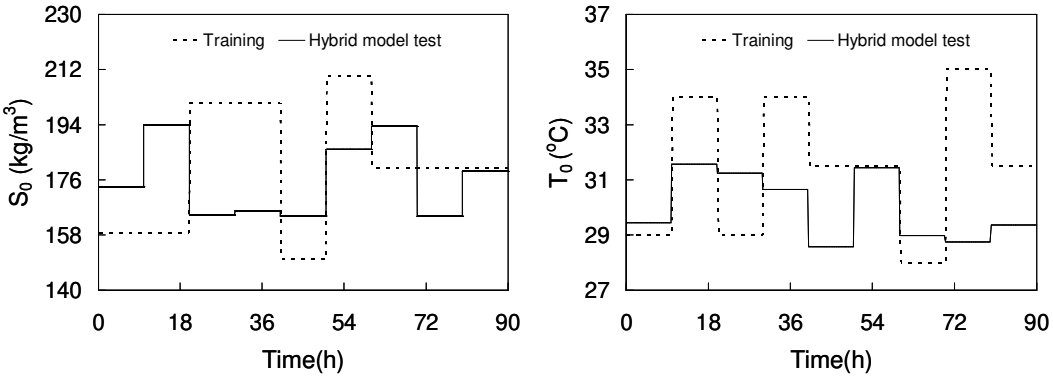


Figure 5.4. Disturbances for the test of the hybrid model

5.6 Discussion

Bioreactors are quite difficult to model because their operation involves microbial growth under constantly changing conditions. Past experience has shown the difficulties of using phenomenological models without kinetic parameters reestimation to describe the dynamic behavior of the process for long periods of operation, when there are changes in substrate and yeast conditions (a common situation in industrial plants). However, the reestimation of kinetic parameters in a phenomenological model is difficult and time consuming, mainly when the kinetics dependence of temperature is taken into account. As the main nonlinearities and uncertainties are in the description of the process kinetic, the use of a hybrid approach whereby the kinetics is described by a FLN can simplify the reestimation step.

In this work, we have shown that a hybrid neural model combining mass balance equations with FLNs is able to describe the process dynamic behavior with a performance similar to that of a phenomenological model. However, the reestimation of the weights of a FLN is much simpler than the reestimation of kinetic parameters in a phenomenological model. The estimation of the FLN weights is a linear optimization problem, so it is very fast and convergence is guaranteed. The use of this approach enables the development of an online reestimation procedure.

5.7 Conclusions

The proposed hybrid model seems to be a quite suitable approach to follow changes in the process which impact the system behavior. The reestimation of the hybrid model is easily done especially because the model is built-up using FLNs. These networks showed to have a good non linear approximation capability, although the estimation of its weights is linear. The proposed approach for hybrid modeling is able to deal with disturbances in transient phases of the process and different yeasts, because the identification procedures is carried out adequately with an appropriate set of data representative of the biotechnological process.

Acknowledgments

The authors acknowledge Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support.

Nomenclature

a	Constant in Eq. 11
A	Constant in Eq. 10
A_i	Area of the i^{th} heat exchanger, m^2
Cp_i	Reagent fluid heat capacity, J/kg·°C
Cp_j	Cooling fluid heat capacity, J/kg·°C
E	Activation energy, J/mol
f	Activation function of the FLN
F_a	Water flow rate, m^3/h
F_{Ci}	Reagent fluid flow rate to the i^{th} heat exchanger, m^3/h
F_i	Feed stream flow rate to i^{th} reactor, m^3/h
F_{Ji}	Cooling fluid flow rate in the i^{th} heat exchanger, m^3/h
F_L	Cell suspension flow from centrifuge, m^3/h
F_{L1}	Cell suspension flow to treatment tank, m^3/h
F_0	Fresh medium flow rate, m^3/h
F_R	Cell recycling flow rate, m^3/h
F_S	Purge flow rate, m^3/h
F_V	Liquid phase flow to rectification column, m^3/h
F_W	Feed stream flow rate, m^3/h
h	Monomials generated by the functional expansion
K_0	Constant in Eq. 11
K_s	Substrate saturation constant, kg/m^3
$LMTD_i$	Logarithmic mean temperature difference for the i^{th} heat exchanger
m	Constant in Eq. 9
m	Number of outputs of the FLN
M	Number of monomials

n	Constant in Eq. 9
n_z	Number of auxiliary inputs
P_i	Product concentration in the i^{th} reactor, kg/m^3
P_{MAX}	Product concentration when cell growth ceases, kg/m^3
P_R	Product concentration in the cells recycle, kg/m^3
P_W	Feed product concentration, kg/m^3
R	Universal gas constant, J/mol·K
RR	Cell recycle rate
S_i	Substrate concentration in the i^{th} reactor, kg/m^3
S_0	Inlet substrate concentration, kg/m^3
S_R	Substrate concentration in the cells recycle, kg/m^3
S_W	Feed Substrate concentration, kg/m^3
t	Time, h
T_{Ci}	Temperature of the reagent fluid in the i^{th} heat exchanger, ^oC
T_i	Temperature in the i^{th} reactor, ^oC
T_{Ji}	Cooling fluid temperature at the i^{th} heat exchanger exit, ^oC
T_{Je}	Inlet cooling fluid temperature in the i^{th} heat exchanger, ^oC
U	Global exchange coefficient, $J/h\cdot{}^oC\cdot m^2$
V_i	Volume of the i^{th} reactor, m^3
V_{Ji}	Cooling fluid volume in the i^{th} heat exchanger, m^3
w	FLN weights
$\mathbf{x_e}$	Input vector of the FLN
X_i	Biomass concentration in the i^{th} reactor, kg/m^3
X_L	Biomass concentration in the heavy phase from centrifuge, kg/m^3
X_{MAX}	Biomass concentration when cell growth ceases, kg/m^3
X_R	Cell recycling concentration, kg/m^3
X_V	Biomass concentration in the light phase to rectification column, kg/m^3
X_W	Feed biomass concentration, kg/m^3
$Y_{P/S}$	Yield of product based on cell growth, kg/kg
$Y_{X/S}$	Limit cellular yield, kg/kg
z	Auxiliary input vector

Greek letters

ΔH Reaction heat, J/kg

ρ_i Reagent fluid density in the i^{th} reactor, kg/m^3

ρ_j Cooling fluid density, kg/m^3

μ_i Specific growth rate in the i^{th} reactor, h^{-1}

μ_{MAX} Maximum specific growth rate, h^{-1}

References

1. Ghose, T.K. and Tyagi, R.D. (1979), Biotechnol. Bioeng. **21**, 1387-1400.
2. Lee, J.M., Pollard, J.F. and Coulman, G.A. (1983), Biotechnol. Bioeng. **25**, 497-511.
3. Atala, D.I.P., Costa, A. C., Maciel Filho, R. and Maugeri Filho, F. (2001), Appl. Biochem. Biotech. **91(3)**, 353-365.
4. Aldiguier, A.S., Alfenor, S., Cameleyre, X., Goma, G., Uribelarrea, J.L., Guillouet, S.E. and Molina-Jouve, C. (2004), Bioproc. Biosyst. Eng. **26**, 217-222.
5. Costa, A.C., Alves, T.L.M., Henriques, A.W.S., Maciel Filho, R. and Lima, E.L. (1998), Comput. Chem. Eng. **22**, 859-862.
6. Harada, L.H.P., Costa, A. C. and Maciel Filho, R. (2002), Appl. Biochem. Biotech. **98(1-9)**, 1009-1024.
7. Psichogios, D.C. and Ungar, L.H. (1992), AIChE J. **38**, 1499-1511.
8. Smets, I.Y., Claes, J.E., November, E.J., Bastin, G.P. and Van Impe, J.F. (2004), J. Proc. Control **14**, 795-805.
9. Nishiwaki, A. and Dunn, I.J. (1999), Biochem. Eng. J. **4**, 37-44.
10. Ricci, M., Martini, S., Bonechi, C., Trabalzini, L., Santucci, A. and Rossi, C. (2004), Chem. Phys. Lett. **387**, 377-382.
11. Patnaik, P.R.(2003), Biochem. Eng. J. **15**, 165-175, 2003.
12. Chen, S. and Billings, S.A. (1992), Int. J. Control **56**, 319-346.
13. Andrietta, S.R. and Maugeri, F. (1994), Adv. Bioprocess Engng. **1**, 47-52.
14. Andrietta, S.R. (1994), PhD Thesis, State University of Campinas, SP, Brazil.

15. Fogler, H.S. (1999), Elements of chemical reaction engineering, 3rd ed., Prentice Hall, New York.
16. Billings, S.A., Chen, S. and Korenberg, M.J. (1989), Int. J. Control **49**, 2157-2189.
17. Milton, J.S. and Arnold, J.C. (1990), Introduction to Probability and Statistics, McGraw Hill, New York.

CAPÍTULO 6. DEVELOPMENT OF A METHODOLOGY FOR THE ESTIMATION OF KINETIC PARAMETERS IN A STRUCTURED MODEL FOR ETHANOL PRODUCTION

Trabalho submetido à revista **Brazilian Journal of Chemical Engineering**.
Elmer Ccopa Rivera, Aline C. Costa, Rubens Maciel Filho.

Abstract

The estimation of kinetic parameters of deterministic models for fermentation processes is usually complex, mainly due to non-linearities, great number of parameters and interactions among them. For that reason, in this work, the optimization was evaluated using a methodology with different optimization techniques. The methodology consists of four steps. First, initial values were found from literature. Subsequently, the potential of global searching of Real-Coded Genetic Algorithm (RGA) was applied for simultaneous estimation of the parameters. In the third step the most significant parameters were identified using Placket–Burman (PB) design. Finally, the Quasi-Newton algorithm (QN) was used for optimization of the most significant parameters, near the global optimum region, as the initial values were already determined by the RGA global searching algorithm. The kinetic model optimized describes satisfactorily the fermentation process as demonstrated by the experimental results.

Keywords

Ethanol fermentation, parameter estimation, optimization techniques.

6.1 Introduction

The lack of robustness of fermentation in the presence of fluctuations in the quality of the raw material, which leads to changes in the kinetic behaviour with impact on yield, productivity and conversion, is among the main current problems related to the alcoholic fermentation process. Changes in the operational conditions are quite common in plants of alcoholic fermentation; they occur not only due to the changes in the quality of the raw material but also due to variations of dominant yeast in the process.

The lack of robustness can be corrected by adjustments in the operational and control parameters of the process when fluctuations occur. In order to accomplish this, it is important that a mathematical model is available to aid in the decision making, mainly when the difficulties of monitoring the key process variables (concentrations of biomass, substrate and product, and temperature) are

taken into account. However, the operational changes described make the prediction of the dynamic behavior of the process with a single model difficult, as they lead to changes in microorganism kinetics. Thus, it would be of great advantage to have a mathematical model that could be easily adapted to changes in operational conditions. This can be accomplished by reestimating the kinetic parameters when necessary.

Estimation of kinetic parameters of deterministic models is usually complex, mainly due to non-linearities, great number of parameters and interactions among them. In biochemical engineering, the most classical method involves the mathematical estimation of model parameters based on the minimization of some cost function built-up with the parameters to be estimated. Several kinetic models have been proposed for the alcoholic fermentation process (Dourado et al., 1986, Andrietta et al., 1994). Many techniques are available for minimizing the error of estimation, generally methods based on gradient search, such as Quasi-Newton Algorithm (QN).

Artificial intelligence (AI), such as Genetic Algorithms (GA), covers a wide range of techniques and tools that facilitate decision making and have often been found to be as powerful and effective as gradient search methods in many engineering applications. These methods have already been successfully applied in the optimization and control of bioprocesses for more than 20 years (Schügerl, 2001). GAs have been successfully utilized for kinetic parameters estimation in biotechnological process (Park et al., 1998, Saíz, 2003).

In this work, the optimization procedure was evaluated using a methodology with different optimization techniques: Real-Coded Genetic Algorithm, Placket–Burman design and Quasi-Newton algorithm. The approach is applicable when the structure of a kinetic model has been set up and the kinetic parameters should be estimated.

The parameter estimation methodology was demonstrated on an alternative dynamic structured model (Stremel, 2001), adapted from Rotboll and Jorgensen (1993) to simulate a fixed bed bioreactor for ethanol production by immobilized *Saccharomyces Cerevisiae*. The model contains 34 kinetic parameters and 9

parameters (KE, KE2, F1, F2, F3, F5, F6, kd, C), related to the glycolitic and respiratory (TCA) paths, which were reestimated.

6.2 Methodology

This section describes a parameter estimation methodology in four steps which was developed in order to estimate kinetic parameters in structured models. First, initial values were found from literature. Secondly, a Real-Coded Genetic Algorithm (RGA) was applied for simultaneous estimation of the parameters and thirdly the most significant parameters were identified using Placket–Burman (PB) design. Finally, the Quasi-Newton algorithm (QN) was used for optimization of the most significant parameters, near the global optimum region, as the initial values were already determined by the RGA global searching algorithm.

6.2.1 Obtaining Initial Parameters Values

Stremel (2001) developed a structured bicompartmental model that describes the intracellular synthetic and structural compounds. The synthetic compartment is responsible for the cellular machinery and conversion of glucose to pyruvate, acetaldehyde and ethanol as well as the structural part of the cell. The stequiometric parameters in the kinetic model are adjusted for glycolitic and respiratory pathways to represent ethanol production. The TCA (tricarboxylic acid) route is included in the model to describe the conversion of intermediates through respiratory mechanism. Moreover, the dynamic intraparticular model includes complex growth kinetic of immobilized cells in a high productivity bioreactor for ethanol production. The kinetic rates representing metabolic pathways resemble the pseudo steady state Michaelis Menten equation. The reactions in equilibrium are catalyzed by specific enzymes such as acetaldehyde dehydrogenase and alcohol dehydrogenase. The model also includes terms for the inhibition by ethanol, substrate, and saturation by the cells inside the pellets. Stremel (2001) also investigated all the parameters and their effects to identify the most significant, but did not use a methodology to optimize simultaneously the parameters of the model.

Initial values for kinetic parameters are required before any model solution can be obtained. Thus, in this work, an initial guess for parameters was taken from Stremel (2001) in order to find their approximate values.

6.2.2 Optimization by Real-Coded Genetic Algorithm

The GA considered in this work is based on the freeware versions written in FORTRAN of a Real-Coded Genetic Algorithm (RGA) developed by Yedder (2002), adapted and enhanced to the specific needs of the parameter optimization.

In the proposed strategy, the chromosome or individual are vectors where each real-value (gene) stands for each one of the unknown parameters in the kinetic model. The stopping criterion is selected when the maximum number of fixed generations is reached. The RGA parameters were adjusted to minimize the number of generations (iterations) required to reach a satisfactory fitness (objective function) value by minimizing Eq. (1), which reflects a good agreement between measured concentration and the concentration computed by de model.

$$E(\theta) = \sum_{n=1}^{np}\left[\frac{(S_n - Se_n)^2}{Se_{\max}^2} + \frac{(P_n - Pe_n)^2}{Pe_{\max}^2}\right] = \sum_{n=1}^{np}\varepsilon_n(\theta) \qquad (1)$$

In this equation Se_n and Pe_n are the measured concentrations of substrate and ethanol at the sampling time n. S_n and P_n are the concentrations computed by the model at the sampling time n. $Se_{\max}$ and $Pe_{\max}$ are the maximum measured concentrations and the term np is number of sampling points. Here $\varepsilon_n(\theta)$ is the error in the output due to the nth sample.

The main technical features of the genetic algorithms are reported in Table 6.1.

Table 6.1. Main technical features used by the genetic algorithm

Option chosen - RGA	Parameters	Param. values
Niching, elitism, barycentric crossover, non-uniform mutation	Individual length (number of parameters in the model)	43
	Population size	10
	Crossover probability	0.9
	Mutation probability	0.03

6.2.3 Identification of Significant Parameters

This step consists in applying Placket-Burman (PB) design to identify the kinetic parameters which are significant on the optimization problem. The Plackett-Burman design is a partial factorial method that allows the testing of multiple independent process variables within a single experiment. The influence of 43 parameters on Eq. (1) was investigated using the methodology of Plackett–Burman (Plackett and Burman, 1946). Four variables were designated as "dummy" variables because no change is made to them, but they are used to give an estimate of the standard error for each factor. Each parameter (factor) is tested at two levels, a high (+) and a low (−) level, which will be determined after the optimization by RGA. The Plackett-Burman design contains a total of 47 trials.

6.2.4 Final Optimization by Quasi-Newton Algorithm

A simultaneous estimation of the kinetic parameters selected in the previous step is performed using the Quasi-Newton Algorithm, whereas all other parameters remained fixed in the global optimum region, calculated by the RGA. The FORTRAN IMSL routine DBCONF was used for this purpose. The straightforward idea is to implement the optimization problem as a nonlinear programming problem (NLP) that can be written as:

Minimize Equation 1

Subjet to $l_p \leq x_p \leq u_p$, p=1,…,5

where x_p is the parameters. The l_p and u_p are specified lower and upper bounds on the parameters, with $l_p \leq u_p$.

6.3. Results and Discussion

The optimized parameters by RGA are shown in Table 6.2 (see Stremel (2001) for details of nomenclature).

By considering the values optimized by RGA as central point in a Plackett-Burman design, the effects for all parameters on Eq. (1) (response), for a 95% confidence level, were calculated. The effects of the parameters are given in Table 6.3 (the five most significant parameters are marked in bold).

Table 6.2. Optimized parameters by RGA

Parameter	Param. values optimized by RGA	Parameter	Param. values optimized by RGA
k1	0.469	k1e	29.262
k2	0.262	k4e	1.264
k3	8.112	k8e	3.615E-03
k4	2.467E-03	k10e	0.892
k5	1418.132	KE	5.092E-02
k6	1.150	KE2	6.038E-02
k8	0.377	k1i	6.435
k9	2.883E-02	m2	3.477
k10	2.275E-03	m2e	3.476
s1	1.213E-02	k4i	9919.791
s2	4.920E-04	k5i	1981.616
s3	8.072E-02	k5r	0.303
s4	1.00E-06	k9i	2716.006
s5	0.358	k10i	504.813
s6	1.801E-02	kd	7.910E-04
s8	2.016E-02	F1	2.249
s9	1.0E-06	F2	1.276
s1e	0.147	F3	0.420
s5e	5.770E-02	F5	2.362
s8e	1.978E-03	F6	1.212
s9e	9.558	C	1.098
s10e	168.26		

A total of five most significant parameters, i.e. greater effects on error, Eq. (1), (KE, F1, F2, F3 and F5) were reestimated by Quasi-Newton Algorithm, which are showed in Table 6.4, whereas all other parameters remained fixed in the global optimum region, in the values set by the RGA. Under such optimal model, the computed profiles for ethanol (*P*) and substrate (*S*) are shown in the Figure 6.1 and Figure 6.2, respectively.

The residual standard deviation (RSD) (Atala, 2001), Eq. (2), written as a percentage of the average of the experimental values was the measurement used for characterizing the quality of the prediction of the model.

$$RSD(\%) = \frac{\left(\frac{1}{np} \left(\sum_{p=1}^{np} (d_p - x_p)^2 \right)^{0.5} \right)}{\bar{d}_p} \times 100 \qquad (2)$$

where x_p and d_p are, respectively, the value predicted by the mathematical model and experimental value, $\bar{d}_p$ is the average of the experimental values and *np* is the number of experimental points.

Table 6.3. Effect estimate on Eq. (1) from results of PB design

Factor	Effect	S.E.	t(4)	P	-95%	95%
Mean	0.081860*	0.000503*	162.6241*	0.000000*	0.080462*	0.083258*
k1	0.000571	0.001007	0.5673	0.600833	-0.002224	0.003366
k2	0.000514	0.001007	0.5107	0.636455	-0.002281	0.003309
k3	0.000470	0.001007	0.4669	0.664869	-0.002325	0.003265
k4	-0.000473	0.001007	-0.4696	0.663097	-0.003268	0.002322
k5	0.000766	0.001007	0.7604	0.489370	-0.002030	0.003561
k6	-0.012757*	0.001007*	-12.6718*	0.000223*	-0.015552*	-0.009962*
k8	0.006145*	0.001007*	6.1039*	0.003645*	0.003350*	0.008940*
k9	-0.001227	0.001007	-1.2185	0.289984	-0.004022	0.001568
k10	0.001798	0.001007	1.7855	0.148728	-0.000998	0.004593
s1	0.000386	0.001007	0.3830	0.721203	-0.002410	0.003181
s2	0.000459	0.001007	0.4557	0.672248	-0.002336	0.003254
s3	-0.000939	0.001007	-0.9324	0.403912	-0.003734	0.001856
s4	0.000494	0.001007	0.4910	0.649146	-0.002301	0.003289
s5	-0.000230	0.001007	-0.2285	0.830464	-0.003025	0.002565
s6	-0.000306	0.001007	-0.3036	0.776569	-0.003101	0.002489
s8	-0.000521	0.001007	-0.5174	0.632173	-0.003316	0.002274
s9	0.000187	0.001007	0.1855	0.861863	-0.002608	0.002982
s1e	-0.000262	0.001007	-0.2601	0.807626	-0.003057	0.002533
s5e	-0.001247	0.001007	-1.2391	0.283046	-0.004043	0.001548
s8e	0.000423	0.001007	0.4202	0.695930	-0.002372	0.003218
s9e	-0.000205	0.001007	-0.2038	0.848458	-0.003000	0.002590
s10e	-0.001174	0.001007	-1.1664	0.308253	-0.003969	0.001621
k1e	-0.002425	0.001007	-2.4092	0.073619	-0.005221	0.000370
k4e	-0.000807	0.001007	-0.8013	0.467855	-0.003602	0.001988
k8e	0.001162	0.001007	1.1543	0.312646	-0.001633	0.003957
k10e	0.000644	0.001007	0.6396	0.557227	-0.002151	0.003439
KE	**0.016200***	**0.001007***	**16.0913***	**0.000087***	**0.013405***	**0.018995***
KE2	0.007588*	0.001007*	7.5369*	0.001660*	0.004793*	0.010383*
k1i	0.001114	0.001007	1.1065	0.330574	-0.001681	0.003909
m2	0.000628	0.001007	0.6235	0.566748	-0.002167	0.003423
m2e	0.000541	0.001007	0.5371	0.619683	-0.002254	0.003336
k4i	-0.000171	0.001007	-0.1696	0.873557	-0.002966	0.002624
k5i	0.000715	0.001007	0.7101	0.516848	-0.002080	0.003510
k5r	0.000108	0.001007	0.1071	0.919866	-0.002687	0.002903
k9i	-0.000721	0.001007	-0.7166	0.513234	-0.003517	0.002074
k10i	-0.000685	0.001007	-0.6808	0.533368	-0.003481	0.002110
kd	0.000506	0.001007	0.5027	0.641591	-0.002289	0.003301
F1	**-0.024644***	**0.001007***	**-24.4788***	**0.000017***	**-0.027439***	**-0.021849***
F2	**0.022353***	**0.001007***	**22.2032***	**0.000024***	**0.019558***	**0.025148***
F3	**0.023282***	**0.001007***	**23.1265***	**0.000021***	**0.020487***	**0.026077***
F5	**0.021074***	**0.001007***	**20.9330***	**0.000031***	**0.018279***	**0.023869***
F6	-0.002826*	0.001007*	-2.8074*	0.048445*	-0.005621*	-0.000031*
C	-0.000778	0.001007	-0.7726	0.482874	-0.003573	0.002017

* Significant for a 95% confidence level.

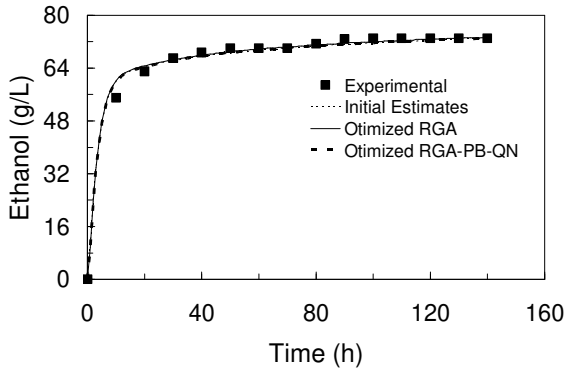


Figure 6.1. Ethanol concentration

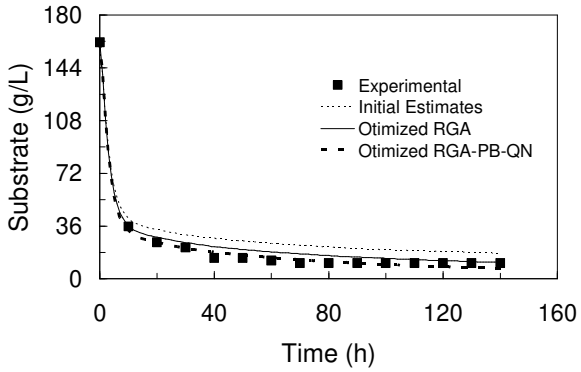


Figure 6.2. Substrate concentration

Table 6.4. Reestimated parameters by QN

Parameter	Parameter values optimized by QN
KE	5.099E-02
F1	2.244
F2	1.276
F3	0.435
F5	2.353

The RSD(%) for the model optimized by only RGA and the methodology proposed in this work (RGA-PB-QN) are shown in Table 6.5. The results obtained using the methodology RGA-PB-QN was better than the results using only RGA. It can be seen that the deviations are 8.6% and 2.3% for concentrations of substrate and ethanol, respectively. In bioprocess engineering, values of RSD(%) below 10% can be considered acceptable (Atala, 2001). Thus, the estimated model was able to fit experimental observations satisfactorily.

Table 6.5. Residual standard deviation, RSD(%) used to characterize the prediction quality of the model

Output variable	RSD(%)	
	RGA	RGA-PB-QN
Ethanol (g/L)	2.4	2.3
Substrate (g/L)	25.7	8.6

6.4. Concluding Remarks

Bioethanol (ethanol from biomass) is an attractive and sustainable energy source to fuel and raw material for chemical industries and is nowadays the largest

fermentation product obtained from the sugar cane. From ethanol it is possible to achieve high quality acetaldehyde, acetic acid, ethyl acetate, ethylene and from them a huge amount of chemicals, including polymers, making the sugar cane very interesting from the environmental point of view and an economically attractive raw material to be used to obtain chemicals. Either sugar cane or bagasse may be used as feedstock, and several biochemical pathways are known.

Compared with growing studies of microorganism populations, few improvements on the development and application of structured models for product formation have appeared. The work of Stremel (2001) evaluated the performance of a structured model for the alcoholic fermentation process.

The major problem with structured models is their large number of parameters, which makes the estimation procedure a difficult task. The use of deterministic optimization methods, such as Quasi-Newton algorithm, to estimate a large number of parameters (in the case studied 43) usually leads to lack of convergence. On the other hand, Genetic Algorithms are well suited to large-scale problems, but have the drawback of slow convergence. In this work it is proposed an estimation methodology in four steps that can be used always that a re-estimation of parameters is necessary.

The first step is to obtain initial values for all parameters in the model and then a Real Coded Genetic algorithm is used to calculate the parameters in the model. The third step is to identify the most significant of the 43 parameters using Plackett and Burman design and finally the most significant (in the case studied 5 parameters) are optimized using a Quasi-Newton Algorithm, which converges much more quickly than RGA to the optimal.

The results have shown that the performance of the model to describe the experimental data (measured as RSD%) is improved using the proposed methodology in comparison with a model whose parameters were only optimized by RGA.

Finally, it should be noted that although the example used in this work had the parameters already near their optimal values, the proposed methodology can

be applied in many other parameter estimation problems and can be used as a reliable tool for optimizing other types of fermentation processes.

Acknowledgements

The authors acknowledge FAPESP and CNPq for financial support.

References

1. Andrietta, S.R., Maugeri, F., (1994). Optimum design of a continuous fermentation unit of an industrial plant for alcohol production. *Advances in Bioprocess Engineering*, **1**, 47-52.
2. Atala, D.I.P., Costa, A.C., Maciel Filho, R., Maugeri, F., (2001). Kinetics of ethanol fermentation with high biomass concentration considering the effect of temperature. *Applied Biochemistry and Biotechnology*, **91-93**(1-9), 353-366.
3. Dourado, A., Goma, G., Albuquerque, U., Sevely, Y., (1987). Modeling and static optimization of the ethanol production in a cascade reactor. I. Modeling. *Biotechnology and Bioengineering*, **29**, 187-194.
4. Park, L.J., Park, C.H., Park, C. and Lee, T., Application of genetic algorithms to parameter estimation of bioprocesses. *Medical and Biological Engineering and Computing*, 1997, **35** 47-49.
5. Plackett, R.L., Burman, J.P. The design of optimun multifactorial experiments. *Biometrika*, 1946, **33** 305-325.
6. Sáiz, J.M.G., Pizarro, C., Vidal, D.G., Evaluation of kinetic models for industrial acetic fermentation: Proposal of a new model optimized by genetic algorithm. *Biotechnology Progress*, 2003, **19** 599-611.
7. Schügerl, K., Progress in monitoring, modeling and control of bioprocess during the last 20 years. *Journal of Biotechnology*, 2001, **85** 149-173.
8. Rotboll, M., Jorgensen, S.B. Validation of a metabolic and Biomass Structured Model for Yeast Fermentation, in: *European Conference on Biotechnology*, Florence, Italy, 1993.

9. Stremel, D.P. Development of alternative structured models for ethanol process production. PhD thesis, State University of Campinas (UNICAMP), Campinas-SP, Brazil, 2001.
10. Yedder, R.B., Optimisation numérique et Contrôle optimal: (applications en chimie moléculaire). PhD thesis, Ecole Nationale des Ponts et Chaussées (ENPC), Paris, France, 2002.

CAPÍTULO 7. DEVELOPMENT OF ADAPTIVE MODELING TECHNIQUES TO DESCRIBE THE TEMPERATURE DEPENDENT KINETICS OF BIOTECHNOLOGICAL PROCESSES

Trabalho aceito para publicação na revista **Biochemical Engineering Journal**, 2007.
Elmer Ccopa Rivera, Aline C. Costa, Rafael R. Andrade, Daniel I. P. Atala, Francisco Maugeri, Rubens Maciel Filho.

Abstract

Bioprocesses are quite difficult and expensive to model, since their operation involves microbial growth under constantly changing conditions, with impact on process kinetics and performance. Hence, there is a need and incentive for the improvement of methods for rapid development of simple, though realistic, mathematical models. In this work the modelling of biotechnological processes is studied with focus on developing methodologies that can be used whenever a re-estimation of parameters is necessary. The ethanol fermentation process is used as a case study. The performance of a hybrid neural model and a balance based model, both considering the effect of temperature on the kinetics, are evaluated not only by their accuracy in describing experimental data, but also by the difficulties involved in the adaptation of their parameters. Experiments are performed to develop the two models and further experiments (using sugar cane molasses from a different harvest and a different production medium) validate the methodologies for re-estimation of kinetic parameters.

Keywords

Bioreactors, modeling, optimization, parameter estimation, temperature effect, artificial intelligence.

7.1 Introduction

The interest in biotechnology-based production of fuels tends to augment with the concern about exhaustion of fossil fuels and the increase in their price. The world meetings make clear that policies for renewable energy are essential to achieve sustainable development in a broad sense. Environmental protection, job creation, alleviation of external debts in developing countries and security of supply are some of the key issues to mention **[1]**. In Brazil, the sugar cane industry keeps the greatest commercial energy production in the world with ethanol and the almost complete use of sugar cane bagasse as fuel.

There are many minor industrial problems associated with the ethanol fermentation processes to be solved nowadays, when optimal operation is a target.

Among them is the lack of the processes robustness in the presence of fluctuations in operational conditions, which leads to changes in the kinetic behavior, with impact on yield, productivity and conversion. These changes are very common in ethanol plants, where they occur not only due to the variations in the quality of the raw material but also due to variations of dominant microorganism in the process.

Another issue in ethanol fermentation processes is the influence of temperature on the kinetics. It is difficult to support a constant temperature during large-scale alcoholic fermentation and variations in temperature affects productivity through changes in kinetics as well as in microorganism lifetime. The temperature in a typical industrial fermentor varies from 33.5ºC during the night to 35 ºC at the end of the day due to fluctuations in the cooling water temperature. In plants with poor temperature control the temperature in the fermentor goes up to 40ºC. Thus, a description of the influence of temperature on kinetics of the microorganism involved is essential for a reliable mathematical modeling to be used in process optimization, control and operation.

During the last years, many studies of the mathematical modeling of ethanol fermentation processes have been carried out **[2, 3]**. However, despite the importance of temperature, there are very few works in the literature taking into account the effect of temperature on the kinetic parameters. Among them is the work of Aldiguier et al. **[4]**, which determines different values for kinetic parameters in different temperatures, but do not determine a temperature function to describe the kinetic parameters. In a recent work, Phisalaphong et al. **[5]** used equations correlating temperature and kinetic parameters to develop a mathematical model describing temperature evolution.

Although balance based mathematical models provide understanding about the process, practical experience has shown that they are only valid for the specific conditions in which they were determined. Another way to deal with the problem is to use hybrid neural models **[6-13]**. As was demonstrated by Psichogios and Ungar **[6]**, it is simple to propose a bioreactor mathematical model through application of mass balances on the process variables. The really difficult part is the mathematical representation of the kinetics. The kinetic rates are most often

poorly understood nonlinear functions, while the corresponding parameters are in general time-varying **[14]**. Literature presents a great number of approximate kinetic models, which take different factors into account, and the proper choice of the mathematical description of these expressions is object of discussion in many works **[3, 15-17]**. Hybrid neural models are, thus, suitable to modeling biotechnological processes, as they combine mass and energy balance equations with neural networks, which represent the unknown kinetics. These models are expected to perform better than "black-box" neural network models **[6, 18-20]**, since generalization and extrapolation are confined only to the uncertain parts of the process and the basic model is always consistent with first principles. Besides, significantly less data are required for their training **[6]**.

In this work an adaptive methodology for hybrid modeling of the effect of temperature on the kinetics of batch fermentation was proposed. The rate expressions for cell growth, substrate consumption and product formation are described by multilayer perceptron neural networks (MLPNN) and the neural network parameters are re-estimated in an adaptive scheme when there are changes in operational conditions and fluctuations in the quality of raw material.

The objective of this work was to present a comparison of methodologies for the adaptive modeling of biotechnological processes. The use of balance based and hybrid models was evaluated considering the accuracy with which they describe experimental data and the difficulties involved in the re-estimation of the kinetics parameters.

7.2 Material and Methods

7.2.1 Experiments

Experiments used to develop the mathematical models (first data set)

Five batch experiments (at 28, 31, 34, 37 and 40^{o}C) were used to estimate the parameters of the proposed models. Details about these experiments are described elsewhere **[21]**. The initial conditions and temperatures of these experiments are given in Table 7.1.

Experiments with changes in operational conditions (second data set)

Another five batch experiments were used to validate the methodologies for re-estimation of kinetic parameters. The microorganism used was *Saccharomyces cerevisiae* cultivated in the Bioprocess Engineering Laboratory in the Faculty of Food Engineering/State University of Campinas and obtained from an industrial fermentation plant. The growth medium for inoculum, materials and analytical methods are the same as described in **[21].** There is a difference in the production medium used, which consisted only of diluted sugarcane molasses, while in the first experiments yeast extract and $MgSO_4$ were added. Sugar cane molasses was from a harvesting period different from that used in the first experiments and so there is a change in the quality of raw material when this data set is compared to the first one. The initial conditions and temperatures of these experiments are given in Table 7.1.

Table 7.1. Initial conditions and temperature of the experiments

	Initial conditions		
Temperature (°C)	Biomass (kg/m^3)	TRS (kg/m^3)	Ethanol (kg/m^3)
First data set			
28	1.90	182	11.7
31	1.90	213	2.60
34	1.60	242	0.60
37	0.60	272	0.00
40	1.20	244	0.00
Second data set			
30	1.20	128	3.20
31.2	0.40	86.8	3.70
34	0.70	119	3.90
36.8	0.30	84.6	4.80
38	1.00	119	5.00

7.3 Mathematical modeling

7.3.1 Balance based modeling

Phenomenological or balance based models comprise the mass balance equations, with microorganism growth, substrate consumption and ethanol formation for a batch reactor described as follows:

$$\frac{dX}{dt} = r_x \tag{1}$$

$$\frac{dS}{dt} = -r_s \tag{2}$$

$$\frac{dP}{dt} = r_p \tag{3}$$

where X is the concentration of cell mass (kg/m^3), S is the concentration of substrate (kg/m^3) and P is the concentration of ethanol (kg/m^3).

Before a balance based model can be used, the rates of cell growth, r_x (kg/[m^3·h]), substrate consumption, r_s (kg/[m^3·h]), and product formation, r_p (kg/[m^3·h]) must be determined. This can be done by choosing among various kinetic models, most of them empirical and based on either Monod's equation or its numerous modifications, which take into account the inhibition of microbial growth by a high concentration of product and substrate **[17]**.

The methodology for the calculation of the kinetic parameters as a function of temperature used in this work is described below:

(1) Determine the appropriate forms of kinetic rates.

Experimental observations have shown that biomass, substrate and product inhibitions are significant for ethanol fermentation **[15]**. Eq. (4) shows the cell growth rate equation, r_x, which includes terms for such types of inhibitions.

$$r_x = \mu_{max} \frac{S}{K_s + S} \exp(-K_i S)\left(1 - \frac{X}{X_{max}}\right)^m \left(1 - \frac{P}{P_{max}}\right)^n X \tag{4}$$

where μ_{max} is the maximum specific growth rate (h^{-1}), K_s the substrate saturation constant (kg/m^3), K_i the substrate inhibition parameter, X_{max} the biomass concentration when cell growth ceases (kg/m^3), P_{max} the product concentration when cell growth ceases (kg/m^3), and m and n are the parameters used to describe cellular and product inhibitions, respectively.

In this study, Luedking-Piret expression **[22]** was used to account for the ethanol formation rate, r_p.

$$r_p = Y_{px} r_x + m_p X \tag{5}$$

where Y_{px} is Luedeking-Piret growth associated constant (kg/kg) and m_p is the Luedeking-Piret non growth- associated constant (kg/[kg·h]).

The substrate consumption rate, r_s, is given by Eq. (6), describing the sugar consumption during fermentation, which leads to cell mass and ethanol formation.

$$r_s = (r_x / Y_x) + m_x X \tag{6}$$

In this equation, Y_x and m_x are the cell yield (kg/kg) and maintenance parameter (kg/[kg·h]), respectively.

(2) Estimate a set of temperature dependent parameters for each temperature considered in the experiments.

Some of the parameters in the kinetic expressions above (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px}) are known to be dependent on temperature **[15]**. Let θ specify the parameters vector, which contains all the temperature-dependent parameters. The objective of the mathematical estimation of model parameters is to find out θ by minimizing the objective function, min $E(\theta)$:

$$E(\theta) = \sum_{n=1}^{np} \left[\frac{(X_n - Xe_n)^2}{Xe_{max}^2} + \frac{(S_n - Se_n)^2}{Se_{max}^2} + \frac{(P_n - Pe_n)^2}{Pe_{max}^2} \right] = \sum_{n=1}^{np} \varepsilon_n^2(\theta) \tag{7}$$

where Xe_n, Se_n and Pe_n are the measured concentrations of cell mass, substrate and ethanol at the sampling time n. X_n, S_n and P_n are the concentrations computed by the model at the sampling time n. Xe_{max}, Se_{max} and Pe_{max} are the maximum measured concentrations and the term np is the number of sampling points. Here $\varepsilon_n(\theta)$ is the error in the output due to the nth sample.

The determination of the feasible region of the total search space in the multiparameter optimization of the deterministic model is complex. For that reason, in a previous work **[21]**, the optimization was evaluated using different optimization techniques. Quasi-Newton algorithm (deterministic technique) was compared to the real-coded genetic algorithm (stochastic technique) to solve the problem of parameter estimation. From the computational results, it was observed that the two algorithms have shown similar good performances. Consequently, it is possible to apply deterministic or stochastic non-linear optimization to solve the estimation problem to obtain the kinetic parameters.

The parameters which are not temperature dependent are fixed in the values given by **Atala et al. [15]** and are: K_s = 4.1 kg/m^3, K_i = 0.002 m^3/kg, m_p = 0.1 kg/[kg·h], m_x = 0.2 kg/[kg·h], m = 1.0 and n = 1.5.

(3) Propose an equation to describe the influence of temperature and fit it to the optimized values obtained for each temperature.

The influence of temperature on μ_{max}, X_{max}, P_{max} and Y_{px}, is non-linear and Eq. (8) can be used to express it:

$$\text{temperature-dependent parameter} = A\exp\left(\frac{B}{T}\right) + C\exp\left(\frac{D}{T}\right) \quad (8)$$

The influence of temperature on Y_x was described by Eq. (9):

$$\text{temperature-dependent parameter} = A\exp\left(\frac{B}{T}\right) \quad (9)$$

In these equations, A, B, C and D are constants, and T is the temperature in ^oC. The Levenberg-Marquardt algorithm is used to determine the best parameters (A, B, C and D) for each temperature dependent kinetic parameter.

In a real industrial process **Andrietta et al. [23]** determined the existence of four groups of kinetic parameters in different harvesting periods in the year. This indicates that, in order to describe the dynamic behaviour of the plant for long periods of time, the balance based model should have its kinetic parameters updated periodically.

7.3.2 Hybrid modeling

Besides phenomenological modeling, another option is the use of a hybrid model in which the temperature-dependent kinetics is described by multilayer perceptron neural networks. In this study, the structure of the hybrid model is derived taking into consideration the mass balances (Eqs. 1-3) for the batch fermentation process, with neural networks describing the rate expressions for cell growth, r_x, substrate consumption, r_s, and product formation, r_p.

A multilayer perceptron neural network consists of three types of layers: an input layer, an output layer and one or more hidden layers. Each layer may have a

different number of neurons, and even a different transfer function. An appropriate architecture would help to achieve higher model accuracy.

For the current study, each rate expression (r_x, r_s, and r_p) was modeled with a MLPNN with four inputs (concentrations of biomass, substrate and ethanol, and temperature), a single hidden layer, described mathematically by Eq. (10), and one output. Tests have shown that the use of four neural networks with one output led to better results than using one neural network with four outputs. Both input and output data were normalized to the range [0, 1].

$$\hat{y} = g[x,\theta] = F\left[\sum_{j=1}^{M} W_j f_j\left(\sum_{l=1}^{N} w_{jl} x_l + w_{j0}\right) + W_0\right] \tag{10}$$

In Eq. (10), θ specifies the parameter vector, which contains all the adjustable parameters of the network; i.e., the weights and biases $\{w_{j,l}\ W_j\}$.

It follows from Cybenko's theorem **[24]** that all continuous functions can be approximated to any desired accuracy with a network of one hidden layer of sigmoidal ($f(x) = 1/(1+\exp^{-x})$) hidden units (nodes) and a layer of linear output nodes. Such structure is used in this work. Nguyen–Widrow initialization algorithm is used for initialization of weights and biases and is subsequently trained with the Levenberg–Marquardt algorithm in Matlab's neural network toolbox. Training was stopped after 1000 epochs.

The input data for the neural networks are available from direct measurement, but the kinetic rates, chosen as the networks outputs, are not directly measurable and must be estimated. One way to perform this task is by discretization of the balance equations (Eqs. 1-3). However, due to the small number of data points available, this approach did not lead to good results. In order to solve this problem, a sigmoidal function was fitted to the experimental data. The sigmoidal-Boltzman, Eq. (11), was used to fit biomass, substrate and product versus time for each temperature. In Eq. (11), y can be the concentration of biomass, substrate or product and a, b, c and d are constants for each temperature. The derivative of Eq. (11), fitted for X, S and P, will give dX/dt, dS/dt and dP/t, respectively. Besides enabling the calculation of the kinetic rates, Eq.

(11) was used to increase the data set for training by interpolation. The fitted sigmoidal function for the data at 34ºC is shown in Figure 7.1. This equation fitted the data accurately for all the temperatures.

$$y = a + \frac{b}{1 - \exp\left[-\frac{c - time}{d}\right]} \tag{11}$$

The appropriate number of nodes to be included in the hidden layer was addressed with the cross-validation technique in order to avoid model over-fitting. This technique splits the sample data into a training sample and a validation sample. Then MLPNN with different numbers of hidden nodes are trained with the training sample, and their performance monitored with the validation sample in terms of the lowest error.

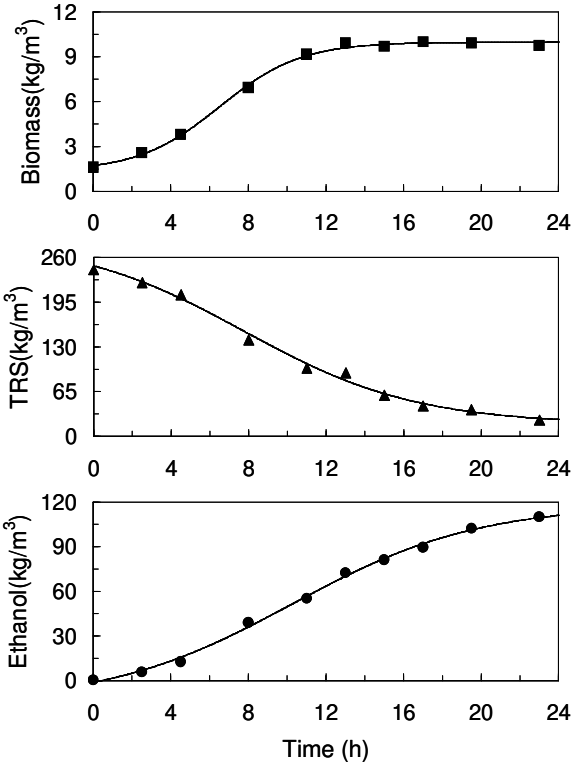


Figure 7.1. Sigmoidal-Boltzman fitting for experimental data at 34ºC. The experimental data are for concentrations of cell mass, X (■); Substrate (Total reducing sugars), S (▲) and ethanol, P (●)

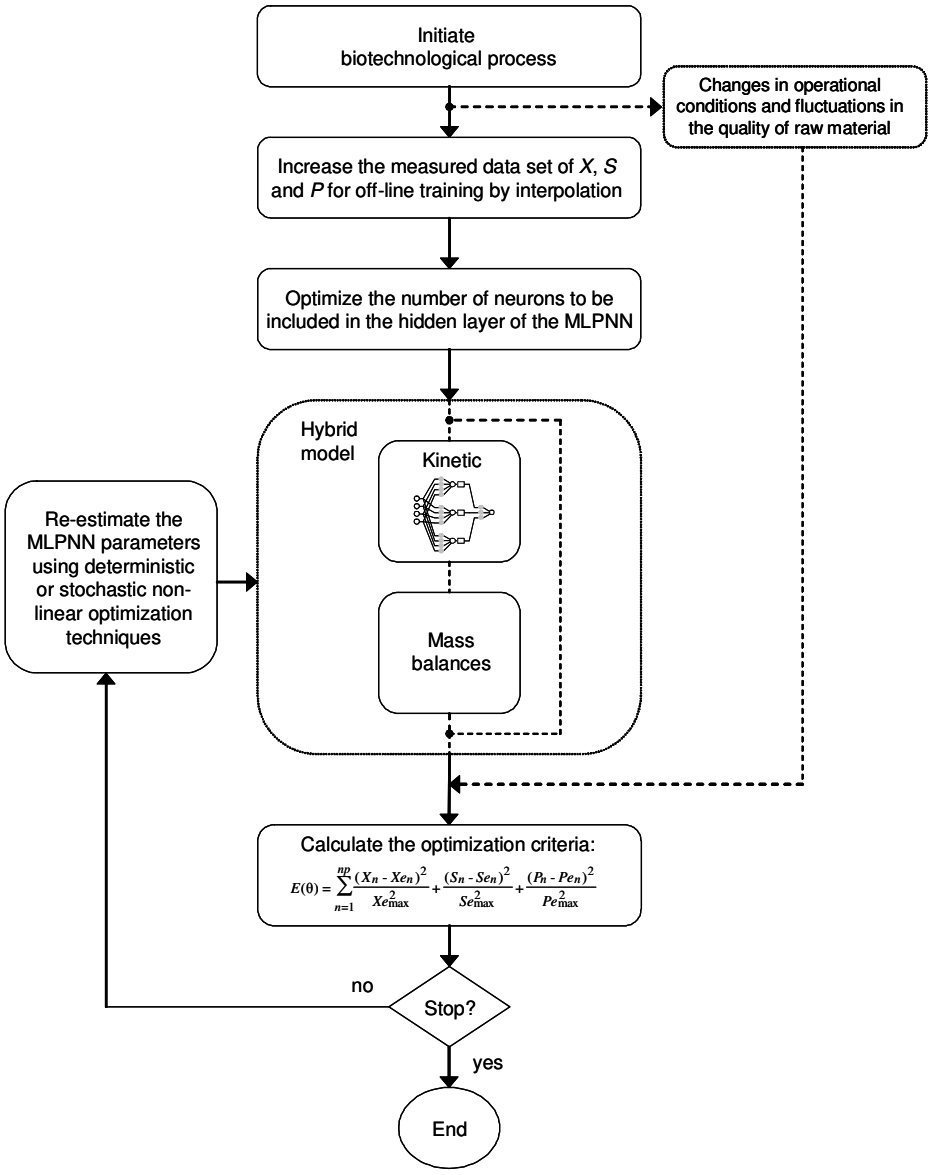


Figure 7.2. Flow diagram showing the adaptive approach that re-estimates the parameters of the neural networks

As the kinetic behavior of ethanol fermentation changes with fluctuations in raw material quality and changes in operational conditions, re-estimation of network weights will be necessary. The re-estimation of the neural network parameters was performed by minimizing the objective function in Eq. (7). In this case, θ specify the parameter vector, which contains all the neural network parameters. The variables X_n, S_n and P_n are the concentrations computed by the hybrid model at the sampling time n. The MLPNN parameters are adapted so that the error between the measured and predicted value of concentrations of cell

mass, substrate and ethanol are acceptable. At this point it is worthwhile mentioning that the parameter re-estimation via hybrid model is less time consuming and easy to be made, especially because it does not require the mapping of feasible regions for the parameter estimation, which is an interesting feature for real-time applications.

The proposed algorithm for adaptation of the neural network parameters in the hybrid model is shown in Figure 7.2. Initially the appropriate neural network architecture, including the initial parameter set, is determined. After this step, if there are changes in operational conditions and/or fluctuations in the quality of raw material, the model can be directly adapted by minimizing Eq. (7), as shown by the dashed lines in the figure. If a minimum is reached, the re-estimation is terminated. If not, the neural network parameters are adapted and a new iteration begins with the hybrid model simulation.

7.4 Results and Discussion

7.4.1 Prediction quality of the models

Experimental observations at five temperatures (28, 31, 34, 37 and 40^{o}C) are first used to estimate the kinetic model parameters. Five other data sets (at 30, 31.2, 34, 36.8 and 38^{o}C and different initial substrate concentrations) are then considered for re-estimating the parameters of the models. The performance of the models can be assessed by using the residual standard deviation (RSD) described as a percentage of the average of the real values, $\bar{y}_i$, as described by **Atala et al. [15]**:

$$\text{RSD}(\%) = \frac{\sqrt{\text{RSD}}}{\bar{y}_i} \times 100 \tag{12}$$

$$\text{RSD} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_{Pi})^2}{n}} \tag{13}$$

where y_i is the experimental observation, y_{Pi} is the value predicted by the balance based model or the hybrid model and n is the number of points. Values of RSD(%) up to 10% indicates an acceptable prediction in biotechnological engineering **[15]**.

7.4.2 Results for balance based model

The model kinetic parameters were estimated using the results of the first five experiments **[21]** and its prediction at 34 ºC is shown in Figure 7.3. Table 7.2 shows the RSD (%) values for all the temperatures. When the same model was used to describe the second data set, the prediction quality is poor, as can be seen in Figure 7.4 for the data at 38ºC. This figure shows that conditions such as molasses harvesting and medium composition affects process performance (kinetics and dynamic behavior). Such changes occur frequently in industrial operations, and this reinforces the importance of adaptation of kinetic parameters. The RSD(%), calculated by Eq. (12), obtained using the balance based model without re-estimation for all the experiments can be seen in Table 7.3. It can be noticed that in many situations the values of RSD(%) are unacceptably high.

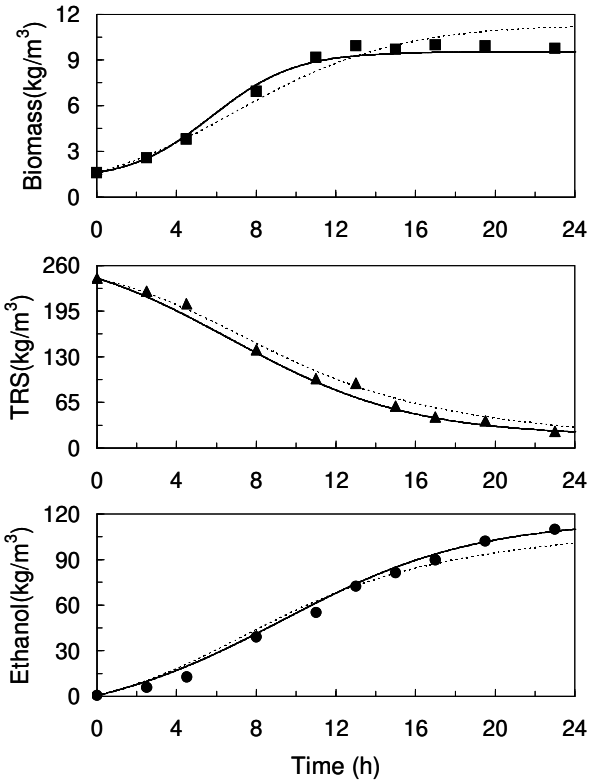


Figure 7.3. Prediction of the original balance based and hybrid models for the original data at 34ºC. The experimental data are for concentrations of cell mass, X (■); substrate (Total reducing sugars), S (▲) and ethanol, P (●). Simulated results of the models were represented by lines (Hybrid Model —; Balance Based Model ---)

Table 7.2. Residual standard deviation (RSD) written as a percentage of average of experimental values, used to characterize the prediction quality of the balance based and hybrid models for the original data set

Output variable	RSD(%)									
	T=28°C		T=31°C		T=34°C		T=37°C		T=40°C	
	BBM	HM	BBM	HM	BBM	HM	BBM	HM	BBM	HM
X	6.67	4.62	4.77	3.23	9.41	4.91	8.98	2.77	6.13	2.86
S	6.61	2.14	6.26	5.31	7.14	8.00	7.08	3.30	3.48	2.55
P	9.20	5.43	6.84	8.33	10.6	6.71	7.60	3.30	10.2	6.68

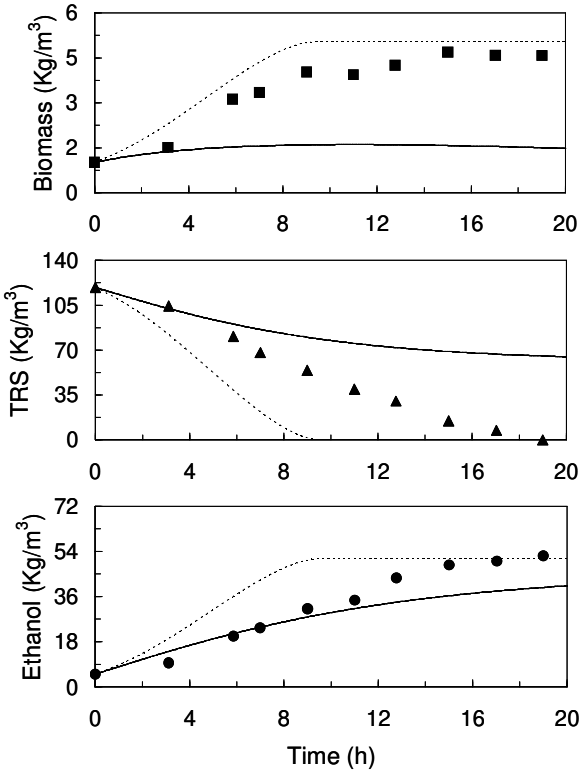


Figure 7.4. Experimental and simulated data for batch fermentation experiments at 38°C. The experimental data are for concentrations of cell mass, *X* (■); substrate (Total reducing sugars), *S* (▲) and ethanol, *P* (●). Simulated results of the models without re-estimation of their parameters were represented by lines (Hybrid Model —; Balance Based Model ---)

Table 7.3. Residual standard deviation (RSD) written as a percentage of average of experimental values, used to characterize the prediction quality of the balance based and hybrid models without parameters re-estimation in the presence of changes in the quality of raw material

Output variable	RSD(%)									
	T=30°C		T=31.2°C		T=34°C		T=36.8°C		T=38°C	
	BBM	HM	BBM	HM	BBM	HM	BBM	HM	BBM	HM
X	62.5	52.2	58.6	191	29.8	125	24.6	114	21.0	65.5
S	30.0	15.9	23.1	36.6	15.8	44.8	24.4	116	60.5	74.4
P	143	286	29.1	72.7	20.6	68.4	31.6	25.1	37.6	24.1

For the re-estimation of the parameters, Eqs. (1-3) were solved using a Fortran program with integration by an algorithm based on the fourth-order Runge-Kutta method **[25]**. The temperature dependent parameters (μ_{max}, X_{max}, P_{max}, Y_x and Y_{px} in Eqs. 4-6) were determined by minimizing Eq. (7) using a Quasi-newton algorithm. The FORTRAN IMSL routine DBCONF was used for this purpose. This procedure was repeated for each temperature considered (30, 31.2, 34, 36.8 and 38°C). The parameters that are not temperature-dependent were not altered.

Table 7.4. Constant values in Eqs. (8) and (9) and coefficient of determination (r^2) of parameters fitting as functions of temperature (in ºC).
Constants valid from 30 to 38ºC

Param.	(r^2)	*A*	*B*	*C*	*D*
μ_{max}	(0.99)	$1.18\times10^{3}\pm$ 1.09×10^{3}	$-244\pm$ 19.9	$-2.65\times10^{4}\pm$ 1.16×10^{4}	$-368\pm$ 33.5
X_{max}	(0.99)	$2.25\times10^{-17}\pm$ 5.91×10^{-18}	$1.25\times10^{3}\pm$ 7.09	$9.54\pm$ 0.0854	$50.9\pm$ 0.316
P_{max}	(0.99)	$5.61\times10^{-19}\pm$ 2.49×10^{-19}	$1.38\times10^{3}\pm$ 11.8	$88.3\pm$ 1.32	$-5.67\pm$ 0.531
Y_{px}	(0.99)	$-2.98\times10^{-16}\pm$ 6.53×10^{-17}	$1.13\times10^{3}\pm$ 5.96	$10.9\pm$ 0.0865	$-3.02\pm$ 0.284
Y_x	(0.99)	$0.0144\pm$ 3.00×10^{-5}	$34.4\pm$ 0.0632		

In order to describe the correlation between temperature and the parameters, the data was smoothed and interpolated and the expressions given by Eqs. (8) and (9) were fitted. The obtained constants (*A*, *B*, *C* and *D*) are shown in Table 7.4. This table also shows the coefficient of determination (r^2) of the optimal parameters fitting as functions of temperature and the confidence intervals. As it was already shown that changes in medium and raw material affects the kinetics of the process, always that these changes occur the adaptation of kinetic parameters is necessary and the values of the constants will be changed. Figure 7.5 shows the behavior predicted by Eqs. (8) and (9) with temperature and the optimized parameters estimated from the experimental data in the range of 30-38°C.

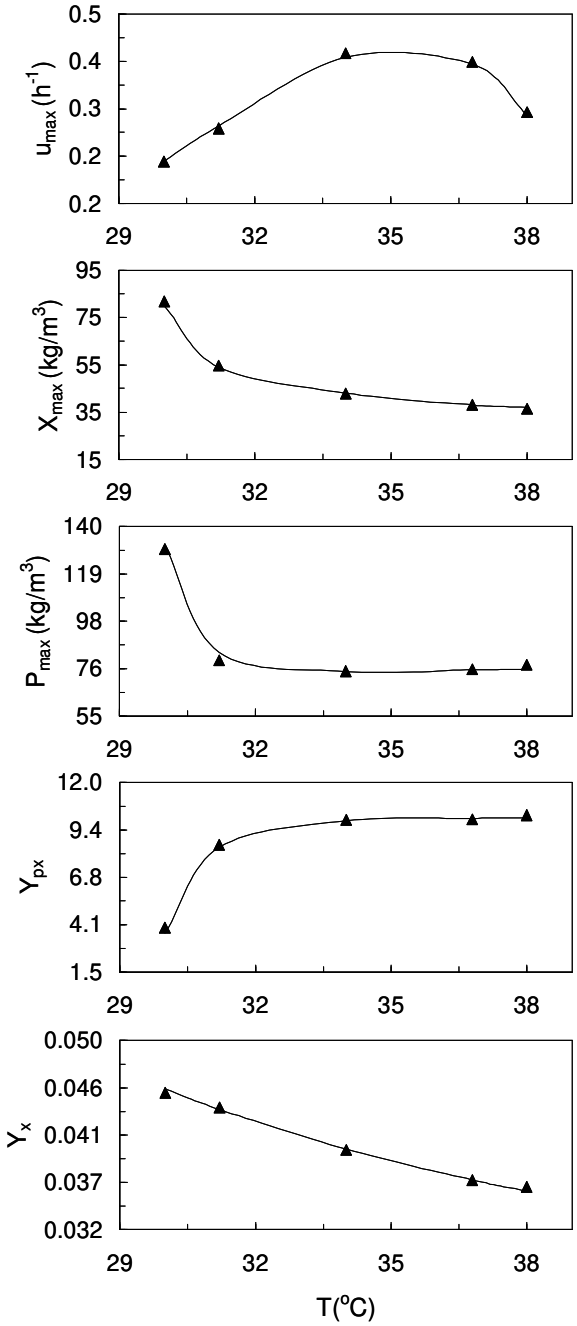


Figure 7.5. Parameters behavior with temperature at 30, 31.2, 34, 36.8 and 38°C

The performance of the model with re-estimated parameters in describing the experimental data is shown in Figures 7.6 to 7.8. In Table 7.5 the RSD (%) of the adapted balance based model for all the experiments is depicted. It can be seen that the model with re-estimation of the temperature dependent kinetic parameters described accurately the experimental data, presenting low values of RSD(%).

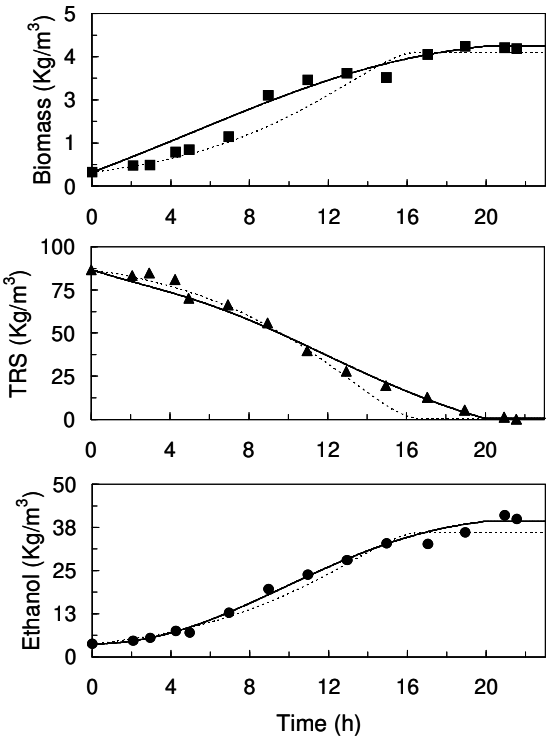


Figure 7.6. Experimental (cell mass, X(■); substrate (Total reducing sugars), S(▲) and ethanol, P(●)) and modeling (Hybrid Model —; Balance Based Model ---) results with parameters re-estimation at 31.2°C

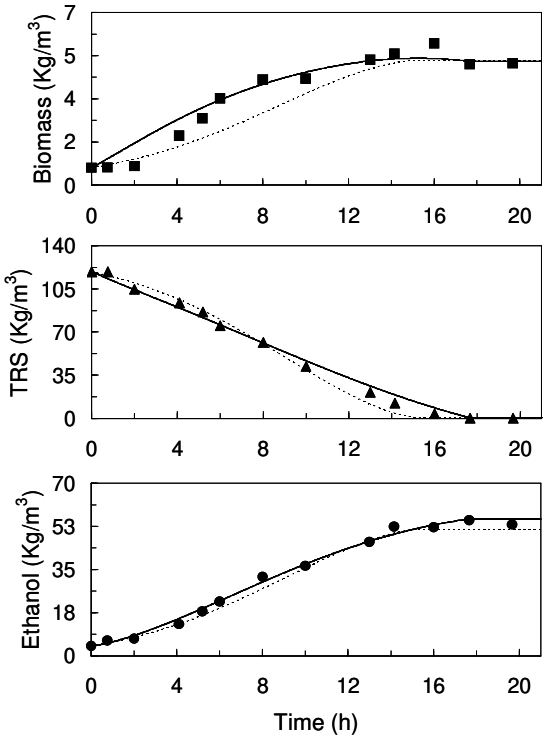


Figure 7.7. Experimental (cell mass, X(■); substrate (Total reducing sugars), S(▲) and ethanol, P(●)) and modeling (Hybrid Model —; Balance Based Model ---) results with parameters re-estimation at 34°C

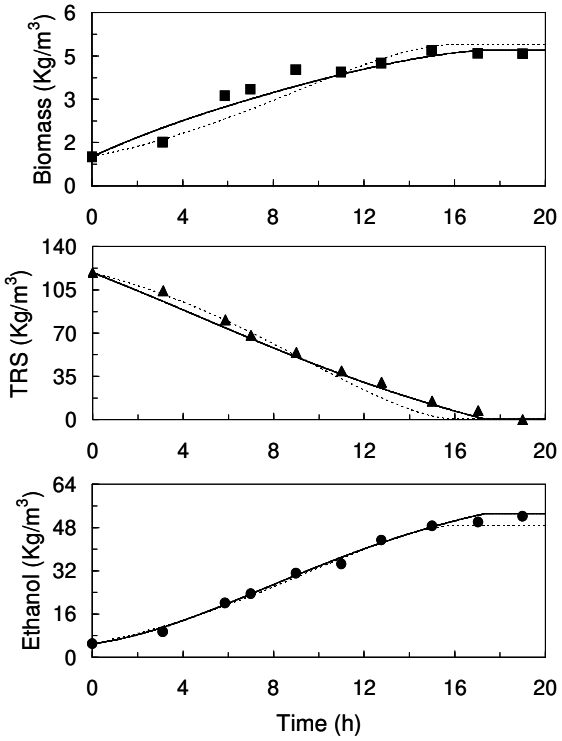


Figure 7.8. Experimental (cell mass, X (■); substrate (Total reducing sugars), S (▲) and ethanol, P (●)) and modeling (Hybrid Model —; Balance Based Model ---) results with parameters re-estimation at 38°C

Table 7.5. Residual standard deviation (RSD) written as a percentage of average of experimental values, used to characterize the prediction quality of the balance based and hybrid models with parameters re-estimation in the presence of changes in the quality of raw material.

Output variable	RSD(%)									
	T=30°C		T=31.2°C		T=34°C		T=36.8°C		T=38°C	
	BBM	HM	BBM	HM	BBM	HM	BBM	HM	BBM	HM
X	10.1	8.63	15.1	12.6	19.0	12.0	10.5	12.3	12.5	8.56
S	24.2	5.88	11.4	8.26	7.66	6.59	13.7	11.4	12.4	8.84
P	13.1	9.26	11.2	6.72	6.92	5.01	4.36	7.70	5.21	4.24

7.4.3 Results for hybrid model

7.4.3.1 Training and validation

The batch experimental observations at five temperatures (28, 31, 34, 37 and 40°C) **[21]** are used to choose the neural network architectures and for the estimation of their weights. The database generated for training (600 discrete-time data) and validation (100 discrete-time data) were presented to the neural network in a randomized sequence.

The MLPNN used has a single hidden layer. The number of hidden nodes was varied from 2 to 8, and the optimal number chosen by the cross-validation criterion with the number of epochs fixed at 1000 for all the studied architectures. The neural network with five hidden nodes for r_x, five hidden nodes for r_s and four hidden nodes for r_p were found to present the lowest RSD(%) for the validation sample. The kinetic rates (r_x, r_s, and r_p) were then modeled with MLPNNs with 31, 31 and 25 scalar parameters (weights and bias), respectively. The results are shown in Figure 7.3 at 34ºC and RSD(%) values in all temperatures are shown in Table 7.2. It can be seen that the hybrid model describes the original data accurately.

The hybrid model trained with the first five batch runs was used to describe the new experimental results. Figure 7.4 shows a comparison of the predicted state profiles for both, hybrid model (solid curves) and balance based model (dashed curves) for the experimental data at 38^{o}C. Table 7.3 shows the values of RSD (%) for the hybrid model without adaptation of the weights for all the new experiments. It can be seen that for most of the cases the values of RSD(%) are very high, which shows the importance of the use of an adaptive scheme.

7.4.3.2 Re-estimation of network weights

The neural network parameters were re-estimated for the new experiments using the methodology described on Figure 7.2. The results are shown in Figures 7.6 to 7.8. It can be noticed that the adapted hybrid model effectively tracks the desired trajectory of experimental observations for concentrations of biomass, substrate and product at 30, 31.2, 34, 36.8 and 38^{o}C. Table 7.5 shows the RSD(%) for the adapted hybrid model for all the experiments.

It is worthwhile mentioning that the hybrid model obtained is able to deal with temperature disturbances in the process and different initial substrate concentrations since the data-driven identification procedure is carried out adequately with a suitable set of data representative of the process.

Tables 7.2, 7.3 and 7.5 can be used for comparison between the two models and between models with and without updating of their parameters. From Table

7.2 it can be noticed that the two models described accurately the experiments without changes in operational conditions, with the hybrid model presenting lower values of RSD(%) in most of the cases. When the models are used to describe experiments with changes in operational conditions without parameters re-estimation, the deterministic model performs better than the hybrid model, as can be seen in Table 7.3. When adaptation of the kinetic parameters is carried out, Table 7.5 shows the predictions of both models are similar and closely match the observed data, with the hybrid model presenting the lowest RSD(%) values in most cases.

7.5 Concluding remarks

The difficulty in modeling biotechnological processes is essentially on the precise description of the kinetics and robust modeling can only be achieved by incorporating reliable procedures to easily update the model when there are changes in operational conditions. An accurate model should also consider the influence of temperature on the process kinetics, as it has been shown that changes in temperature have great influence on the kinetic behavior.

The performances of a balance based model and a hybrid model to describe the dynamic behaviour of the ethanol fermentation process were assessed. When using the non adapted models to describe data in the presence of changes in operational conditions, the balance based model presented better performance. However, even this model presented unacceptably high values of RSD(%), which indicated the importance of re-estimating the kinetic parameters. The adapted models presented similar performances.

Comparing the adaptation procedures, the re-estimation of the network weights was simpler than the re-estimation of the kinetic parameters of the balance based model. Even considering that the rate equations and the functions that describe the parameters dependence with temperature are known in the balance based model, the estimation problem is complex and time consuming. This suggests that using a balance based model in a situation where frequent re-estimation is necessary could be a limitation.

The updating of the hybrid model, however, is straightforward. The structure of the neural network (number of layers and of neurons in each layer) was fixed and the weights were re-estimated. The use of this approach enables the implementation of an on-line re-estimation procedure.

This work has shown that the combination of MLPNNs with mass balances was able to describe the fermentative process with remarkably high prediction accuracy. The values of RSD(%) were lower than that of a detailed phenomenological model for most of the conditions. This approach can partly eliminate the difficulties of having to specify completely the complex structure of the kinetics of a fermentation process. Although this could not be a very significant advantage for a process well studied and known as the alcoholic fermentation, it can make a great difference for less known biotechnological processes, as it enables a rapid determination of a mathematical description that can be used for on-line optimization, soft sensor and control.

It should be emphasized that the application of the adaptation procedures (for the balance based and the hybrid model) presented in this work to other biochemical processes is straightforward.

Acknowledgements

The authors acknowledge Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support.

Nomenclature

f activation function of the MLPNN

K_i substrate inhibition parameter (m^3/kg)

K_s substrate saturation parameter (kg/m^3)

m parameter used to describe cellular inhibition

m_p Luedeking-Piret non-growth-associated constant (kg/[kg·h])

m_x maintenance parameter (kg/[kg·h])

n parameters used to describe product inhibitions

P product concentration (kg/m^3)

P_{max} product concentration when cell growth ceases (kg/m^3)

r_p rate of product formation (kg/[m^3·h])

r_s rate of substrate consumption (kg/[m^3·h])

r_x rate of growth (kg/[m^3·h])

S substrate concentration (kg/m^3)

T temperature into the fermentor (^{o}C)

wj_0 bias of the neuron (j) in the hidden layer of the MLPNN

wjl weights connecting the neuron (l) in the input layer and the neuron (j) in hidden layer of the MLPNN

W_0 bias of the neuron in the output layer of the MLPNN

W_j weights connecting the neuron (j) in the hidden layer and the neuron in the output layer of the MLPNN

X biomass concentration (kg/m^3)

X_{max} biomass concentration when cell growth ceases (kg/m^3)

Y_{px} Luedeking-Piret growth-associated constant (kg/kg)

Y_x limit cellular yield (kg/kg)

$\hat{y}$ outputs of the MLPNN

Greek letter

μ_{max} maximum specific growth rate (h^{-1})

δ inputs of the MLPNN

θ parameter vector of the MLPNN

References

1. Goldemberg, J., Coelho, S.T., Nastari, P.M. and Lucon, O., Ethanol learning curve – the Brazilian experience. *Biomass Bioenerg.*, **26**, (2004), 301-304.
2. Ghose, T.K. and Tyagi, R.D., Rapid ethanol fermentation of cellulose hydrolysate. I. Batch versus continuous system. *Biotechnol. Bioeng.*, **21**, (1979), 1387-1400.
3. Lee, J.M., Pollard, J.F. and Coulman, G.A., Ethanol fermentation with cell recycling: Computer simulation. *Biotechnol. Bioeng.*, **25**, (1983), 497-511.

4. Aldiguier, A.S., Alfenor, S., Cameleyre, X., Goma, G., Uribelarrea, J.L., Guillouet, S.E. and Molina-Jouve, C., Synergistic temperature and ethanol effect on *Saccharomyces cerevisiae* dynamic behaviour in ethanol bio-fuel production. *Bioproc. Biosyst. Eng.*, **26**, (2004), 217-222.
5. Phisalaphong, M., Srirattana N. and Tanthapanichakoon W., Mathematical modeling to investigate temperature effect on kinetic parameters of ethanol fermentation. *Biochem. Eng. J.*, **28**, (2006), 36-43.
6. Psichogios, D.C. and Ungar, L.H., A hybrid neural network-first principles approach to process modeling. *AICHE J.*, **38**, (1992), 1499–1511.
7. Costa, A.C., Alves, T.L.M., Henriques, A.W.S., Maciel Filho, R. and Lima, E.L., An adaptive optimal control scheme based on hybrid neural modelling. *Comput. Chem. Eng.*, **22**, (1998), 859-862.
8. Harada, L.H.P., Costa, A.C. and Maciel Filho, R., Hybrid neural modeling of bioprocess using functional link networks. *Appl. Biochem. Biotech.*, **98-100**, (2002), 1009-1024.
9. Georgieva, P., Meireles, M.J. and Feyo de Azevedo, S., Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena. *Chem. Eng. Sci.*, **58**, (2003), 3699-3713.
10. Patnaik, P.R., An integrated hybrid neural system for noise filtering, simulation and control of a fed-batch recombinant fermentation. *Biochem. Eng. J.*, **15**, (2003), 165-175.
11. Oliveira, R., Combining first principles modelling and artificial neural network: a general framework. *Comput. Chem. Eng.*, 28, (2004), 755-766.
12. Ccopa Rivera, E. Mantovaneli, I., Costa, A.C. and Maciel Filho, R., Hybrid modeling for continuous production of bioethanol. Proceedings of the 16th European Symposium on Computer Aided Process Engineering, **21**, (2006a), 613-618, July 9-13, Garmisch-Partenkirchen, Germany.
13. Zuo, K., Cheng, H.P., Wu, S.C. and Wu, S.C., A hybrid model combining hydrodynamic and biological effects for production of bacterial cellulose with a pilot scale airlift reactor. *Biochem. Eng. J.*, **29**, (2006), 81-90.

14. Smets, I.Y., Claes, J.E., November, E.J., Bastin, G.P. and Van Impe, J.F., Optimal adaptive control of (bio)chemical reactors: past, present and future. *J. Process Contr.*, **14**, (2004), 795-805.
15. Atala, D.I.P., Costa, A.C., Maciel, R. and Maugeri, F., Kinetics of ethanol fermentation with high biomass concentration considering the effect of temperature. *Appl. Biochem. Biotech.*, **91-93**, (2001), 353-366.
16. Ricci, M., Martini, S., Bonechi, C., Trabalzini, L., Santucci, A. and Rossi, C., Inhibition effects of ethanol on the kinetics of glucose metabolism by *S. cerevisiae*: NMR and modelling study. *Chem. Phys. Lett.*, **387**, (2004), 377-382.
17. Blanco, M., Peinado, A.C. and Mas, J., Monitoring alcoholic fermentation by joint use of soft and hard modelling methods. *Anal. Chim. Acta*, **556**, (2006), 364-373.
18. Can van, H.J.L., te Braake, H.A.B., Hellinga, C., Luyben, K.Ch.A.M. and Heifnen, J.J. Strategy for dynamic process modeling based on neural networks and macroscopic balances. *AIChE J.*, **42**, (1996), 3403–3418.
19. Braake, te, H.A.B., Can van, H.J.L. and Verbruggen, H.B. Semi-mechanistic modeling of chemical processes with neural networks. *Eng. Appl. Artif. Intel.*, **11**, (1998), 507–515.
20. Zorzetto, L. F. M., Maciel Filho, R., and Wolf-Maciel, M. R. Process modelling development through artificial neural networks and hybrid models. *Comput. Chem. Eng.*, **24**, (2000), 1355–1360.
21. Ccopa Rivera, E., Costa, A.C., Atala, D.I.P., Maugeri, F., Wolf, M.R. and Maciel Filho, R., Evaluation of optimization techniques for parameter estimation: Application to ethanol fermentation considering the effect of temperature. *Process Biochem.*, **41**, (2006b), 1682-1687.
22. Luedeking, R. and Piret, E.L., Transient and steady states in continuous fermentation. Theory and experiment. *J. Biochem. Microbiol. Technol. Eng.*, **1**, (1959), 431-450.
23. Andrietta, S.R., Migliari, P.C. and Andrietta, M.G.S., Classificação de cepas de levedura de processos industriais de fermentação alcoólica utilizando

capacidade fermentativa. *STAB - Açúcar Álcool e Subprodutos*, **17**, (1999), 54-59.

24. Cybenko, G. Approximation by superpositions of a sigmoidal function, *Math. Control Signal*, **2**, (1989), 303-314.
25. Petzold, L.R., Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. Report SAND80-8230, Sandia National Laboratories, Livermore, California, 1980.

CAPÍTULO 8. ETHYL ALCOHOL PRODUCTION OPTIMIZATION BY COUPLING GENETIC ALGORITHM AND MULTILAYER PERCEPTRON NEURAL NETWORK

Trabalho publicado na revista **Applied Biochemistry and Biotechnology**, v. 129-132, p. 969-984, 2006.
Elmer Ccopa Rivera, Aline C. Costa, Maria R. Wolf Maciel, Rubens Maciel Filho

Abstract

In this present article, genetic algorithms and multilayer perceptron neural network (MLPNN) have been integrated in order to reduce the complexity of an optimization problem. A data-driven identification method based on MLPNN and optimal design of experiments is described in detail. The non-linear model of an extractive ethanol process, represented by a MLPNN, is optimized using real-coded and binary-coded genetic algorithms to determine the optimal operational conditions. In order to check the validity of the computational modeling, the results were compared with the optimization of a deterministic model, whose kinetic parameters were experimentally determined as functions of the temperature.

Keywords

Alcoholic fermentation process; modeling; artificial intelligence; design of experiments; penalty function.

8.1 Introduction

Demand for ethyl alcohol as a renewable fuel for automotive industries continues to be high as countries are pushing greater energy self-sufficiency. As a large tropical country, Brazil has a high potential for the use of biomass **(1)**. Sugarcane products are today the most economically important biomass source for energy cogeneration. One of the main cost contributive parameters for ethanol production from biomass is the cost of the raw material. Such cost can be reduced if the conversion efficiency of the raw material is maximized. This would also avoid the additional energy required to remove water from ethanol, which is one of the largest costs involved in the production of ethanol to be used as a fuel additive. The productivity, conversion and yield can be optimized both at the biochemical level and at the level of process operation **(2)**. Some studies have used mathematical models based on fundamental mass balances and kinetic equations with experimentally determined parameters to investigate the influence of operating variables on the conversion, productivity, and yield **(3, 4)**.

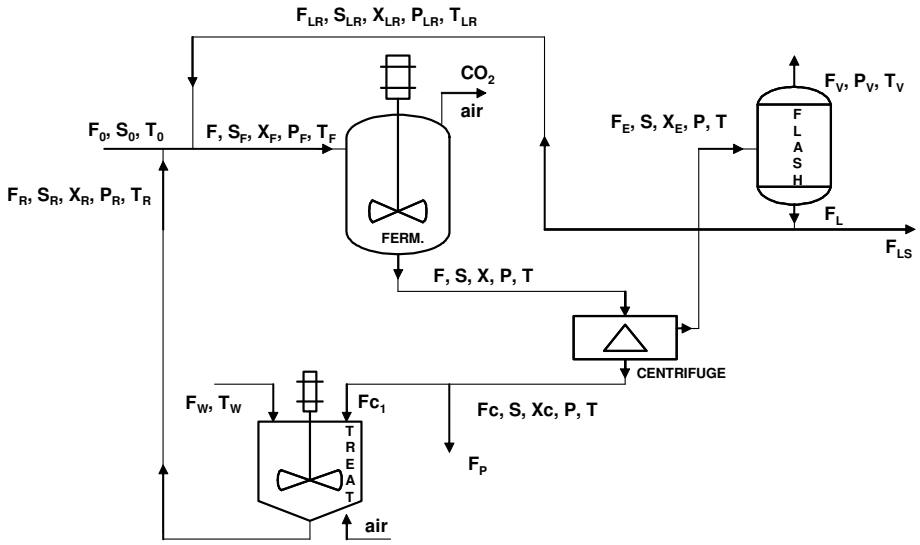


Figure 8.1. Extractive alcoholic fermentation scheme

Artificial intelligence (AI), such as artificial neural networks and genetic algorithms (GA), covers a wide range of techniques and tools that facilitate decision making. These methods have already been successfully applied in the optimization and control of bioprocesses for more than 20 yr **(5)**. Moreover, recent studies in chemical and biochemical research showed that these tools can often be favorably combined especially for modeling, optimization **(6, 7)** and control **(8)**.

This work focuses on the process operation aspects using model-based optimization of an extractive alcoholic fermentation process. Silva et al. **(9)** have shown that a scheme combining a fermentor with a vacuum flash vessel presents several positive features and better performance than a conventional industrial process **(10)**. The objective in optimizing this process is to maximize productivity, although maintaining a high conversion of substrates to ethanol in the fermentor. A comparison is made between the performances of two models when the process is optimized using a real-coded genetic algorithm and a binary-coded genetic algorithm. In real coded genetic algorithm, the decision variables are coded in real numbers, unlike the binary numbers used in the binary coded genetic algorithms. The first model used is a deterministic model, whose kinetics parameters were

experimentally determined as functions of the temperature, the second is a nonlinear model represented by a multilayer perceptron neural network (MLPNN).

A methodology based on AI was used to determine the best conditions from the interactive effect of four variables: inlet substrate concentration, cells recycle rate, residence time and flash recycle.

8.2 Methods

8.2.1 A neural network mathematical model for the ethanol production process

The process to be optimized is shown in Figure 8.1 **(9)**. The process consists of four interlinked units: the fermentor (ethanol production unit), the centrifuge (cell separation unit), the cell treatment unit, and the vacuum flash vessel (ethanol-water separation unit). The mathematical model is built up of five ordinary differential equations derived from mass and energy balances on the fermentor; algebraic equations for the mass balance on the centrifuge, cell treatment unit, and the purge, as well as the mass and energy balances on the flash tank. Details of the model can be found in Costa et al. **(11)**. The objective of the optimization is to maximize productivity and conversion, which are strongly influenced by the inlet substrate concentration, S_0; cell recycle rate, R; residence time, t_r; and flash recycle rate, r **(11)**.

In order to obtain simulation data for the training of the neural network, the inputs (S_0, R, t_r and r) are disturbed according to a design of experiment within the operational intervals (80, 280), (0.2, 0.5), (1.0, 2.5) and (0.2, 0.6), respectively. These intervals were defined based on prior knowledge of the process **(11)**.

A detailed description of the design of experiments theory can be found in **(12)**. In the present work a full factorial design 2^4+star configuration with a central point was used (Table 8.1). In order to have a large amount of training data, eight different values of ΔS_0, Δt_r, ΔR and Δr were used, as shown in Table 8.2, so that eight factorial designs 2^4+star configuration were simulated. The data of productivity and conversion are shown in Figure 8.2 as functions of S_0 and R.

Table 8.1. Coded factor levels and equations to determine the real values for the input variables used in the training of the neural networks

Coded factor level	S_0	t_r (h)	R	r
Level +2	(+2)·ΔS_0(i)+180	(+2)·Δt_r(i)+1.75	(+2)·ΔR(i)+0.35	(+2)·Δr(i)+0.4
Level +1	(+1) ·ΔS_0(i)+180	(+1)·Δt_r(i) +1.75	(+1)·ΔR(i)+0.35	(+1)·Δr(i)+0.4
Central point (0)	180	1.75	0.35	0.4
Level -1	(-1) ·ΔS_0(i)+180	(-1)·Δt_r(i)+1.75	(-1)·ΔR(i)+0.35	(-1)·Δr(i)+0.4
Level -2	(-2) ·ΔS_0(i)+180	(-2)·Δt_r(i)+1.75	(-2)·ΔR(i)+0.35	(-2)·Δr(i)+0.4

Table 8.2. Values of ΔS_0, Δt_r, ΔR and Δr to simulate the data set, used to generate table 8.1

i	ΔS_0(i) (kg/m^3)	Δt_r(i) (h)	ΔR(i)	Δr(i)
1	50	0.075	0.375	0.1
2	45	0.0675	0.3375	0.09
3	40	0.06	0.3	0.08
4	35	0.0525	0.2625	0.07
5	30	0.045	0.225	0.06
6	25	0.0375	0.1875	0.05
7	20	0.03	0.15	0.04
8	10	0.015	0.075	0.02

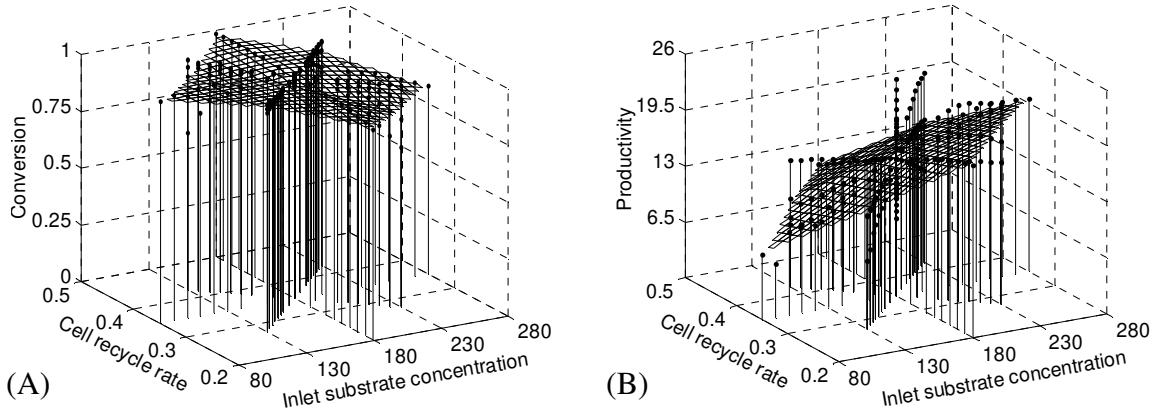


Figure 8.2. Spatial distribution of the data set for identification (the partially collected data points would be located only on the mesh surface if only one full factorial 2^4+star were done): (A) Conversion and (B) Productivity as functions of cells recycle rate and inlet substrate concentration

8.2.2 Building the multilayer perceptron neural network model

In the last years many works **(6, 13)** have been proposed that use artificial neural networks as a substitute for deterministic models. The substitution of the deterministic model by an equivalent MLPNN at the optimization step presents the advantage of high speed processing.

One of the main current problems related to the alcoholic fermentation process is the lack of robustness of the fermentation in the presence of fluctuations in the quality of the raw material, which leads to changes in the kinetic behavior and influences yield, productivity, and conversion. These changes make the prediction of the process dynamic behavior with a single mathematical model difficult.

In order to take the kinetic changes into account, the kinetic parameters of the model should be re-estimated. However, frequent reestimation of these parameters in deterministic models is usually difficult, mainly due nonlinearities, great number of parameters and interactions among them. The use of neural network models is a good option in this case, because the changes in the kinetic behavior can be assessed by updating the network weights.

A MLPNN consists of three types of layers: an input layer, an output layer, and one or more hidden layers, as shown in Figure 8.3 and described mathematically by Eq. 1. Each layer may have a different number of neurons, and even a different transfer function. The most used transfer functions are the sigmoidal, the hyperbolic tangent and the linear transfer functions. More details about neural networks can be found in the works of Chen et al. (6) and Chang and Hung (13). An appropriate architecture would help to achieve higher model accuracy.

$$\hat{y}_1 = g_1[\delta, \theta] = F_1[\sum_{j=1}^{3} W_{1j} f_j(\sum_{l=1}^{4} w_{jl}\delta_l + w_{j0}) + W_{10}] \quad (1)$$

In Eq. 1, θ specifies the parameters vector, which contains all the adjustable parameters of the network; i.e., the weights and biases $\{w_{j,l}\ W_{1,j}\}$.

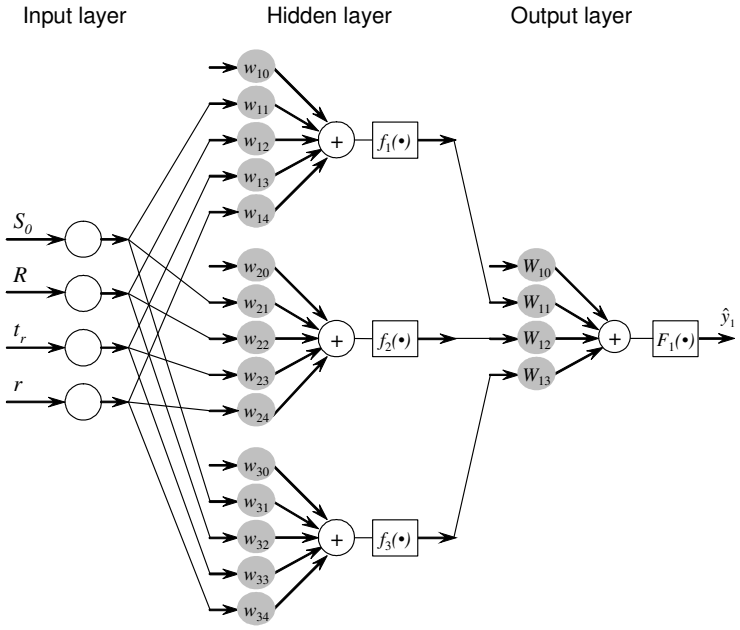


Figure 8.3. Representation of a MLPNN with one hidden layer

In this work, conversion and productivity were each one modeled with a MLPNN with four inputs (inlet substrate concentration, S_0, cell recycle rate, R, residence time, t_r and flash recycle rate, r) and one hidden layer.

The appropriate number of nodes to include in the hidden layer was addressed with the cross-validation technique in order to avoid model over-fitting. This technique splits the sample data into a training sample and a validation sample. Then MLPNN with different numbers of hidden nodes are trained with the training sample, and their performance monitored with the validation sample in terms of the lowest normalized mean square error **(14)**, given by Eq. 2. In addition, the quality of the prediction of the neural models can be also characterized using the correlation coefficient **(15)**, given by Eq. 3.

The Normalized Mean Square Error (NMSE):

$$NMSE = \frac{1}{\sigma^2 np}\sum_{n=1}^{np}\left(d_p - x_p\right)^2 \tag{2}$$

where σ^2 denoted the sample variance of the desired outputs in the test set, x_p and d_p are, respectively, the network and desired outputs and np is the number of patterns tested.

The Correlation Coefficient (COR):

$$\mathrm{COR} = \left(1 - \frac{\mathrm{SEE}}{\mathrm{S_{TT}}}\right) \times 100 \tag{3}$$

where $\mathrm{SEE} = \sum_{p=1}^{np} (d_p - x_p)^2$ and $\mathrm{S_{TT}} = \sum_{p=1}^{np} (d_p - \bar{d}_p)^2$.

In the present work, one half of the data sample was used for training and the other half was intended for validation. The parameters were adjusted using the Levenberg–Marquardt algorithm in Matlab's neural network toolbox.

After the neural networks for conversion and productivity were trained and validated, their parameters were fixed, and the neural network models operated in parallel (Figure 8.4) to simulate a function, *f* (*conv*, *prod*), that was used as the fitness criteria in the ensuing genetic algorithm optimization. In order to accomplish this task, the neural model was written in Fortran to be used by the optimization algorithm, also written in Fortran.

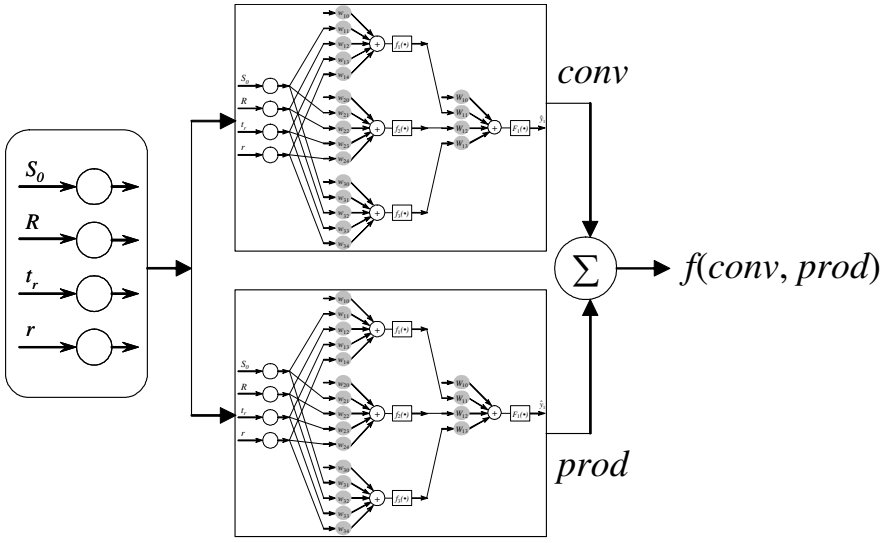


Figure 8.4. Combination of the neural network models (hybrid model) for the optimization of the fermentation process

8.2.3 Optimization by Genetic Algorithms

Genetic Algorithm (GA) is used in this work to optimize a given objective function *f* over a given search space. A population of individuals undergoes some artificial Darwinian evolution (genetic inheritance and Darwinian strife for survival)

based on the fitness of each individual. The fitness of an individual is directly related to the value of the objective function of this individual. The manipulation is done by the genetic operators that work on the chromosomes, in which the parameters of possible solutions are encoded. In each generation of the GA, the new solutions replace the solutions in the population that are selected for deletion. The GA considered in the present article is based on the freeware versions written in Fortran of a real-coded genetic algorithm (RGA) developed by Yedder **(16)** and a binary-coded genetic algorithm (BGA) developed by Carroll **(17),** adapted to the specific needs of the process optimization. These algorithms are well suited to large-scale problems because they do not demand the computation of any Hessian matrix or its inverse, thus having relatively small memory and processing requirements. Also, it ensures the convergence of the optimization procedure (with a second-order rate).

The formulation of objective functions is one of the crucial steps in the application of optimization to a practical problem. In this work, the objective function was formulated as a nonlinear programming (NLP) of the following type:

$$\begin{aligned} &\text{Optimize} \quad f(\vec{x}) \\ &\text{Subjet to} \quad g_i(\vec{x}) \le 0,\ i = 1,\ldots,p, \\ &\qquad\qquad h_j(\vec{x}) = 0,\ j = p+1,\ldots,m, \\ &\qquad\qquad l_k \le x_k \le u_k,\ k = 1,\ldots,n, \end{aligned} \tag{4}$$

where $\vec{x}$ is a vector of n decision variables ($x_1,\ldots,x_n$), $f(\vec{x})$ is the objective function, $g_i(\vec{x})$ is the ith inequality constraints, and $h_j(\vec{x})$ is the jth equality constraints. The l_k and u_k are specified lower and upper bounds on the variables with $l_k \le u_k$.

A general way to deal with constraints, whatever the optimization method, is by penalty function method **(18)**. This technique is the most common approach in the genetic algorithms community, but there are no general guidelines on designing. Some suggestions for genetic algorithms are given in **(19)**. The penalty method of nonlinear programming transforms a constrained problem into a sequence of unconstrained problem. The method used in this work, belong to the

second category of constrained handling methods described by **(19)**. In this approach, for handling inequality constraints, the penalty function $P(\vec{x})$ is defined as the sum of the objective function $f(\vec{x})$ and a penalty term which depends on the constraints violation $g_i(\vec{x})$ and $h_j(\vec{x})$:

$$\left.\begin{array}{ll} \text{Optimize} & f(\vec{x}) \\ \text{Subjet to} & g_i(\vec{x}) \le 0,\ i = 1,\ldots,p, \\ & h_j(\vec{x}) = 0,\ j = \\ & p+1,\ldots,m, \\ & l_k \le x_k \le u_k,\ k = 1,\ldots,n, \end{array}\right\} \quad \text{Optimize: } P(f, g_i, h_j, C) \qquad (5)$$

where $P(f, g_i, h_j, C)$ is a penalty function, and C is a positive penalty parameter. After the penalty function is formulated, it is minimized for a series of values of increasing C-values, which force the sequence to approach the optimum of the constrained problem.

The most significant advantages of the GA are that it avoids the initial guess selection problem and provides a systematic scanning of the whole population and several acceptable local solutions. Thus, it is important to properly define constraints on the possible parameter values. First, constraints limit the search space and thus speed up the optimization. Second, physically impossible values for the parameters are avoided, which improves the reliability of the approach. Additional constraints depend on the type of model that is used. In this work the parameters satisfy such conditions as: $80<S_0<280$ kg/m^3, $0.2<R<0.5$, $1.0<t_r<2.5$ h and $0.2<r<0.6$. Costa et al. **(4)** already used these values successfully in a previous study.

8.3 Results

8.3.1 Choice of the neural network architecture

The number of hidden nodes was varied from 2 to 8 and, using the cross-validation criterion, the network with three hidden nodes was found to present the lowest normalized mean square error (Eq. 2) for the validation sample. The results can be seen in Figure 8.5.

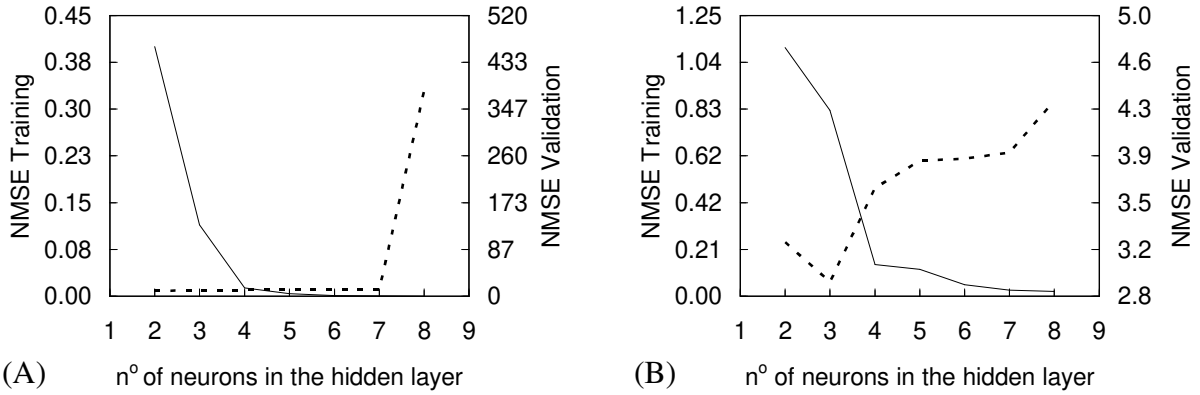


Figure 8.5. Training (solid line) and validation (dashed line) error as functions of the number of neurons in the hidden layer: (a) Productivity and (b) conversion

Productivity and conversion were then modeled both with a MLPNN with 19 scalar parameter, which is calculated using Eq. 6.

$$\text{number of parameters} = (n_i + 1)n_h + (n_h + 1)n_o \quad (6)$$

where n_i = 4, n_h = 3 and n_o = 1 are the number of neurons in the input, hidden and output layers, respectively.

Table 8.3 shows the quality of the prediction for productivity and conversion for training and validation. As can be seen, in both cases by increasing the number of hidden neurons, the training accuracy improves, as indicated by the smaller NMSE and COR values approaching 1. However, using the cross-validation criterion, three neurons in the hidden layer appeared to be the optimal architecture to prevent over-fitting.

Table 8.3. Normalized Mean Square Error (NMSE) and Correlation Coefficient (COR) for productivity and conversion

Neurons	NMSE				COR			
	Productivity (kg/m^3.h)		conversion		Productivity (kg/m^3.h)		conversion	
	train.	valid.	train.	valid.	train.	valid.	train.	valid.
2	0.4	10.0	1.1	3.2	99.06	97.59	98.72	95.99
3	0.1	11.0	0.8	2.9	99.33	97.11	99.13	92.53
4	0.01	11.0	0.1	3.6	99.96	96.76	99.68	92.44
5	0.004	13.2	0.1	3.9	99.97	96.48	99.87	80.31
6	0.001	12.5	0.05	3.9	99.99	96.18	99.88	91.98
7	0.0005	12.0	0.03	3.9	99.99	96.16	99.95	73.19
8	0.0003	380.1	0.02	4.3	99.99	16.26	99.97	82.82

8.3.2 Computation of an optimal solution using genetic algorithms and deterministic model

The optimization was conducted with the deterministic steady state model of the process, which can be found in Costa et al. **(11)**. It consists of the steady state mass and energy balances for the fermentor and all the other process units (see Fig. 1). The equations that describe productivity and conversion can also be found in Costa et al. **(11)**. The optimization problem is to maximize productivity. Thus, the objective function can be formulated as follows:

Maximize *prod* (7)

subject to the equality constraints described by the steady state mass and energy balances for the fermentor and to the inequality constrains:

$conv > 0.99$ (8)

$28 < T < 40^{o}C$ (9)

$80 < S_0 < 280$ kg/m^3 (10)

$0.2 < R < 0.5$ (11)

$1.0 < t_r < 2.5$ h (12)

$0.2 < r < 0.6$ (13)

Because the objective in Yedder's real-coded genetic algorithm driver **(16)** is to minimize the objective function, it is necessary to map the formulated problem (Eqs. 7-13) into unconstrained objective function given by Eq. 14, following the penalty function method for handling inequality constraints.

The equality constraints were handled by converting them as inequality constraints as $g_j(\vec{x}) \equiv \lambda - \left|h_j(\vec{x})\right| \geqslant 0$ for all *j*, where λ (a small positive value) is set to 10^{-3}, in order to allow some room for the search algorithm to work on.

$$\min P(\vec{x}, R) = f(\vec{x}) + R_1 \left|g_1(\vec{x})\right|^2 + R_2 \left(\sum_{j=1}^{5} (10^{-3} - \left|h_j(\vec{x})\right|)^2\right) \quad (14)$$

where $f(\vec{x}) = 1/|prod|$, $g_1(\vec{x}) = 0.99 - |conv|$ and $h_j(\vec{x})$ are the steady state mass and energy balances for the fermentor. The parameters $(\vec{x})$ satisfy such

conditions as: 80 kg/m^3 < S_0 < 280 kg/m^3, 0.2 < R < 0.5, 1.0 h < t_r < 2.5 h and 0.2 < r < 0.6.

The optimization variables are: concentration of viable cells, X_v, dead cells, X_d, substrate, S, product, P and Temperature, T, as well as the variables used by Costa et al. **(11):** S_0, R, t_r e r.

With the best penalty parameters, R_1= 0.02 and R_2= 1250 for constraints, the calculated values were S_0 = 216.29 kg/m^3, R = 0.41, t_r = 1.77 h and r = 0.48. The optimal solution to this problem, using RGA, is: productivity = 13.1 kg/(m^3.h) and conversion = 0.99.

Carroll's binary-coded genetic algorithm **(17)** maximizes the objective function. In consequence, the problem (Eqs. 7-13) was formulated into the penalty function given by Eq. 15.

$$\max P(\vec{x}, R) = f(\vec{x}) + R_1(\left|g_1(\vec{x})\right|^2)^{-1} + R_2(\sum_{j=1}^{5}(10^{-3} - \left|h_j(\vec{x})\right|)^2)^{-1} \quad (15)$$

where $f(\vec{x}) = |prod|$, $g_1(\vec{x}) = 0.99 - |conv|$ and $h_j(\vec{x})$ are the steady state mass and energy balances for the fermentor. Here, the parameters ($\vec{x}$) also satisfy conditions given by Eqs. 9-13.

Using BGA the following values are obtained: S_0 = 135.86 kg/m^3, R = 0.28, t_r = 1.96 h and r = 0.40. The optimal solution to this problem with the best penalty parameter, R_1= 97.0 and R_2= 1.0 for constraints, are: productivity = 13.1 kg/(m^3.h) and conversion = 0.99.

Kalil et al. **(3)** optimized the industrial conventional process designed by Andrietta and Maugeri **(10)** and obtained productivity of 12 kg/(m^3.h), conversion of 0.99 and yield of 0.86. Although in the formulated problem only one inequality restriction was considered (conv > 0.99), the values obtained for yield: 0.87 and 0.90 for RGA and BGA, respectively, were better than the optimized conventional process of Kalil et al. **(3)**. The optimization using genetic algorithms led to productivity a little higher than when the extractive process was optimized using successive quadratic programming by Costa et al. **(4)**.

8.3.3 Computation of an optimal solution using genetic algorithms and neural network model

Because the objective of the optimization study was to maximize productivity, although maintaining a high conversion, the neural network models were structured (see Figure 8.4) to explicitly estimate these output variables from the inlet variables of interest (S_0, R, t_r and r). This structure models the mass and energy balances for the fermentor and all the other process units. After the solution of the optimization problem, the optimal values of S_0, R, t_r and r were used in the deterministic model to determined if the MLPNN predictions for optimal productivity and conversion present deviations from the values calculated by the deterministic model.

The optimization problem is the function f (*prod*, *conv*) transformed in the penalty function form given by Eq. 16 and Eq. 17 for RGA and BGA, respectively:

$$\min P(\vec{x}, C) = f(\vec{x}) + C_1 \left| g_1(\vec{x}) \right|^2 \tag{16}$$

where $f(\vec{x}) = 1/|prod|$ and $g_1(\vec{x}) = 0.99\text{-}|conv|$.

In this problem, such as in the previous case, a constraint in conversion is included (*conv*>0.99). With the best penalty parameter, C_1= 260 for constraint, the calculated values were S_0 = 131.93 kg/m^3, R = 0.28, t_r = 1.20 h and r = 0.54.

The optimal solution to this problem, using RGA, is: productivity = 13.0 kg/(m^3.h) and conversion = 0.99. Using the values calculated for S_0, R, t_r and r in the deterministic model it was obtained productivity of 13.2 kg/(m^3.h) and conversion of 0.99.

$$\max P(\vec{x}, C) = f(\vec{x}) + C_1 (\left| g_1(\vec{x}) \right|^2)^{-1} \tag{17}$$

where $f(\vec{x}) = |prod|$ and $g_1(\vec{x}) = 0.99\text{-}|conv|$.

Using BGA the following values are calculated: S_0 = 161.4 kg/m^3, R = 0.44, t_r = 1.60 h and r = 0.37. The optimal solution to this problem with the best penalty parameter, C_1= 97 for constraint, is: productivity = 13.0 kg/(m^3.h) and conversion = 0.99. Using the values calculated for S_0, R, t_r and r in the deterministic model it was found productivity of 13.0 kg/(m^3.h) and conversion of 0.99.

It is possible to notice that, when the optimization problem is solved using genetic algorithms and the neural network models, the values of temperature and concentrations in the fermentor do not have to be considered as optimization variables and so, the number of the optimization variables is smaller than using deterministic model.

Table 8.4 shows the optimization results of productivity and conversion calculated by the RGA and BGA on the deterministic and MLPNN models. It also shows the comparison among the optimization variables. The results of the optimization methods are expected to lead to different parameters values since they are numerical procedures with expected imprecision in finding the global optimum. Different mathematical models may lead to different predictions, however, a well trained neural network is able to capture the process dynamic behavior. It is worthwhile mentioning that all the parameters values in Table 8.4 are within the accepted range for industrial process operation and this was taken into account in the definition of the constraints for the optimization problem.

Table 8.4. Optimization variable values obtained through RGA and BGA on the deterministic and MLPNN models

	Neural-RGA	Neural-BGA	Deterministic-RGA	Deterministic-BGA
S_0 (kg/m^3)	131.93	161.40	216.29	135.86
R	0.28	0.44	0.41	0.28
t_r (h)	1.20	1.60	1.77	1.96
r	0.54	0.37	0.48	0.40
X_v (kg/m^3)	-	-	36.66	24.26
X_d (kg/m^3)	-	-	3.18	2.69
S (kg/m^3)	-	-	2.39	2.06
P (kg/m^3)	-	-	39.31	37.99
T (ºC)	-	-	32.98	35.18
prod (kg/m^3.h)	13.2	13.0	13.1	13.1
conv	0.99	0.99	0.99	0.99

8.4 Discussions

This work focused on modeling and optimization on the level of process operation of an extractive alcoholic fermentation. The computational intelligence techniques are sought to efficiently combine all available knowledge and to direct the development towards an improved process operation strategy. The genetic

algorithms have shown a good capability to determine the best conditions of conversion and productivity using both the deterministic and the neural network models.

The NN model was developed using a data set generated from the deterministic model, what means the data are not noisy. However, this may not be considered a limitation, because the procedure may be used when experimental data are available. It can be observed that the values for productivity and conversion obtained using the MLPNN models are similar to that obtained using the deterministic model. MLPNN modeling is a powerful and flexible tool and its use in the alcoholic fermentation process is advantageous because updating the neural network parameters is a simpler procedure than re-estimating kinetic parameters of the deterministic models. Frequent re-estimation of kinetic parameters (or in the case of this work updating the neural network parameters) is necessary owing to changes in the dynamic behavior caused by fluctuations in the quality of the raw material, changes in microorganism metabolism, and variation of the dominant yeast strains present in the fermentation process, among other factors.

Normally the optimization methods based on GA require a large computer burden compared to the conventional optimization methods. Neural network models led to a mathematical representation of lower order than the deterministic approach, which intuitively makes the optimization procedure based on hybrid algorithms (NN-RGA and NN-BGA) to be significantly quicker. In fact, the neural network model requires only four optimization variables, while the deterministic model needs nine.

Experimental design is used to obtain data for the training of the neural network in order to guarantee that the region of interest is covered by the training data. This is a very important issue to be addressed as neural network models have no (or very limited) extrapolation properties. The penalty function optimization using RGA and BGA was done in terms of constraint violation to check the feasibility of the solution. In the case of infeasible solution, the one having smaller constraint violation is preferred. In real coded genetic algorithm, the decision variables are coded in real numbers unlike the binary numbers as in the

case of binary coded genetic algorithms. The most important feature of the RGAs is their capacity to exploit local continuities, and the corresponding one of the BGA is their capacity to exploit the discrete similarities **(20)**. According to Goldberg **(21)**, BGA are less efficient when applied to multi-dimensional, high precision or continuous problems. The bit strings can become very long and the search space blows up. In RGA the chromosome/individual is simply a vector where each real-valued element (gene) stands for an unknown parameter of the model.

Although the best result was the one obtained using the rigorous model, the values for productivity and conversion obtained using the MLPNN model are acceptable and can partly eliminate the difficulties of having to specify completely the structure of an alcoholic fermentation model. It is especially promising in situations where rigorous model is excessively difficult computationally, time-consuming or costly.

Acknowledgements

The authors acknowledge FAPESP (process number 03/03630-3) for financial support.

Nomenclature:

F	Feed stream flow rate (m^3/h)
F_c	Cell suspension flow from centrifuge (m^3/h)
F_{c_1}	Cell suspension flow to treatment tank (m^3/h)
F_E	Light phase flow rate to flash tank (m^3/h)
F_L	Liquid outflow from the vacuum flash tank (m^3/h)
F_{LR}	Liquid phase recycling flow rate (m^3/h)
F_{LS}	Liquid phase flow to rectification column (m^3/h)
F_0	Fresh medium flow rate (m^3/h)
F_R	Cell recycling flow rate (m^3/h)
F_V	Vapor outflow from the vacuum flash tank (m^3/h)
F_w	Water flow rate (m^3/h)

P Product concentration into the fermentor (kg/m^3)

P_F Feed product concentration (kg/m^3)

P_{LR} Product concentration in the light phase from centrifuge (kg/m^3)

P_R Product concentration in the cells recycle (kg/m^3)

P_V Product concentration in the vapor phase from the flash tank (kg/m^3)

$R = F_R / F$ Cells recycle rate

$r = F_{LR} / F_L$ Flash recycle rate

S Substrate concentration into the fermentor (kg/m^3)

S_F Feed substrate concentration (kg/m^3)

S_{LR} Substrate concentration in the light phase from centrifuge (kg/m^3)

S_0 Inlet substrate concentration (kg/m^3)

S_R Substrate concentration in the cells recycle (kg/ m^3)

T Temperature into the fermentor (^{o}C)

T_F Feed temperature (^{o}C)

T_{LR} Light phase temperature (^{o}C)

T_0 Inlet temperature of the fresh medium (^{o}C)

T_R Cells recycle temperature (^{o}C)

T_V Temperature of vapor from the flash tank (^{o}C)

T_W Water temperature (^{o}C)

X_c Biomass concentration in the heavy phase from centrifuge (kg/m^3)

X_d Dead biomass concentration into the fermentor (kg/m^3)

X_E Biomass concentration in the light phase flow rate to flash tank (kg/m^3)

X_F Feed biomass concentration (kg/m^3)

X_{LR} Biomass concentration in the light phase from centrifuge (kg/m^3)

X_R Cell recycling concentration (kg/m^3)

$X_t = X_v + X_d$ Total biomass concentration into the fermentor (kg/m^3)

X_v Viable biomass concentration into the fermentor (kg/m^3)

References

1. Goldemberg, J., Coelho, S.T., Nastari, P.N. and Lucon, O. (2004), *Biomass Bioenerg*. 26, 301-304.
2. Parekh, S., Vinci, V.A. and Strobel, R.J. (2000), *Appl. Microbiol. Biot.* 54, 287–301.
3. Kalil, S.J., Maugeri, F. and Rodrigues, M.I. (2000), *Process Biochem*. 35, 539-550.
4. Costa, A.C. and Maciel Filho, R. (2004), *App. Biochem. Biotech*. 114/1-3, 485-496.
5. Schügerl, K. (2001), *J. Biotechnol.* 85, 149–173.
6. Chen, L., Nguang, S.K., Chen, X. D. and Li, X.M. (2004), *Biochem. Eng. J.* 22, 51-61.
7. Nandi, S., Badhe, Y., Lonari, J., Sridevi, U., Rao, B.S., Tambe, S.S. and Kulkarni, B.D. (2004), *Chem. Eng. J.* 97, 115-129.
8. Ahmad, A.L., Azid, I.A., Yusof, A.R. and Seetharamu, K.N. (2004), *Comput. Chem. Eng.* 28, 755-766.
9. Silva, F.L.H., Rodrigues, M.I. and Maugeri, F. (1999), *J. Chem. Tech. Biotechnol.* 74, 176-182.
10. Andrietta, S.R. and Maugeri, F. (1994), In *Adv. Bioprocess Engng*., 47-52.
11. Costa, A.C., Atala, D.I.P., Maciel Filho, R. and Maugeri Filho, F. (2001), *Process Biochemistry.* 37-2, 125-137.
12. Montgomery, Douglas C. (1997). *Design and Analysis of Experiments*, 4th ed., John Wiley and Sons, New York.
13. Chang, J. and Hung, B. (2002), *Ind. Eng. Chem. Res.* 41, 2716-2727.
14. Geva, A.B. (1998), *IEEE Trans. neural net*. 9-5, 1471-1482.
15. Milton, J.S. and Arnold, J.C. (1990), *Introduction to probability and statistics*, McGraw Hill, New York.
16. Yedder, R.B. (2002), Ph.D. thesis, ENPC, Paris, France.
17. Carroll, D.L. (1998), FORTRAN Genetic Algorithm Driver, Last accessed date 04-04-2005. University of Illinois, USA (carroll@cuaerospace.com).

18. Edgar, T. F. and Himmelblau, D. M. (1998), *Optimization of Chemical Processes*, 2nd ed., McGraw-Hill, New York.
19. Michalewicz, Z. and Schoenauer, M. (1996), *Evol. Comp*. 4, 1-32.
20. Herrera, F., Lozano, M. and Verdegay, J.L. (1998), *A. I. Review*. 12, 265-319.
21. Goldberg, D. E. (1989), *Genetic algorithms in search, optimization and machine learning.* Addisson-Wesley Reading, Massachussets.

CAPÍTULO 9. CONSIDERAÇÕES FINAIS

Prefácio

Este capítulo apresenta os principais resultados obtidos neste trabalho de tese. Como complemento, uma discussão é apresentada sobre a tendência da modelagem e otimização de bioprocessos, e a proposta de possíveis extensões deste trabalho e novas direções a serem tomadas.

9.1 Conclusões Gerais

Este trabalho de tese apresenta aplicações de modelos matemáticos, metodologias sistemáticas, além de ferramentas computacionais para o desenvolvimento e a otimização de bioprocessos. Estes campos de pesquisa ainda não são aplicados extensamente na indústria de bioprocessos por várias razões. Entretanto, espera-se que uma aproximação baseada em modelo para o desenvolvimento destes processos reduzirá o tempo de desenvolvimento e renderá uma compreensão melhor da dinâmica do processo. Além disso, em muitos casos, a otimização de bioprocessos pode fornecer benefícios econômicos significativos, já que estes processos são caracterizados por diferenças significativas de preço entre reagentes e produtos. As metodologias propostas nesta pesquisa foram aplicadas ao desenvolvimento e otimização na produção de bioetanol.

A modelagem de bioprocessos é muito complexa, desde que sua operação envolve o crescimento microbiano na presença das flutuações constantes. Conseqüentemente, há uma necessidade para o aperfeiçoamento dos métodos para um tempo de processamento rápido, mas também realista, de modelos matemáticos simples. Inicialmente, técnicas de otimização representativas de algoritmos pertencentes a sistemas computacionais bio-inspirados e de algoritmos determinísticos foram avaliadas para estimar os parâmetros do modelo cinético do processo de fermentação alcoólica em batelada para a produção de etanol usando *Saccharomyces cerevisiae*. Observações experimentais a diferentes temperaturas foram usadas para formular o problema de estimação dos parâmetros do modelo proposto, os quais foram caracterizados por funções de correlação que adotam dependência da temperatura. Sendo que a determinação da região factível do

espaço total de busca na otimização multi-paramétrica de um modelo determinístico é complexa, a otimização foi avaliada usando técnicas diferentes de otimização. Os modelos cinéticos otimizados usando um Algoritmo Quasi-Newton e um Algoritmo Genético de Codificação Real (RGA) descreveram satisfatoriamente o processo de fermentação em batelada, como demonstrado pelos resultados experimentais. Dos resultados computacionais, observou-se que o Algoritmo Genético de Codificação Real forneceu uma solução alternativa. A vantagem mais significativa do Algoritmo Genético é que evita o problema da estimativa inicial das variáveis de decisão, fornece diversas soluções locais aceitáveis e são bem sucedidos com problemas de grande porte, desde que o RGA não exija o cálculo computacional de gradientes.

As mudanças nas condições operacionais são muito comuns nas plantas de fermentação alcoólica; ocorrem não somente devido às variações na qualidade da matéria-prima, mas também devido às variações da levedura dominante no processo. Portanto, tem-se que a dificuldade principal em técnicas baseadas em modelo, no que se refere à definição de estratégias operacionais, controle e otimização, é a obtenção de um modelo confiável. Assim, foi proposto o uso de modelos híbridos para o processo de produção contínua em plantas de bioetanol. E, continuando com as aplicações de sistemas computacionais bio-inspirados, a cinética dos modelos foi descrita por redes neurais artificiais (RNA) para simplificar a etapa de re-estimação de parâmetros. Redes neurais do tipo Multilayer Perceptron (MLP) para um processo extrativo de fermentação alcoólica e as redes do tipo Functional Link Networks (FLN) para um processo de fermentação alcoólica com múltiplos estágios foram usadas para a modelagem da cinética. Foi demonstrado que os modelos híbridos, que acoplam equações de balanço com redes neurais do tipo MLP ou redes neurais do tipo FLN, podem descrever adequadamente o comportamento dinâmico do processo com um desempenho similar àquele obtido por um modelo fenomenológico. Pode-se concluir com isso, que os modelos híbridos podem ser muito confiáveis para serem usados em futuros estudos de controle de bioprocessos.

Posteriormente, a abordagem híbrida foi testada simulando mudanças operacionais e cinéticas no comportamento dinâmico do processo fermentativo. Os parâmetros da rede neural, que representam a cinética do processo, foram re-estimados por uma aproximação integral usando um Algoritmo Genético de Código Real (RGA) e um Algoritmo Genético de Código Binário (BGA). Os resultados mostraram que uma escolha apropriada da arquitetura da RNA e o uso adequado de algoritmos de busca global resultam em um modelo híbrido adaptativo robusto que descreve satisfatoriamente o comportamento dinâmico do processo.

O interesse na investigação em diferentes áreas de possíveis formas de interação entre algoritmos pertencentes a sistemas computacionais bio-inspirados e métodos tradicionais de engenharia têm crescido de forma expressiva nos últimos tempos. É evidente que o mesmo deveria ocorrer com os bioprocessos, mais especificamente na proposta de metodologias de otimização. Assim, considerando a complexidade do problema de estimação de parâmetros cinéticos de modelos determinísticos para processos de fermentação, foi proposta uma nova metodologia que usa uma combinação de diferentes técnicas de otimização.

A primeira etapa é obter valores iniciais para todos os parâmetros do modelo. Conseqüentemente, um algoritmo RGA é usado para a estimação simultânea dos parâmetros. Na terceira etapa, os parâmetros mais significativos foram identificados usando o planejamento Plackett-Burman (PB). Finalmente, os parâmetros mais significativos são otimizados usando o algoritmo Quasi-Newton (QN), que converge para o ótimo muito mais rápido que o RGA.

O modelo cinético otimizado descreveu satisfatoriamente o processo de fermentação como demonstrado pelos resultados experimentais. Por outro lado, a metodologia proposta pode ser aplicada em muitos outros problemas de estimação de parâmetros e pode ser usado como uma ferramenta de confiança para otimizar outros tipos de processos fermentativos.

Com a finalidade de aperfeiçoar as abordagens de estimação de parâmetros em modelos fenomenológicos e híbridos, para situações com resultados experimentais, foi estudada a modelagem de processos biotecnológicos com foco no desenvolvimento das metodologias que podem ser usadas sempre que uma re-

estimação de parâmetros seja necessária. O processo de fermentação alcoólica é usado como um estudo de caso. Dois modelos são propostos: um modelo determinístico e um modelo híbrido, onde o efeito da temperatura na cinética é considerado.

Ao usar os modelos não adaptados para descrever os dados na presença de mudanças nas condições operacionais, o modelo fenomenológico apresentou melhor desempenho. Comparando os procedimentos de adaptação, a re-estimação dos parâmetros da rede neural no modelo híbrido foi muito mais simples do que a re-estimação dos parâmetros cinéticos de um modelo fenomenológico. Este resultado sugeriu que usando um modelo fenomenológico em uma situação onde a re-estimação freqüente seja necessária, poderia ser uma limitação. A atualização do modelo híbrido, entretanto, foi direta. Desta forma, puderam-se eliminar em parte as dificuldades de ter que especificar completamente a estrutura complexa da cinética de um processo de fermentação, o qual não poderia ter uma vantagem muito significativa para um processo bem estudado e conhecido como a fermentação alcoólica, mas pode fazer uma diferença grande para processos biotecnológicos menos conhecidos, permitindo uma determinação rápida de uma descrição matemática que possa ser usada para otimização on-line, software sensor ou controle.

Diversos pesquisadores têm constatado que técnicas de sistemas computacionais bio-inspirados como as redes neurais artificiais e os algoritmos genéticos são, em muitos aspectos, complementares. Assim, estas técnicas foram combinadas para melhorar a estratégia de operação de bioprocessos.

O modelo não-linear de um processo de fermentação alcoólica extrativa, representado por redes neurais do tipo Multilayer Perceptron (MLP) é otimizado, usando um Algoritmo Genético de Código Real (RGA) e um Algoritmo Genético de Código Binário (BGA), para encontrar condições ótimas de operação de modo que a conversão seja maximizada aos limites definidos. A fim de verificar a validade do modelo MLP, os resultados foram comparados à otimização de um modelo determinístico, onde os parâmetros cinéticos foram determinados experimentalmente como funções da temperatura. Observou-se que os valores

obtidos para a produtividade e a conversão usando o modelo neural são similares àqueles obtidos usando o modelo determinístico. Também, é importante destacar que o modelo neural foi desenvolvido usando uma série de dados gerados a partir do modelo determinístico, o que significa que não são dados ruidosos. Porém, isto não pode ser considerado uma limitação, desde que a metodologia proposta pode ser usada quando dados experimentais estão disponíveis. Por outro lado, sabe-se que as metodologias de otimização que utilizam algoritmos genéticos requerem uma grande carga computacional comparado aos métodos convencionais de otimização. O modelo neural conduziu a uma representação matemática de uma ordem mais baixa, comparada à aproximação determinística, o qual improvisa intuitivamente o procedimento de otimização baseado em algoritmos híbridos (MLP-RGA e MLP-BGA) para ser significativamente mais rápida. De fato, o modelo da rede neural requer somente quatro variáveis de otimização, enquanto o modelo determinístico necessita nove.

9.2 Sugestões para trabalhos futuros

Para tratar a falta de robustez nos processos de fermentação, é importante realizar uma monitoração eficiente do processo de fermentação alcoólica, se possível em tempo real, de algumas variáveis e inferênciar em outras, que são difíceis ou caras de serem medidas, fazendo-se ajustes nas condições de operação e controle para manter o processo operando nas condições ideais.

Como na maioria dos bioprocessos as variáveis importantes na fermentação alcoólica são principalmente concentrações, sabe-se que as medidas destas em tempo real costumam ser caras e podem introduzir atrasos significativos, além do alto custo e relativa falta de robustez dos equipamentos de medida em ambientes industriais. Existe uma grande necessidade de se estudar formas para realizar estas medidas com precisão e baixo custo **(Araúzo-Bravo et al., 2004)**.

Uma abordagem que pode ser estudada, em substituição a um equipamento caro, são os chamados *software sensors*, algoritmos usados para inferir variáveis de difícil medida a partir de medidas simples e de baixo custo (variáveis secundárias). Esta é uma área de pesquisa bastante promissora, não só para

monitoração da fermentação alcoólica, mas para a monitoração de bioprocessos em geral **(Araúzo-Bravo et al., 2004)**.

Assim, dando continuidade à aplicação de redes neurais artificias, utilizadas neste trabalho para descrever a cinética da fermentação dentro de uma abordagem híbrida, sugere-se o desenvolvimento de software sensors para determinação de concentrações de biomassa, substrato e produto a partir de medidas secundárias (turbidez, pH, vazão de CO2 e temperatura) e usando um modelo híbrido neuronal que combina as equações de balanço de massa com redes neuronais que descrevem a cinética.

Ainda em relação ao processo de fermentação alcoólica extrativa, sabe-se que apresenta uma grande quantidade de aspectos positivos e vem despertando interesses do setor industrial. O uso do evaporador flash possibilita usar altas concentrações de açúcares no meio de alimentação do reator, o que tem como conseqüência maior produção de etanol, reduzindo o custo da destilação e menor produção de vinhaça. A literatura apresenta diversos modelos termodinâmicos para o estudo de equilíbrio de fases que levam em conta diferentes fatores, sendo a escolha apropriada da descrição analítica destas expressões, objeto de discussão em muitos trabalhos. Entretanto, na prática, a entrada do tanque flash consiste de uma mistura de uma solução complexa, pois o caldo de fermentação contém açúcares (sacarose, glicose e frutose), glicerol e etanol em solução, bem como uma suspensão de biomassa, resíduos do próprio melaço e gás carbônico (**Atala, 2004**). Uma maneira de se lidar com este problema é o uso de modelos híbridos que incorporam redes neuronais para descrever a termodinâmica do equilíbrio de fases no evaporador flash (**Chen et al., 2004**).

Através da revisão bibliográfica, foram reveladas importantes vantagens da síntese de produtos químicos usando cana-de-açúcar como matéria-prima, merecendo maior foco a descrição de duas destas: a primeira baseia-se na fundamentação de que os processos biotecnológicos possibilitam a pesquisa e a exploração de inúmeras rotas de obtenção de produtos de alto valor agregado, formando caminhos paralelos e alternativos para a produção dos mesmos produtos de interesse, denotando assim a possibilidade de enfrentar o problema de

esgotamento das matérias-primas disponíveis atualmente; a segunda é fundamentada no argumento ambiental, onde a diminuição acentuada da exploração dos recursos naturais, proporcionada pelos processos de síntese biotecnológica, teria uma conseqüência equivalente à qualidade de vida e do meio ambiente, e inversamente proporcional à degradação ambiental. Dentro deste contexto, uma outra sugestão refere-se à ampliação de estudos referentes ao desenvolvimento de um modelo estruturado, o qual descreverá o metabolismo microbiano.

O trabalho de **Stremel (2001)**, que focaliza as rotas de piruvato e acetaldeído para a produção de etanol, pode ser utilizado como base para a elaboração do modelo. O modelo desenvolvido será projetado para descrever a fermentação alcoólica aeróbia para o estado estacionário e dinâmico. Com estas informações, será possível a definição de estratégias de operação levando-se em conta os aspectos fisiológicos da célula de microrganismo que produz o etanol, através de alterações nas variáveis operacionais do fermentador.

Referências Bibliográficas

1. Aráuzo-Bravo, M.J., Cano-Izquierdo, J.M., Gómez-Sánchez, E., López-Nieto, M.J., Dimitriadis, Y.A. e López-Coronado, J., Automatization of a penicillin production process with soft sensors and adaptive controller based on neuro fuzzy systems. *Control Engineering Practice*, 2004, **12** 1073-1090.
2. Atala, D. I. P. Montagem, Instrumentação, Controle e Desenvolvimento Experimental de um Processo de Fermentação Alcoólica Extrativo. Campinas - SP., 2004. Tese (Doutor em Engenharia de Alimentos). Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas.
3. Chen, L., Hontoir, Y., Huang, D., Zhang, J. e Morris, A.J., Combining first principles with black-box techniques for reaction systems. *Control Engineering Practice*, 2004, **12** 819-826.

4. Stremel, D.P. Desenvolvimento de modelos estruturados alternativos para o processo de produção de etanol. Campinas - SP., 2001. Tese (Doutor em Engenharia Química). Faculdade de Engenharia Química, Universidade Estadual de Campinas.

Apêndice A. Estudo do comportamento dinâmico de bioprocessos utilizando análise estatística.

A.1 Introdução

Em processos biotecnológicos, um grande número de fatores pode influenciar o rendimento, produtividade e conversão. Normalmente, não é evidente qual desses fatores são os mais importantes. Deve-se, portanto, recorrer a técnicas de análise multivariada como a do planejamento experimental. Assim, é possível, além da determinação dos fatores importantes do processo, identificar as estruturas de controle para estratégias de operação diferenciadas. O planejamento fatorial é uma técnica estatística multivariada que agrupa variáveis que estão altamente correlacionadas. Quando existe um grande número de variáveis pode-se aplicar um planejamento fatorial fracionado ou um planejamento proposto por **Plackett e Burman (1946)**. Os planejamentos Plackett-Burman permitem estimar todos os $k = n - 1$ efeitos principais (onde n representa o número de ensaios) com variância mínima.

O planejamento fatorial vem sendo aplicado com grande freqüência na área de processos biotecnológicos para a otimização operacional devido às inúmeras variáveis que estes processos envolvem. Entretanto, a utilização do planejamento Plackett-Burman está mais voltada à elaboração de meios de cultura. Alguns trabalhos bem sucedidos são citados a seguir: **Grado e Chandra (1998)** utilizaram o planejamento fatorial para comparar o impacto relativo de vários parâmetros nos custos de produção de etanol. Verificaram que uma análise de sensibilidade utilizando planejamento fatorial pode ser utilizada como um guia para reduzir custos no produto final. **Kalil et al. (2000)** otimizaram as condições operacionais de um processo de fermentação alcoólica usando planejamento fatorial e análise de superfície de resposta, de modo a maximizar o rendimento e a produtividade. **Rodrigues et al. (2002)** propuseram uma aproximação para ajustar o controlador GPC (*Generalized Predictive Control*) nas configurações adaptativas e não adaptativas aplicadas a um processo de produção de penicilina utilizando o planejamento fatorial completo, o qual fornecia o conjunto ótimo de parâmetros. **Burkert et al. (2004)** utilizaram com sucesso o planejamento

fatorial e a metodologia de superfície de resposta para estudar os efeitos das concentrações de fontes de carbono e nitrogênio na produção de lípase por *Geotrichum* sp. **Bennamoun et al. (2004)** aplicaram o planejamento PB de 12 ensaios (7 variáveis reais e 4 variáveis inertes) para a otimização da produção de α-amylase por *Aspergillus oryzae Ahlburg (Cohen)*. **Vaidya, et al. (2003)** usaram o planejamento PB para a formulação de um meio nutritivo para otimizar a produção de chitinase por *Alcaligenes xylosoxydans.*

Neste apêndice estão colocados os resultados das primeiras investigações da análise do comportamento dinâmico das variáveis do processo extrativo de fermentação alcoólica contínua utilizando técnicas de planejamento fatorial. Foi realizado o mapeamento da dinâmica do processo, visando avaliar correlações multivariáveis e a quantificação dos efeitos principais e de interação, além de identificar as estruturas de controle mais viáveis para estratégias de operação diferenciadas. Neste estudo também é realizada a análise da influência dos parâmetros cinéticos no processo.

A.2 Análise Estatística

A.2.1 Planejamento fatorial completo

O planejamento fatorial inicia-se especificando os níveis permissíveis superiores e inferiores de cada uma das *n* variáveis. A faixa permissível é então discretizada em níveis eqüidistantes. A região limitada pelos níveis das variáveis é denominada espaço do planejamento, cujos vértices determinam um cubo *n*-dimensional. Se cada uma das variáveis for especificada somente nos limites superior e inferior (dois níveis), o planejamento fatorial é chamado de fatorial completo 2^n. Neste tipo de planejamento, costuma-se identificar os níveis, superior e inferior com os valores (+1) e (-1), respectivamente. Analogamente, um planejamento fatorial completo 3^n é feito especificando os limites superior, central e inferior (três níveis: {-1, 0, 1}) para cada uma das *n* variáveis.

O planejamento fatorial pode ser utilizado para ajustar modelos quadráticos. Um modelo quadrático pode melhorar significativamente o processo de otimização quando o modelo linear mostra-se insatisfatório para o ajuste, devido à interação

entre as variáveis e a curvatura da superfície. A construção de um modelo quadrático de n variáveis requer ao menos $(n+1)\times(n+1)/2$ avaliações da resposta.

Um planejamento fatorial completo realizado em dois níveis foi proposto para verificar e quantificar os efeitos individuais e de interações a partir de perturbações simultâneas das variáveis manipuladas (fatores) nas variáveis do processo (resposta). Os valores para as condições normais de operação, bem como os níveis inferior e superior das variáveis são mostrados na Tabela A.1. Como pode-se observar nesta tabela as variáveis: concentração do meio de alimentação S_0 (kg/m^3), razão de reciclo celular R, vazão do meio de alimentação F_0 (m^3/h), razão de reciclo do tanque flash r e temperatura do meio de alimentação T_0 (^{o}C) foram perturbadas em 10% sobre os valores normais das condições de operação. A matriz de coeficientes de contraste para o planejamento 2^5, cinco variáveis em dois níveis, é apresentada na Tabela A.2.

Nas Figuras A.1, A.2, A.3 e A.4, pode-se observar os perfis dos efeitos principais e efeitos de interação entre as variáveis de entrada (S_0, R, F_0, r e T_0) nas variáveis de saída: concentração de etanol no fermentador P (kg/m^3), concentração de substrato no fermentador S (kg/m^3), concentração de microrganismos no fermentador X (kg/m^3), e temperatura no fermentador T (^{o}C), respectivamente. Estas variáveis de saída foram calculadas como funções do tempo até atingir um novo estado estacionário (neste caso 10 horas).

Tabela A.1. Valores dos fatores utilizados no planejamento fatorial 2^5

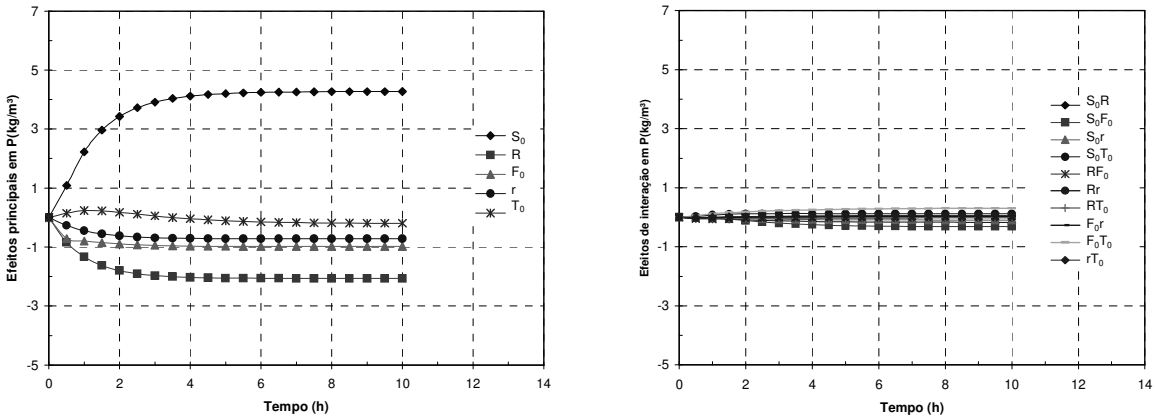
	S_0	R	F_0	r	T_0
Condições normais de operação	103	0.346	100	0.2	30
Nível superior (+10%)	113.3	0.3806	110	0.22	33
Nível inferior (-10%)	92.7	0.3114	90	0.18	27

De forma geral, pode-se observar que a quantificação dos efeitos de interação entre as variáveis de entrada (S_0, R, F_0, r e T_0) nas variáveis de saída P, S, X e T mostrados nas Figuras A.1b, A.2b, A.3b e A.4b, respectivamente, não são significativos para variações $\pm$ 10% quando comparados aos efeitos principais (Figuras A.1a, A.2a, A.3a e A.4a). A partir das Figuras A.1a, A.2a, A.3a e A.4a, foi possível identificar relações de forte influência, baixa influência e influência

desprezível dos efeitos das variáveis de entrada (S_0, R, F_0, r e T_0) nas variáveis de saída P, S, X e T, mostradas na Tabela A.3. Esta tabela pode ser usada para determinar as melhores estruturas para um controle eficiente do processo. Por exemplo, a entrada R influência, principalmente, na concentração de microrganismos no fermentador, apresentando pouca influência em P, S e T. Assim, uma malha que manipule R, e controle X pode ser considerada desacoplada das outras malhas. Deste modo, se uma perturbação desvia a concentração de microrganismos no fermentador do seu *set-point*, o controle desta concentração através da manipulação da razão de reciclo de microrganismos R não afeta as outras variáveis de saída.

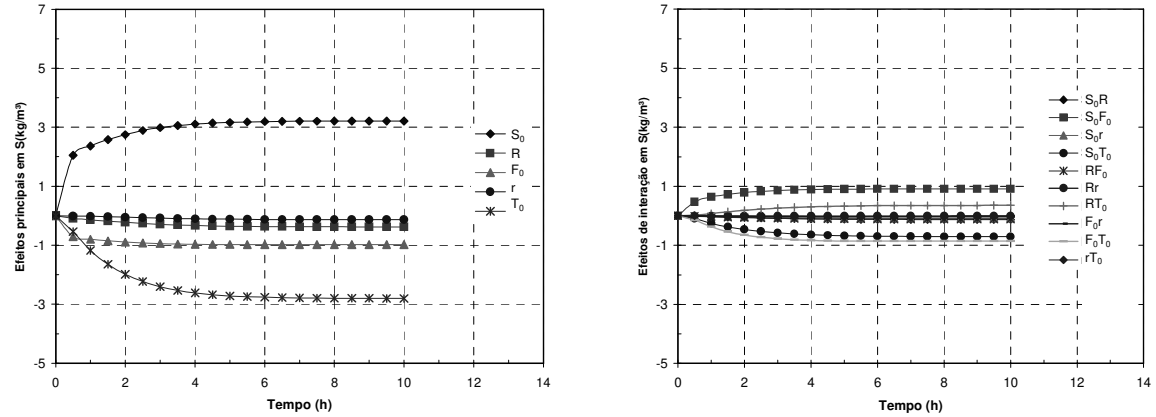
Tabela A.2. Matriz de coeficientes de contraste para o planejamento fatorial 2^5

Ensaio	S_0	R	F_0	r	T_0
1	-1	-1	-1	-1	-1
2	+1	-1	-1	-1	-1
3	-1	+1	-1	-1	-1
4	+1	+1	-1	-1	-1
5	-1	-1	+1	-1	-1
6	+1	-1	+1	-1	-1
7	-1	+1	+1	-1	-1
8	+1	+1	+1	-1	-1
9	-1	-1	-1	+1	-1
10	+1	-1	-1	+1	-1
11	-1	+1	-1	+1	-1
12	+1	+1	-1	+1	-1
13	-1	-1	+1	+1	-1
14	+1	-1	+1	+1	-1
15	-1	+1	+1	+1	-1
16	+1	+1	+1	+1	-1
17	-1	-1	-1	-1	+1
18	+1	-1	-1	-1	+1
19	-1	+1	-1	-1	+1
20	+1	+1	-1	-1	+1
21	-1	-1	+1	-1	+1
22	+1	-1	+1	-1	+1
23	-1	+1	+1	-1	+1
24	+1	+1	+1	-1	+1
25	-1	-1	-1	+1	+1
26	+1	-1	-1	+1	+1
27	-1	+1	-1	+1	+1
28	+1	+1	-1	+1	+1
29	-1	-1	+1	+1	+1
30	+1	-1	+1	+1	+1
31	-1	+1	+1	+1	+1
32	+1	+1	+1	+1	+1



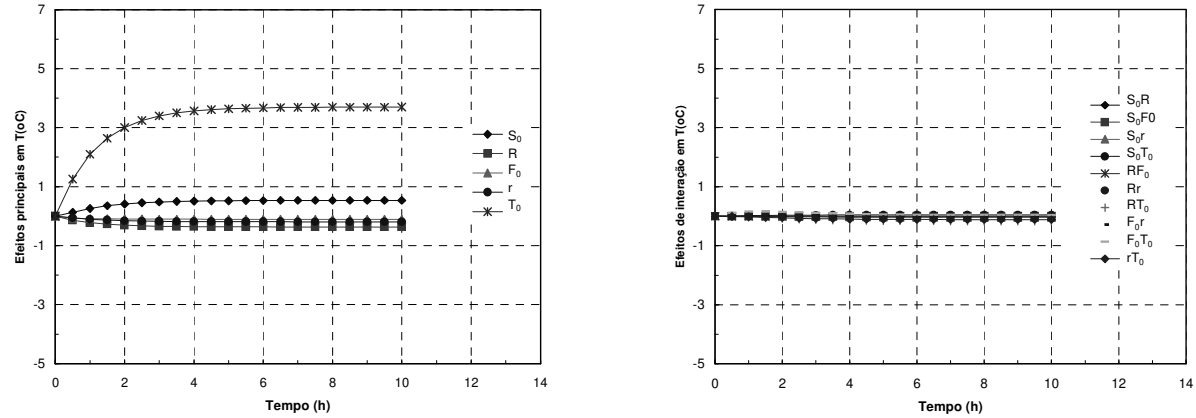
(a) (b)

Figura A.1. (a) Efeitos principais em *P* (b) Efeitos de interação entre as variáveis de entrada em *P*



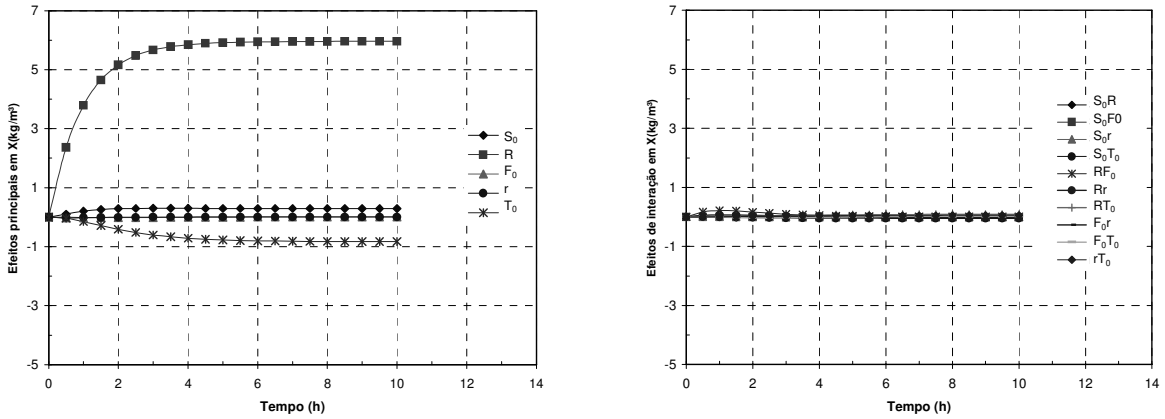
(a) (b)

Figura A.2. (a) Efeitos principais em *S* (b) Efeitos de interação entre as variáveis de entrada em *S*



(a) (b)

Figura A.3. (a) Efeitos principais em *X* (b) Efeitos de interação entre as variáveis de entrada em *X*



(a) (b)

Figura A.4. (a) Efeitos principais em T (b) Efeitos de interação entre as variáveis de entrada em T

Tabela A.3. Influência dos efeitos das variáveis de entrada (S_0, R, F_0, r e T_0) nas variáveis de saída P, S, X e T

	P	S	X	T
S_0	Forte influência	Forte influência	Desprezível	Baixa influência
R	Baixa influência	Baixa influência	Forte influência	Desprezível
F_0	Baixa influência	Baixa influência	Desprezível	Desprezível
R	Baixa influência	Desprezível	Desprezível	Desprezível
T_0	Desprezível	Forte influência	Baixa influência	Forte influência

Tabela A.4. Valores dos fatores utilizados no planejamento Plackett-Burman

Parâmetro		**Condições normais de operação**	**Nível superior (+10%)**	**Nível inferior (-10%)**
μ_{max}	(h^{-1})	0.42	0.46	0.38
X_{max}	(kg/m^3)	63.05	69.36	56.75
P_{max}	(kg/m^3)	111.38	122.52	100.24
Y_X	(kg/kg)	4.99E-02	5.49E-02	4.49E-02
Y_{PX}	(kg/kg)	8.82	9.70	7.94
K_S	(kg/m^3)	3.67E-03	4.04E-03	3.30E-03
K_i	(m^3/kg)	0.40	0.44	0.36
m_p	(kg/kg.h)	3.99E-17	4.39E-17	3.59E-17
m_x	(kg/kg.h)	4.1	4.51	3.69
m	(cte.)	0.1	0.11	0.09
n	(cte.)	0.2	0.22	0.18
K_{dP}	(m^3/kg)	1	1.1	0.9
K_{dT}	(h^{-1})	1.5	1.65	1.35
ρ	(kg/m^3)	390	429	351
γ	(kg/m^3)	0.78	0.858	0.702

A.2.2 Planejamento Plackett-Burman

O planejamento Plackett-Burman (PB) é muito utilizado na seleção dos fatores mais importantes do processo. Os planejamentos mais usuais propostos neste método são para 12, 20, 24, 28 e 36 ensaios, sendo conveniente realizar quatro ensaios a mais que o número de variáveis a serem estudadas (denominadas variáveis inertes).

Dando seqüência ao estudo do comportamento dinâmico do processo extrativo de fermentação alcoólica contínua, é feita a análise de sensibilidade dos parâmetros presentes nas equações das taxas cinéticas. Os valores para as condições normais de operação, bem como os níveis, inferior e superior das variáveis são mostrados na Tabela A.4. Como pode-se observar nesta tabela, todos os parâmetros cinéticos foram perturbados em 10% sobre os valores normais das condições de operação.

Para o processo em estudo neste trabalho, havia 15 parâmetros cinéticos, neste caso um planejamento completo 2^{15} corresponderia a 32.768 ensaios, e mesmo sendo um trabalho de simulação, esse número de experimentos inviabiliza sua realização. Por isso, foi usado um planejamento Plackett-Burman de 20 ensaios. A Tabela A.5 apresenta a matriz de coeficientes de contraste para o planejamento PB para 20 ensaios. As variáveis inertes (VI_n) são usadas para a determinação do erro padrão e, assim, definir os fatores estatisticamente significativos através dos efeitos principais de cada uma delas (**McIntyre et al., 1996**). O cálculo dos efeitos é determinado com a equação:

$$\mathbf{X}^t\mathbf{y} \tag{A.1}$$

onde **X** é a matriz completa de coeficientes de contraste e **y** o vetor resposta. Os efeitos resultam da divisão de cada elemento do vetor coluna $\mathbf{X}^t\mathbf{y}$ por 2^{k-1} (k é o número de fatores). Os efeitos correspondentes às variáveis inertes (ef_{VIn}) são utilizados para calcular o erro padrão (SE) dos fatores:

$$SE = \sqrt{\frac{\sum_{n=1}^{n}(ef_{VIn})^2}{n}} \tag{A.2}$$

A significância de cada fator pode ser determinada usando a distribuição *t-Student*. Neste planejamento, o *valor t* tabelado para 95% de confiança e 14 graus de liberdade é 2.1447. Este valor permitiu determinar o intervalo de confiança através do valor *tf*:

$$tf = valor\ t \cdot SE \tag{A.3}$$

$$\text{-95\% Intervalo de confiança} = ef_i - tf \tag{A.4}$$

$$\text{+95\% Intervalo de confiança} = ef_i + tf \tag{A.5}$$

onde ef_i é o efeito de cada fator (parâmetro cinético).

Os efeitos cujo intervalo de confiança não incluem zero, serão os estatisticamente significativos.

As Tabelas A.6, A.7, A.8 e A.9 apresentam os efeitos dos parâmetros cinéticos sobre *P*, *S*, *X* e *T*, respectivamente, para um nível de confiança de 95%.

Tabela A.5. Matriz de coeficientes para o planejamento Plackett-Burman para 20 ensaios

En	μ_{max}	X_{max}	P_{max}	Y_X	Y_{PX}	K_S	K_i	m_p	m_x	m	n	K_{dP}	K_{dT}	ρ	γ	VI_1	VI_2	VI_3	VI_4
1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1
2	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1
3	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1
4	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1
5	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1
6	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1
7	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1
8	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1
9	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1
10	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1
11	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1
12	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1
13	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1
14	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1
15	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1
16	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1
17	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1
18	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1
19	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1
20	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

VI_n são as variáveis inertes.

Pela análise das Tabelas A.6, A.7, A.8 e A.9, pode-se verificar que os parâmetros cinéticos (fatores) μ_{max}, X_{max}, P_{max}, Y_X, K_i, m_p e m_x têm influência em todas as respostas (P, S, X e T). O fator K_S só influi significativamente sobre a concentração do produto no fermentador (P) e sobre a concentração de microrganismos (X). Os fatores Y_{PX}, K_{dT} e ρ só influem significativamente sobre P. Assim, dentro da faixa estudada verifica-se que os fatores K_{dP}, m, n e γ, não exercem efeitos estatisticamente significativos para as respostas de interesse.

Para uma melhor visualização, a Figura A.5, apresenta os efeitos de cada fator sobre as variáveis P, S, X e T.

De uma maneira geral, pode-se afirmar que existe uma influência significativa dos parâmetros cinéticos (perturbados em 10% sobre os valores normais das condições de operação do processo) no comportamento dinâmico do processo.

O planejamento Plackett-Burman mostrou-se eficiente e prático para estabelecer os fatores estatisticamente significativos. Mesmo se tratando de obtenção de resultados por simulação, a redução do número de ensaios foi bastante significativa. Os resultados do planejamento Plackett-Burman mostraram a importância da disponibilidade de um bom modelo cinético para o projeto eficaz de bioprocessos.

Tabela A.6. Estimativa dos efeitos sobre a concentração de produto, P (kg/m^3)

Fator	Efeito	Erro padrão	t(14)	p	-95%	+95%
Méd.	30.1734*	0.0666*	453.1530*	0.0000*	29.8878*	30.4590*
μ_{max}	0.8564*	0.1332*	6.4308*	0.0000*	0.5708*	1.1420*
X_{max}	0.7094*	0.1332*	5.3270*	0.0001*	0.4238*	0.9950*
P_{max}	0.4544*	0.1332*	3.4122*	0.0042*	0.1688*	0.7400*
Y_X	3.6464*	0.1332*	27.3814*	0.0000*	3.3608*	3.9320*
Y_{PX}	4.3028*	0.1332*	32.3104*	0.0000*	4.0172*	4.5884*
K_S	-0.6316*	0.1332*	-4.7428*	0.0003*	-0.9172*	-0.3460*
K_i	-0.1196	0.1332	-0.8981	0.3843	-0.4052	0.1660
m_p	0.7832*	0.1332*	5.8812*	0.0000*	0.4976*	1.0688*
m_x	-0.8614*	0.1332*	-6.4684*	0.0000*	-1.1470*	-0.5758*
m	-0.6166*	0.1332*	-4.6301*	0.0004*	-0.9022*	-0.3310*
n	-0.5272*	0.1332*	-3.9588*	0.0014*	-0.8128*	-0.2416*
K_{dP}	0.2608	0.1332	1.9584	0.0704	-0.0248	0.5464
K_{dT}	0.0788	0.1332	0.5917	0.5635	-0.2068	0.3644
ρ	-0.3176*	0.1332*	-2.3849*	0.0318*	-0.6032*	-0.0320*
γ	-0.0816	0.1332	-0.6127	0.5499	-0.3672	0.2040

* Significância para um nível de confiança de 95%

Tabela A.7. Estimativa dos efeitos sobre a concentração de substrato, S (kg/m^3)

Fator	Efeito	Erro padrão	t(14)	p	-95%	+95%
Méd.	5.1704*	0.2154*	23.9989*	0.0000*	4.2463*	6.0945*
μ_{max}	-2.0070*	0.4309*	-4.6579*	0.0004*	-2.9311*	-1.0829*
X_{max}	-1.8057*	0.4309*	-4.1907*	0.0009*	-2.7298*	-0.8816*
P_{max}	-1.5580*	0.4309*	-3.6157*	0.0028*	-2.4821*	-0.6338*
Y_X	2.4616*	0.4309*	5.7129*	0.0001*	1.5375*	3.3857*
Y_{PX}	0.5428	0.4309	1.2598	0.2283	-0.3813	1.4670
K_S	1.3940*	0.4309*	3.2352*	0.0060*	0.4699*	2.3181*
K_i	0.0886	0.4309	0.2056	0.8400	-0.8355	1.0127
m_p	0.0237	0.4309	0.0550	0.9570	-0.9004	0.9478
m_x	0.1234	0.4309	0.2863	0.7788	-0.8008	1.0475
m	1.8526*	0.4309*	4.2995*	0.0007*	0.9285*	2.7767*
n	1.1960*	0.4309*	2.7756*	0.0149*	0.2718*	2.1201*
K_{dP}	-0.4589	0.4309	-1.0651	0.3049	-1.3830	0.4652
K_{dT}	-0.1593	0.4309	-0.3697	0.7172	-1.0834	0.7648
ρ	0.5635	0.4309	1.3078	0.2120	-0.3606	1.4876
γ	0.0849	0.4309	0.1971	0.8466	-0.8392	1.0090

Tabela A.8. Estimativa dos efeitos sobre a concentração de biomassa, X (kg/m^3)

Fator	Efeito	Erro padrão	t(14)	p	-95%	+95%
Méd.	33.6727*	0.0091*	3718.5788*	0.0000*	33.6338*	33.7115*
μ_{max}	0.0727*	0.0181*	4.0142*	0.0013*	0.0339*	0.1115*
X_{max}	0.0617*	0.0181*	3.4069*	0.0043*	0.0229*	0.1005*
P_{max}	0.0535*	0.0181*	2.9541*	0.0105*	0.0147*	0.0923*
Y_X	0.3391*	0.0181*	18.7240*	0.0000*	0.3003*	0.3779*
Y_{PX}	0.0025	0.0181	0.1380	0.8922	-0.0363	0.0413
K_S	-0.0569*	0.0181*	-3.1418*	0.0072*	-0.0957*	-0.0181*
K_i	0.0021	0.0181	0.1160	0.9093	-0.0367	0.0409
m_p	0.0009	0.0181	0.0497	0.9611	-0.0379	0.0397
m_x	-0.0751*	0.0181*	-4.1468*	0.0010*	-0.1139*	-0.0363*
m	-0.0623*	0.0181*	-3.4400*	0.0040*	-0.1011*	-0.0235*
n	-0.0431*	0.0181*	-2.3798*	0.0321*	-0.0819*	-0.0043*
K_{dP}	0.0191	0.0181	1.0546	0.3095	-0.0197	0.0579
K_{dT}	0.0041	0.0181	0.2264	0.8242	-0.0347	0.0429
ρ	-0.0247	0.0181	-1.3639	0.1941	-0.0635	0.0141
γ	-0.0041	0.0181	-0.2264	0.8242	-0.0429	0.0347

Tabela A.9. Estimativa dos efeitos sobre a temperatura, T (°C)

Fator	Efeito	Erro padrão	t(14)	p	-95%	+95%
Méd.	32.5624*	0.0091*	3561.1256*	0.0000*	32.5232*	32.6016*
μ_{max}	0.0880*	0.0183*	4.8120*	0.0003*	0.0488*	0.1272*
X_{max}	0.0788*	0.0183*	4.3089*	0.0007*	0.0396*	0.1180*
P_{max}	0.0670*	0.0183*	3.6637*	0.0026*	0.0278*	0.1062*
Y_X	-0.0780*	0.0183*	-4.2652*	0.0008*	-0.1172*	-0.0388*
Y_{PX}	0.0022	0.0183	0.1203	0.9060	-0.0370	0.0414
K_S	-0.0612*	0.0183*	-3.3465*	0.0048*	-0.1004*	-0.0220*
K_i	-0.0044	0.0183	-0.2406	0.8134	-0.0436	0.0348
m_p	0.0032	0.0183	0.1750	0.8636	-0.0360	0.0424
m_x	-0.0106	0.0183	-0.5796	0.5714	-0.0498	0.0286
m	-0.0802*	0.0183*	-4.3855*	0.0006*	-0.1194*	-0.0410*
n	-0.0526*	0.0183*	-2.8763*	0.0122*	-0.0918*	-0.0134*
K_{dP}	0.0204	0.0183	1.1155	0.2834	-0.0188	0.0596
K_{dT}	0.0070	0.0183	0.3828	0.7076	-0.0322	0.0462
ρ	-0.0256	0.0183	-1.3998	0.1833	-0.0648	0.0136
γ	-0.0040	0.0183	-0.2187	0.8300	-0.0432	0.0352

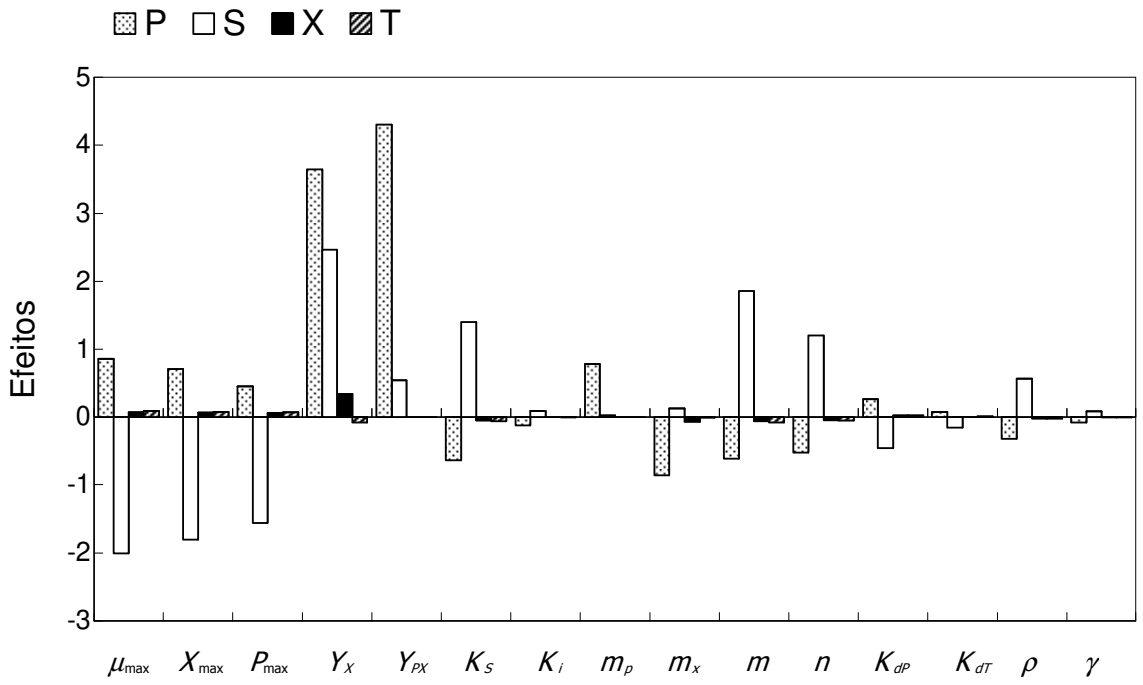


Figura A.5. Efeito dos parâmetros cinéticos (fatores) sobre P, S, X e T

Referências Bibliográficas

1. Bennamoun, L., Meraihi, Z., Dakhmouche, S. Utilisation de la planification expérimentale pour l'optimisation de la production de l'α-amylase par A*spergillus oryze Ahlburg (Cohen) 1042.72* cultivé sur milieu à base de déchets d'oranges. *Journal of Food Engineering*, 2004,**64** 257-264.
2. Burkert, J.F.M., Maugeri, F., Rodrigues, M.I. Optimization of extracellular lípase production by *Geotrichum* sp. using factorial design. *Bioresource Technology*, 2004, **91** 77-84.
3. Costa, A.C., Atala, D.I.P., Maugeri, F., Maciel Filho, R. Factorial design and simulation for the optimization and determination of control structures for an extractive alcoholic fermentation. *Process Biochemistry*, 2001, **37** 125-137.
4. Grado, S.C., Chandra, M.J. A factorial design analysis of a biomass to ethanol production system. *Biomass and Bioenergy*, 1998, **15** 115-124.
5. Kalil, S.J., Maugeri, F., Rodrigues, M.I. Response surface analysis and simulation as a tool for bioprocess design and optimization. *Process Biochemistry*, 2000, **35** 539-550.
6. McIntyre, J.J., Bull, A.T., Bunch, A.W. Vancomycin production in batch and continuous culture. *Biotechnology and Bioengineering*, 1996, **49** 412-420.
7. Plackett, R.L., Burman, J.P. The design of optimun multifactorial experiments. *Biometrika*, 1946, **33** 305-325.
8. Rodrigues, J.A.D., Toledo, E.C.V., Maciel Filho, R. A tuned approach of the preditive-adaptatie GPC controller applied to a fed-batch bioreactor using complete factorial design. *Computers and Chemical Engineering*, 2002, **26** 1493-1500.
9. Vaidya, R., Vyas, P., Chhatpar, H.S. Statistical optimization of medium components for the production of chitinase by *Alcaligenes xylosoxydans*. *Enzyme and Microbial Technology*, 2003, **33** 92-96.

Apêndice B. Otimização do processo extrativo de fermentação alcoólica usando Programação Quadrática Sucessiva (SQP)

Com a finalidade de comparar os resultados de otimização do processo de fermentação alcoólica extrativa utilizando Algoritmos Genéticos (RGA e BGA), expostos no capítulo 8, a Tabela B.1 apresenta os resultados de otimização dos modelos não-lineares do processo, compreendendo um modelo neural e um modelo determinístico otimizados utilizando um algoritmo de otimização determinístico: Programação Quadrática Sucessiva (SQP). A rotina DNCONF da biblioteca IMSL do Fortran foi utilizada para implementar a otimização do processo fermentativo.

Tabela B.1. Valores das variáveis de otimização obtidas usando algoritmos RGA, BGA e SQP nos modelos neural e deterministico do processo extrativo de fermentação alcoólica

	Neural-RGA*	Neural-BGA*	Neural-SQP	Determ.-RGA*	Determ.-BGA*	Determ.-SQP
S_0 (kg/m^3)	131.93	161.40	130.0	216.29	135.86	280.0
R	0.28	0.44	0.26	0.41	0.28	0.5
t_r (h)	1.20	1.60	1.07	1.77	1.96	1.0
r	0.54	0.37	0.59	0.48	0.40	0.55
X_v (kg/m^3)	-	-	-	36.66	24.26	47.34
X_d (kg/m^3)	-	-	-	3.18	2.69	6.42
S (kg/m^3)	-	-	-	2.39	2.06	1.91
P (kg/m^3)	-	-	-	39.31	37.99	29.52
T (ºC)	-	-	-	32.98	35.18	31.28
prod (kg/m^3.h)	13.2	13.0	13.0	13.1	13.1	13.2
conv	0.99	0.99	0.99	0.99	0.99	0.99

*Resultados apresentados no Capítulo 8.

Inserindo implicitamente os modelos determinístico e neural, do processo fermentativo na rotina SQP, foi possível atingir a convergência do problema. O problema de otimização formulado foi o mesmo descrito no Capítulo 8, ou seja, maximizar a produtividade (*prod*) da fermentação para encontrar condições ótimas de operação de modo que a conversão (*conv*) seja maximizada aos limites definidos. O tempo computacional gerado foi pequeno. Entretanto, percebou-se que os resultados da otimização são dependentes das estimativas iniciais e dos limites superiores e inferiores das variáveis de decisão. Usando algoritmos genéticos foi possível atingir a convergência do problema sem a presença desses inconvenientes.

Apêndice C.

No Capítulo 3, devem ser observadas as seguintes correções:

Na primeira coluna das Tabelas 3.2 e 3.3, onde se observa a expressão Y_x, deve estar a expressão Y_{px} e vice-versa. Após esta correção, na Tabela 3.3, a fila que corresponderia a Y_x deveria apresentar os seguintes valores (não arredondados), para o modelo otimizado pelo Algoritmo Quasi-Newton (QN): A=53539; B=-35,2691; C=53540; D= -35,2695.

No título da Figura 3.1, onde se observa a frase *solid lines*, deve estar a frase *dashed lines* e vice-versa.