



TÉCNICAS DE IDENTIFICAÇÃO VOLTADAS PARA A OTIMIZAÇÃO DE PROCESSOS EM TEMPO REAL

Maurício Melo Câmara

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Química, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Química.

Orientadores: José Carlos Costa da Silva Pinto
André Domingues Quelhas

Rio de Janeiro
Junho de 2018

TÉCNICAS DE IDENTIFICAÇÃO VOLTADAS PARA A OTIMIZAÇÃO DE
PROCESSOS EM TEMPO REAL

Maurício Melo Câmara

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA QUÍMICA.

Examinada por:

Prof. José Carlos Costa da Silva Pinto, D.Sc.

Dr. André Domingues Quelhas, D.Sc.

Prof. Tito Livio Moitinho Alves, D.Sc.

Dr. Ismael Humberto Ferreira dos Santos, D.Sc.

Dr. Angelo Ernani Maia Ciarlini, D.Sc.

Dr. Mario Cesar Mello Massa de Campos, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

JUNHO DE 2018

Câmara, Maurício Melo

Técnicas de Identificação Voltadas para a Otimização de Processos em Tempo Real/Maurício Melo Câmara. – Rio de Janeiro: UFRJ/COPPE, 2018.

XII, 283 p. 29, 7cm.

Orientadores: José Carlos Costa da Silva Pinto

André Domingues Quelhas

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Química, 2018.

Referências Bibliográficas: p. 196 – 261.

1. Otimização em tempo real. 2. Monitoramento de processos. 3. Reconciliação de dados. 4. *Big data*. I. Pinto, José Carlos Costa da Silva *et al*. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química. III. Título.

Perhaps the final irony is that it is the most successful automated systems, with rare need for manual intervention, which may need the greatest investment in human operator training (BAINBRIDGE, 1983).

Agradecimentos

Vinícius Uemura, muito obrigado por me encaminhar a Fábio Toniolo; e Fábio Toniolo, muito obrigado por me encaminhar a José Carlos. Cada um de vocês tem uma contribuição especial em minha vinda para o Rio de Janeiro, meu ingresso no PEQ/COPPE/UFRJ e para a existência dessa tese.

Zé Carlos, muito obrigado!!! É um privilégio ter tido a oportunidade de ser seu orientado e poder receber parte de sua enriquecedora experiência de vida, pessoal e profissional.

André Domingues Quelhas, muito obrigado por compartilhar seus imensos conhecimentos e experiência, e por apresentar valiosas visões sobre a otimização em tempo real. Isso também é um privilégio.

À minha esposa Marina, e à família sempre presente, meus sinceros agradecimentos pelo amor, paciência e apoio para a realização desse trabalho.

Thiago de Sá Feital e Rafael Marinho Soares, muito obrigado, vocês são importantes parceiros nessa caminhada.

Caríssimos colegas e amigos do LMSCP, antigos e atuais, baianos, sergipanos, mineiros, *etc.*, flamenguistas, botafoguences, *etc.*, especialmente os frequentadores e membros da sala G-130, muito obrigado!

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

TÉCNICAS DE IDENTIFICAÇÃO VOLTADAS PARA A OTIMIZAÇÃO DE PROCESSOS EM TEMPO REAL

Maurício Melo Câmara

Junho/2018

Orientadores: José Carlos Costa da Silva Pinto
André Domingues Quelhas

Programa: Engenharia Química

Este trabalho revisa o tema da otimização estacionária em tempo real e apresenta os elementos das diversas metodologias e implementações relatadas na literatura. É mostrado o papel central da identificação do processo, da qualidade do modelo e das medidas disponíveis na técnica de otimização em tempo real. Especial atenção é dada a aspectos básicos da operação de processos, como a capacidade de identificação do processo, dadas as variáveis medidas disponíveis, e o adequado tratamento aos aspectos numéricos envolvidos. Considerando processos com disponibilidade massiva de dados, propôs-se uma estratégia para monitoramento de processos de grande escala com base no uso de modelos empíricos combinados a métodos de controle estatístico de processos, que indicam a qualidade do modelo e são eficazes para detecção e diagnóstico de falhas. Processos com baixa disponibilidade de informações também são considerados com o desenvolvimento de um procedimento de identificação *on-line*, que visa lidar com a dificuldade de construção e atualização de modelos para uso em procedimentos aplicados em tempo real, quando diversas grandezas importantes não estão disponíveis. A partir de uma avaliação de desempenho de diferentes técnicas empregadas para a resolução de problemas de otimização, são mostradas evidências das possíveis vulnerabilidades presentes na resolução do problema de reconciliação de dados, mesmo com modelos rigorosos e dados medidos adequados. Os estudos contribuem para a área de otimização em tempo real, na medida em que se inserem nas etapas dos sistemas RTO comerciais. Também representam inovações como técnicas de suporte à tomada de decisão com amplo potencial de aplicação, na medida em que podem ser implementadas com investimentos mais baixos e não requerem a complexidade exigida por sistemas RTO.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

IDENTIFICATION TECHNIQUES FOR REAL-TIME PROCESS OPTIMIZATION

Maurício Melo Câmara

June/2018

Advisors: José Carlos Costa da Silva Pinto

André Domingues Quelhas

Department: Chemical Engineering

This thesis reviews the subject of static real-time optimization and presents the elements of various methodologies and industrial implementations reported in the literature. The central role of process identification, model accuracy and availability of measurements in real-time optimization is shown. Special attention is given to the basic aspects of the practical reality of processes operation, such as the identification of processes, given the available measured variables, and the appropriate treatment of the involved numerical aspects. Considering industrial processes with massive availability of data, a strategy was proposed to monitor large-scale processes based on the use of empirical models combined with methods of statistical process control, indicating the quality of the model and being effective for fault detection and diagnosis. Processes with low availability of information are also considered with the development of an on-line identification procedure, which aims to deal with the difficulty of building and updating models for use in real-time procedures when several important quantities are not available. The performance assessment of different techniques used to solve the optimization problem revealed possible vulnerabilities to solve data reconciliation problems, even when rigorous models and adequate measured variables are available. The studies contribute to the subject of real-time optimization as they represent components of commercial RTO systems. They also represent innovative decision support techniques with broader application, as they can be implemented with lower investments and do not demand the complexity required by RTO systems.

Sumário

Prefácio	xi
1 Introdução	1
1.1 Motivação	1
1.2 Objetivos	7
2 Revisão da Literatura	9
2.1 Evolução do Tema	9
2.2 Otimização no Sistema de Automação e Controle de Processos	13
2.3 Estratégias de Otimização em Tempo Real	18
2.3.1 Abordagem em Duas Etapas	21
2.3.1.1 Estabilidade	22
2.3.1.2 Critério Local de Adequação do Modelo	24
2.3.2 Método ISOPE e Suas Extensões	26
2.3.3 Métodos de Adaptação dos Modificadores	30
2.3.4 Condições Suficientes de Viabilidade e Otimalidade	36
2.3.5 Otimização por Regulação	38
2.3.5.1 Controle Auto-Otimizado	39
2.3.5.2 Controle Preditivo Econômico baseado em Modelo	44
2.4 Sistemas RTO	49
2.4.1 Estrutura e Funcionamento	49
2.4.2 Detecção de Estado Estacionário	50
2.4.2.1 Alternativas para o controle de execução do sistema RTO	54
2.4.3 Detecção de Erros Grosseiros, Reconciliação de Dados e Estimção de Parâmetros	55
2.4.4 Otimização Econômica	62
2.4.5 Algoritmos de Otimização	64
2.4.6 Aspectos das Incertezas e seu Efeito na Solução do Sistema RTO	68
2.4.6.1 Análise dos Resultados	69

2.4.6.2	Estratégias de Imposição das Restrições	70
2.4.6.3	Otimização sob Incertezas	72
2.4.7	Projeto e Avaliação de Desempenho de Sistemas RTO	77
2.4.8	Aplicações	85
3	Processos de Elevada Dimensão	88
3.1	Introdução	90
3.2	Detecção e Diagnóstico de Falhas com Métodos Baseados em Dados .	92
3.3	Procedimento Proposto	94
3.3.1	Modelagem	94
3.3.2	Índices de Detecção e Limites de Controle	95
3.3.3	Função Alarme e Visualização por Diagrama de Cores	99
3.3.4	Extensão do Conjunto de Dados com Variáveis Calculadas . .	100
3.3.5	Algoritmo de Monitoramento do Processo	102
3.4	Aplicação em Dados Industriais Reais	104
3.4.1	Avaliação do Monitoramento com o Conjunto de Dados Original	106
3.4.2	Avaliação do Monitoramento com o Conjunto de Dados Es- tendido	109
3.5	Conclusões	114
4	Processos com Baixa Disponibilidade de Informações	116
4.1	Introdução	118
4.2	Estratégia para Monitoramento de SSF	123
4.2.1	Modelagem da SSF de Hidrolisado de Milho	124
4.2.2	Descrição do Modelo	126
4.2.2.1	Modelo de Sacarificação	126
4.2.2.2	Modelo de Fermentação	128
4.2.2.3	Modelo do Processo	129
4.2.3	Estimação de Parâmetros e Implementação do Modelo	133
4.2.4	Procedimento <i>On-line</i> de Identificação	135
4.2.5	Dados Industriais Reais	138
4.3	Resultados	139
4.3.1	Estimação de Parâmetros e Validação do Modelo	139
4.3.2	Monitoramento do Processo	142
4.3.2.1	Avaliação dos Parâmetros de Ajuste	143
4.3.2.2	Desempenho do Sensor Virtual	145
4.4	Conclusão	150

5	Aspectos Numéricos e Abordagens de Resolução	152
5.1	Introdução	153
5.2	Visão Geral da Reconciliação de Dados em Aplicações Industriais . . .	157
5.2.1	Formulação do Problema	157
5.2.2	Métodos de Solução	162
5.2.3	Estratégias de Solução	165
5.2.4	Algoritmos de Otimização para Programação Não Linear . . .	167
5.2.4.1	Abordagem de Solução	171
5.3	Avaliação Numérica de Diferentes Configurações para a Resolução da Reconciliação de Dados Não Linear em Estado Estacionário	173
5.3.1	Configurações Testadas	174
5.4	Resultados	179
5.4.1	Abordagem de Solução	179
5.4.2	Estimativa Inicial e Estimação de Gradientes	180
5.4.3	Formulação do Problema e Algoritmo de Otimização	183
5.4.4	Comentários Sobre a Métrica de Desempenho	187
5.5	Conclusões	189
6	Conclusões	192
	Referências Bibliográficas	196
A	Questões Preliminares sobre a Formulação do Problema	262
A.1	Programação Matemática	262
A.2	Otimização de Processos Baseada em Modelo	266
A.2.1	Otimização Estática	266
A.3	Presença de Incertezas	270
A.3.1	Incerteza de Medição: Ruídos Aleatórios	272
A.3.2	Incerteza de Modelagem: Erros Estruturais do Modelo e Er- ros Paramétricos	275
A.4	Otimização sob Incertezas	276
A.4.1	Programação Estocástica	276
A.4.2	Otimização Robusta	280

Prefácio

O presente texto é resultado da pesquisa de tese que surgiu como proposta de continuação à tese de doutorado de André Domingues Quelhas (QUELHAS, 2013) e foi realizada no Laboratório de Modelagem, Simulação e Controle de Processos (LMSCP) do Programa de Engenharia Química (PEQ) da COPPE/UFRJ contando com o apoio das empresas parceiras Petrobras, EMC e Global Yeast. André fez um trabalho primoroso, ao analisar de forma teórica os fundamentos da otimização em tempo real e aspectos da implementação prática desses sistemas na indústria, dando especial atenção à estratégia de otimização *on-line* baseada em modelos de processo estacionários (RTO) conhecida como estratégia em duas etapas. Outros trabalhos relacionados ao tema, também desenvolvidos no LMSCP e influentes na presente pesquisa, foram os trabalhos de GOMES (2007), MAGALHÃES (2010), JESUS (2011) e TEIXEIRA (2013).

Com base no trabalho de QUELHAS (2013), foi possível identificar aspectos técnicos, diferentes formulações e detalhes de implementações de sistemas RTO que ainda carecem de atenção na literatura, oferecendo oportunidades de pesquisa. Entre tais oportunidades, destaca-se a questão relativa ao controle de execução de sistemas RTO, predominantemente vinculado à detecção de estacionariedade dos sinais de processo, que atua como filtro no sistema; e o emprego de técnicas de otimização sob incertezas, no escopo da otimização em tempo real aplicada a processos químicos, temas esses que nortearam o início do presente trabalho. Todavia, embora de inegável valor científico, o valor prático desses estudos pode ser anulado (ou dificilmente percebido) por conta de alguma das vulnerabilidades de sistemas RTO. Além disso, visto que sistemas RTO somente podem estar presentes em processos com elevado nível de automação, esse valor somente estaria acessível, em tese, a esses processos.

Mostram-se mais relevantes, estando presentes em diversos processos, a despeito do nível de automação existente, os aspectos básicos da realidade prática da operação de processos, como a capacidade de identificação do processo dadas as variáveis medidas disponíveis (seja em processos com muitas ou poucas variáveis medidas) e o adequado tratamento aos aspectos numéricos envolvidos. Diante disso, técnicas de suporte à tomada de decisão, em especial as técnicas de monitoramento *on-line*,

tornaram-se o foco principal da presente pesquisa, mas sem perder o vínculo com a otimização de processos em tempo real. Tais técnicas podem ser utilizadas para realizar a otimização de processos em tempo real, mas não são dependentes da existência de um sistema RTO. Sendo o monitoramento um dos requisitos para aplicação de técnicas de controle e otimização, a aplicação de técnicas de monitoramento em processos com baixo nível de automação representa um importante avanço para a modernização e aplicação de técnicas mais avançadas.

A apresentação dessa tese de doutorado foi dividida em quatro partes. Na primeira parte, representada pelo Capítulo 1, é feita uma introdução, tendo o objetivo primário de apresentar a relevância científica e prática do tema e os objetivos da pesquisa.

A revisão da literatura é apresentada na segunda parte, contida no Capítulo 2, com uma discussão sobre o estado da arte da otimização em tempo real. Faz-se uma discussão sobre aspectos gerais, estratégias de abordagem, características e classificação de sistemas RTO. Constam no Apêndice A a apresentação de características e terminologia dos problemas de programação matemática utilizadas na introdução.

Na terceira parte do trabalho, representada pelos Capítulos 3, 4 e 5, são apresentadas as principais contribuições dessa tese e os trabalhos realizados, que envolvem aplicações reais no âmbito de projetos de pesquisa e desenvolvimento provenientes de parcerias entre a COPPE/UFRJ e empresas que apoiam atividades relacionadas ao tema da otimização em tempo real. Questões relacionadas à identificação de processos (representação do sistema real com auxílio de modelos), detecção da condição de operação (estacionária ou dinâmica, normal ou em falha) e outras condições de adequação do modelo ao processo e técnicas de otimização do processo foram investigadas.

Na última seção, Capítulo 6, apresentam-se as conclusões mais importantes da pesquisa, bem como aspectos discutidos ao longo do texto que podem servir de referência e inspiração para trabalhos futuros.

Por fim, deve ficar registrado que esta pesquisa foi financiada pela CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, e por empresas parceiras.

Capítulo 1

Introdução

1.1 Motivação

Redução de custos de produção e aumento de lucro operacional são temas de grande relevância em qualquer processo de manufatura, principalmente em ambientes de alta competição. Em alguns setores, como o petroquímico, é reconhecido já há algum tempo que pequenas melhorias nas condições de operação podem afetar significativamente as receitas (RUCKAERT *et al.*, 1978). Adicionalmente, outros objetivos, como o aumento da eficiência e restrições com relação à qualidade do produto, segurança do processo e impactos ambientais, podem requerer significativas modificações do projeto e dos procedimentos e condições de operação. Logo, toda ferramenta ou método que se proponha a lidar com a melhoria do processo desperta grande interesse. Uma das ferramentas mais importantes da engenharia para esse propósito é a otimização (EDGAR *et al.*, 2001).

A aplicação das teorias de otimização em problemas de Engenharia Química é estudada e desenvolvida pela área conhecida como *Process Systems Engineering* (PSE)¹, que nasceu com o objetivo de aplicar as técnicas de otimização no projeto ótimo de operações unitárias, mas evoluiu e expandiu seu escopo, incluindo hoje a operação de processos, bem como a integração entre projeto e operação, e reunindo conhecimentos da engenharia química, da matemática aplicada e das ciências da computação para o desenvolvimento de metodologias relacionadas ao processo de tomada de decisão. PSE pode ser melhor entendido como Engenharia de Sistemas para Processos e pode ser definido como a área da Engenharia Química relacionada com metodologias usadas para tomada de decisão e aplicadas para indicar como planejar, como projetar, como operar e como controlar operações unitárias, processos

¹ Até onde se sabe, o termo PSE apareceu pela primeira vez em 1963, como título para o simpósio CEP Symp. Series No. 46, Vol. 59. (STEPHANOPOULOS e REKLAITIS, 2011; TAKAMATSU, 1983). Vale ressaltar que o termo *computer-aided process engineering* (CAPE) não é considerado sinônimo, sendo PSE uma área muito mais ampla (KLATT e MARQUARDT, 2009).

de produção e indústrias químicas. Dessa forma, o desenvolvimento e aplicação de métodos e ferramentas para modelagem, simulação, monitoramento, controle e otimização na análise e projeto de equipamentos e processos são a competência central da área PSE, usados para lidar com a complexidade e os múltiplos objetivos do processo de tomada de decisão em ambientes industriais (GANI *et al.*, 2000; KLATT e MARQUARDT, 2009; PONTON, 1995; TAKAMATSU, 1983).

Decisões relacionadas à operação podem ser determinadas e implementadas por meio de sistemas de controle e automação² em diferentes níveis, como processo, planta industrial ou corporação (EDGAR *et al.*, 2001). Usualmente, o sistema de controle e automação de processos é composto por camadas (Figura 1.1), que se dispõem de forma hierárquica e se distinguem pelo tempo de resposta às perturbações, pelos modelos utilizados, pelos objetivos de controle, pelos algoritmos de solução e, por fim, pelos diferentes conjuntos de usuários. Essas camadas podem ser conceitualmente divididos em³ : planejamento, programação, otimização de processo, controle supervisorio (geralmente técnicas de controle avançado, como controle preditivo multivariável não linear com restrições) e controle regulatório (usualmente controles monovariáveis, como o controle proporcional integral derivativo - PID) (ELLIS *et al.*, 2014).

Nas camadas mais altas, o planejamento e a programação da produção são responsáveis pela coordenação das atividades em longos horizontes de tempo, como meses, semanas e dias, estabelecendo as metas de produção, de forma a atender as restrições que envolvem a capacidade produtiva, as vendas, o transporte e a distribuição. Enquanto isso, objetivos de operação, como nível de produção e qualidade de produto, são implementados pela camada de controle, que confere estabilidade e rejeição de efeitos de alguns tipos de perturbações, mantendo as variáveis controladas em seus *setpoints* especificados e operando continuamente em uma frequência que abrange de minutos a segundos.

A otimização de processos está na camada intermediária, normalmente conectando esses dois níveis, o que lhe confere relevância técnica e estratégica. *a priori*, a camada de otimização usualmente consiste no nível mais baixo em que aspectos econômicos podem ser considerados explicitamente, sendo capaz de traduzir os objetivos das camadas superiores em informações que a camada de controle seja capaz de

² A ciência e a engenharia de Controle e Automação também são descritas pelo termo *Automática* (BOTTURA, 2013). O objetivo clássico da automação é substituir as intervenções humanas no controle, planejamento e solução de problemas por dispositivos automáticos e computadores (BAINBRIDGE, 1983).

³ SHOBRYE e WHITE (2002) afirmam que planejamento, programação e controle são ferramentas fundamentais na atividade industrial, devendo existir em todo ambiente de produção, e que a hierarquia entre elas se forma pelo fluxo natural de decisões em uma organização. Esse trabalho oferece uma discussão profunda sobre a função do planejamento, programação e controle de uma atividade industrial, trazendo um resumo de suas características, práticas típicas em empresas, tecnologias disponíveis e desafios relacionados à melhoria e integração entre essas atividades.

processar. Usualmente, a conexão fornecida pela camada de otimização se dá pelas variáveis controladas, calculando os novos *setpoints* que serão implementados pelas camadas inferiores. Adicionalmente, variáveis manipuladas que não foram associadas a sistemas de controle constituem graus de liberdade disponíveis e, portanto, podem ser incorporadas a essa camada de otimização. Dessa forma, os efeitos das perturbações de processo são atenuados pelos controles supervisão e regulatório, ficando a cargo da otimização a detecção e determinação de mudanças nas condições ótimas de operação, que consistem em perturbações de frequências mais baixas, como alterações na qualidade, composição e custo da matéria-prima (KOOKOS, 2005; MARLIN e HRYMAK, 1997; SKOGESTAD, 2000a, 2004; TRIERWEILER, 2014; YE *et al.*, 2014).

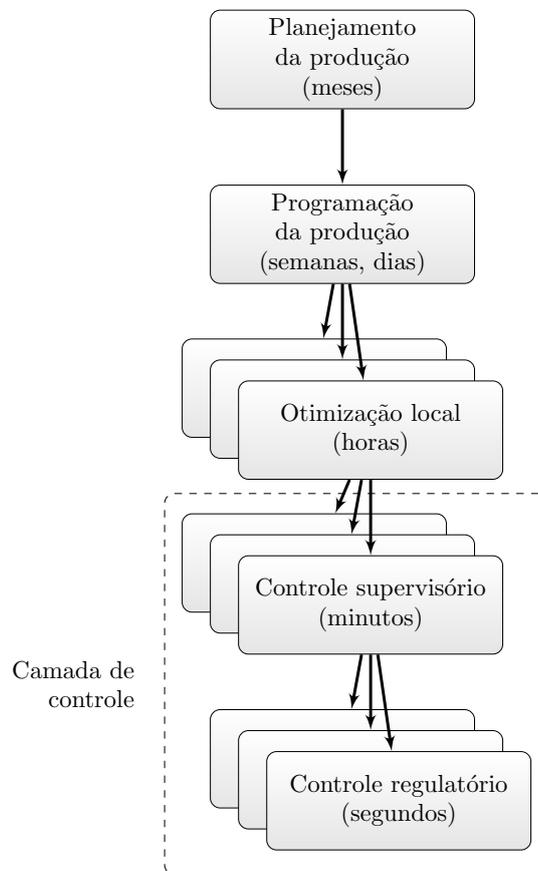


Figura 1.1: Hierarquia usual em um sistema de controle, adaptado de SKOGESTAD (2004).

Nesse contexto, a otimização de processos consiste em uma metodologia sistemática para determinação das melhores condições operacionais, apresentando-se como uma ferramenta de auxílio ao processo de tomada de decisão. Na camada de otimização de processos, a determinação de *setpoints* é formulada como um problema de otimização, *i.e.*, formulada como um modelo de programação matemática, composto por uma função objetivo, um modelo do processo e restrições pertinentes às

variáveis de operação, em que técnicas de otimização são usadas para encontrar as soluções. Comumente a otimização de processos é utilizada para aumento da rentabilidade econômica, sendo referida como otimização econômica. Nesse caso, a função objetivo consiste em um índice de desempenho do processo, geralmente um índice econômico, como rentabilidade operacional, na forma de custo ou lucro de produção em um determinado período de tempo, em que usualmente ignoram-se os custos fixos e contabilizam-se apenas efeitos variáveis como custos de matéria-prima, custos de produção e preço de venda dos produtos. Modelos mecânicos baseados em primeiros princípios têm tido preferência sobre modelos empíricos para a descrição matemática do processo, pois são válidos em uma faixa operacional mais ampla e fornecem variáveis e parâmetros com significado físico. Considerando processos contínuos, o modelo adotado é geralmente do tipo estacionário, porquanto a intenção é que o processo seja operado em estado estacionário, a não ser que ocorram mudanças de *setpoint*. Mas a validade do modelo estacionário deve ser testada, pois depende da frequência das perturbações de entrada e do tempo característico do processo, *i.e.*, se a resposta dinâmica é rápida ou lenta (JANG *et al.*, 1987; NAYSMITH e DOUGLAS, 1995; SEBORG *et al.*, 2010).

Embora a otimização econômica possa parecer uma solução definitiva, na prática ela é afetada pela presença de variadas fontes de incerteza, que comprometem a confiabilidade dos resultados. Mesmo que o processo seja operado de maneira a apresentar comportamento estacionário, parâmetros e outras variáveis importantes variam ao longo da operação em virtude de desgaste natural dos equipamentos, incrustações, perda de atividade de catalítica, entre outras causas, bem como em virtude de perturbações de processo que possuem comportamento variante no tempo e a mudanças nas condições climáticas (HOFFERBER e WEBB, 1978). Por outro lado, do ponto de vista do mercado, demanda, disponibilidade e preço de produtos e insumos flutuam em função de diversas outras variáveis. Não fossem as incertezas, dada a disponibilidade de um modelo do processo, seria possível utilizá-lo em uma otimização *off-line* e implementar os valores das variáveis de entrada em malha aberta (BONVIN e SRINIVASAN, 2013).

Essas fontes de incerteza, que podem ser estacionárias ou dinâmicas, determinísticas ou estocásticas, e em grande parte nada ou pouco controláveis, geralmente alteram o ponto ótimo de operação e as restrições, limitando a possibilidade de alcançar o ótimo econômico por meio de métodos *off-line* (JANG *et al.*, 1987). Assim, a execução repetida da otimização econômica para a contínua reavaliação e alteração das condições de operação com base nas medidas do processo, de forma a maximizar a rentabilidade sem violar restrições operacionais, constitui uma estratégia sedutora, que emergiu como otimização baseada em medições e tem sido designada na literatura como otimização em tempo real (*real-time optimization*)

(CUTLER e PERRY, 1983), otimização *on-line* (BAMBERGER e ISERMANN, 1978; CHEN e JOSEPH, 1987; JANG *et al.*, 1987) e controle otimizado (*optimizing control*) (ARKUN e STEPHANOPOULOS, 1980).

É usual que estudos da literatura incorporem o uso de modelo de processo estacionário à definição do termo *otimização em tempo real* (*e.g.*, ZHANG *et al.* (2001b)). Porém, buscando maior clareza na definição dos conceitos, admite-se nesse texto que problemas de otimização em tempo real baseados em modelos estacionários devam ser designados apenas pela sigla RTO, enquanto aqueles baseados em modelos dinâmicos devam ser designados por DRTO (do termo em inglês *Dynamic Real-Time Optimization*) (BONVIN e SRINIVASAN, 2013).

Até onde se pode determinar, a otimização em tempo real vem sendo tema de pesquisa na academia desde os anos 1960 (DURBECK, 1965, *apud* ROBERTS, 1978) e prontamente teve sua importância reconhecida pela indústria (YOULE e DUNCANSON, 1970, *apud* ELLIS *et al.*, 1988), tornando-se mais tarde uma necessidade da visão industrial (CUTLER e PERRY, 1983; KLEINSCHRODT III e JONES, 1996) e fazendo parte de diversas tecnologias patenteadas (CULLICK e JOHNSON, 2007; KRIST *et al.*, 1996; LU, 2013; SHAPIRO *et al.*, 2015). Inúmeros estudos contribuíram para o desenvolvimento da otimização em tempo real (DARBY *et al.*, 2011; FRANÇOIS, 2014; FRANÇOIS e BONVIN, 2013a; NAYSMITH e DOUGLAS, 1995), que tornou-se um amplo tópico de pesquisa, envolvendo muitas áreas, como modelagem, estatística, estimação de parâmetros, estimação de estados, reconciliação de dados, controle preditivo baseado em modelos (*model-based predictive control* – MPC) e programação matemática (MARCHETTI, 2009). Uma quantidade substancial de aplicações industriais já foi relatada (BAILEY *et al.*, 1993; BASAK *et al.*, 2002; LAUKS *et al.*, 1992; MARLIN e HRYMAK, 1997), muitas delas na indústria petroquímica (ROTAVA e ZANIN, 2005; WHITE, 1998, 2005; YOUNG, 2006). Vale ressaltar, todavia, que a otimização em tempo real pode não ser vantajosa nos casos em que as condições operacionais ótimas da planta sejam pouco sensíveis às variações esperadas nas condições de processo e de mercado (MERCANGÖZ e DOYLE, 2008).

Grande parte da literatura da área dedica-se a sistemas RTO em que as medidas são usadas para adaptar continuamente o modelo do processo, sendo esse método conhecido como RTO em duas etapas (do termo em inglês *two-step* ou *two-phase* RTO). Reconhecida como a primeira abordagem amplamente disponível de um método de otimização em tempo real, além de ser também o mais comum, consiste em adaptar os parâmetros do modelo com base nos desvios observados entre valores preditos e dados medidos (1^a etapa) e utilizar o modelo atualizado para determinar os valores das variáveis de decisão que conferem o melhor desempenho econômico ao processo (2^a etapa) (FRANÇOIS e BONVIN, 2013a; KIPARISSIDES *et al.*, 1994;

ROBERTS, 1979). Além disso, a estratégia em duas etapas é a forma como sistemas RTO estão implementados em *softwares* comerciais típicos, como ROMeO 5.1 (Invensys, Houston, TX) e AspenPlus 7.1 (Aspentech, Burlington, MA) (QUELHAS *et al.*, 2013). Uma implementação típica do sistema RTO em duas etapas possui essencialmente os seguintes módulos: i) identificação de estado estacionário; ii) detecção de erros grosseiros, reconciliação de dados e estimação de parâmetros; iii) otimização econômica; e iv) análise dos resultados (NAYSMITH e DOUGLAS, 1995; TRIERWEILER, 2014).

Apesar da popularidade, uso em ferramentas comerciais e emprego industrial, implementações da estratégia em duas etapas podem apresentar algumas vulnerabilidades, notadamente: i) uso de modelos imperfeitos, ou não adequados, do processo; ii) conhecimento incompleto e incerto das informações; iii) caracterização insuficiente das incertezas das informações corrompidas; e iv) questões relacionadas às técnicas de otimização (QUELHAS *et al.*, 2013).

Uma das fragilidades mais abordadas na literatura trata da falta de adequação do modelo para estimar com precisão os gradientes da planta, *i.e.*, do processo real. Essa questão já foi há muito identificada e tem servido como motivação para o desenvolvimento de outras estratégias de RTO e critérios de adequação de modelos (FORBES *et al.*, 1994; MARCHETTI *et al.*, 2009; ROBERTS, 1978). Ainda que contribuições significativas tenham sido feitas no campo teórico, com garantias de viabilidade e otimalidade quando o algoritmo converge para uma solução, muitas discussões admitem casos idealizados com total disponibilidade de informações, desconsiderando como os algoritmos se comportariam em aplicações práticas, nas quais dificilmente os dados existentes permitem a determinação precisa das funções e suas derivadas (BUNIN, 2014).

Um desafio importante existente na etapa de identificação (modelagem) de processos industriais é a elevada dimensão do número de variáveis, equipamentos, cenários de operação, tipos de perturbações possíveis, dentre outros. Procedimentos em tempo real para aplicação em sistemas industriais de grande porte impõem exigências de desempenho computacional e precisão de técnicas numéricas que amplificam ainda mais essa dificuldade. Se simplificações forem feitas diante dessas dificuldades para viabilizar a estratégia, certamente influenciarão o desempenho final do sistema RTO (BIEGLER *et al.*, 1985).

Embora a dimensão do problema em análise impacte significativamente a complexidade do procedimento de solução, influenciando a estratégia e a abordagem usadas na implementação de procedimentos em tempo real, as dificuldades de modelagem e implementação de procedimentos em tempo real não é eliminada pela redução da dimensão do problema. Diversos processos industriais são regidos por fenômenos que nem sempre dispõem de tecnologias de sensoriamento e medição ma-

duras e adequadas ao ambiente fabril. Assim, a dificuldade de medição de diversas grandezas importantes a alguns processos também se mostra como um desafio na etapa de identificação de modelos fenomenológicos para uso em procedimentos em tempo real (FRALEIGH *et al.*, 2003).

Por fim, há casos em que uma representação matemática precisa do processo e um conjunto de informações adequado estão disponíveis. No entanto, a disponibilidade de um modelo fenomenológico associado a um conjunto adequado de variáveis medidas não é suficiente para garantir uma resolução satisfatória do problema de programação matemática, seja na etapa de atualização do modelo, seja na etapa de otimização econômica (FORBES *et al.*, 1994; ROBERTS, 1978).

1.2 Objetivos

No âmbito de projetos de pesquisa e desenvolvimento provenientes de parcerias entre a COPPE/UFRJ e empresas que apoiam atividades relacionadas ao tema da otimização em tempo real, desafios relacionados à identificação de processos (representação do sistema real com auxílio de modelos), detecção da condição de operação (estacionária ou dinâmica, normal ou em falha) e solução do problema de otimização mostraram-se importantes. Essas atividades suportam diversas ferramentas de auxílio ao processo de tomada de decisão e estão presentes em sistemas RTO, podendo exercer forte influência sobre o desempenho final desses sistemas, sobretudo nos *softwares* comerciais dos grandes provedores de tecnologia que fazem uso da estratégia em duas etapas (CÂMARA *et al.*, 2016; QUELHAS *et al.*, 2013).

Dado que muitos estudos focaram no efeito da qualidade do modelo sobre a etapa de otimização, pretendeu-se dar um passo atrás para considerar como a disponibilidade dos dados influencia a identificação do processo. O modelo do processo pode ter seu uso ampliado para atividades além da otimização, como detecção de falhas de operação (VENKATASUBRAMANIAN *et al.*, 2003). Visto que falhas podem revelar-se por meio de dinâmica na trajetória das variáveis medidas, o próximo estado estacionário pode ser uma parada não programada. Isso é especialmente relevante para aplicações reais, pois mesmo que o modelo esteja completamente aderente ao processo, pode ser mais interessante detectar e corrigir uma falha na operação que progredir com a otimização do processo. Adicionalmente, aspectos numéricos importantes influenciam a solução dos problemas de otimização, mas poucos trabalhos na literatura dedicaram-se a caracterizar esses efeitos de maneira detalhada.

Nesse sentido, no contexto genérico do problema de otimização em duas etapas representado na Figura 1.2, que inclui uma etapa de identificação (atualização) do processo, uma etapa de análise de adequação do modelo ao processo (alarmes) e uma etapa de otimização do processo com base no modelo identificado, o presente

trabalho teve como objetivo abordar parte dos desafios mencionados nos seguintes termos:

1. Desenvolver uma estratégia para monitoramento de processos industriais de grande escala com base no uso de modelos empíricos por meio de alarmes que atestam a adequação do modelo ao processo;
2. Desenvolver um procedimento de identificação *on-line* para monitoramento de processos com baixa disponibilidade de informações com base no uso de um modelo fenomenológico;
3. Caracterizar o desempenho de diferentes técnicas numéricas empregadas na resolução do problema de otimização associado às etapas de *Identificação* do processo e de *Otimização* da Figura 1.2.

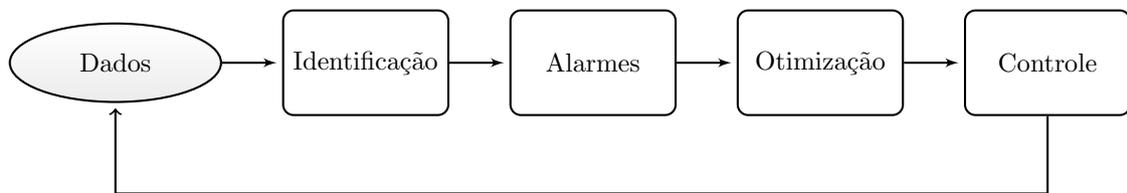


Figura 1.2: Fluxograma alternativo com as etapas envolvidas em uma implementação real de sistema RTO.

Capítulo 2

Revisão da Literatura

2.1 Evolução do Tema

Já há algum tempo que o interesse em melhorar o retorno econômico de processos industriais motivou a aplicação de procedimentos sistemáticos que envolvem técnicas de otimização para determinar condições de operação ótimas e bem definidas, em lugar de procedimentos heurísticos ou soluções *ad hoc*. Logo foi identificado que a natureza dinâmica das perturbações de processo, parâmetros e outras variáveis importantes geralmente impõem sérias limitações para que se alcance o resultado previsto, quando a técnica é executada *off-line*. Assim, não demorou para que fosse reconhecida a importância de métodos capazes de avaliar continuamente as condições operacionais e alterar as variáveis para buscar a máxima produtividade econômica, *i.e.*, métodos de otimização *on-line* (CUTLER e PERRY, 1983; GARCIA e MORARI, 1981; JANG *et al.*, 1987).

A otimização em tempo real tem como uma de suas raízes o problema de determinação experimental do extremo de uma função (HOTELLING, 1941). Dentre as primeiras iniciativas para lidar com esse problema, pode ser citada a metodologia de superfície de resposta, associada principalmente ao uso da técnica de planejamento de experimentos (BOX, 1954; BOX e WILSON, 1951). O procedimento consiste em construir um modelo empírico com a informação obtida por meio dos experimentos, geralmente um polinômio de primeira ou segunda ordem⁴, que é então utilizado para identificar o conjunto de valores que conduza à melhor direção do gradiente da função de desempenho. Assim que o ponto corrente é atualizado, repete-se o procedimento. Também conhecidos como métodos de busca direta, esses métodos de otimização *on-line* também envolvem outras estratégias, como a implementação direta de algoritmos de busca desenvolvidos na teoria de otimização (SPENDLEY

⁴ O ajuste com o polinômio equivale a admitir que a função das respostas possa ser localmente representada por sua expansão em série de Taylor, ignorando termos de ordem superior (BOX e YOULE, 1955).

et al., 1962), bem como métodos de busca aleatória (BROOKS, 1958). Algumas dessas técnicas foram comparadas no trabalho de EDLER *et al.* (1970).

Outro método de busca direta que ganhou bastante popularidade em processos industriais recebeu o nome de Operação Evolucionária (do termo em inglês *Evolutionary Operation* - EVOP) (BOX, 1957; BOX e DRAPER, 1969). O método EVOP avalia o desempenho de cada ponto em um conjunto de experimentos realizados na planta em torno da condição de operação corrente. Com base em testes estatísticos, a condição que tenha conferido o melhor desempenho é implementada e o procedimento é repetido. Cada experimento (ou ponto) do conjunto pode ser realizado de maneira aleatória ou em alguma ordem particular, sendo que o método é adequado tanto para processos descontínuos, quanto para contínuos. No caso de processos contínuos, deve-se permitir que o processo alcance o estado estacionário após a implementação de mudanças nas condições operacionais, *i.e.*, nas variáveis independentes (HIMMELBLAU, 1970).

Embora operacionalmente mais trabalhoso, o método EVOP é simples, fácil de implementar e ser mantido, e especialmente útil e adequado quando existe pouco conhecimento sobre a operação do processo. Ademais, toda mudança de operação apontada pelo EVOP, em tese, produz melhoria de desempenho (FORBES, 1994). Além disso, à medida que o método avança para o estado ótimo de operação, grande quantidade de informação é adquirida a partir dos experimentos e pode ser utilizada de forma inteligente em outras aplicações. Existem registros de implementações bem sucedidas do método EVOP na indústria química (HUNTER e KITTRELL, 1966; RYAN, 2011), como por exemplo em processos de polimerização (CAVALCANTI e PINTO, 1997). Uma discussão sobre EVOP e outras estratégias de otimização de processos por meio de experimentação podem ser encontradas em HIMMELBLAU (1970) e RYAN (2011).

Todos os métodos de busca direta que não usam qualquer informação *a priori* do processo são pouco atrativos por conta da grande quantidade de experimentações e perturbações necessárias para encontrar o ponto ótimo da planta, as quais aumentam geometricamente com o número de variáveis. Adicionalmente, experimentos em plantas são caros na maior parte dos processos. Sendo procedimentos essencialmente estacionários, as medidas de cada experimento devem ser registradas apenas depois que o regime do processo tenha se estabelecido, o que torna o método lento, especialmente se forem muitas as variáveis manipuladas. Em processos com resposta dinâmica lenta, sua implementação pode ser inviável. Além disso, ruídos de medição podem dificultar a determinação confiável da direção do gradiente da função de desempenho na região de busca (ELLIS *et al.*, 1988; NAYSMITH e DOUGLAS, 1995). Assim, questões relacionadas à quantidade de experimentações, tempo para a aquisição dos dados, velocidade de convergência e precisão do método, capacidade

de lidar com ruídos de medida e competência para tratar sistemas multivariáveis de grande dimensão motivaram o desenvolvimento de outros métodos de otimização *on-line*, os quais adotaram como estratégia a introdução de informação *a priori* por meio de um modelo do processo (GARCIA e MORARI, 1981).

Mais próximos do que se entende como procedimento *on-line* e com o objetivo de otimizar a operação estacionária da planta, alguns trabalhos propuseram um procedimento adaptativo, com o uso de um modelo dinâmico para fornecer informação estacionária do processo e contribuir com a identificação do ponto ótimo. No trabalho de BAMBERGER e ISERMANN (1978), uma abordagem que combina a identificação recursiva de um modelo dinâmico empírico e otimização estacionária foi aplicada *on-line* ao sistema de resfriamento de vapor de escape em um processo termelétrico em escala piloto. A proposta, que estendeu a técnica de controle adaptativo para um sistema RTO, consiste em estimar os parâmetros do modelo dinâmico com base nas medidas transientes do processo, no lugar de utilizar apenas dados estacionários, o que é especialmente interessante para processos com resposta dinâmica lenta. Uma vez que o modelo dinâmico tenha sido identificado, sua versão estacionária pode ser usada na otimização para determinar os valores das variáveis manipuladas que promovam ganho econômico.

Outras abordagens também seguiram a estratégia de identificação *on-line* de um modelo dinâmico empírico (BHATTACHARYA e JOSEPH, 1982; GARCIA e MORARI, 1981; MCFARLANE e BACON, 1989). Em resumo, esses métodos envolvem as etapas: i) aquisição de dados da planta; ii) identificação do modelo dinâmico empírico; e iii) otimização do índice de desempenho, utilizando a versão estacionária do modelo dinâmico identificado. GOLDEN e YDSTIE (1989) apresentaram uma breve revisão de algumas das abordagens que envolvem o uso de um modelo dinâmico identificado para otimização estacionária do processo e estenderam a proposta de BAMBERGER e ISERMANN (1978) para incorporar informações de modelagem fenomenológica no esquema adaptativo. Embora supere diversas dificuldades e permita a aferição direta dos gradientes do processo, o método pode apresentar problemas com a sintonia de parâmetros presentes na implementação e com imprecisões na etapa de identificação, principalmente em problemas de dimensão elevada (BHATTACHARYA e JOSEPH, 1982). Embora cada um dos métodos use um algoritmo diferente para determinar o ponto ótimo de operação, todos se baseiam em experimentos na planta, por meio de perturbações nas variáveis manipuladas para atualizar o modelo *on-line*, tendo sido classificados como método baseados em perturbação (ZHANG e FORBES, 2006).

Em vista das vantagens do uso de informações *a priori* sob a forma de um modelo, mesmo no caso de um modelo dinâmico empírico, perceberam-se muitos benefícios com o uso de modelos fenomenológicos como estratégia para prever o

comportamento do processo e fornecer melhores estimativas da direção do gradiente da função objetivo (GOLDEN e YDSTIE, 1989). O desenvolvimento do modelo do processo, com o uso de balanços de conservação de massa, energia e momento e de expressões termodinâmicas para relações de equilíbrio e cinética de reações, resulta em um conjunto de equações algébricas e/ou diferenciais que permitem simular o estado do processo. Os parâmetros do modelo são determinados por conhecimento *a priori* do projeto de instalação da planta (*e.g.*, volumes e áreas), por dados da literatura (*e.g.*, constantes de equilíbrio termodinâmica e velocidades específicas de reação), ou por técnicas de estimação com base em dados da operação e medidas experimentais (FRANÇOIS, 2014). Por isso, defende-se que seu uso forneça uma descrição mais completa em relação a modelos empíricos, pois é capaz de capturar as principais características da operação, descrever o sistema em uma faixa mais ampla das variáveis e atribuir interpretação física aos parâmetros (JANG *et al.*, 1987). Se o modelo for preciso e capaz de prever com melhor qualidade o comportamento local da planta, as condições ótimas de operação podem ser encontradas em poucas etapas. Com efeito, alguns trabalhos mostram a superioridade das abordagens baseadas em modelo frente àquelas puramente empíricas (CHEN e JOSEPH, 1987; YIP e MARLIN, 2004).

O método mais antigo e popular baseado em modelo é conhecido como abordagem em *duas etapas*, que pode ser aplicado tanto em problemas de otimização estacionária quanto otimização dinâmica (FRANÇOIS e BONVIN, 2013a). A primeira etapa efetua a atualização do modelo com a informação mais recente da planta, em que a diferença entre os dados do processo e as estimativas calculadas é minimizada para obter o melhor ajuste do modelo aos dados, necessária porque o processo muda ao longo do tempo. Na sequência, o modelo atualizado é usado na etapa de otimização, que fornece o conjunto de valores para as variáveis de decisão que, *a priori*, promovem a melhoria do desempenho do processo de forma iterativa, até alcançar a operação ótima. A origem da abordagem em duas etapas na otimização em tempo real não é precisa; enquanto o método já inspirava os estudos de ROBERTS (1978, 1979) e referências neles citadas, alguns autores atribuem o método a JANG *et al.* (1987) e CHEN e JOSEPH (1987). O método foi proposto por JANG *et al.* (1987) em um problema de otimização dinâmica, ao inserir mudanças de implementação na decomposição formulada por BAMBERGER e ISERMANN (1978) e MUKAI *et al.* (1981). CHEN e JOSEPH (1987), por sua vez, utilizaram a abordagem de JANG *et al.* (1987) na otimização estacionária em tempo real de um reator industrial de produção de óxido de etileno. Além disso, alguns livros fornecem referência teórica para a introdução de estruturas hierárquicas em métodos de otimização (LASDON, 1970; MESAROVIC *et al.*, 1970).

O uso de modelos fenomenológicos em estado estacionário e o uso da abordagem

de otimização em duas etapas constituem dois importantes elementos, cuja associação surgiu por volta dos anos 1980 e teve essencial contribuição dos seguintes fatores: i) avanços tecnológicos na área da computação (com os graduais aumentos de memória e velocidade de processamento, aliado à maior acessibilidade desses equipamentos, em virtude da constante redução de custos); ii) avanços no desenvolvimento de métodos numéricos capazes de lidar com problemas esparsos de grande dimensão e da modelagem por equações abertas; iii) estudos teóricos que forneceram os fundamentos da síntese de estruturas de controle; e iv) desenvolvimento da tecnologia de controle preditivo baseado em modelo (DARBY *et al.*, 2011). Do ponto de vista teórico e científico, esses elementos motivaram o desenvolvimento de diversos estudos na área de otimização em tempo real, que culminaram em importantes avanços, como o surgimento de novos métodos e algoritmos. Do ponto de vista prático, essa combinação veio a constituir a forma padrão de implementações de sistemas RTO na indústria (CUTLER e PERRY, 1983; DARBY *et al.*, 2011; MARLIN e HRYMAK, 1997; NAYSMITH e DOUGLAS, 1995; TRIERWEILER, 2014; WHITE, 1997), sendo o método disponível nos principais *softwares* comerciais de uso industrial corrente, a saber ROMeO (Invensys, Houston, TX) e AspenPlus (Aspentech, Burlington, MA) (QUELHAS *et al.*, 2013).

Antes de apresentar os avanços no campo teórico e seus desdobramentos, bem como os detalhes das implementações reais, convém apresentar algumas contribuições dos estudos que fundamentaram a síntese de estruturas de controle e direcionaram as implementações práticas dos sistemas RTO.

2.2 Otimização no Sistema de Automação e Controle de Processos

Ao contrário de métodos que envolvem modificações no processo (*e.g.*, esquemas de integração energética), a otimização em tempo real é uma estratégia que busca fazer o melhor uso da configuração existente, explorando os graus de liberdade disponíveis para determinar a melhor condição de operação. Como a planta está constantemente exposta a perturbações externas que afetam o processo, mudam suas restrições, alteram a condição de operação ideal ou forçam a operação fora da região de operação permitida, a busca pelas melhores condições operacionais deve ser realizada de forma contínua, em conjunto com alguma forma de controle, para manter a operação no ponto ótimo e evitar a violação de restrições. Isso é alcançado de maneira mais eficiente por meio de sistemas de automação e controle, em que o acompanhamento contínuo das melhores condições de operação e a sua implementação quando tais perturbações ocorrem é denominado controle otimizado (GARCIA e MORARI, 1984).

O propósito principal de um sistema de controle em processos químicos é fornecer uma estrutura dinâmica entre variáveis medidas e manipuladas que cumpra determinados objetivos de maneira (quase) contínua (MORARI *et al.*, 1980). Dadas as diversas fontes de incerteza, seja por conta de erros de modelagem ou pela presença de perturbações não medidas, as medidas do processo consistem na única forma de determinar o estado da planta e a informação nelas contida é utilizada em todos os níveis. Conhecendo o estado corrente da operação, *i.e.*, os valores das respostas do sistema (variáveis de saída) e das variáveis não controladas (por medição ou estimação), a unidade de controle calcula e implementa decisões sobre os valores das variáveis controladas com base nos objetivos de controle (BRDYŚ e TATJEWSKI, 2005).

Esses objetivos de controle podem ser classificados em duas categorias: uma relacionada à viabilidade operacional, *i.e.*, relacionada às diversas dependências existentes entre as variáveis de processo; e outra relacionada aos aspectos financeiros. Essa segunda categoria de objetivos é considerada no sistema apenas se houver graus de liberdade disponíveis para adaptar as condições operacionais após o cumprimento do primeiro conjunto de objetivos. Dessa forma, a otimização do processo torna-se parte integrante do sistema de controle, o qual é composto por: um conjunto de variáveis controladas, usadas para atingir um conjunto de objetivos específicos; um conjunto de variáveis medidas para os propósitos de controle; um conjunto de variáveis manipuladas; e uma estrutura que forneça conexão entre variáveis medidas e manipuladas (MORARI *et al.*, 1980).

Entre as estratégias existentes para o desenvolvimento sistemático e organizado de estruturas de controle para processos químicos, aquelas baseadas no conceito de estrutura hierárquica de controle têm sido as mais relevantes (BRDYŚ e TATJEWSKI, 2005; FINDEISEN *et al.*, 1978; MORARI *et al.*, 1980) e podem ser identificadas como o padrão adotado pela indústria (QIN e BADGWELL, 2003; ROTAVA e ZANIN, 2005; TATJEWSKI, 2002; YOUNG, 2006). A estratégia consiste essencialmente em decompor o objetivo de controle original em uma sequência de subtarefas mais simples, diferentes e estruturadas de maneira hierárquica. As subtarefas criadas são mais bem definidas, facilitando o projeto do sistema ao simplificar a identificação e implementação de ferramentas técnicas específicas, contribuindo também para a melhoria do desempenho e confiabilidade do controle.

Os métodos de decomposição podem ser classificados como funcionais, espaciais ou temporais (BRDYŚ e TATJEWSKI, 2005). O método funcional é aplicado ao processo como um todo e conduz a uma estrutura vertical, em multicamadas, cujas camadas apresentam dependência hierárquica. A decomposição espacial é baseada em uma divisão da tarefa de controle (ou de uma tarefa parcial, *e.g.*, dentro de uma camada da estrutura multicamadas) em subtarefas locais de mesma

função, mas relacionadas a regiões distintas, espacialmente isoladas, do processo. Esse tipo conduz a uma estrutura multigrupos ou multiníveis, efetuando uma decomposição horizontal no processo, dividindo-o em grupos de equipamentos, que são coordenados por um elemento central que permite a interação entre os grupos (MORARI *et al.*, 1980). Dessa forma, a otimização é feita para cada grupo separadamente, de acordo com o elemento coordenador. Diversas abordagens utilizam essa estrutura para o projeto do sistema de controle, como controle otimizado descentralizado (MORARI *et al.*, 1980), controle otimizado para plantas integradas (ARKUN e STEPHANOPOULOS, 1981), controle otimizado de sistemas interconectados (BRDYŚ e TATJEWSKI, 2005) e controle global (LARSSON e SKOGESTAD, 2000; LUYBEN *et al.*, 1997; STEPHANOPOULOS e NG, 2000).

O método de decomposição temporal, por sua vez, é aplicado aos casos em que existe diferença significativa entre a taxa de variação das variáveis de estado ou das perturbações do sistema, também conduzindo a uma estrutura em multicamadas. Nesse método, o objetivo de controle é formulado como um problema de otimização dinâmico e a estrutura em multicamadas formada (conforme mostrado na Figura 1.1) pode ser interpretada como um elemento natural das atividades de controle em uma planta química típica, estando baseada na forma de lidar com as incertezas (BRDYŚ e TATJEWSKI, 2005; MORARI *et al.*, 1980).

A síntese da estrutura de controle que segue o método de decomposição temporal estabelece como objetivo a otimização de um índice de desempenho, sem violar as diversas restrições de operação. Uma vez que perturbações de natureza aleatória estão presentes e exercem influência significativa, a otimização deve considerar essas perturbações, resultando em um problema de controle estocástico não linear, na prática intratável. Uma forma conveniente de manipular o problema é dividir as incertezas em duas componentes, uma estacionária e outra não estacionária, o que define implicitamente uma escala de tempo em dois níveis. A componente estacionária das perturbações varia rapidamente, de forma que o valor estimado é essencialmente zero depois de um curto período de tempo. Essa componente pode ser desprezada na otimização feita para longos horizontes de tempo, sendo tarefa do controle regulatório suprimir sua influência. Por outro lado, a componente não estacionária representa perturbações persistentes ou periódicas, que afetam o processo no longo prazo e, portanto, devem ser consideradas na otimização.

Admitindo que o processo esteja em estado pseudo estacionário (*i.e.*, um estado em que perturbações de alta frequência se anulam no curto prazo), e considerando um período de tempo grande o suficiente para a estimativa das perturbações ser aproximadamente constante e para a dinâmica da planta ser desprezível, é possível reformular o problema original de controle estocástico em um problema de otimização estática (MORARI *et al.*, 1980). Essa abordagem conduz a uma decom-

posição vertical das atividades de controle com base na escala de tempo associada à frequência das perturbações em cada nível, cuja camada mais baixa é responsável pela rejeição de perturbações rápidas, garantindo que o processo seja operado em um regime pseudo estacionário. A camada mais alta, por sua vez, é responsável pela determinação dos *setpoints* ótimos que suprimem perturbações de baixa frequência.

Portanto, o sistema de otimização em tempo real corresponde ao problema de controle otimizado, que consiste na determinação e aplicação automatizada de *setpoints* ou valores de referência para as camadas inferiores do sistema de controle, com base em um critério de desempenho que vincula as condições operacionais ao resultado econômico do processo. Desse modo, o sistema de otimização representa uma malha de controle mais externa, em que geralmente o critério de desempenho é uma relação que confere valores monetários ao uso de determinado conjunto de graus de liberdade da operação do processo.

A otimização estática deve funcionar em uma frequência inferior à das camadas acima, mas superior à ocorrência de perturbações externas e variações (“envelhecimento”, desgaste) do processo (ELLIS *et al.*, 1988; QUELHAS, 2013). Além de ter origem na estratégia de decomposição das ações de controle, a otimização estática e o uso de modelos estacionários são justificados também com o argumento de que o desempenho econômico é determinado principalmente pela operação estacionária. Ademais, o efeito do desempenho do controle dinâmico pode ser parcialmente considerado na análise econômica, incluindo um termo de erro do controle como uma perturbação adicional (SKOGESTAD, 2000b). Apesar disso, se as mudanças nas perturbações ou nas características do processo não apresentarem variação lenta, quando comparadas à dinâmica do processo, a otimização estacionária em tempo real não fornece valores ótimos; esse é um caso importante do ponto de vista prático, pois processos contínuos são quase sempre projetados para operar em estado estacionário. Adicionalmente, nem sempre consideram-se na etapa de projeto as características dinâmicas do processo (BRDYŚ e TATJEWSKI, 2005).

Embora seja uma ferramenta muito atraente, sistemas RTO não são necessários ou adequados em todos os casos e cabe analisar sua relevância para a melhoria do desempenho econômico (SKOGESTAD, 2004). Conforme mencionado anteriormente, o primeiro objetivo do sistema de controle em um processo químico é garantir que a operação seja segura, que os produtos atendam às especificações de qualidade e que a produção atinja as metas estabelecidas, sempre respeitando as diversas restrições de operação existentes. Dessa forma, se não houver graus de liberdade disponíveis após o cumprimento do primeiro objetivo, não é possível melhorar a operação com o RTO (SKOGESTAD, 2000b). Na situação em que existem graus de liberdade disponíveis, haverá sempre alguma oportunidade de melhoria com a otimização em tempo real se i) a métrica de desempenho do processo apresentar dependência significativa

com tais graus de liberdade, ii) perturbações ocorrerem com frequência tal que exija um ajuste automatizado das variáveis manipuladas (MARLIN e HRYMAK, 1997), e iii) se as perturbações existentes mudarem significativamente o ponto ótimo ao longo da operação. Deve-se analisar também a relação entre a operação ótima e o conjunto de restrições ativas. Se as restrições ativas não mudarem e houver variáveis que se encaixem no esquema de controle auto-otimizável, então o RTO proporciona pouco benefício (SKOGESTAD, 2004). Ainda, se a operação ótima consistir em operar sob as restrições, que por sua vez podem ser determinadas pelo MPC, então o RTO não é necessário (DARBY *et al.*, 2011). A maximização da produção de combustível de aviação na destilação do petróleo e da pressão em um reator de craqueamento catalítico fluidizado, quando compressores e sopradores estão operando em sua capacidade máxima, são exemplos de casos em que a otimização pode ser realizada exclusivamente pela camada de controle, *i.e.*, o ponto ótimo de operação está nas restrições (ROTAVA e ZANIN, 2005). Há casos em que a otimização *off-line* (ou manual) é suficiente para a melhoria do desempenho, podendo ser realizada, *e.g.*, a partir de procedimentos de operação padrão (MARLIN e HRYMAK, 1997; SKOGESTAD, 2004).

A característica chave que aponta para o uso do RTO consiste na variabilidade do ponto ótimo de operação em virtude de mudanças nas condições operacionais, nas características da matéria-prima, no conjunto de restrições e nas condições de mercado. Além disso, o RTO complementa o controle supervisor, geralmente representado pelo MPC, seja por capturar o comportamento não linear, seja por considerar um conjunto maior da planta em seu escopo, *i.e.*, considerar um envelope maior. Nessas situações, mesmo se o ponto ótimo estiver nas restrições, é possível que apenas o RTO seja capaz de encontrar os novos valores ótimos para o conjunto de restrições (DARBY *et al.*, 2011).

Não obstante, a estrutura em multicamadas apresenta algumas desvantagens, cuja principal consiste na necessidade em aguardar que a planta atinja o regime permanente antes de coletar as medidas e realizar a otimização (ENGELL, 2007). O atraso na otimização ocorre após uma mudança nas variáveis manipuladas ou ocorrência de perturbações. Semelhantemente ao que ocorre com os métodos diretos, a convergência do sistema ao ponto ótimo pode ser lenta. Se as perturbações forem medidas e as mudanças nas condições de operação forem conhecidas, então o último modelo estacionário pode ser usado na otimização sem que a planta alcance o novo estado estacionário (SEQUEIRA *et al.*, 2002), mas isso dificilmente acontece. A estratégia que combina RTO e MPC em uma única camada, conhecida como MPC econômico (*economic model predictive control* - EMPC) (DE GOUVÊA e ODLOAK, 1998), pode constituir uma solução alternativa para essa desvantagem, conforme discutido na Seção 2.3.5.2.

Além disso, a otimização no sistema RTO não considera o desempenho do processo na transição de uma condição de operação para outra. Assim, para os processos que precisam responder rapidamente a variações e perturbações, o EMPC ou um sistema RTO dinâmico (ou DRTO) podem constituir soluções alternativas mais adequadas e vantajosas. Enquanto o RTO estático fornece o conjunto de valores ótimos do regime estacionário da operação, o DRTO fornece uma trajetória de mudanças do ponto de operação. Um sistema DRTO não precisa que condições estacionárias sejam atingidas ou aplicadas, sendo sua formulação e solução semelhante à abordagem usada para os controles preditivos não lineares (TRIERWEILER, 2014).

2.3 Estratégias de Otimização em Tempo Real

Um sistema RTO resolve continuamente um problema de otimização estática, estabelecido como: dado um processo contínuo operando em estado estacionário, observado por um conjunto de variáveis de saída \mathbf{y}_p , e um conjunto de graus de liberdade \mathbf{u} (que podem ser *setpoints* do controle supervisor), determinar os valores de \mathbf{u}_p que correspondem ao extremo da métrica de desempenho da planta φ_p e que não violem o conjunto de restrições \mathbf{g}_p . A formulação desse problema, discutido em detalhes na Seção A.2.1, é escrito matematicamente como:

$$\begin{aligned} \mathbf{u}_p^* := & \arg \min_{\mathbf{u}} J_p := \varphi_p(\mathbf{u}, \mathbf{y}_p) \\ & \text{sujeito a } \mathbf{g}_p(\mathbf{u}, \mathbf{y}_p) \leq \mathbf{0} \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned} \quad , \quad (2.1)$$

em que $\varphi_p : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}$ é a função escalar a ser minimizada; \mathbf{u}_p^* é o argumento que minimiza $J_p(\cdot)$, *i.e.*, a solução do problema; $\mathbf{u} \in \mathbb{R}^{n_u}$ é o vetor de n_u variáveis de decisão (ou de entrada); $\mathbf{y}_p \in \mathbb{R}^{n_y}$ é o vetor de n_y variáveis de saída; $g_{p,i} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}$, $i = 1, \dots, n_g$ é o conjunto de funções de restrição de desigualdade; e \mathbf{u}^L e \mathbf{u}^U são os limites das variáveis de decisão.

Diferentes técnicas de implementação da otimização em tempo real estacionária estão disponíveis na literatura, sendo identificadas como métodos de otimização baseada em medições (FRANÇOIS e BONVIN, 2013a; KADAM *et al.*, 2005; SRINIVASAN *et al.*, 2003). Não são considerados nesse conjunto de métodos RTO os métodos puramente experimentais, *i.e.*, que não utilizam nenhuma informação *a priori* na forma de um modelo do processo (FRANÇOIS, 2014). Fazem parte desse último grupo os métodos de superfície de resposta inspirados em planejamento de experimentos (GEORGAKIS, 2009; MYERS *et al.*, 2009) e os métodos que não usam informações de derivadas, como forma de contornar o problema de estimação dos gradientes (BOX e DRAPER, 1969; CONN *et al.*, 2009; HOLMES, 2003).

Métodos RTO podem ser classificados de acordo com uma característica específica do problema ou da formulação (CHACHUAT *et al.*, 2009; ELLIS *et al.*, 1988; FRANÇOIS *et al.*, 2012; GARCIA e MORARI, 1981; ZHANG e FORBES, 2006). Nas propostas mais antigas, o critério que se destaca diz respeito ao uso de um modelo fenomenológico na estratégia, classificando os métodos em diretos (esquemas baseados diretamente nas medições obtidas) e indiretos (esquemas que empregam modelos fenomenológicos) (ELLIS *et al.*, 1988; GARCIA e MORARI, 1981).

Não obstante, a essência das estratégias de otimização em tempo real consiste em usar as medidas da planta para compensar o efeito das incertezas no desempenho do processo. Dessa forma, outra proposta de classificação que surge naturalmente consiste em adotar como critério a forma com que os métodos usam as medidas para determinar as mudanças nas variáveis manipuladas (CHACHUAT *et al.*, 2009; FRANÇOIS e BONVIN, 2013a; ZHANG e FORBES, 2006). Nessa classificação, os métodos podem ser divididos em três categorias: i) esquemas de adaptação do modelo, em que as medidas são empregadas para atualizar o modelo do processo que é usado no problema de otimização; ii) esquemas de adaptação dos modificadores, em que a formulação do problema de otimização é alterada; iii) esquemas de adaptação direta, em que as variáveis manipuladas são modificadas diretamente por uma estratégia *feedback* apropriada.

Outra proposta de classificação enfatiza a formulação do problema, distinguindo os métodos que envolvem a resolução numérica de um problema de programação matemática como explícitos (representados pelos esquemas de adaptação do modelo e dos modificadores), e como implícitos aqueles métodos que reformulam o problema de otimização como um problema de controle e determinam mudanças nas variáveis de entrada de forma direta, no lugar de métodos numéricos de otimização (representados pelos esquemas de adaptação direta) (FRANÇOIS e BONVIN, 2014; FRANÇOIS *et al.*, 2012).

Há ainda outra classificação feita com base no uso do modelo na otimização (BONVIN *et al.*, 2001; MARCHETTI *et al.*, 2009). Adotando essa proposta de classificação, os métodos de RTO podem ser divididos em:

- (a) *Métodos com atualização de modelo* são caracterizados pela utilização de um modelo fenomenológico do processo e pelo uso das medidas da planta para atualizar esse modelo por meio de procedimentos de estimação de parâmetros. A atualização do modelo é sucedida pela otimização econômica. Fazem parte desse grupo a abordagem em duas etapas (CHEN e JOSEPH, 1987; JANG *et al.*, 1987), bem como o método ISOPE (do termo em inglês *Integrated System Optimization and Parameter Estimation*) e suas extensões (BRDYŚ e TATJEWSKI, 2005; ROBERTS, 1995).

(b) *Métodos sem atualização de modelo*: envolvem métodos com modelos fixos e métodos que não dependem de um modelo.

(b1) *Métodos de modelo fixo* empregam um modelo nominal do processo. De maneira similar aos métodos com atualização do modelo, esses métodos são iterativos e envolvem a resolução de um problema de programação matemática não linear. Todavia, diferem no fato de que as medidas são usadas para atualizar determinadas variáveis que modificam a função objetivo e as restrições do problema de otimização, melhorando a representação da planta na iteração mais recente. Nessa categoria encontram-se os métodos de adaptação das restrições (CHACHUAT *et al.*, 2008; FORBES e MARLIN, 1994), atualização de desvios e gradientes (GAO e ENGELL, 2005b), atualização dos modificadores (COSTELLO *et al.*, 2014; MARCHETTI, 2013; MARCHETTI *et al.*, 2009) e o método que implementa condições suficientes de viabilidade e otimalidade (do termo em inglês *sufficient conditions for feasibility and optimality* - SCFO) (BUNIN *et al.*, 2013b,c).

(b2) *Métodos livres de modelo* não dependem de um modelo para realizar a otimização *on-line*, mas envolvem a determinação de uma lei de controle que corresponde à solução do problema de otimização econômica. Fazem parte dessa classe os métodos de busca das restrições ativas (SRINIVASAN *et al.*, 2001), busca das condições necessárias de otimalidade (dos termos em inglês *necessary conditions of optimality* - NCO e *NCO tracking*) (FRANÇOIS *et al.*, 2005; SRINIVASAN e BONVIN, 2007), controle auto-otimizado (do termo em inglês *self-optimizing control* - SOC) (SKOGESTAD, 2000a,b) e controle extremal (do termo em inglês *extremum-seeking control*) (ARIYUR e KRSTIC, 2003; GUAY e ZHANG, 2003).

Comparações entre diferentes métodos RTO estão disponíveis na literatura por meio de estudos simulados (CHACHUAT *et al.*, 2009; ELLIS *et al.*, 1988; FRANÇOIS *et al.*, 2012; GAO e ENGELL, 2005a; MENDOZA *et al.*, 2016; YIP e MARLIN, 2004; ZHANG e FORBES, 2006).

Há ainda a proposta de integração entre otimização econômica e controle em uma única camada, que não é considerada nas classificações propostas mas que igualmente se enquadra como método RTO. O método EMPC foi proposto por pesquisadores brasileiros (DE GOUVÊA, 1997; DE GOUVÊA e ODLOAK, 1998) e enquadra-se no grupo de métodos de otimização por regulação.

2.3.1 Abordagem em Duas Etapas

Um dos primeiros métodos de otimização em tempo real, a abordagem em duas etapas pode ser aplicada também na otimização de sistemas dinâmicos. A abordagem em duas etapas envolve a resolução de dois problemas de programação matemática: um para a atualização do modelo (primeira etapa) e outro para a otimização econômica (segunda etapa), resolvidos separadamente. Em virtude dos avanços tecnológicos na área da computação, passou a ser usado em aplicações com modelos fenomenológicos, crescendo em interesse, popularidade e número de aplicações, tornando-se o padrão de implementações de sistemas RTO (CUTLER e PERRY, 1983; KLEINSCHRODT III e JONES, 1996).

A etapa de atualização do modelo pode ser entendida como uma recalibração do modelo estacionário, tendo como objetivo obter previsões mais acuradas do comportamento mais recente da planta. Tipicamente, essa etapa consiste em um problema de estimação de parâmetros, em que os valores do (sub)conjunto de parâmetros $\boldsymbol{\theta}$ do modelo são estimados por meio das observações \mathbf{y}_{ms} das variáveis de saída da planta. A segunda etapa corresponde ao problema de otimização econômica, que incorpora o modelo atualizado como uma restrição do problema de programação matemática. Uma vez que o sistema RTO atua toda vez que a planta atinge um estado estacionário, os dois problemas podem ser formulados da seguinte maneira para a iteração k (FRANÇOIS e BONVIN, 2013a):

$$\begin{aligned} \text{Identificação:} \quad \boldsymbol{\theta}_k^* := & \arg \min_{\boldsymbol{\theta}} J^{id} := \|\mathbf{y}_p(\mathbf{u}_k) - \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta})\| \\ & \text{sujeito a } \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^{n_\theta} \end{aligned} \quad (2.2)$$

$$\begin{aligned} \text{Otimização:} \quad \mathbf{u}_{k+1}^* := & \arg \min_{\mathbf{u}} J^{opt} := \phi[\mathbf{y}(\mathbf{u}, \boldsymbol{\theta}_k^*), \mathbf{u}] \\ & \text{sujeito a } \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}_k^*) \leq \mathbf{0} \end{aligned} \quad (2.3)$$

em que $J^{id}(\cdot)$ é a função objetivo do problema de identificação e estimação de parâmetros; $J^{opt}(\cdot)$ é a função objetivo do problema de otimização econômica, sendo nesse caso o índice de desempenho $\phi(\cdot)$; e $\boldsymbol{\theta}$ é o vetor de parâmetros incertos pertencente ao conjunto Θ de números reais contido no espaço de dimensão n_θ . No problema de estimação de parâmetros (2.2), \mathbf{u}_k é o conjunto de graus de liberdade aplicados à planta na iteração k do RTO e $\mathbf{y}_p(\mathbf{u}_k)$ são as medidas obtidas no estado estacionário. Os parâmetros $\boldsymbol{\theta}_k^*$ correspondem à solução do problema de estimação de parâmetros e são os valores utilizados pelo modelo para fornecer a previsão da planta na etapa de otimização seguinte.

O problema de otimização econômica (2.3) é formulado de acordo com a Equação (A.7) e, uma vez obtida a solução \mathbf{u}_{k+1}^* , os valores dos graus de liberdade da

planta aplicados no próximo instante (\mathbf{u}_{k+1}), podem corresponder diretamente à solução obtida,

$$\mathbf{u}_{k+1} = \mathbf{u}_{k+1}^* \quad ,$$

ou podem ser determinados a partir de relações funcionais convenientes, como um filtro exponencial de primeira ordem:

$$\mathbf{u}_{k+1} := (\mathbf{I} - \mathbf{K}) \mathbf{u}_k + \mathbf{K} \mathbf{u}_{k+1}^* \quad , \quad (2.4)$$

em que \mathbf{K} é uma matriz diagonal de ganhos com $K_{i,j} \in [0, 1] \forall i, j \in [1, \dots, n_u]$. O uso de filtros é motivado pelos riscos de que a aplicação direta de \mathbf{u}_{k+1}^* seja muito agressiva e conduza o processo a uma condição inviável. Nessa filtragem, também conhecida como coeficiente de relaxamento ou coeficiente de ganho (BRDYŠ e TATJEWSKI, 2005), o valor de $K = 1$ representa ausência de filtragem, enquanto $K = 0$ representa filtragem total.

A formulação apresentada nas Equações (2.2) e (2.3) é determinística e admite que toda a incerteza (incluindo as perturbações do processo) pode ser representada pelo conjunto de parâmetros $\boldsymbol{\theta}$, o que pode representar uma simplificação importante. Ademais, o problema de atualização do modelo é reduzido a um problema de estimação de parâmetros e a formulação apresentada na Equação (2.2) incorpora outras duas importantes simplificações: i) admite-se que não há incertezas nas medidas e que as variáveis de saída são perfeitamente conhecidas, de forma que $\mathbf{y}_{ms} = \mathbf{y}_p$; e ii) utiliza-se a distância euclidiana ou uma soma ponderada de desvios quadráticos entre os valores medidos e os valores preditos pelo modelo como critério de otimização (FRANÇOIS e BONVIN, 2013a).

Essas simplificações podem não ser adequadas à maioria das implementações reais de sistemas RTO, que incluem outros elementos no sistema em malha fechada, como procedimentos de reconciliação de dados e de diagnóstico dos resultados calculados (DARBY *et al.*, 2011; TRIERWEILER, 2014). Além disso, a resolução do problema de identificação pode apresentar problemas, caso não exista informação suficiente nos dados medidos. Por isso, alguns trabalhos se dedicaram a propor métodos capazes de identificar e garantir a existência de informação suficiente nos dados para a etapa de estimação de parâmetros, podendo envolver o uso de múltiplos conjuntos de dados (YIP e MARLIN, 2002) e procedimentos de planejamento de experimentos (PFAFF *et al.*, 2006; YIP e MARLIN, 2003).

2.3.1.1 Estabilidade

A estabilidade da abordagem em duas etapas, bem como a estabilidade local no ponto ótimo verdadeiro (representado aqui como \mathbf{u}_p^*), foram discutidas por QUE-

LHAS (2013) e FORBES e MARLIN (1996), respectivamente.

Um sistema RTO pode ser considerado como um sistema recursivo (ou iterativo), que opera em um conjunto de variáveis manipuladas \mathbf{u}_k , de maneira que a estabilidade do sistema RTO é determinada pelos seus componentes (etapas de estimação e otimização) e pela interação do sistema com o processo real. Assim, o sistema RTO pode ser representado como um mapeamento não linear sobre as variáveis manipuladas $\mathbf{u}_{k+1} = \Gamma(\mathbf{u}_k)$ e as condições de convergência podem ser formuladas a partir do Teorema de Ponto Fixo de Banach (ou Teorema da Contração) (KREYSZIG, 1989), enunciado como:

Teorema 1. (Teorema do Ponto Fixo de Banach) *Seja o espaço métrico completo (\mathbb{U}, d) com $\mathbb{U} \subset \mathbb{R}^{n_u}$ e seja $\Gamma : \mathbb{U} \rightarrow \mathbb{U}$ uma contração em \mathbb{U} . Assim, as seguintes condições são verificadas:*

- Γ admite um ponto fixo único em \mathbb{U} , dado por \mathbf{u}_∞ , i.e., $\Gamma(\mathbf{u}_\infty) = \mathbf{u}_\infty$;
- para qualquer $\mathbf{u}_0 \in \mathbb{U}$, a sequência definida por $a_0 = \mathbf{u}_0$ e $a_{k+1} = \Gamma(a_k)$ converge para \mathbf{u}_∞ ;
- a expressão

$$d(\mathbf{u}_\infty, \mathbf{u}_k) \leq \frac{\alpha^k}{1 - \alpha} d(\mathbf{u}_1, \mathbf{u}_0) \quad (2.5)$$

em que $\alpha \in \mathbb{R} \subset [0; 1)$, fornece uma estimativa de erro a priori válida para a iteração k .

Dessa forma, interpretando a estabilidade do sistema RTO como um caso particular do problema genérico do ponto fixo de uma sequência iterativa, as propriedades do teorema anterior podem ser usadas para prever a condição de estabilidade do sistema e também para avaliar a velocidade de convergência (QUELHAS, 2013).

A avaliação da estabilidade local da abordagem em duas etapas no ponto ótimo da planta segue uma abordagem similar e também passa por reconhecer os elementos do sistema como mapeamentos não lineares. Dessa maneira, é possível mostrar que o procedimento será estável no ponto ótimo verdadeiro da planta \mathbf{u}_p^* diante de perturbações pequenas e arbitrárias nas variáveis de entrada \mathbf{u} se (FORBES e MARLIN, 1996):

$$\left\| \frac{d\mathbf{u}^*}{d\boldsymbol{\theta}} \Big|_{\bar{\boldsymbol{\theta}}} \frac{d\boldsymbol{\theta}}{d\mathbf{y}_p} \Big|_{\mathbf{y}_p(\mathbf{u}_p^*)} \frac{d\mathbf{y}_p}{d\mathbf{u}} \Big|_{\mathbf{u}_p^*} \right\| < 1 \quad (2.6)$$

em que $\frac{d\mathbf{u}^*}{d\boldsymbol{\theta}} \Big|_{\bar{\boldsymbol{\theta}}}$ corresponde à sensibilidade paramétrica do problema de otimização baseada em modelo; $\frac{d\boldsymbol{\theta}}{d\mathbf{y}_p} \Big|_{\mathbf{y}_p(\mathbf{u}_p^*)}$ corresponde à sensibilidade dos parâmetros estimados com relação às variáveis medidas; e $\frac{d\mathbf{y}_p}{d\mathbf{u}} \Big|_{\mathbf{u}_p^*}$ corresponde à sensibilidade das variáveis medidas com relação às variáveis manipuladas. Deve ser enfatizado que esse critério

permite identificar se o método é capaz de retornar ao ótimo verdadeiro \mathbf{u}_p^* diante de desvios arbitrariamente pequenos do próprio ponto \mathbf{u}_p^* . Desse modo, o critério de estabilidade tem validade local, não garantindo a convergência do método para o ótimo verdadeiro após perturbações finitas e suficientemente distantes desse ponto.

2.3.1.2 Critério Local de Adequação do Modelo

A principal preocupação no estudo de estratégias RTO consiste na capacidade do método em determinar o ótimo verdadeiro do processo. Tendo encontrado esse ponto, a atenção se volta à velocidade com que a condição ótima de operação é alcançada a partir das condições iniciais, *i.e.*, a velocidade de convergência da técnica (ELLIS *et al.*, 1988). O efeito da seleção de modelos na abordagem em duas etapas (FORBES e MARLIN, 1996) e a convergência para o ótimo da planta em sistemas afetados por erros de modelagem (BIEGLER *et al.*, 1985; FORBES *et al.*, 1994) foram estudados.

Um modelo é considerado adequado para uso se for capaz de determinar o ponto ótimo verdadeiro da planta \mathbf{u}_p^* . Embora a etapa de estimação de parâmetros e a etapa de otimização econômica sejam conduzidas separadamente, a solução do problema de otimização depende dos parâmetros estimados, de maneira que essa interação deve ser considerada e analisada (ROBERTS, 1979).

A abordagem em duas etapas apresentará bom desempenho se i) houver pouco efeito de erros de modelagem, e ii) os dados medidos tiverem informação suficiente para a estimação dos parâmetros do modelo. Em outras palavras, se o modelo for estruturalmente correto e os parâmetros forem identificáveis, devem existir valores $\bar{\theta}$ para os parâmetros para os quais um ponto estacionário ótimo \mathbf{u}^* determinado pelo método coincida com o ponto ótimo verdadeiro \mathbf{u}_p^* . Nesse caso, os valores $\bar{\theta}$ podem não representar os valores verdadeiros dos parâmetros, especialmente na presença de erro estrutural, quando a ideia de “verdadeiro” perde significado (MARCHETTI *et al.*, 2009). Na presença de erros de modelagem pronunciados, não é possível determinar a convergência do sistema, ou em que ponto o algoritmo irá convergir. Nessa situação, o objetivo do problema de estimação de parâmetros pode estar totalmente desconectado do problema de otimização, de maneira que minimizar $J^{id}(\mathbf{y}_p)$ pode não contribuir em nada com as propriedades de viabilidade e otimalidade da solução obtida na otimização econômica (CHACHUAT *et al.*, 2009; FRANÇOIS e BONVIN, 2013a).

Para que o procedimento em duas etapas seja capaz de determinar o ótimo verdadeiro da planta, é necessário que o modelo seja capaz de prever com acurácia os gradientes das variáveis de saída da planta em relação às variáveis manipuladas (DURBECK, 1965, *apud* ROBERTS, 1978; FOORD, 1974, *apud* ROBERTS, 1978; ELLIS *et al.*, 1988), de forma que a adaptação permita a equivalência entre

as condições de Karush-Kuhn-Tucker (KKT) do modelo e da planta. Adicionalmente, considerando que a etapa de atualização do modelo seja um procedimento de estimação de parâmetros, a abordagem em duas etapas será capaz de determinar os valores ótimos (e desconhecidos) das variáveis de entrada \mathbf{u}_p^* se forem simultaneamente satisfeitas as condições suficientes de segunda ordem para os problemas de programação matemática de ambas as etapas (Equações 2.2 e 2.3) (FORBES e MARLIN, 1994; FORBES *et al.*, 1994). Essas condições, que receberam o nome de critério de adequabilidade de modelos, são (FORBES e MARLIN, 1996):

$$\frac{\partial J^{id}}{\partial \boldsymbol{\theta}} [\mathbf{y}_p(\mathbf{u}_p^*), \mathbf{y}(\mathbf{u}_p^*, \bar{\boldsymbol{\theta}})] = \mathbf{0} \quad (2.7)$$

$$\frac{\partial^2 J^{id}}{\partial \boldsymbol{\theta}^2} [\mathbf{y}_p(\mathbf{u}_p^*), \mathbf{y}(\mathbf{u}_p^*, \bar{\boldsymbol{\theta}})] > \mathbf{0} \quad (\text{positiva definida}) \quad (2.8)$$

$$G_i(\mathbf{u}_p^*, \bar{\boldsymbol{\theta}}) = 0, \quad i \in A(\mathbf{u}_p^*) \quad (2.9)$$

$$G_i(\mathbf{u}_p^*, \bar{\boldsymbol{\theta}}) < 0, \quad i \notin A(\mathbf{u}_p^*) \quad (2.10)$$

$$\nabla_r J^{opt}(\mathbf{u}_p^*, \bar{\boldsymbol{\theta}}) = \mathbf{0} \quad (2.11)$$

$$\nabla_r^2 J^{opt}(\mathbf{u}_p^*, \bar{\boldsymbol{\theta}}) > \mathbf{0} \quad (\text{positiva definida}) \quad (2.12)$$

em que $A(\mathbf{u}_p^*)$ representa o conjunto ativo, $\nabla_r J^{opt}$ e $\nabla_r^2 J^{opt}$ são o gradiente reduzido e a matriz Hessiana reduzida⁵ da função objetivo da otimização econômica, respectivamente.

As condições apresentadas, envolvendo tanto relações de igualdade como de desigualdade que dependem de $\bar{\boldsymbol{\theta}}$, evidenciam a relação de dependência entre os dois problemas (2.2) e (2.3). Como \mathbf{u}_p^* é um ponto fixo correspondente ao ótimo da planta, não pode ser considerado grau de liberdade.

As condições (2.7) e (2.8) são suficientes para que $\bar{\boldsymbol{\theta}}$ seja um mínimo local do problema de estimação de parâmetros em \mathbf{u}_p^* , enquanto as condições (2.9) a (2.12) representam condições suficientes para que \mathbf{u}_p^* seja um mínimo local do problema de otimização econômica com os parâmetros $\bar{\boldsymbol{\theta}}$. Se todas essas condições forem satisfeitas, o ótimo verdadeiro da planta \mathbf{u}_p^* corresponderá ao ótimo local fornecido pelo modelo para $\boldsymbol{\theta} = \bar{\boldsymbol{\theta}}$ e as condições (2.7) a (2.12) são suficientes para a adequabilidade do modelo. Todavia, essas condições não são necessárias. Quando a Hessiana reduzida é semidefinida, pode haver a situação em que \mathbf{u}_p^* corresponde ao ótimo do modelo, mas a condição (2.12) não é verificada. Além disso, é importante observar que a igualdade imposta pela condição (2.7) consome todos os n_θ graus de liberdade, tornando as condições de adequabilidade um conjunto sobredeterminado, que dificilmente será satisfeito na presença de erros de modelagem. Em outras palavras,

⁵ A matriz Hessiana reduzida é definida como $\nabla_r^2 J^{opt} = \mathbf{Z}^T (\partial^2 \mathcal{L} / \partial \mathbf{u}^2) \mathbf{Z}$, em que \mathbf{Z} representa o espaço nulo da matriz Jacobiana das restrições ativas e \mathcal{L} representa a função de Lagrange do problema de otimização (FRANÇOIS e BONVIN, 2013a).

convergir para $\bar{\theta}$ não garante a obtenção de \mathbf{u}_p^* (FRANÇOIS, 2014; MARCHETTI *et al.*, 2009).

Apesar de fornecer importantes relações teóricas, há muitas limitações para o uso do critério de adequabilidade. O critério foi concebido para verificar se a otimização econômica oferece condições para determinar \mathbf{u}_p^* para o conjunto de parâmetros estimados, mas nada se pode inferir sobre a influência desse conjunto no procedimento de estimação. Além disso, a validade é local em torno do ponto ótimo verdadeiro \mathbf{u}_p^* , não havendo garantia que outros pontos ótimos referentes a diferentes condições operacionais possam ser encontrados. Por fim, além de dificuldades no cálculo do gradiente e da Hessiana reduzidos, \mathbf{u}_p^* é desconhecido e, portanto, o critério de adequabilidade de modelo não pode ser usado para selecionar modelos, tampouco para apontar quais parâmetros estimar (QUELHAS, 2013).

Em resumo, a abordagem em duas etapas é projetada para lidar apenas com erros paramétricos (GAO e ENGELL, 2005b); a presença de erros estruturais de modelagem geralmente impedem que uma predição das variáveis de saída seja suficiente para que as condições necessárias de otimalidade da planta (NCO) sejam satisfeitas (MARCHETTI *et al.*, 2009). Conforme mostram diversos trabalhos na literatura, a etapa de estimação de parâmetros na presença de erros estruturais pode ser completamente ineficiente para a determinação do ótimo verdadeiro, podendo ser prejudicial e conduzir o processo a uma condição de pior desempenho que o mesmo processo sem intervenção do sistema RTO (AGARWAL, 1997; GAO e ENGELL, 2005a; MARCHETTI, 2009; QUELHAS *et al.*, 2013). Assim, além dos ajustes feitos com base nas respostas do modelo e da planta, ficou demonstrada a necessidade de introduzir ajustes com base nas NCOs do modelo e da planta, incorporando informação sobre gradientes no método RTO (BIEGLER *et al.*, 1985; FORBES *et al.*, 1994; ROBERTS, 1978).

2.3.2 Método ISOPE e Suas Extensões

Em resposta às deficiências da abordagem em duas etapas, notadamente com relação à falta de ajuste do modelo para também reproduzir as condições de KKT da planta, uma abordagem em duas etapas modificada foi proposta, recebendo mais tarde o nome de sistema integrado de otimização e estimação de parâmetros (do termo em inglês *Integrated System Optimization and Parameter Estimation* - ISOPE) (ROBERTS, 1979, 1995; ROBERTS e WILLIAMS, 1981). O método ISOPE mantém a estrutura em duas etapas, mas utiliza a interação entre os problemas de estimação de parâmetros e otimização econômica para incorporar informação sobre os gradientes da planta, *i.e.*, informação sobre o gradiente das variáveis de saída em relação às variáveis manipuladas. Essa informação é introduzida na função objetivo econô-

mica na forma de uma constante, deduzida a partir das condições de otimalidade de primeira ordem para a planta (condições A.8 a A.11), garantindo que o ponto de convergência do algoritmo corresponda a um ponto KKT da planta.

O método ISOPE originalmente proposto não considera restrições que dependam das variáveis de saída \mathbf{y} e pode ser considerado como uma simples modificação da abordagem em duas etapas, em que o problema de estimação de parâmetros permanece inalterado e a função objetivo do problema de otimização econômica contém termos extras $\boldsymbol{\lambda}_k^\top \mathbf{u}$ (ROBERTS, 1979). Além disso, admite-se que o número de variáveis medidas e de parâmetros do modelo seja o mesmo, o que permite determinar parâmetros $\boldsymbol{\theta}_k$ para os quais as respostas do modelo sejam idênticas aos valores medidos, *i.e.*,

$$\mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta}) = \mathbf{y}_p(\mathbf{u}_k) \rightarrow \boldsymbol{\theta}_k \quad . \quad (2.13)$$

Essa hipótese representa uma condição de qualificação do modelo (ZHANG e ROBERTS, 1991), sendo encontrada em toda a literatura relacionada ao método (BRDYŚ e TATJEWSKI, 2005; ROBERTS, 1995). A otimização econômica modificada segue a seguinte formulação:

$$\begin{aligned} \mathbf{u}_{k+1}^* := \arg \min_{\mathbf{u}} \quad & J^{opt} := \phi[\mathbf{u}, \mathbf{y}(\mathbf{u}, \boldsymbol{\theta}_k)] + \boldsymbol{\lambda}_k^\top \mathbf{u} \\ \text{sujeito a} \quad & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned} \quad , \quad (2.14)$$

em que a variável $\boldsymbol{\lambda}_k$ é denominada modificador. Dessa maneira, admitindo que os gradientes experimentais estejam disponíveis, *i.e.*, que seja possível estimar $\partial \mathbf{y}_p(\mathbf{u}_k) / \partial \mathbf{u}$, o modificador $\boldsymbol{\lambda}_k$ é calculado como (MARCHETTI *et al.*, 2009):

$$\boldsymbol{\lambda}_k^\top := \frac{\partial \phi}{\partial \mathbf{y}}[\mathbf{u}_k, \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta}_k)] \left[\frac{\partial \mathbf{y}_p(\mathbf{u}_k)}{\partial \mathbf{u}} - \frac{\partial \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta}_k)}{\partial \mathbf{u}} \right] . \quad (2.15)$$

Adicionalmente, um filtro exponencial de primeira ordem, similar ao da Equação (2.4), pode ser usado para determinar o próximo ponto de operação \mathbf{u}_{k+1} .

Estudos posteriores dedicaram-se a análises de convergência e otimalidade do método (BRDYŚ e ROBERTS, 1987; BRDYŚ *et al.*, 1987), bem como a estender sua formulação para considerar restrições de desigualdade envolvendo as variáveis de saída do processo (BRDYŚ *et al.*, 1986; ZHANG e ROBERTS, 1991).

Embora o método ISOPE seja capaz de determinar o ponto ótimo verdadeiro da planta mesmo na presença de erros estruturais de modelagem, sua utilização somente será possível se os gradientes experimentais da planta puderem ser estimados, o que constitui a principal desvantagem da técnica e uma das maiores dificuldades da sua implementação prática. Por isso, a literatura também dedica bastante esforço ao desenvolvimento e comparação de métodos para a determinação *on-line* das derivadas do processo (BUNIN *et al.*, 2013e; FRANÇOIS *et al.*, 2012; GAO e ENGELL,

2005b; MANSOUR e ELLIS, 2003; ZHANG e FORBES, 2006).

A técnica originalmente proposta no método ISOPE para a determinação dos gradientes experimentais envolve a aplicação de métodos de perturbação estacionários, em que pequenas perturbações são realizadas na planta e os dados estacionários obtidos são usados para o cálculo numérico das derivadas (BRDYŚ *et al.*, 1987; ROBERTS, 1979). Nesse caso, a estratégia mais direta consiste em estimar os gradientes aplicando técnicas de diferenças finitas diretamente na planta, nas quais cada uma das n_u variáveis manipuladas deve ser perturbada em torno do ponto operacional. Por exemplo, aplicando a técnica de diferenças finitas progressivas, a derivada da variável de resposta y_j com relação a i -ésima variável manipulada u_i (na iteração k do RTO) é estimada como (COSTELLO, 2015):

$$\left(\frac{\partial y_{p,j}}{\partial u_i} \right)_k = \frac{y_{p,j}(\mathbf{u}_k + \delta \mathbf{u}_i) - y_{p,j}(\mathbf{u}_k)}{\|\delta \mathbf{u}_i\|}, \quad (2.16)$$

em que $\delta \mathbf{u}_i$ é um vetor alinhado com a i -ésima variável de entrada.

No entanto, essas técnicas são custosas, visto que o estado estacionário deve ser atingido a cada perturbação, principalmente em problemas com muitas variáveis (assim como ocorre nos métodos de busca direta discutidos anteriormente). Essa dificuldade motivou o surgimento de propostas que envolvem versões com informações estatísticas e com identificação de modelos dinâmicos. Esse último teve como objetivo incorporar o procedimento de BAMBERGER e ISERMANN (1978) ao algoritmo ISOPE, de forma que um modelo dinâmico identificado é usado para estimar as derivadas no regime estacionário (ZHANG e ROBERTS, 1990). Não obstante, essas propostas também se mostraram difíceis de implementar, em virtude de limitações relacionadas à presença de ruído e à acurácia das informações obtidas a partir das medidas transientes.

Avanços significativos para a estimação dos gradientes experimentais foram alcançados com o trabalho de BRDYŚ e TATJEWSKI (1994), por meio da proposta do método *ISOPE dual*. Os autores propuseram uma implementação diferente da técnica de diferenças finitas, estimando os gradientes sem causar perturbações adicionais. No método ISOPE dual, o problema de otimização econômica incorpora uma restrição para a região de busca do próximo ponto de operação como estratégia para garantir a existência de perturbação suficiente nas variáveis medidas, para a estimação dos gradientes na iteração seguinte, o que resulta em uma estrutura *dual* (BRDYŚ e TATJEWSKI, 2005; TADEJ e TATJEWSKI, 2001). Desse modo, a geração de perturbações para a determinação dos gradientes experimentais ocorre simultaneamente à atualização do ponto de operação e não se baseia na perturbação das variáveis, mas em dados estacionários anteriores em conjunto com o dado mais recente.

Matematicamente, dado o ponto de operação \mathbf{u}_k , o número n_u de pontos operacionais anteriores $\mathbf{u}_{k-n_u}, \dots, \mathbf{u}_{k-1}$, e os valores correspondentes de medidas da planta $\mathbf{y}_p(\mathbf{u}_{k-n_u}), \dots, \mathbf{y}_p(\mathbf{u}_{k-1}), \mathbf{y}_p(\mathbf{u}_k)$, uma estimativa $\hat{\boldsymbol{\beta}}_k$ do gradiente $\partial \mathbf{y}_p(\mathbf{u}_k) / \partial \mathbf{u}$ pode ser determinada por (BRDYŚ e TATJEWSKI, 2005):

$$\frac{\partial \mathbf{y}_p(\mathbf{u}_k)}{\partial \mathbf{u}} \approx \hat{\boldsymbol{\beta}}_k := \mathbf{Y}_p(\mathbf{u}_k) \mathbf{U}^{-1}(\mathbf{u}_k) \quad (2.17)$$

com

$$\mathbf{U}(\mathbf{u}_k) := [\mathbf{u}_k - \mathbf{u}_{k-1} \ \dots \ \mathbf{u}_k - \mathbf{u}_{k-n_u}]^\top \in \mathbb{R}^{n_u \times n_u} \quad (2.18)$$

$$\mathbf{Y}(\mathbf{u}_k) := [\mathbf{y}_p(\mathbf{u}_k) - \mathbf{y}_p(\mathbf{u}_{k-1}) \ \dots \ \mathbf{y}_p(\mathbf{u}_k) - \mathbf{y}_p(\mathbf{u}_{k-n_u})]^\top \in \mathbb{R}^{n_y \times n_u} \quad (2.19)$$

em que a Equação (2.17) exige a não singularidade da matriz $\mathbf{U}(\mathbf{u}_k)$. Adicionalmente, a introdução de um limite superior $\kappa(\mathbf{u})$ ao número de condicionamento de $\mathbf{U}(\mathbf{u}_k)$ auxilia o bom desempenho do método, pois força que a inclusão do próximo ponto operacional não resulte em mau condicionamento da matriz na iteração seguinte (MARCHETTI *et al.*, 2010):

$$\kappa(\mathbf{u}) := \frac{\varrho_{max}[\mathbf{U}(\mathbf{u})]}{\varrho_{min}[\mathbf{U}(\mathbf{u})]} \leq \delta \quad (2.20)$$

em que $\delta \in (0; 1)$ e ϱ_{max} e ϱ_{min} denotam o maior e o menor valor singular de $\mathbf{U}(\mathbf{u})$, respectivamente. Assim, a expressão para o cálculo do modificador no método ISOPE dual passa a ser (MARCHETTI, 2009):

$$\boldsymbol{\lambda}_k^\top := \frac{\partial \phi}{\partial \mathbf{y}} [\mathbf{u}_k, \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta}_k)] \left[\hat{\boldsymbol{\beta}}_k - \frac{\partial \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta}_k)}{\partial \mathbf{u}} \right], \quad (2.21)$$

em conjunto com a seguinte formulação para o problema de otimização econômica:

$$\begin{aligned} \mathbf{u}_{k+1}^* := \arg \min_{\mathbf{u}} \quad & J^{opt} := \phi[\mathbf{u}, \mathbf{y}(\mathbf{u}, \boldsymbol{\theta}_k)] + \boldsymbol{\lambda}_k^\top \mathbf{u} \\ \text{sujeito a} \quad & \kappa(\mathbf{u}_{k-n_u+1}, \dots, \mathbf{u}_k, \bar{\mathbf{u}}) \leq \delta \\ & \bar{\mathbf{u}} = (\mathbf{I} - K) \mathbf{u}_k + K \mathbf{u} \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned}, \quad (2.22)$$

em que $\bar{\mathbf{u}}$ corresponde ao valor filtrado das variáveis manipuladas.

A Equação (2.22) pode adicionalmente incorporar na função objetivo um termo de regularização, $\rho \|\mathbf{u}_k - \mathbf{u}\|^2$ com $\rho > 0$ (BRDYŚ e TATJEWSKI, 2005), que confere propriedades de otimalidade ao método, mesmo para funções objetivo não convexas. O método ISOPE com esse termo de regularização foi chamado de ISOPE aumentado (BRDYŚ *et al.*, 1987).

Uma última contribuição ao método ISOPE que merece destaque foi proposta por

TATJEWSKI (2002), ao mostrar que a condição de qualificação do modelo (Equação 2.13) pode ser satisfeita sem que haja atualização dos parâmetros do modelo. A eliminação da etapa de estimação dos parâmetros é possível com a introdução de um termo de deslocamento $\mathbf{a}_k := \mathbf{y}_p(\mathbf{u}_k) - \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta})$ no problema de otimização econômica modificado, rescrevendo a Equação (2.14) como (MARCHETTI *et al.*, 2009):

$$\mathbf{u}_{k+1}^* := \underset{\mathbf{u}}{\arg \min} \quad J^{opt} := \phi[\mathbf{u}, \mathbf{y}(\mathbf{u}, \boldsymbol{\theta}) + \mathbf{a}_k] + \boldsymbol{\lambda}_k^\top \mathbf{u} \quad , \quad (2.23)$$

sujeito a $\mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U$

com

$$\boldsymbol{\lambda}_k^\top := \frac{\partial \phi}{\partial \mathbf{y}}[\mathbf{u}_k, \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta}) + \mathbf{a}_k] \left[\frac{\partial \mathbf{y}_p(\mathbf{u}_k)}{\partial \mathbf{u}} - \frac{\partial \mathbf{y}(\mathbf{u}_k, \boldsymbol{\theta})}{\partial \mathbf{u}} \right] \quad . \quad (2.24)$$

Por remover a necessidade de ajuste dos parâmetros, essa formulação, que também foi posteriormente usada por GAO e ENGELL (2005b), deixa de ser bem representada pelo nome ISOPE.

2.3.3 Métodos de Adaptação dos Modificadores

No estudo de sistemas de controle, o uso de modelos matemáticos de processo contribui com a determinação de ações de controle mais apropriadas, notadamente por permitir que o efeito das restrições seja antecipado (BROSILOW e ZHAO, 1988). Se as restrições dependem do comportamento da planta, não é possível admitir que seus valores sejam conhecidos com acurácia e o auxílio do modelo para estimar tais valores não garante que essas restrições sejam satisfeitas (GAO e ENGELL, 2005b). Por isso, diversos sistemas de controle baseados em modelo foram desenvolvidos, incorporando desvios calculados como a diferença entre medições da planta e predições do modelo, para considerar o efeito das perturbações (CUTLER e RAMAKER, 1980; GARCIA e MORARI, 1982; RICHALET *et al.*, 1978). Entre as vantagens observadas, o uso desses desvios confere importantes propriedades aos algoritmos de controle discretos aplicados a sistemas lineares, como ausência de *offset* e robustez em relação aos erros de modelagem (GARCIA e MORARI, 1985). Desse modo, houve interesse em incorporar o método de atualização dos desvios nos métodos RTO.

FORBES e MARLIN (1994) foram os primeiros a propor a inserção de termos de desvio em um método RTO, e mais especificamente, termos de desvio para ajuste das restrições. A abordagem de atualização dos desvios (do termo em inglês *bias-update approach*) baseia-se na utilização das medidas para atualizar diretamente as restrições do problema de programação matemática, de tal modo que as restrições são deslocadas com base nos termos de desvios e se aproximam dos valores medidos na iteração mais recente. De modo similar à abordagem em duas etapas, o modelo

de processo faz parte do problema NLP que é resolvido repetidamente, sendo usado no procedimento de otimização para estimar valores das restrições. Os termos de desvio equivalem, na prática, a fatores de correção (ou termos modificadores), sendo obtidos a partir da diferença entre valores medidos e preditos pelo modelo para as restrições. Assim, os parâmetros do modelo permanecem fixos, de maneira que os parâmetros ajustáveis no método correspondem aos termos de desvio. As restrições consideradas envolvem tanto restrições de igualdade, que correspondem ao modelo do processo, quanto a restrições de desigualdade, que correspondem tipicamente a limites de operação.

Embora tenham restringido seu estudo aos problemas com número equivalente de variáveis de decisão e restrições ativas, FORBES e MARLIN (1994) aplicaram o critério local de adequabilidade de modelo de FORBES *et al.* (1994) ao esquema de atualização dos desvios, mostrando que a adequabilidade do modelo é determinada pelos gradientes da função objetivo e das restrições ativas estimados por meio do modelo. Assim, o uso do teste local de adequabilidade permite determinar a faixa máxima admissível para os erros nos parâmetros (fixos) que garante a convergência do método para o ponto ótimo verdadeiro.

Posteriormente, CHACHUAT *et al.* (2008) estenderam a análise da abordagem de atualização dos desvios, fornecendo justificativas teóricas para uso desse método em casos em que há efeito moderado de incertezas, discutindo também outros aspectos, como detecção das restrições ativas e questões de convergência. O método, chamado de esquema de adaptação das restrições (do termo em inglês *constraint-adaptation scheme*), apresenta rápida convergência para soluções localizadas nas restrições do problema de otimização, requerendo poucas iterações do sistema RTO. Além disso, a abordagem não envolve o uso de gradientes experimentais da função objetivo e restrições, tornando o método menos sensível aos ruídos de medição, quando comparado a outros métodos de otimização em tempo real. Embora o uso de um modelo fixo evite a necessidade de atualização do modelo (ao contrário da abordagem em duas etapas) e o método não necessite estimativas dos gradientes da planta (ao contrário do método ISOPE), a qualidade do modelo pode ocasionar convergência a pontos subótimos. Mesmo com a perda de otimalidade, contudo, o resultado sempre será um ponto viável (CHACHUAT *et al.*, 2008).

Outra proposta significativa foi apresentada por GAO e ENGELL (2005b). Com base no método ISOPE com modelo fixo de TATJEWSKI (2002), os autores propuseram um novo método de lidar com restrições dependentes do processo, em que, juntamente com os termos de correção das restrições usados no método de atualização dos desvios (CHACHUAT *et al.*, 2008; FORBES e MARLIN, 1994), termos de correção dos gradientes são incluídos nas restrições. Desse modo, não apenas os valores das restrições preditos pelo modelo coincidem com os valores medidos, mas

também os gradientes das restrições. Tais termos de correção dos gradientes são determinados pela diferença entre valores experimentalmente estimados e preditos para os gradientes das restrições, similares aos termos no método ISOPE, inclusive na necessidade de prover estimativas para os gradientes da função objetivo e das restrições, a partir das medidas da planta. Assim, a convergência do algoritmo resulta em um ponto que satisfaz as NCOs da planta. Além disso, GAO e ENGELL (2005b) propuseram um novo método de estimação de gradientes baseado na técnica de BRDYŚ e TATJEWSKI (1994).

Essa habilidade de usar estimativas das NCO da planta para determinar o ponto ótimo verdadeiro, mesmo na presença de erros estruturais de modelagem e perturbações não medidas, motivou o surgimento de uma metodologia única. A adaptação direta do problema de otimização por termos modificadores foi generalizada e formalizada em uma abordagem única no trabalho de MARCHETTI *et al.* (2009) como método de adaptação dos modificadores. O trabalho também analisou condições de convergência local e adequabilidade do modelo para a abordagem, bem como estratégias alternativas de filtragem e modificação. Adicionalmente, foram apresentadas as conexões com as abordagens anteriores, notadamente com o método ISOPE e com a abordagem de GAO e ENGELL (2005b), verificando que o método de adaptação dos modificadores pode ser apresentado como uma generalização do método ISOPE.

O método de adaptação dos modificadores (do termo em inglês *modifier-adaptation scheme* - MA) conserva a importante característica de usar uma parametrização e um critério de atualização que são concebidos para satisfazer as condições necessárias de otimalidade. Expressando a diferença entre as NCO medidas (ou estimadas) da planta e as NCO preditas pelo modelo, os termos modificadores são adaptados a cada iteração, modificando as restrições e função objetivo do problema de otimização (MARCHETTI *et al.*, 2009). Além disso, o uso de um modelo de processo fixo ao longo das iterações do sistema RTO confere uma característica pragmática ao procedimento, na medida em que há um desacoplamento com o possível sentido físico associado à adaptação dos modificadores (QUELHAS, 2013).

No contexto do método de adaptação dos modificadores, o conjunto de variáveis manipuladas \mathbf{u}_k é aplicado ao processo na iteração k e, tendo a planta alcançado o regime estacionário, comparam-se os valores medidos da função objetivo, restrições e seus respectivos gradientes com os valores preditos pelo modelo, que irão modificar o problema de otimização para determinação do ponto ótimo \mathbf{u}_{k+1}^* na iteração seguinte. Admitindo que as variáveis de resposta $\mathbf{y}(\mathbf{u}, \boldsymbol{\theta})$ possam ser representadas explicitamente em termos de \mathbf{u} e $\boldsymbol{\theta}$, o problema de otimização é formulado como

(FRANÇOIS, 2014):

$$\begin{aligned} \mathbf{u}_{k+1}^* := \arg \min_{\mathbf{u}} \quad & J^{opt}(\mathbf{u}, \boldsymbol{\theta}) := \phi(\mathbf{u}, \boldsymbol{\theta}) + \epsilon_k^\phi + \boldsymbol{\lambda}_k^\phi (\mathbf{u} - \mathbf{u}_k) \\ \text{sujeito a} \quad & \mathbf{G}^m(\mathbf{u}, \boldsymbol{\theta}) := \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) + \epsilon_k^G + \boldsymbol{\lambda}_k^G (\mathbf{u} - \mathbf{u}_k) \leq \mathbf{0} \end{aligned}, \quad (2.25)$$

com

$$\epsilon_k^\phi := \phi_p(\mathbf{u}_k) - \phi(\mathbf{u}_k, \boldsymbol{\theta}) \in \mathbb{R} \quad (2.26)$$

$$\epsilon_k^G := \mathbf{G}_p(\mathbf{u}_k) - \mathbf{G}(\mathbf{u}_k, \boldsymbol{\theta}) \in \mathbb{R}^{n_g} \quad (2.27)$$

$$\boldsymbol{\lambda}_k^\phi := \frac{\partial \phi_p}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial \phi}{\partial \mathbf{u}}(\mathbf{u}_k, \boldsymbol{\theta}) \in \mathbb{R}^{n_u} \quad (2.28)$$

$$\boldsymbol{\lambda}_k^G := \frac{\partial \mathbf{G}_p}{\partial \mathbf{u}}(\mathbf{u}_k) - \frac{\partial \mathbf{G}}{\partial \mathbf{u}}(\mathbf{u}_k, \boldsymbol{\theta}) \in \mathbb{R}^{n_u \times n_g}, \quad (2.29)$$

em que (CHACHUAT *et al.*, 2009):

- O escalar ϵ_k^ϕ representa o modificador de ordem zero para o valor da função objetivo. Com exceção dos casos em que se deseja melhorar a predição do valor da função objetivo, esse modificador é geralmente ignorado, pois não altera o valor mínimo da função.
- O vetor linha ϵ_k^G representa o modificador de ordem zero para os valores das restrições, sendo determinado pela diferença entre valores experimentais $\mathbf{G}_p(\mathbf{u}_k)$ e preditos $\mathbf{G}(\mathbf{u}_k, \boldsymbol{\theta})$ das restrições. Considerar apenas ϵ_k^G na adaptação da Equação (2.25) consiste no método de adaptação das restrições (CHACHUAT *et al.*, 2008; FORBES e MARLIN, 1994).
- O vetor linha $\boldsymbol{\lambda}_k^\phi$ representa o modificador de primeira ordem para os gradientes da função objetivo, sendo determinado pela diferença entre valores experimentais $(\partial \phi_p / \partial \mathbf{u})|_{\mathbf{u}_k}$ e preditos $(\partial \phi / \partial \mathbf{u})|_{\mathbf{u}_k}$ dos gradientes da função objetivo. A adaptação da função objetivo por meio de um modificador calculado com informações de gradiente equivale à proposta de ROBERTS (1979), que resultou no método ISOPE, embora ainda estivesse inserido no contexto da abordagem em duas etapas. Foi apenas com o trabalho de TATJEWSKI (2002) que a etapa de estimação de parâmetros foi eliminada do método.
- A matriz $\boldsymbol{\lambda}_k^G$ representa o modificador de primeira ordem para os gradientes das restrições, sendo determinado pela diferença entre valores experimentais $(\partial \mathbf{G}_p / \partial \mathbf{u})|_{\mathbf{u}_k}$ e preditos $(\partial \mathbf{G} / \partial \mathbf{u})|_{\mathbf{u}_k}$ dos gradientes das restrições. A estratégia de incluir termos modificadores de primeira ordem para as funções de desigualdade dependentes do processo, em conjunto com o termo de desvio das restrições ϵ_k^G , foi proposta por GAO e ENGELL (2005b). Desse modo, é

possível obter correspondência entre valores medidos e preditos, tanto para as restrições, quanto para seus gradientes.

É conveniente representar os gradientes e modificadores como vetores de dimensão $n_c = (n_u + 1)(n_g + 1)$ na forma

$$\begin{aligned} \mathbf{\Lambda} &:= \left[\boldsymbol{\epsilon}^\phi, \boldsymbol{\epsilon}^{G_1}, \dots, \boldsymbol{\epsilon}^{G_{n_g}}, \boldsymbol{\lambda}^{\phi^\top}, \boldsymbol{\lambda}^{G_1^\top}, \dots, \boldsymbol{\lambda}^{G_{n_g}^\top} \right]^\top \\ \mathbf{C} &:= \left[\phi, \mathbf{G}_1, \dots, \mathbf{G}_{n_g}, \frac{\partial \phi}{\partial \mathbf{u}}, \frac{\partial \mathbf{G}_1}{\partial \mathbf{u}}, \dots, \frac{\partial \mathbf{G}_{n_g}}{\partial \mathbf{u}} \right]^\top \end{aligned} .$$

Assim, as Equações (2.26) a (2.29) podem ser rescritas como

$$\mathbf{\Lambda}_k = \mathbf{C}_p(\mathbf{u}_k) - \mathbf{C}(\mathbf{u}_k, \boldsymbol{\theta}) \quad . \quad (2.30)$$

Além disso, as variáveis de decisão implementadas na próxima iteração correspondem diretamente à solução da Equação (2.25), *i.e.*, $\mathbf{u}_{k+1} := \mathbf{u}_{k+1}^*$.

No método MA, os parâmetros $\boldsymbol{\theta}$ permanecem constantes e as variáveis de ajuste no método passam a ser os termos modificadores $\mathbf{\Lambda}$. A estratégia de adaptação mais simples equivale a implementar a modificação total representada por $\mathbf{\Lambda}_k$. Como isso pode ocasionar modificações excessivas e aumentar a sensibilidade do método a ruídos de medida, uma estratégia mais adequada consiste em filtrar o valor dos modificadores com um filtro exponencial de primeira ordem:

$$\mathbf{\Lambda}_k = (\mathbf{I} - K)\mathbf{\Lambda}_{k-1} + K\mathbf{\Lambda}_k \quad , \quad (2.31)$$

em que $K \in \mathbb{R}^{n_c \times n_c}$ é uma matriz de ganhos. A matriz de ganhos K pode ser encarada como um novo grau de liberdade e utilizada para forçar a convergência do método (FRANÇOIS, 2014).

O vetor de grandezas \mathbf{C}_p , determinado a partir de medidas experimentais, está diretamente relacionado às condições de KKT, sendo na prática determinado a partir das respostas do processo \mathbf{y}_p . Os valores experimentais das restrições $\mathbf{G}_p(\mathbf{u})$ são diretamente estimados de $\mathbf{g}(\mathbf{u}, \mathbf{y}_p)$, que força a viabilidade das soluções, mas não garante a determinação correta do conjunto ativo (CHACHUAT *et al.*, 2009). A determinação dos gradientes experimentais da função objetivo e restrições, $(\partial \phi_p / \partial \mathbf{u})|_{\mathbf{u}_k}$ e $(\partial \mathbf{G}_p / \partial \mathbf{u})|_{\mathbf{u}_k}$, necessários para a implementação dos modificadores, constituem a principal dificuldade para a implementação do método. Tais valores podem ser inferidos a partir das medidas das repostas da planta $\mathbf{y}_p(\mathbf{u})$ e seus gradientes estimados

$(\partial \mathbf{y}_p / \partial \mathbf{u})|_{\mathbf{u}_k}$ como (MARCHETTI *et al.*, 2009):

$$\begin{aligned} \left. \frac{\partial \phi_p(\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{u}_k} &= \frac{\partial \phi}{\partial \mathbf{u}} [\mathbf{u}_k, \mathbf{y}_p(\mathbf{u}_k)] + \frac{\partial \phi}{\partial \mathbf{y}} [\mathbf{u}_k, \mathbf{y}_p(\mathbf{u}_k)] \frac{\partial \mathbf{y}_p(\mathbf{u}_k)}{\partial \mathbf{u}} \\ \left. \frac{\partial \mathbf{G}_p(\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{u}_k} &= \frac{\partial \mathbf{g}}{\partial \mathbf{u}} [\mathbf{u}_k, \mathbf{y}_p(\mathbf{u}_k)] + \frac{\partial \mathbf{g}}{\partial \mathbf{y}} [\mathbf{u}_k, \mathbf{y}_p(\mathbf{u}_k)] \frac{\partial \mathbf{y}_p(\mathbf{u}_k)}{\partial \mathbf{u}} \quad . \end{aligned}$$

Assim como na abordagem em duas etapas, o modelo usado no método MA deve possuir certas propriedades para ser utilizado. Um modelo será considerado adequado para o método MA se existirem valores dos modificadores $\bar{\mathbf{\Lambda}}$ para os quais o algoritmo seja capaz de convergir para o ótimo da planta \mathbf{u}_p^* . Ao contrário das condições obtidas para a abordagem em duas etapas, o critério de adequabilidade para o método MA é mais simples. Como o modelo é fixo e não existe a etapa de estimação de parâmetros, as condições (2.7) e (2.8) não são necessárias. Por outro lado, em virtude da definição dos modificadores, as condições (2.9) a (2.11) ficam automaticamente satisfeitas. Dessa maneira, a única condição suficiente restante para a adequabilidade do modelo é a equação equivalente à condição (2.12), *i.e.*,

$$\nabla_r^2 J^{opt}(\mathbf{u}_p^*, \boldsymbol{\theta}) > \mathbf{0} \quad .$$

Essa condição indica que, se a matriz Hessiana reduzida da função objetivo $J^{opt}(\cdot)$ for positiva definida em \mathbf{u}_p^* , então estará garantida a existência de valores dos modificadores, de modo que \mathbf{u}_p^* seja solução do problema de otimização modificado (2.25) (MARCHETTI *et al.*, 2009). Embora a condição de adequabilidade para o método MA seja mais facilmente satisfeita que no método em duas etapas, a adequabilidade é uma propriedade do modelo que não pode ser verificada *a priori*, uma vez que sua verificação exige o conhecimento do conjunto de restrições ativas e, especialmente, do ótimo da planta \mathbf{u}_p^* (FRANÇOIS, 2014).

O método de adaptação dos modificadores foi implementado em estudos simulados de relevância prática, notadamente em sistemas de geração de energia (COSTELLO, 2015; COSTELLO *et al.*, 2013b; FRANÇOIS *et al.*, 2015; SERRALUNGA *et al.*, 2013), mas também teve sua validação experimental em um sistema de três tanques (MARCHETTI *et al.*, 2009) e na operação ótima de pilhas a combustível de óxido sólido (BUNIN *et al.*, 2012b). Muitos aspectos do método MA foram investigados recentemente, dentre os quais destacam-se:

- desenvolvimento de estratégia para condicionamento do modelo por meio de aproximações convexas, que permitem satisfazer as condições de adequabilidade *a priori* e sem conhecer o ótimo verdadeiro \mathbf{u}_p^* , garantindo a convergência para o ótimo da planta (FRANÇOIS e BONVIN, 2013b);

- uma vez que a disponibilidade dos gradientes experimentais envolvidos no cálculo dos modificadores é a principal barreira para a implementação do método MA em sistemas reais, muitos trabalhos investigam diferentes abordagens para a estimação dos gradientes (BUNIN *et al.*, 2012a, 2013e; COSTELLO *et al.*, 2015; FRANÇOIS e BONVIN, 2014; MARCHETTI *et al.*, 2010; MARCHETTI, 2013; RODGER, 2010; RODGER e CHACHUAT, 2011; SERRALUNGA *et al.*, 2013), bem como a utilização de medidas transientes do processo para estimação dos gradientes (FRANÇOIS e BONVIN, 2014; RODRÍGUEZ-BLANCO *et al.*, 2017) e estratégias de reformulação do algoritmo que dispensam a estimação dessas grandezas (NAVIA *et al.*, 2013);
- estratégias que reduzem a ocorrência de violação das restrições, evitando que a convergência siga trajetórias inviáveis (BUNIN *et al.*, 2011; NAVIA *et al.*, 2012), e reformulações que tornam o método mais adequado para problemas com grande número de graus de liberdade (COSTELLO, 2015; SERRALUNGA *et al.*, 2013);
- extensão do método para sua aplicação tanto em processos descontínuos (SERRALUNGA *et al.*, 2014) quanto transientes (COSTELLO, 2015);
- generalização do método MA (*generalized modifier-adaptation algorithm*) para problemas em que os graus de liberdade da planta e do modelo são diferentes, tanto para sistemas sem restrições (COSTELLO *et al.*, 2013a) quanto para sistemas com restrições em malha fechada (COSTELLO *et al.*, 2014);
- versão generalizada do método MA para incluir modificadores de segunda ordem, bem como desenvolvimento de condições suficientes de convergência para o algoritmo MA de várias ordens (FAULWASSER e BONVIN, 2014);
- proposta de um algoritmo de adaptação dos modificadores com convergência global, embora o desenvolvimento tenha sido baseado em casos idealizados, sem considerar como o algoritmo se comportaria em aplicações reais na presença de diversas questões limitadoras, como valores imprecisos das funções e indisponibilidade das derivadas experimentais (BUNIN, 2014);
- proposta de algoritmo para casos com elevado nível de ruído, utilizando aproximações quadráticas (GAO *et al.*, 2016).

2.3.4 Condições Suficientes de Viabilidade e Otimalidade

Embora o objetivo dos métodos RTO seja determinar os valores dos graus de liberdade que conduzam a planta ao seu máximo desempenho sem violar restrições,

notadamente aquelas relacionadas a limites de segurança, nenhum método apresenta garantias para satisfação estrita das restrições. Nesse contexto, BUNIN *et al.* (2013b, 2014) propuseram um conjunto de condições capazes de garantir a propriedade de viabilidade estrita para qualquer algoritmo RTO. A ideia básica do conjunto de condições suficientes de viabilidade e otimalidade (SCFO) é garantir que o conjunto de valores implementados na iteração seguinte, *i.e.*, \mathbf{u}_{k+1} , sempre pertença a uma região que garanta localmente a redução do valor da função objetivo e a satisfação das restrições. Assim, a satisfação dessas condições fornece garantias teóricas de que uma sequência de experimentos converge para uma solução ótima local do Problema (2.1), com resultados viáveis a cada iteração.

Como nenhum método RTO impõe todas as SCFO, uma formulação capaz de forçar essas condições envolve, primeiramente, projetar o conjunto de valores das variáveis de decisão fornecidos por um método RTO no espaço localmente viável e descendente, de modo que a convergência dentro de uma região viável é imposta sem reduzir sobremaneira a velocidade de convergência, todavia à custa de hipóteses difíceis de serem satisfeitas na prática. O problema de projeção é formulado como (BUNIN *et al.*, 2013b):

$$\begin{aligned} \bar{\mathbf{u}}_{k+1}^* := \arg \min_{\mathbf{u}} \quad & J^{opt} := \|\mathbf{u} - \mathbf{u}_{k+1}^*\|_2^2 \\ \text{sujeito a} \quad & \nabla g_{p,j}(\mathbf{u}_k)^\top (\mathbf{u} - \mathbf{u}_k) \leq -\delta_{g,j}, \quad \forall j : g_{p,j}(\mathbf{u}_k) \geq -\epsilon_j \\ & \nabla \phi_p(\mathbf{u}_k)^\top (\mathbf{u} - \mathbf{u}_k) \leq -\delta_\phi \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned} \quad (2.32)$$

em que \mathbf{u}_{k+1}^* é o resultado fornecido por um método RTO; $\bar{\mathbf{u}}_{k+1}^*$ é o resultado projetado em um espaço viável e descendente; $\nabla \phi_p$ e $\nabla g_{p,j}$ são as derivadas da função objetivo e funções de desigualdade com relação às variáveis de entrada \mathbf{u} , respectivamente; e δ_ϕ e $\delta_{g,j} > \mathbf{0}$ são necessários para aproximar os espaços locais estritamente descendentes. Assim, \mathbf{u}_{k+1} é determinado pela filtragem adequada dos resultados projetados,

$$\mathbf{u}_{k+1} := \mathbf{u}_k + K_k (\bar{\mathbf{u}}_{k+1}^* - \mathbf{u}_k) \quad (2.33)$$

em que K_k é calculado de acordo com a condição de viabilidade, que limita o tamanho do passo de adaptação para que o resultado permaneça dentro do espaço viável, e da condição de diminuição monotônica da função objetivo, que limita o tamanho do passo de adaptação de forma a não mudar a direção local de descida.

Dadas as notórias limitações impostas pelas hipóteses admitidas para implementação das SCFO, especialmente estimativas precisas das derivadas do processo, BUNIN *et al.* (2013c) discutiram aspectos práticos das condições e como elas podem ser usadas em uma implementação real, mas mesmo assim limitando a discussão aos casos em que as variáveis manipuladas são perfeitamente conhecidas e livres de

ruído e sem considerar degradação do processo, *i.e.*, $\phi_p(\mathbf{u})$ e $\mathbf{G}_p(\mathbf{u})$ permanecem constantes ao longo de todas as iterações. Como as condições não podem ser implementadas da forma como são formuladas, uma reformulação da Equação (2.32) foi proposta, utilizando relaxações das desigualdades estritas por meio de recuos (ou afastamentos, *back-offs*) individuais, resultando em uma projeção robusta. Isso permite a implementação das condições SCFO em dado problema ou em conexão com algum método RTO. Para sua implementação isolada, o algoritmo pode ser entendido e aplicado como um método RTO. Alternativamente, quando conectado com outro método em um sistema RTO, pode ser usado para modificar o conjunto de resultados obtido pelo método RTO para garantir resultados que não violem restrições e melhorem o valor da função objetivo a cada iteração. Essa estratégia de acoplamento das condições SCFO a outros métodos RTO foi avaliada por JUN *et al.* (2015), que estudaram o desempenho da abordagem em duas etapas, adaptação dos modificadores e adaptação das restrições em um estudo simulado que implementa o sistema RTO em um processo de lixiviação de ouro. O acoplamento das SCFO melhorou significativamente os valores da função objetivo econômica do processo frente a incertezas paramétricas, erros de modelagem e perturbações de processo, quando comparados aos resultados obtidos pelos métodos RTO sem as SCFO. A reformulação, juntamente com toda a teoria envolvida, foi implementada e está disponível como um *script* de código aberto, podendo ser obtido com facilidade na internet (BUNIN *et al.*, 2013d). Uma aplicação rudimentar das SCFO a um problema sem restrições de ajuste iterativo de controles apresentou resultados positivos para o uso da teoria (BUNIN *et al.*, 2013a).

2.3.5 Otimização por Regulação

A depender das características do processo que se deseja otimizar, é possível que a velocidade de atualização dos *setpoints* seja baixa, consistindo em uma das grandes desvantagens de sistemas RTO. Conforme discutido anteriormente, isso ocorre porque o sistema RTO executa suas atividades apenas com informações coletadas no regime estacionário. Nesses casos, o tempo transcorrido até que o sistema RTO transmita o novo conjunto de *setpoints* para serem implementados pela camada de controle pode deteriorar significativamente o desempenho da camada de otimização. Logo, se mudanças ocorrerem em escalas de tempo inferiores ao período de amostragem do RTO, associar um sistema RTO ao controle supervisor não proporcionará qualquer benefício (ENGELL, 2007). Nesse contexto, a estratégia de uma classe dos métodos livres de modelo consiste em evitar a resolução do problema de programação matemática e reinterpretar o problema de otimização como um problema de controle, em que as variáveis de decisão são modificadas diretamente por meio

de uma lei de controle adequada, que implicitamente corresponde à resolução do problema de otimização. Em outras palavras, a estratégia consiste em traduzir objetivos econômicos em objetivos de controle (MORARI *et al.*, 1980), tendo resultado no conceito conhecido como controle auto-otimizado (SKOGESTAD, 2000a). Outros métodos, por sua vez, seguem uma filosofia similar, com foco na camada de controle supervísório, introduzindo elementos econômicos nos controles avançados com o objetivo de integrar RTO e (N)MPC, dando origem ao controle preditivo econômico (ELLIS *et al.*, 2014).

2.3.5.1 Controle Auto-Otimizado

O conceito de controle auto-otimizado foi proposto por SKOGESTAD (2000a,b) como uma estratégia para construir uma estrutura de controle *feedback* capaz de fornecer um desempenho econômico aceitável do processo, mantendo variáveis controladas (*controlled variables* - CVs) em *setpoints* constantes e evitando a necessidade de realizar nova otimização econômica quando perturbações ocorrem. Como um controle integral é capaz de manter uma CV em determinado *setpoint* no estado estacionário, a adequada seleção das CVs é o principal foco do SOC (FRANÇOIS e BONVIN, 2013a; YE *et al.*, 2013). Assim, busca-se no contexto do SOC selecionar um conjunto adequado de CVs que minimizem a perda econômica em relação ao ótimo verdadeiro (ou, por outro ponto de vista, que forcem o desempenho ótimo da planta) quando controladas em *setpoints* constantes. É natural que essa estratégia ocasione alguma perda, mas faz parte da abordagem o desafio de selecionar um conjunto de variáveis controladas, um conjunto de variáveis manipuladas e um conjunto de *setpoints* para os quais as perdas de desempenho do processo sejam aceitáveis, fazendo com que a planta seja “auto-otimizada” a despeito das perturbações (YE *et al.*, 2014).

A proposta do SOC se mostra atrativa, pois permite aumentar significativamente a velocidade de convergência do algoritmo para o ponto ótimo. A otimização por regulação pode ser genericamente representada como (CHACHUAT *et al.*, 2009):

$$\mathbf{c} = \Gamma_{sel} [\mathbf{y}_p(\mathbf{u}_k)] \quad (2.34)$$

$$\mathbf{u}_{k+1} = \Gamma_{soc} [\mathbf{c}(\mathbf{u}_k)] \quad , \quad (2.35)$$

em que \mathbf{c} representa variáveis controladas apropriadas, cuja referência corresponde a \mathbf{c}^{sp} ; $\Gamma_{sel}(\cdot)$ é um mapeamento que descreve a seleção das variáveis controladas, incluindo detecção de restrições ativas; e $\Gamma_{soc} : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_u}$ é o mapeamento que descreve o controle multivariável. SKOGESTAD (2000b) propôs o uso de um modelo estacionário do processo para selecionar as CVs ou combinações de medidas que resultem em bom desempenho no método SOC. O método SOC, associado a essa

metodologia de seleção, foi aplicado em problemas simulados de grande dimensão (DE ARAÚJO *et al.*, 2007; LARSSON *et al.*, 2003, 2001), mostrando que a perda de desempenho econômico ocasionada pela regulação em *setpoints* fixos em lugar da otimização estática pode ser baixa para estruturas de controle bem determinadas. LARSSON *et al.* (2001) comentaram que a abordagem sistemática pode resultar em esquemas de controle nada óbvios para um engenheiro de controle treinado, mas comprovadamente viáveis de acordo com simulações dinâmicas, mostrando que a aplicação do SOC garante perdas menores que 5% no desempenho.

O melhor conjunto de CVs para um controle auto-otimizado pode ser obtido a partir do método exato local por meio de um NLP (HALVORSEN *et al.*, 2003). No entanto, por simplicidade, a prática mais comum consiste em selecionar a melhor combinação linear das medidas \mathbf{y}_p para as CVs, podendo ser genericamente expressas como

$$\mathbf{c} = \mathbf{H} \mathbf{y}_p \quad , \quad (2.36)$$

em que $\mathbf{c} \in \mathbb{R}^{n_u}$ são as CVs do controle com *setpoints* ótimos nominais \mathbf{c}_p^* ; $\mathbf{H} \in \mathbb{R}^{n_u \times n_y}$ é uma matriz constante, chamada de matriz de combinação. De acordo com essa construção, selecionar CVs implica determinar a matriz de combinação, \mathbf{H} , sendo dois os critérios comumente usados para sua determinação, a saber: a perda de desempenho resultante do pior cenário (HALVORSEN *et al.*, 2003) e a perda de desempenho média com base em perturbações uniformemente distribuídas (KARIWALA *et al.*, 2008). Esses critérios consideram perdas causadas por perturbações e erros de implementação. A matriz de combinação pode ser definida como o argumento que minimiza a função de perda definida por um dos critérios. De acordo com esse resultado, CVs podem ser selecionadas tanto como variáveis medidas individuais, quanto por combinações das variáveis medidas. Normalmente, a adoção de combinações que envolvem medidas como CVs resulta em menores perdas econômicas, quando comparadas àquelas obtidas pelo uso direto de medidas individuais.

Diversas metodologias foram propostas para a seleção de CVs adequadas ao método SOC (ALSTAD e SKOGESTAD, 2007; ALSTAD *et al.*, 2009; HALVORSEN *et al.*, 2003; JÄSCHKE e SKOGESTAD, 2012, 2013; KARIWALA, 2007; KARIWALA *et al.*, 2008; YE *et al.*, 2013). Com base em combinações lineares das variáveis medidas, um método de seleção de variáveis, denominado de método do espaço nulo, foi proposto por ALSTAD e SKOGESTAD (2007), que seleciona o conjunto de CVs que resultam em perda nula com relação às perturbações analisadas. Posteriormente, o método foi estendido por ALSTAD *et al.* (2009) para considerar perdas causadas por ruídos de medida. Também com base em combinações lineares, KARIWALA (2007) e KARIWALA *et al.* (2008) desenvolveram abordagens basea-

das na decomposição em valores característicos, para minimizar os critérios do pior cenário de perda e da perda média, respectivamente. No método de JÄSCHKE e SKOGESTAD (2013), por sua vez, uma estratégia de determinação da matriz de combinação com base apenas em dados de processo, especificamente, dados históricos e gradientes estimados por perturbações na planta, foi proposta. Por ser baseado puramente em dados, o método dispensa a otimização *off-line* de um modelo para selecionar CVs. Apesar de não precisar de um modelo do processo, esse método exige grandes quantidades de dados, que representam uma barreira para a aplicação prática desse método.

Semelhantemente aos problemas de projeto de outras estruturas de controle, a seleção de CVs constitui um problema combinatorial. Encontrar um subconjunto de medidas globalmente ótimo torna-se inviável quando o número de variáveis medidas é grande, aumentando consideravelmente as combinações possíveis para as CVs. Motivado por essas questões, algoritmos *branch-and-bound* bidirecionais foram desenvolvidos e aplicados com diferentes critérios de seleção de CVs, especificamente, para a regra de mínimo valor singular (CAO e KARIWALA, 2008), para o pior cenário de perda de desempenho (KARIWALA e CAO, 2009) e para a perda de desempenho média (KARIWALA e CAO, 2010). Adicionalmente, YELCHURU e SKOGESTAD (2012) propuseram o uso de programação quadrática inteira mista para a seleção de CVs, com base na minimização da perda média de desempenho.

Todos os métodos SOC admitem que o conjunto de restrições ativas permanece o mesmo, quando perturbações ocorrem. Todavia, uma das principais dificuldades práticas ocorre justamente quando há mudanças das restrições ativas, pois isso altera o número de graus de liberdade. Esse tipo de evento usualmente exige a reconfiguração da estrutura de controle, para que continue fornecendo perdas em níveis aceitáveis. Para problemas com diferentes conjuntos de restrições ativas, CAO (2004) propôs um método que usa uma abordagem de controle em cascata, em que um laço interno força as restrições ativas, quando necessário, enquanto um laço externo modifica o *setpoint* do laço interno, que controla as CVs. Assim, a estrutura permite que a condição ótima e as restrições ativas sejam automaticamente alternadas entre si, de acordo com a perturbação que afeta o processo, de modo que ambas sejam satisfeitas. Seguindo outra abordagem, HU *et al.* (2012) propuseram uma metodologia para determinar uma estrutura de controle que não precise ser reconfigurada, mas permaneça capaz de satisfazer todas as restrições em um conjunto definido de perturbações e erros de implementação, selecionando CVs por meio da minimização do critério da perda de desempenho médio local. MANUM e SKOGESTAD (2012) estenderam o método do espaço nulo para problemas com mudanças no conjunto de restrições ativas. Por meio da aplicação do método do espaço nulo em diferentes regiões, a metodologia identifica mudanças no conjunto de restrições e

implementa a estrutura de controle correspondente ao conjunto ativo mais recente, alternando entre as matrizes de combinação de cada região. Essa abordagem apresenta bom desempenho quando o número de regiões é pequeno e o processo se move continuamente de um ponto de operação a outro.

Embora seja conveniente controlar combinações lineares das variáveis medidas, a convergência ao ponto ótimo da planta é dificilmente garantida. Motivada por essa questão, a abordagem conhecida como busca de NCO seleciona como variáveis controladas as duas componentes das condições necessárias de otimalidade, *i.e.*, as restrições ativas \mathbf{G}_p^a e os gradientes reduzidos $\nabla_r \phi_p$:

$$\mathbf{c} = [\mathbf{G}_p^a \quad \nabla_r \phi_p]^\top \quad , \quad (2.37)$$

tendo como *setpoints* correspondentes $\mathbf{c}_p^* = \mathbf{0}$. Assim, as variáveis de entrada do processo são manipuladas diretamente para que as NCO sejam satisfeitas (CHACHUAT *et al.*, 2008, 2009), que corresponde, implicitamente, a resolver o problema de otimização econômica para a planta no sentido das NCOs de primeira ordem (FRANÇOIS e BONVIN, 2013a). Assim, essas variáveis precisam ser estimadas *on-line*. Duas classes de abordagem entram nessa categoria, a saber: controle extremal (ARIYUR e KRSTIC, 2003; GUAY e ZHANG, 2003) e busca de NCO (FRANÇOIS *et al.*, 2005; SRINIVASAN e BONVIN, 2007; SRINIVASAN *et al.*, 2008). O método de busca de NCO foi originalmente desenvolvido para processos batelada, fazendo uso da informação produzida ao longo de repetidas bateladas para estimar as NCOs, de modo que a operação da planta possa convergir para a condição ótima (FRANÇOIS *et al.*, 2005). Posteriormente, o método foi estendido para processos contínuos (SRINIVASAN *et al.*, 2008).

A principal diferença entre SOC e busca de NCO está relacionada à implementação e o funcionamento dos métodos. Ao implementar o método SOC, primeiro utiliza-se alguma metodologia para a seleção das CVs de acordo com as medições disponíveis para um dado conjunto de perturbações, enquanto a execução *on-line* é realizada pelo controle *feedback* que mantém as CVs selecionadas em seus respectivos *setpoints*, os quais são determinados *off-line*. Por sua vez, o método de busca de NCO estima restrições ativas e gradientes reduzidos em tempo real com as medidas disponíveis, de modo que o valor ótimo das variáveis manipuladas é calculado explicitamente de acordo com determinadas regras de adaptação. Assim, enquanto SOC concentra-se em CVs ótimas, busca de NCO determina a solução ótima explicitamente, embora essa diferença não seja fundamental, já que variáveis manipuladas também podem ser selecionadas como CVs no método SOC ou incluídas nas combinações das medidas que compõem as CVs. Dessa maneira, SOC e busca de NCO focam em diferentes aspectos do RTO (YE *et al.*, 2013). Além disso, esses dois mé-

todos RTO, assim como os métodos de adaptação dos modificadores, dependem da estimação *on-line* de gradientes experimentais para controle. No SOC, aproximações dos gradientes estão *implicitamente* embutidas nos modelos construídos para as CVs, os quais são determinados *off-line* de acordo com as técnicas mencionadas. Isso confere à abordagem SOC rápida convergência e, assim, capacidade para rejeitar perturbações de alta frequência. No entanto, como CVs são determinadas com base em linearizações do modelo de processo em torno de um ponto operacional, seu desempenho é ótimo apenas localmente. Na busca de NCO, os gradientes são avaliados tanto por abordagens livres de modelo, quanto por abordagens baseadas em modelo (FRANÇOIS *et al.*, 2012). Nas abordagens livres de modelo, os gradientes são determinados experimentalmente *on-line*, *e.g.*, por aplicação direta do método de diferenças finitas na planta por meio de perturbações (SRINIVASAN *et al.*, 2003, 2008). Geralmente, esse procedimento *on-line* de estimação tem validade global para o ótimo, embora seja lento em virtude do tempo necessário para que a planta atinja o regime estacionário. Para melhorar a velocidade de convergência, os gradientes também podem ser determinados por abordagens baseadas em modelo, como a técnica conhecida como *neighboring-extremal control* (NEC) (GROS *et al.*, 2009). Nessas abordagens, a construção do modelo ocorre *off-line*, permitindo que o controle possa agir com rapidez. Todavia, a abordagem NEC também apresenta somente validade local. Alguns trabalhos na literatura fornecem comparações entre o método SOC e busca de NCO (FRANÇOIS *et al.*, 2012; JÄSCHKE e SKOGESTAD, 2011), bem como comparações entre SOC e outros métodos RTO (MANUM e SKOGESTAD, 2012; SCHULTZ e FARENZENA, 2015).

Motivados pelas características de rápida convergência do método SOC e da otimalidade global da busca de NCO, YE *et al.* (2013) propuseram uma abordagem que combina ambos os métodos, selecionando as CVs por meio de aproximações para as condições necessárias de otimalidade. Como a aproximação proposta é feita ao longo de toda a faixa de operação, as CVs apresentam um desempenho auto-otimizado com caráter global, enquanto outros métodos SOC são locais, válidos apenas na vizinhança do ponto de operação nominal em virtude da linearização. Assim, o método resulta em perdas aceitáveis em uma região maior da operação, quando comparado a outras abordagens locais, embora a precisão da aproximação possa não ser satisfatória para faixas de operação muito extensas.

Recentemente, novas estruturas de controle foram propostas, combinando SOC e RTO. JÄSCHKE e SKOGESTAD (2011) propuseram uma estrutura na qual a camada de controle corresponde ao método SOC, com CVs selecionadas de acordo com o método do espaço nulo (ALSTAD e SKOGESTAD, 2007), enquanto a camada do RTO corresponde ao método de busca de NCO (SRINIVASAN *et al.*, 2003). Nessa abordagem, no entanto, os *setpoints* das CVs devem ser perturbados

para que seja possível estimar os gradientes. Na proposta de YE *et al.* (2014), o método de busca de NCO também é combinado ao método SOC para a construção de uma nova estrutura de controle, que integra o método SOC a um método RTO. Na proposta, as NCO foram usadas como variáveis controladas e um critério estatístico de não otimalidade foi usado para decidir quando a variável controlada deveria ser atualizada. A metodologia está baseada no conceito de adaptação das variáveis controladas (*controlled variable adaptation* - CVA) e detecção de não otimalidade. Na estratégia CVA, as CVs são determinadas e adaptadas com base em um algoritmo de regressão que aproxima as NCO, permitindo que o SOC seja adaptado às mudanças nas condições de operação. A não otimalidade é detectada com base em técnicas de monitoramento estatístico de processos (T^2 e SPE), em que a não otimalidade é tratada como um tipo de falha de processo, sendo usada como pré-requisito para a ativação da CVA. Diferentemente das estruturas de controle tradicionais, a abordagem de YE *et al.* (2014) resulta em uma camada de controle capaz de estabilizar o processo e, simultaneamente, conduzir a operação a uma condição ótima de operação com rapidez. A camada de otimização, por sua vez, permite estimar as perturbações, resolver o problema de otimização econômica, transferir *setpoints* à camada de controle e, além disso, adaptar as CVs, se um estado de não otimalidade for detectado. Na recente metodologia de GRACIANO *et al.* (2015), há associação não somente entre RTO e SOC, mas também com MPC, na estrutura de controle. A estratégia proposta se aplica aos casos em que novas restrições são ativadas entre atualizações do RTO e também visa a abordar o problema da lenta convergência do RTO. A abordagem é similar àquela de MANUM e SKOGESTAD (2012), para lidar com mudanças nas restrições ativas, mas não determina os conjuntos de CVs para todas as restrições ativas *a priori*. Em contraste, essas informações são calculadas *on-line* pelo RTO e apenas para a região de operação mais recente. Assim, o RTO é usado para determinar as CVs e seus *setpoints*, enquanto a camada de controle detecta quando uma restrição se torna ativa, controlando as CVs para o novo ótimo, sem ter que aguardar a execução do RTO, o que é denominado de zona de controle (*zone control*).

2.3.5.2 Controle Preditivo Econômico baseado em Modelo

Em virtude da operação estacionária da camada de otimização, o intervalo de amostragem é muito maior nessa camada que aquele usado na camada de controle supervisão, *i.e.*, a execução do sistema RTO é menos frequente que a execução do MPC. Como o MPC não incorpora informações econômicas, não é capaz de detectar perturbações ou outras influências que alterem as condições ótimas da operação durante períodos transientes (ENGELL, 2007; MARLIN e HRYMAK, 1997; SIROLA e EDGAR, 2012). Além das questões envolvidas na baixa frequência de execução

do sistema RTO, outras dificuldades surgem na aplicação prática do MPC. Mesmo que o modelo estacionário usado no sistema RTO seja capaz de determinar o ótimo verdadeiro do processo (FORBES e MARLIN, 1994; FORBES *et al.*, 1994), seus resultados podem não ser consistentes com o modelo usado no MPC, que frequentemente emprega modelos lineares construídos por linearização de um modelo fenomenológico em torno do ponto estacionário desejado ou por técnicas de identificação. A maioria das aplicações industriais de MPC ainda aplica modelos empíricos lineares, construídos a partir de testes específicos na planta (DARBY e NIKOLAOU, 2012; DARBY *et al.*, 2009). Se houver disparidade entre os modelos usados nas diferentes camadas, é possível que pontos de operação calculados pelo sistema RTO sejam inalcançáveis pela camada de controle, resultando frequentemente em *offset* entre a operação do processo e a operação estacionária indicada pelo RTO. Além disso, mudanças bruscas nos *setpoints* determinadas pelo sistema RTO podem resultar em instabilidade do MPC (YING e JOSEPH, 1999).

De modo a aproximar as frequências de execução do sistema RTO e da camada de controle supervisorio e calcular *setpoints* consistentes com o modelo usado no MPC, alguns trabalhos propuseram uma estratégia de atualização dos *setpoints* transmitidos pelo sistema RTO para que sejam consistentes com o controle supervisorio (ŁAWRYŃZUK *et al.*, 2008; MORSHEDI *et al.*, 1985; MUSKE, 1997; YOUSFI e TOURNIER, 1991). Para atingir esse objetivo, um controlador estacionário foi acrescentado entre o sistema RTO e o MPC, formando um esquema de controle denominado de MPC em dois estágios. O MPC em dois estágios divide a camada de controle supervisorio em dois níveis, que operam com a mesma frequência de amostragem, em que o controle estacionário atualiza continuamente os *setpoints* usados pelo MPC por meio da resolução de um problema de programação linear ou quadrático (LP-MPC ou QP-MPC), convertendo o resultado estacionário ótimo determinado pela camada de otimização em um *setpoint* viável para a camada de controle. Uma análise detalhada das propriedades dessa estrutura é apresentada por YING e JOSEPH (1999).

Outra estratégia proposta para lidar com essas questões é conhecida como controle preditivo econômico (EMPC) e consiste em integrar a otimização econômica à camada de controle, incorporando um índice de desempenho na formulação do controle (N)MPC, permitindo assim que o problema de otimização estática não linear seja simultaneamente resolvido com o problema de controle preditivo multivariável, resultando em uma estratégia de uma camada (DE GOUVÊA, 1997; DE GOUVÊA e ODLOAK, 1998; ENGELL, 2007; HELBIG *et al.*, 2000; RAWLINGS e AMRIT, 2009; ZANIN *et al.*, 2002). Essa abordagem não força o processo a operar em um estado estacionário pré-especificado e tem recebido mais atenção, por permitir melhoria de desempenho por meio da operação transiente. Noções tradicionais de

estabilidade (*e.g.*, estabilidade assintótica de um estado estacionário) podem não se aplicar a sistemas EMPC em malha fechada, tornando o projeto rigoroso desses sistemas mais desafiador. Embora o uso de problemas de controle ótimo (*optimal control problem* - OCP) com uma função econômica não constitua um conceito novo, bem como formulações do MPC com termos econômicos, a estabilidade em malha fechada e o desempenho do processo sob ação do controle EMPC foi considerado e provado apenas recentemente (ELLIS *et al.*, 2014).

O controle preditivo baseado em modelo (MPC), também chamado de controle de horizonte retrocedente (*receding horizon control*), é uma técnica *on-line* de controle formulada como um problema de programação matemática que otimiza um índice de desempenho ao longo de um horizonte futuro por meio das previsões de um modelo dinâmico nominal do processo, de acordo com as restrições existentes (MAYNE, 2014; MORARI e LEE, 1999; QIN e BADGWELL, 2003; RAWLINGS e MAYNE, 2009). O principal objetivo do MPC convencional é conduzir a planta a um ponto de operação estacionário (*e.g.*, fornecido pelo sistema RTO), ou implementar uma trajetória ótima (*e.g.*, fornecida pelo sistema DRTO), calculados na camada de otimização. Para considerar o conflito entre a velocidade de resposta do sistema em malha fechada e o tamanho das variações causadas pela ação de controle, necessárias para gerar essa resposta, o MPC é tipicamente formulado com uma função objetivo quadrática, que penaliza tanto os desvios das variáveis de estado quanto das variáveis de entrada em relação aos valores ótimos em estado estacionário ao longo do horizonte de previsão. Especificamente, o MPC é escrito como um problema de otimização dinâmica (ELLIS *et al.*, 2014):

$$\begin{aligned}
& \underset{\mathbf{u} \in S(\Delta)}{\text{Minimizar}} & J_{\text{MPC}}[\mathbf{x}(t), \mathbf{u}(t)] & := \int_0^{t_N} [\mathbf{x}(t)^\top Q_c \mathbf{x}(t) + \mathbf{u}(t)^\top R_c \mathbf{u}(t)] \, dt \\
& \text{sujeito a} & \dot{\mathbf{x}}(t) & = f[t, \mathbf{x}(t), \mathbf{u}(t)] \\
& & \mathbf{x}(0) & = \mathbf{x}(t_k) \\
& & \mathbf{g}[\mathbf{x}(t), \mathbf{u}(t)] & \leq \mathbf{0}, \quad \forall t \in [0; t_N]
\end{aligned} \tag{2.38}$$

em que $\mathbf{u} \in S(\Delta)$ é a trajetória das variáveis de entrada ao longo do horizonte de previsão e $S(\Delta)$ denota a família de funções constantes, contínuas por partes, com período Δ . $Q_c > 0$ e $R_c > 0$ são matrizes positivas definidas de ajuste, que determinam o compromisso entre velocidade de resposta e o custo da ação de controle. A trajetória dos estados $\mathbf{x}(t)$ é a evolução do vetor de estados predita pelo modelo dinâmico nominal, de acordo com o perfil constante por partes $\mathbf{u} \in S(\Delta)$ determinado pelo MPC. As condições iniciais no modelo dinâmico $\mathbf{x}(0) = \mathbf{x}(t_k)$ são obtidas em cada instante de amostragem k por meio das medidas mais recentes. Quando é possível medir os estados, tem-se $\mathbf{y}_{ms} = \mathbf{x}_k = \mathbf{x}(t_k)$. As restrições $\mathbf{g}(\cdot)$ são restrições de processo impostas ao perfil das variáveis de entrada, as quais são tipicamente

restrições pontuais, usualmente escritas como:

$$\mathbf{g}[\mathbf{x}(t_j), \mathbf{u}(t_j)] \leq \mathbf{0}, \quad j = 0, 1, \dots, N \quad . \quad (2.39)$$

Quando o horizonte de predição N é finito, sabe-se que a formulação do MPC, de acordo com a Equação (2.38), pode apresentar instabilidade (MAYNE *et al.*, 2000). Nesse contexto, várias restrições e variações da função objetivo podem ser feitas para garantir estabilidade do sistema em malha fechada quando N é finito. Em processos que exibem forte comportamento não linear e dinâmica complexa, a substituição de modelos lineares por não lineares consiste no NMPC.

A abordagem EMPC integra o sistema RTO e o (N)MPC em uma mesma camada, que na prática consiste na formulação de um único problema de programação matemática. Essa integração ocorre por meio da modificação da função objetivo do MPC,

$$J_{\text{EMPC}}[\mathbf{x}(t), \mathbf{u}(t)] := J_{\text{MPC}} + \phi[\mathbf{x}(t_N), \mathbf{u}_s] \quad (2.40)$$

que considera tanto os termos quadráticos do MPC quanto a função de desempenho econômico, $\phi(\cdot)$, na função objetivo. Assim, o EMPC possui a seguinte formulação geral (ELLIS *et al.*, 2014):

$$\begin{aligned} & \underset{\mathbf{u} \in S(\Delta), \mathbf{u}_s}{\text{Minimizar}} && J_{\text{EMPC}}[\mathbf{x}(t), \mathbf{u}(t)] \\ & \text{sujeito a} && f[\mathbf{x}(t_N), \mathbf{u}_s] = \mathbf{0}, \quad \mathbf{u}_s \in \mathbb{U} \\ & && \mathbf{x}(0) = \mathbf{x}(t_k) \\ & && \mathbf{g}[\mathbf{x}(t), \mathbf{u}(t)] \leq \mathbf{0}, \quad \forall t \in [0; t_N) \end{aligned} \quad , \quad (2.41)$$

em que as variáveis de decisão do problema de otimização incluem tanto a trajetória das variáveis de entrada ao longo do horizonte de predição $\mathbf{u} \in S(\Delta)$, quanto os valores dessas variáveis no estado estacionário \mathbf{u}_s . A restrição $f[\mathbf{x}(t_N), \mathbf{u}_s] = \mathbf{0}$ impõe que a trajetória predita para os estados $\mathbf{x}(t)$ convirja para um estado estacionário admissível.

Muitos trabalhos têm contribuído com o desenvolvimento do EMPC, tanto por meio de formulações gerais (ADETOLA e GUAY, 2010; ALAMO *et al.*, 2012; BECERRA *et al.*, 1998; DE SOUZA *et al.*, 2010), quanto por meio da aplicação em problemas práticos (DE GOUVÊA e ODLOAK, 1998; NATH e ALZEIN, 2000; ZANNIN *et al.*, 2002). Uma revisão bastante abrangente sobre o tema pode ser encontrada no trabalho de ELLIS *et al.* (2014), que faz parte da recente edição especial do periódico *Journal of Process Control* dedicada ao tema EMPC (CHRISTOFIDES e EL-FARRA, 2014). ELLIS *et al.* (2014) apresentaram as diversas formulações propostas, comentando questões de estabilidade da malha fechada e desempenho em sistemas não lineares e comparando algumas abordagens por meio de um exemplo

simulado. Como apenas um modelo é usado para os objetivos de otimização e controle supervisorio, não há risco de inconsistências entre os diferentes modelos usados em cada camada (ENGELL, 2007). Todavia, o uso de modelos preditivos rigorosos baseados em primeiros princípios requer a solução de um problema de controle ótimo em tempo real (DIEHL, 2001). Complexidades computacionais, exigência de técnicas de identificação *on-line* para processos não lineares e falta de estabilidade e resultados robustos para sistemas não lineares são importantes limitações para a implementação prática do NMPC (CAMACHO e BORDONS, 1999; KOUVARITAKIS e CANNON, 2001) e, conseqüentemente, do EMPC.

Vale destacar que o conceito da estratégia do EMPC resulta em parte do crescente interesse e necessidade por operações dinâmicas, motivado pela inerente característica dinâmica dos processos, mas especialmente por uma produção cada vez mais segmentada e flexível, seguindo o modelo de produção *just-in-time* (ENGELL, 2007). Nesse contexto, também surgiram abordagens que não se baseiam na otimização da operação estacionária como melhor estratégia de otimização de processos. Propostas que substituem modelos estacionários por modelos dinâmicos na camada de otimização resultam no sistema RTO dinâmico (DRTO), em que o processo é intencionalmente operado em regime transiente (ALMEIDA NT e SECCHI, 2011; BACKX *et al.*, 1998, 2000; HELBIG *et al.*, 2000; KADAM *et al.*, 2003; KADAM e MARQUARDT, 2007; KADAM *et al.*, 2005; TOSUKHOWONG *et al.*, 2004; WÜRTH *et al.*, 2009). O trabalho de ALMEIDA NETO (2011) analisou e explorou diversos aspectos desses sistemas, com especial atenção a identificação e solução de inviabilidades, realizando também o desenvolvimento de um sistema de monitoramento, análise e diagnóstico do sistema DRTO. O sistema DRTO tem estrutura similar à do EMPC, no sentido de que a operação é dinâmica e que ambos os problemas de programação matemática que caracterizam esses sistemas são problemas de otimização dinâmica, que minimizam um índice de desempenho econômico sujeito a um modelo dinâmico do processo. No entanto, a principal diferença entre DRTO e EMPC é a finalidade de cada um, *i.e.*, o papel que cada sistema desempenha na hierarquia de controle. O DRTO não é usado diretamente como controle, mas sim como um elemento que traduz informações econômicas em informações de processo, transmitindo-as à camada de controle, assim como o sistema RTO. Adicionalmente, poucos estudos foram dedicados ao tratamento teórico da estabilidade em malha fechada do sistema DRTO com a camada de controle. Por outro lado, o EMPC é implementado como controle *feedback*, em que a formulação é dedicada à estabilidade do sistema em malha fechada, admitindo que o modelo dinâmico seja consistente com o modelo da camada de otimização (ELLIS *et al.*, 2014).

2.4 Sistemas RTO

Conforme discutido, o sistema RTO é um sistema de controle adaptativo baseado em modelo, que processa os dados medidos da planta juntamente com informações de mercado e determina continuamente a condição de operação que maximiza um índice de desempenho econômico. Esse sistema atua na camada superior da hierarquia do sistema de controle e automação, fornecendo *setpoints* para a camada de controle, que por sua vez implementa modificações nas variáveis manipuladas de forma a manter a operação do processo o mais próximo possível da condição ótima de operação (EDGAR, 2004; ENGELL, 2007; KRIST *et al.*, 1996). Os métodos de otimização estacionária em tempo real, ou métodos RTO, foram concebidos para detectar a condição ótima de operação e rastrear suas mudanças, rejeitando perturbações de baixa frequência. Todavia, implementações reais desses métodos revelam outros desafios, além daqueles já conhecidos em decorrência dos erros de modelagem, como a presença de ruído aleatório e erros grosseiros nos dados medidos. Por isso, sistemas RTO industriais tipicamente incorporam outros componentes além do método RTO.

2.4.1 Estrutura e Funcionamento

Em vista das diferenças nas concepções dos métodos de otimização estacionária em tempo real, diferentes requisitos e informações devem estar disponíveis em um sistema RTO. O método RTO mais comum em sistemas industriais (e provavelmente o único em processos contínuos de grande escala) é a abordagem em duas etapas, sendo o método implementado em *softwares* comerciais típicos (*e.g.*, ROMeo[®] 5.1 – Invensys Systems, Inc.; Aspen Plus[®] 7.1 – Aspen Technology, Inc.) (QUELHAS *et al.*, 2013). Os demais métodos existentes não são implementados com a mesma facilidade que o método em duas etapas, pois requerem a disponibilidade de informações ainda difíceis de serem obtidas em aplicações reais, especialmente gradientes experimentais, cuja dificuldade de obtenção aumenta significativamente quanto maior for o número de variáveis de entrada (MARCHETTI *et al.*, 2009). Por essa razão e pela evolução histórica da área de otimização em tempo real, grande parte da literatura técnica está, de alguma forma, relacionada à abordagem em duas etapas. Alguns trabalhos sobre implementações reais acabam transmitindo a impressão de que essa abordagem é única viável para a otimização em tempo real estacionária. Ademais, *softwares* comerciais são usualmente baseados em estruturas muito básicas (CÂMARA *et al.*, 2016; QUELHAS *et al.*, 2013) e não incorporam nem mesmo melhorias colaterais vinculadas a esse método, como planejamento de experimentos (YIP e MARLIN, 2003) ou estratégias automáticas de diagnóstico (ZHANG *et al.*, 2001b).

O método RTO é o principal componente do sistema, responsável por resolver o

problema de otimização econômica que determina a condição de operação que conduz o processo ao melhor desempenho econômico. A inclusão de outros componentes no sistema tem como objetivo fornecer o suporte necessário para que valores de referência adequados sejam transmitidos pelo RTO ao sistema de controle. Portanto, um sistema RTO engloba outras responsabilidades, além da otimização econômica em aplicações industriais.

A estrutura tradicionalmente encontrada em sistemas RTO industriais (*i.e.*, sistemas baseados na abordagem em duas etapas), seja em implementações reais (HAN-MANDLU *et al.*, 1985; LID e STRAND, 1997; STRAND e VELAND, 1997) ou em trabalhos da literatura (FORBES e MARLIN, 1996; MANSOUR e ELLIS, 2008; SE-QUEIRA *et al.*, 2002; ZHANG e FORBES, 2000), tipicamente compreende quatro atividades (CUTLER e PERRY, 1983; DARBY *et al.*, 2011; NAYSMITH e DOUGLAS, 1995; TRIERWEILER, 2014), a saber: i) validação dos dados do processo, ii) atualização do modelo, iii) otimização econômica, e iv) análise e diagnóstico dos resultados. A probabilidade de sucesso da otimização depende da qualidade da execução de cada uma dessas atividades. Toda iniciativa feita para melhorar a precisão e acurácia de qualquer uma das etapas será refletida em melhoria da qualidade das decisões sugeridas pelo sistema.

Em virtude do papel fundamental dos dados na otimização, o ciclo de execução do sistema RTO inicia com a validação do conjunto de sinais medidos. Essa atividade comumente envolve detecção (ou identificação) de estado estacionário, detecção de erros grosseiros e reconciliação de dados (BHAT e SARAF, 2004). As etapas de atualização do modelo e otimização econômica correspondem ao método RTO em duas etapas, conforme apresentado na Seção 2.3.1 (p. 21). Por fim, métodos de análise e diagnóstico dos resultados são importantes para garantir que seja significativa a mudança de *setpoint* correspondente à melhoria de desempenho prevista, pois é possível que a propagação de incertezas ao longo do sistema induza variações nos resultados que, quando implementados, não trarão a melhoria de desempenho apontada pelo RTO. Além disso, uma nova verificação de estacionariedade antes da implementação dos valores otimizados das variáveis de decisão se faz necessária, pois os resultados calculados perderão validade caso a planta esteja em operação transiente ou em um ponto estacionário diferente daquele no início da execução do sistema RTO. Detalhes adicionais sobre o funcionamento desses principais componentes e sua integração ao sistema serão discutidos a seguir.

2.4.2 Detecção de Estado Estacionário

Uma vez que o modelo usado para descrever o comportamento do processo é estacionário, é importante garantir que os sinais medidos também demonstrem esse

comportamento, para que seja coerente seu uso com o modelo da planta. Em outras palavras, não é consistente calcular, com base em um modelo estacionário, a condição de operação correspondente ao melhor desempenho econômico se a planta está em um estado transiente. Assim, a primeira etapa da validação geralmente emprega metodologias de detecção (ou identificação) de estado estacionário (*steady-state detection* - SSD), que analisam se a variabilidade de um subconjunto das informações medidas satisfaz a hipótese de estacionariedade, exigida para assegurar a validade do modelo do processo.

A detecção de estado estacionário atua como controle de execução do sistema RTO. Caso o estado estacionário seja detectado, o sistema encaminha o conjunto de dados para ser usado na reconciliação de dados, em que o valor mais provável das variáveis medidas é estimado juntamente com as variáveis não medidas, utilizando o modelo do processo (*i.e.*, equações de balanços de massa e energia) (ENGELL, 2007). Se o teste de SSD para o conjunto de dados não confirmar que existe um estado estacionário, o sistema aguarda que a planta estabilize sua operação e dados válidos estejam disponíveis (WHITE, 1997). Nesse sentido, mesmo que o RTO seja projetado com modelos estacionários, sua conexão à planta (dinâmica) confere a ele uma natureza dinâmica (DARBY *et al.*, 2011). Além disso, a detecção de estado estacionário é dificultada pelo fato de que diferentes seções de uma planta podem estar em diferentes estados (MARLIN e HRYMAK, 1997). Assim, um teste de estacionariedade muito rigoroso pode implicar que o sistema RTO será executado poucas vezes, reduzindo ainda mais a frequência de amostragem do RTO.

Há ao menos quatro decisões importantes relacionadas a essa etapa: a definição do conjunto de medidas usado no teste (*i.e.*, quais variáveis medidas serão usadas), o tamanho da janela de dados, o método de detecção propriamente dito e a tolerância assumida.

Métodos de detecção de estado estacionário tipicamente avaliam propriedades estatísticas, como média e variância ou tendência linear para dado conjunto de dados em um intervalo de tempo (ou janela de dados), comparando-os com intervalos anteriores por meio de testes estatísticos. Regiões estacionárias podem ser identificadas por comparação dos valores médios dos sinais ao longo de duas janelas de tempo consecutivas (SCHLADT e HU, 2007), por comparação da variância dos sinais ao longo de uma janela de tempo contra limites de referência (KIM *et al.*, 2008), por análise da variabilidade da variância entre duas janelas de tempo consecutivas (BHAT e SARAF, 2004; BROWN e RHINEHART, 2000; CAO e RHINEHART, 1995) ou ainda por verificação de tendências lineares em uma janela de tempo (MAHULI *et al.*, 1992; MORENO, 2010; ÖNÖZ e BAYAZIT, 2003).

Embora sejam relatados em poucos trabalhos na área, há também testes não paramétricos, *i.e.*, testes que não requerem qualquer informação *a priori* ou hipótese

a respeito da distribuição de probabilidade dos sinais. Conforme foi mostrado no trabalho de ÖNÖZ e BAYAZIT (2003), que compararam o desempenho entre o teste paramétrico *t* de Student e o teste não paramétrico de Mann-Kendall para detecção de tendências, testes paramétricos podem ser pouco eficazes quando as hipóteses admitidas não são válidas. Assim, existe grande potencial a ser explorado com o uso de testes não paramétricos para detecção de estado estacionário.

Outra abordagem para a detecção de estacionariedade consiste em fazer transformações com os sinais medidos. A esse grupo pertencem métodos que realizam a transformada de Fourier (BARTUSIAK e NICHOLSON, 1998) e transformada *wavelet* (CASTRO *et al.*, 2009; CAUMO *et al.*, 2006; FLEHMIG *et al.*, 1998; FLEHMIG e MARQUARDT, 2006; JIANG *et al.*, 2003). Essas transformações lineares projetam os dados em uma base ortogonal no domínio frequência, em que a detecção de estacionariedade pode envolver simples comparações da densidade espectral (*power spectral density* - PSD) dos sinais com valores de referência (BARTUSIAK e NICHOLSON, 1998). Outra estratégia consiste em representar trechos do sinal por meio de aproximações polinomiais da base ortogonal, associando a estacionariedade com o caso particular em que o ajuste corresponde a um polinômio de ordem zero (FLEHMIG *et al.*, 1998). A representação dos sinais em bases ortogonais também pode estar combinada com outra técnica, como sugerido por JIANG *et al.* (2003), que propuseram uma regra de combinação, similar à de NARASIMHAN *et al.* (1987), que pondera a importância de cada sinal por meio de valores arbitrários. Essa estratégia multivariável foi posteriormente estendida por CAUMO *et al.* (2006) por meio do uso da técnica de análise em componentes principais (*principal component analysis* - PCA), em que a determinação de estacionariedade para uma variável latente, construída pela técnica PCA, permite agregar informações multivariáveis para o teste.

MEJÍA *et al.* (2010) propuseram dois métodos SSD e compararam seu desempenho com outros quatro métodos consolidados na literatura, concluindo que o método baseado na estimação do valor absoluto das primeira e segunda derivadas fornece os melhores resultados. Embora a ideia seja simples, em que o estado estacionário é representado pelo valor nulo das derivadas, dificuldades de implementação acabam surgindo devido aos ruídos de medida e erros grosseiros (TRIERWEILER, 2014). Nesses casos, podem ser usadas técnicas de filtragem para remover perturbações de alta frequência, bem como suprimir efeitos dinâmicos.

Outra abordagem do problema de estacionariedade consiste em focar nos efeitos que os resultados da detecção de estado estacionário provocam no desempenho do sistema. Tal abordagem é pouco comum na literatura, sendo encontrada no âmbito de reconciliação de dados (ÉRIC POULIN *et al.*, 2010) e no contexto da otimização em tempo real (FLEHMIG e MARQUARDT, 2008). O trabalho de FLEHMIG

e MARQUARDT (2008) corresponde, possivelmente, ao único trabalho na literatura que relata a ausência de conexão entre os conceitos embutidos nos testes de estacionariedade e seus efeitos nos resultados da otimização econômica. FLEHMIG e MARQUARDT (2008) propuseram uma metodologia para analisar a confiabilidade da inferência de tendências nas variáveis não medidas, a partir da tendência observada nas variáveis medidas, permitindo determinar a acurácia com a qual as inferências são feitas. Esse método é especialmente adequado ao problema enfrentado pelos métodos de SSD, cujo objetivo consiste em atestar a estacionariedade do processo (que inclui variáveis não medidas) a partir dos dados de um subconjunto das variáveis medidas. Com base em um modelo dinâmico do processo, que confere a relação entre variáveis medidas e não medidas, o método proposto permite determinar um limite superior para o erro de inferência da tendência nas variáveis medidas, que garanta a existência de tendência nas variáveis não medidas. O procedimento auxilia na definição de um limite de subotimalidade da função objetivo econômica em sistemas RTO ocasionada por desvios de estado estacionário. Adicionalmente, a metodologia pode ser empregada para determinar a acurácia com que diferentes conjuntos de variáveis medidas expressam a condição de estacionariedade, ou outro tipo de tendência, sendo especialmente alinhado ao objetivo de selecionar um conjunto de dados medidos em métodos de SSD.

A despeito das diversas propostas existentes para esse fim, a detecção de estado estacionário presente em sistemas reais é bastante simplificada. Com métodos univariáveis, as opções oferecidas consistem em: i) comparação da variância total de um sinal e a variância da diferença entre dois pontos sucessivos dentro desse sinal, ii) análise da frequência do sinal e iii) comparação dos valores médios de duas metades de uma janela de tempo (CÂMARA *et al.*, 2016). Como a parametrização é definida pelo usuário, o fundamento estatístico eventualmente presente pode ficar comprometido.

QUELHAS (2013) investigou as soluções tipicamente usadas para detecção de estacionariedade em sistemas RTO comerciais, comparando diversos testes de estacionariedade existentes e fazendo uma análise de sua *utilidade*, em que *utilidade* é o efeito do método de SSD refletido na função densidade de probabilidade (PDF) de um índice de desempenho que avalia a atuação do sistema RTO, de acordo com algum critério. Tal abordagem focou nas consequências finais para as quais os métodos SSD são utilizados, de maneira equivalente à abordagem de FLEHMIG e MARQUARDT (2008). Por meio de exemplos simulados, foi mostrado que os métodos de SSD representam abordagens restritas à morfologia e variabilidade dos sinais, *i.e.*, abordagens desenvolvidas sob a perspectiva do sinal em si mesmo e de forma desconectada do processo. Nesse sentido, a estacionariedade apontada pelo método não apresenta qualquer correlação real com os valores do índice de desempenho do

sistema RTO.

O descompromisso entre a definição do método de SSD e sua parametrização (comumente uma variável estocástica a serviço do teste de hipóteses estatístico) e o índice de desempenho do sistema de RTO é evidenciado pela falta de sensibilidade observada em relação a, por exemplo: i) o conjunto de sinais usado no método de SSD; ii) influência dos ruídos de medida em todos os testes e condições avaliadas; iii) parametrização interna do modelo de processo; iv) aspectos descritivos dos sinais, especialmente, potência do sinal; e por último, mas não menos importante, v) variações sofridas por variáveis relacionadas ao problema dinâmico, mas que inexistem na apresentação do problema estacionário, que também afetam o índice de desempenho do sistema RTO, mas passam despercebidas ao método de SSD. Isto implica que a determinação da utilidade de um método de detecção de estacionariedade requer o conhecimento do modelo dinâmico do processo, ainda que o sistema RTO seja projetado para a otimização da operação estacionária (QUELHAS, 2013). Tal desconexão entre o método SSD e o índice de interesse último do sistema RTO impede que se determine *a priori* qual é a relação existente entre o nível de confiança estatística usado na parametrização do método de SSD e a incerteza associada à estimativa do índice de desempenho ótimo do processo.

2.4.2.1 Alternativas para o controle de execução do sistema RTO

QUELHAS (2013) mostrou como um nível de subotimalidade, definido como a diferença entre o valor da função objetivo econômica obtido pelo sistema RTO e o valor da mesma função correspondente ao ótimo verdadeiro da planta, *i.e.*, $J^{opt}(\mathbf{u}_p^*) - J^{opt}(\mathbf{u}_{k+1}^*)$, depende da potência dos sinais contidos nos dados das variáveis medidas, do conhecimento do modelo dinâmico do processo e dos parâmetros do modelo. Assim, o problema da detecção de estado estacionário no contexto do sistema RTO consiste, na realidade, na definição de parâmetros que não se atenam apenas à morfologia do sinal, mas que correlacionem bem níveis de tolerância com um índice de desempenho do sistema RTO (ou um grau de subotimalidade esperado) para as previsões da função objetivo econômica. Uma vez que tais dependências são ignoradas por métodos baseados na morfologia do sinal, os quais também não atentam ao modelo interno do problema, fica evidente a necessidade de métodos alternativos aos métodos de SSD para o controle de execução do sistema RTO para superar as deficiências apresentadas.

Embora a metodologia proposta por FLEHMIG e MARQUARDT (2008) seja capaz de inferir o desvio da estacionariedade, não diretamente expresso na morfologia dos dados oriundos das medições, auxiliando na tarefa de fixar a tolerância para os desvios observados, essa proposta foi baseada no uso de um modelo linearizado e não supôs a presença das restrições de desigualdade no problema de otimização

econômica, idealização essa que limita as condições de implementação da metodologia.

QUELHAS (2013) propôs um método alternativo aos métodos de SSD comumente usados como controle de execução do sistema RTO, denominado de teste de adequação (ou de adequabilidade) dos sinais. O nome é motivado pela correlação permanente entre a metodologia e o conceito de *utilidade* estabelecida desde a concepção, bem como de seu caráter pragmático. Ademais, conforme discutido anteriormente, o conceito de estacionariedade perde seu apelo frente ao descompromisso dos métodos SSD com os resultados da função objetivo econômica no contexto do sistema RTO. O método de detecção de adequabilidade proposto resume os seguintes procedimentos: i) descrição dos sinais de uma janela de dados por meio de características pertinentes; e ii) síntese de um veredicto de adequabilidade baseado nas características de todos os sinais disponíveis.

A metodologia foi avaliada em um estudo simulado, considerando o seguinte conjunto de características descritivas: momentos estatísticos (média, desvio-padrão), coeficientes de aproximação polinomial e descrição temporal do espectro de potência do sinal. Foram avaliados quatro casos diferentes, em que foram variados a potência do sinal e o desvio-padrão percentual do ruído gaussiano presente no sinal medido. Os resultados mostraram que a metodologia foi bem sucedida em correlacionar o desempenho realizado com o predito, sendo capaz de prever o efeito dos sinais sobre o desempenho do sistema de otimização. A abordagem também foi caracterizada em termos de seu potencial de falhas, dadas as escolhas feitas para sua configuração e uso, sob o ponto de vista do emprego real do método de adequabilidade. Nesse caso, o método foi capaz de determinar com qual probabilidade a seleção produzida pelo método de adequabilidade produziria resultados contrários ao interesse econômico do processo, a despeito da seleção de sinais estar baseada na melhoria do desempenho, mostrando-se uma ferramenta poderosa para discriminação de estruturas no projeto de sistemas RTO.

2.4.3 Detecção de Erros Grosseiros, Reconciliação de Dados e Estimação de Parâmetros

As técnicas de detecção de erros grosseiros (*gross error detection* - GED) e reconciliação de dados (*data reconciliation* - DR) fazem parte de uma metodologia abrangente para redução de erros de medição conhecida como retificação ou condicionamento de dados, especialmente importante para processos industriais (NARASIMHAN e JORDACHE, 1999; NAYSMITH e DOUGLAS, 1995). A retificação de dados tem por objetivo melhorar a acurácia dos dados, fornecendo estimativas mais prováveis de seus valores verdadeiros por meio da remoção de ruídos aleatórios e não aleató-

rios, sendo imprescindível aos sistemas de controle e automação que dependem da qualidade dos dados coletados na planta (JOHNSTON e KRAMER, 1995; MAH *et al.*, 1976).

Qualquer dado medido contém erros de medição. Em virtude das limitações com relação à precisão e acurácia dos instrumentos de medida, a presença de pequenos erros aleatórios é inevitável. Esses erros de natureza aleatória, que também podem ser causados por outras fontes, como flutuações na rede elétrica, interferências na transmissão e conversão de sinais, mudanças na condição ambiente, *etc.*, podem ser adequadamente caracterizados e representados por meio de métodos estatísticos e distribuições de probabilidade. Nesse sentido, a reconciliação de dados é uma técnica que tem por objetivo estimar os valores mais prováveis dos dados medidos, de modo a satisfazer leis de conservação e outras restrições sabidamente válidas para o sistema de interesse (CROWE, 1996). A principal diferença entre a reconciliação de dados e outras técnicas de filtragem é que a reconciliação faz uso explícito de um modelo do processo para obter estimativas das variáveis. Assim, os valores reconciliados são mais acurados e representam uma melhor estimativa do valor verdadeiro das variáveis medidas, com o benefício adicional de serem consistentes com o modelo do processo (NARASIMHAN e JORDACHE, 1999). Ademais, a reconciliação de dados também é capaz de estimar valores de variáveis que não são medidas, desde que haja redundância nos dados e se tal variável for observável à luz do modelo (CROWE, 1989).

A formulação geral para o problema de reconciliação de dados estacionária é escrita como um problema de programação não linear na forma (MEYER *et al.*, 1993):

$$\begin{aligned} \hat{\mathbf{z}} := & \arg \min_{\mathbf{z}} J_{\text{DR}}[\mathbf{z}_{ms}, \mathbf{z}_m] \\ & \text{sujeito a } \mathbf{f}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \\ & \mathbf{h}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \\ & \mathbf{g}(\mathbf{z}, \boldsymbol{\theta}) \leq \mathbf{0} \end{aligned}, \quad (2.42)$$

em que $J_{\text{DR}}(\cdot)$ é a função objetivo; $\hat{\mathbf{z}}$ é o vetor de valores reconciliados, *i.e.*, estimativas do valor verdadeiro, das variáveis medidas; $\mathbf{z}_{ms} \in \mathbb{R}^{n_{zm}}$ é o vetor de dados das variáveis medidas \mathbf{z} ; $\boldsymbol{\theta} \in \mathbb{R}^{n_{\theta}}$ são parâmetros do modelo; e $\mathbf{f}(\cdot)$, $\mathbf{h}(\cdot)$ e $\mathbf{g}(\cdot)$ são os vetores de funções que representam as equações do modelo, restrições de igualdade e desigualdade, respectivamente. Admite-se implicitamente que o modelo é perfeito. O vetor de variáveis do processo \mathbf{z} pode ser particionado em um vetor de variáveis de saída ($\mathbf{y} \in \mathbb{R}^{n_y}$) e um vetor de variáveis de entrada ($\mathbf{u} \in \mathbb{R}^{n_u}$), tal que $\mathbf{z} = [\mathbf{y}, \mathbf{u}]^T$. O vetor de variáveis medidas $\mathbf{z}_m \in \mathbb{R}^{n_{zm}}$ compreende geralmente as variáveis de saída (\mathbf{y}), mas pode incluir também as variáveis de entrada (\mathbf{u}), sendo definido pela matriz de mapeamento das medidas \mathbf{A}_m , com $\mathbf{A}_m(j, i) = 1$ se a variável j é medida e se o

valor medido está presente em $\mathbf{z}_{ms}(i)$, de modo que $\mathbf{z}_m = \mathbf{A}_m \mathbf{z} = \mathbf{A}_m [\mathbf{y}, \mathbf{u}]^\top$ (LID e SKOGESTAD, 2008a). A depender das variáveis a serem estimadas pela Equação (2.42), diferentes nomenclaturas podem ser dadas ao problema⁶ (KIM *et al.*, 1991).

Um importante elemento da formulação do problema de DR é a definição da função objetivo J_{DR} . Ela pode ser baseada em uma abordagem determinística seguindo técnicas clássicas de regressão (ÖZYURT e PIKE, 2004), sendo definida, por exemplo, como um estimador de mínimos quadrados ponderados (WLS) WHITE (1997):

$$J_{DR} = (\mathbf{z}_{ms} - \mathbf{z}_m)^\top \mathbf{Q}^{-1} (\mathbf{z}_{ms} - \mathbf{z}_m) \quad , \quad (2.43)$$

em que $\mathbf{Q} \in \mathbb{R}^{n_{zm} \times n_{zm}}$ é a matriz de ponderação.

Outra abordagem para formular o problema e definir a função objetivo é baseada em teoria de probabilidade utilizando o método de máxima verossimilhança, para a qual admite-se que as medições de \mathbf{z}_m contêm erros aleatórios, de acordo com:

$$\mathbf{z}_{ms} = \mathbf{z}_m + \boldsymbol{\varepsilon} \quad , \quad (2.44)$$

em que $\boldsymbol{\varepsilon} \in \mathbb{R}^{n_{zm}}$ é o vetor dos erros de medição, com média nula e variância conhecida, apresentando as seguintes propriedades: $E[\boldsymbol{\varepsilon}] = \mathbf{0}$, $Var[\boldsymbol{\varepsilon}] = \boldsymbol{\sigma}_z^2$ e $Cov[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top] = \mathbf{V}_z$. Por meio desse método, a Equação (2.42) torna-se um problema de maximização, em que a função objetivo (ou função de verossimilhança) corresponde à distribuição de probabilidade conjunta das variáveis medidas \mathbf{z}_m . Assim, admitindo que as variáveis medidas sejam estatisticamente independentes, a função objetivo pode ser escrita como o produto das PDFs individuais das variáveis medidas, *i.e.*,

$$J_{DR} := p(\mathbf{z}_{ms}; \mathbf{z}_m, \mathbf{V}_z) = \prod_{i=1}^{n_{zm}} p_i(z_{ms,i}; z_{m,i}, \sigma_i^2) \quad . \quad (2.45)$$

A vantagem da abordagem probabilística é que ela fornece estimativas com propriedades importantes e desejáveis, com estimativas eficientes (com mínima variância teórica possível) e não tendenciosas (flutuando em torno dos valores verdadeiros) (BARD, 1974).

É importante salientar, adicionalmente, que o estimador de mínimos quadrados

⁶ Entre as formulações possíveis, o problema de estimação de estados para sistemas dinâmicos difere não apenas em termos do conjunto de variáveis de decisão, que considera apenas os estados, mas também na hipótese de que o modelo seja perfeito. Em técnicas de estimação de estados, como o filtro de Kalman, a formulação também inclui nas equações do modelo uma variável aleatória para contabilizar ruídos do processo (NARASIMHAN e JORDACHE, 1999; ROMAGNOLI e SÁNCHEZ, 2000; STANLEY e MAH, 1977), chamado de ruído dinâmico ou ruído multiplicativo (AGUIRRE, 2007).

ponderados e a abordagem estatística possuem uma conexão próxima. Na situação em que os erros de medida sejam independentes e sigam distribuição Normal com média nula e variância \mathbf{V}_z , $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V}_z)$, é possível mostrar que o estimador de verossimilhança mostrado da Equação (2.45) corresponde ao estimador de mínimos quadrados ponderado pela matriz de covariância dos erros, *i.e.*, equivale à Equação (2.43) com $\mathbf{Q} = \mathbf{V}_z$ (ANDERSON *et al.*, 1978; BRITT e LUECKE, 1973; SCHWAAB e PINTO, 2007; VALKÓ e VAJDA, 1987). Assim, os problemas de DR são tipicamente formulados de acordo com a Equação (2.42), tendo como função objetivo o estimador de mínimos quadrados ponderados, amparado nas hipóteses de independência e normalidade dos erros de medida. Desse modo, formulado com base no método de máxima verossimilhança, o problema de DR determina os valores reconciliados mais prováveis das variáveis medidas (JOHNSTON e KRAMER, 1995).

Após a resolução do problema de DR, os valores reconciliados das variáveis medidas são usados para atualizar os parâmetros do modelo. A estimação de parâmetros pode seguir a formulação apresentada na Equação (2.2) ou pode ser feita simultaneamente com a DR. O problema de reconciliação de dados com estimação simultânea de parâmetros (RDESP), conhecido na literatura como *error-in-variables method* (EVM), foi motivado pela constatação de que erros presentes nas variáveis independentes podem fornecer estimativas tendenciosas no problema de estimação de parâmetros (KIM *et al.*, 1990; VALKÓ e VAJDA, 1987). No método EVM, erros de medida em todas as variáveis são considerados no problema, que fornece estimativas consistentes com o modelo do processo, tanto para os parâmetros quanto para os dados reconciliados. O problema de RDESP, considerando erros normais e independentes, pode ser formulado como

$$\begin{aligned} \{\hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}\} := & \arg \min_{\mathbf{z}, \boldsymbol{\theta}} J_{\text{RDESP}} := (\mathbf{z}_{ms} - \mathbf{z}_m)^\top \mathbf{V}^{-1} (\mathbf{z}_{ms} - \mathbf{z}_m) \\ & \text{sujeito a } \mathbf{f}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \\ & \mathbf{h}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \\ & \mathbf{g}(\mathbf{z}, \boldsymbol{\theta}) \leq \mathbf{0} \end{aligned}, \quad (2.46)$$

cujos resultados são estimativas para os parâmetros $\hat{\boldsymbol{\theta}}$ e variáveis reconciliadas $\hat{\mathbf{z}}$. Diversos trabalhos investigaram o desempenho da formulação RDESP, que se mostrou superior a outras técnicas de estimação de parâmetros (MACDONALD e HOWAT, 1988). Além disso, algoritmos e técnicas de otimização também foram avaliadas na resolução do problema, mostrando que a qualidade dos resultados pode ser sensível a erros de linearização causados por técnicas de otimização (KIM *et al.*, 1990).

Geralmente, apenas um conjunto dentre todos os parâmetros existentes no modelo são selecionados para atualização. A estimação dos parâmetros pode ser limi-

tada pela excitação insuficiente dos dados (*i.e.*, quantidade de informação insuficiente), mas também por correlações originadas pela estrutura do modelo, em que o problema fica mal condicionado. Em situações práticas, o mau condicionamento do problema de RDESP pode ocorrer por ausência de medidas (ocasionadas, *e.g.*, por falhas em sensores) ou mudanças na condição de operação (YIP e MARLIN, 2002). QUELHAS *et al.* (2013) relataram problemas de estimabilidade no sistema RTO de um processo industrial de produção de etileno, cujo desenvolvimento foi descrito por JESUS (2011), quando a atualização dos parâmetros era feita simultaneamente com o procedimento de DR. Conforme observado na planta, tais problemas estavam relacionados à alta correlação entre certos pares de variáveis e à baixa sensibilidade da função objetivo à modificação de algumas das variáveis de entrada. Esse tipo de dificuldade é tratado na literatura como problema de identificabilidade paramétrica, podendo ser dividido em estrutural, quando a estrutura do modelo não permite a determinação de um conjunto único de valores dos parâmetros, e prático, quando a qualidade e quantidade das medidas disponíveis não permitem a estimação (BELLMAN e ÅSTRÖM, 1970; COBELLI e DISTEFANO, 1980). Por isso, diversas técnicas podem ser usadas para auxiliar na seleção dos parâmetros a serem atualizados (ALBERTON *et al.*, 2013; BOTELHO *et al.*, 2012; GRACIANO *et al.*, 2014). Nesse sentido, MILETIC e MARLIN (1998) propuseram uma metodologia *on-line* de identificação da ocorrência de mau condicionamento do problema de RDESP em conjunto com estratégias de correção, como a redefinição do conjunto de parâmetros estimados. A metodologia de diagnóstico foi introduzida como um novo elemento do sistema, usando a técnica de decomposição em valores singulares para calcular valores de condicionamento para a matriz de covariância dos parâmetros atualizados.

As formulações apresentadas anteriormente foram construídas admitindo que os erros de medida seguem distribuição Normal com média nula. Todavia, outros tipos de erros causados por eventos não aleatórios, com pouca ou nenhuma ligação com o valor medido, também podem estar presentes nos dados. Conhecidos como erros grosseiros, esses erros podem ter causa relacionada ao processo de medição (*e.g.*, falha do instrumento, falta de calibração, amostragem não representativa) ou ao processo (*e.g.*, vazamentos) (MCBRAYER e EDGAR, 1995). Em sistemas RTO, especialmente, dados que representem uma operação não estacionária podem constituir erros grosseiros (MANSOUR e ELLIS, 2008). Erros grosseiros podem ainda ser divididos em desvios sistemáticos (também chamados permanentes ou *bias*), quando os valores medidos são consistentemente mais altos ou mais baixos em relação aos valores verdadeiros, e valores espúrios (ou também *outliers*), quando o comportamento das medidas pode ser considerado anormal e não segue qualquer distribuição estatística representativa do conjunto dos dados (CHEN *et al.*, 1998a). Embora menos comum, a presença de erros grosseiros no conjunto de dados invalida

as hipóteses admitidas sobre os erros de medida e pode afetar significativamente a DR, cujos resultados deixam de ser estimativas confiáveis dos valores verdadeiros das variáveis. Portanto, erros grosseiros devem ser identificados e eliminados (ou corrigidos) do conjunto de dados (BAGAJEWICZ e JIANG, 1998).

O objetivo de técnicas de GED consiste em não apenas identificar a presença de erros grosseiros, mas também estimar sua magnitude, aplicando correções até que a falha seja retificada (MANSOUR e ELLIS, 2008). Muitos trabalhos na literatura foram dedicados à detecção de erros grosseiros. Os primeiros estudos sobre métodos de GED conduziram ao desenvolvimento de quatro testes estatísticos: teste global, teste nodal, teste de medida e teste da razão probabilística generalizada (NARASIMHAN e MAH, 1989). O problema de GED foi continuamente estudado e aplicado para identificar erros grosseiros individualmente. Para a detecção de múltiplos erros grosseiros, metodologias apropriadas foram desenvolvidas, envolvendo basicamente três estratégias clássicas: eliminação serial, compensação serial e compensação simultânea (NARASIMHAN e JORDACHE, 1999). Na estratégia da eliminação serial, os erros grosseiros são detectados sequencialmente, um a um. Após a remoção do erro grosseiro em cada passo de busca, realiza-se novamente o procedimento de DR, até que nenhum erro grosseiro seja encontrado. No método da compensação serial, os erros grosseiros também são detectados sequencialmente, um a um; diferentemente do primeiro, uma vez identificadas as magnitudes dos erros, as medidas são corrigidas, ao invés de eliminadas, continuando até que todos os erros grosseiros tenham sido identificados e corrigidos. A estratégia simultânea, por sua vez, busca detectar todos os erros grosseiros presentes nos dados simultaneamente, em uma única iteração. Outras abordagens gerais para detecção de erros grosseiros incluem técnica PCA (TONG e CROWE, 1995), redes neuronais (TERRY e HIMMELBLAU, 1993) e análise de *clusters* (CHEN e ROMAGNOLI, 1998). É importante ressaltar que a detecção de erros grosseiros admite que o modelo é válido, de modo que as diferenças entre os dados e valores calculados causados por erros de modelagem podem ser confundidos com erros grosseiros, os quais podem não ser facilmente removidos por exclusão dos valores medidos (MARLIN e HRYMAK, 1997).

Outra abordagem muito interessante consiste no uso de funções objetivo que correspondam a funções de verossimilhança capazes de considerar tanto os erros aleatórios, quanto os erros grosseiros. TJOA e BIEGLER (1991) foram os primeiros a propor esse tipo de estratégia, que remove a suposição de que os erros de medida seguem distribuição Normal e permite que a GED seja combinada ao problema de DR (ou RDESP). Especificamente, TJOA e BIEGLER (1991) propuseram o uso da distribuição bivariada, conhecida como Normal contaminada, em lugar da função de mínimos quadrados ponderados, mostrando que os erros grosseiros presentes no conjunto de dados são adequadamente substituídos por valores reconciliados, sem

precisar de um procedimento iterativo. Posteriormente, JOHNSTON e KRAMER (1995) propuseram um método similar para retificação de dados de sistemas estáticos, com base na Teoria de Probabilidade e no método de máxima verossimilhança, introduzindo o uso de estimadores robustos na função objetivo do problema de DR. Outra contribuição do trabalho de JOHNSTON e KRAMER (1995) consistiu no desenvolvimento de uma técnica que faz uso dos dados históricos da planta para determinar a probabilidade *a priori* dos estados. Sendo capaz de capturar as relações entre as variáveis (*e.g.*, balanços de massa e energia, e relações termodinâmicas) e suas correlações estatísticas, a técnica de retificação de dados apresentada permite obter estimativas de variáveis reconciliadas sem um modelo do processo.

A retificação de dados por meio da estatística robusta pode ser aplicada a ambos os tipos de problema, tanto estáticos, quanto dinâmicos; tanto de DR, quanto de RDESP; e sua formulação segue o mesmo procedimento apresentado anteriormente. Nesse caso, a função objetivo $J_{DR}(\cdot)$ ou $J_{RDESP}(\cdot)$ das Equações (2.42) e (2.46), respectivamente, corresponde a um estimador robusto. Em vista da viabilidade e melhor desempenho com o uso do estimador robusto relatado por JOHNSTON e KRAMER (1995), diversos tipos de estimadores robustos e seus desempenhos no contexto da DR foram avaliados (ALBUQUERQUE e BIEGLER, 1996; ARORA e BIEGLER, 2001; ÖZYURT e PIKE, 2004; PRATA *et al.*, 2008a, 2010), demonstrando o potencial da aplicação da estatística robusta nesses problemas (HUBER e RONCHETTI, 2009).

ÖZYURT e PIKE (2004) compararam o desempenho de diferentes estimadores robustos na DR de sistemas estáticos para cinco casos da literatura e dois casos de processos industriais, a saber: a planta de ácido sulfúrico, primeiramente estudada por CHEN *et al.* (1998b); e um processo de alquilação em uma refinaria americana. Os resultados indicaram que os estimadores de Cauchy e Hampel apresentaram melhores desempenhos para DR e GED. PRATA *et al.* (2008a) apresentaram uma análise comparativa entre doze estimadores robustos para um problema de reconciliação de dados dinâmico não linear, mostrando que o estimador robusto baseado na distribuição de Welsch apresentou o melhor desempenho global, usando como critério a redução total do erro (válido apenas para exemplos simulados). Adicionalmente, o trabalho de PRATA *et al.* (2008a) representou o primeiro estudo de avaliação de desempenho de estimadores robustos em sistemas dinâmicos. PRATA *et al.* (2009b) apresentou uma extensa lista de aplicações industriais de técnicas de DR publicadas na literatura, enquanto PRATA *et al.* (2010) citou vários trabalhos que usaram estimadores robustos para DR e GED.

No contexto da otimização estática em tempo real, técnicas de retificação de dados já foram relatadas em sistemas RTO industriais. Dentre eles, destaca-se o trabalho de CHEN *et al.* (1998b), que possivelmente corresponde ao primeiro trabalho a

incorporar técnicas combinadas de DR e GED em sistemas RTO e, adicionalmente, avaliar tais métodos com dados industriais. No estudo, o desempenho da distribuição Normal contaminada e da distribuição Lorenziana foram avaliadas com dados industriais de uma planta de ácido sulfúrico, apontando que a distribuição Normal contaminada é adequada para descrever pequenos erros grosseiros, enquanto a distribuição Lorenziana é indicada quando erros grosseiros superam valores correspondentes a 30 vezes o desvio padrão. Os trabalhos de LID *et al.* (2002) e FABER *et al.* (2006, 2007) também utilizam a estratégia combinada de DR e GED. LID *et al.* (2002) utilizaram a distribuição Normal contaminada na DR de um sistema RTO implementado em uma rede de vinte trocadores de calor em uma refinaria da Statoil. FABER *et al.* (2006) implementaram um sistema RTO para minimização da carga térmica em uma planta piloto de purificação de gás de coqueria, cuja DR e GED se baseia no estimador robusto apresentado por LI *et al.* (2000) e MINGFANG *et al.* (2000). De forma similar, FABER *et al.* (2007) utilizaram o mesmo estimador para minimizar o efeito negativo dos erros grosseiros sobre as estimativas das variáveis e parâmetros em um sistema RTO de um processo industrial de purificação de gás. Outros trabalhos também relatam aplicações de DR em sistemas RTO industriais (BUSSANI *et al.*, 1995; CHIARI *et al.*, 1997; LEE *et al.*, 1998; PIERUCCI *et al.*, 1996), mas não fornecem detalhes sobre o uso de técnicas GED.

Em suma, erros grosseiros podem ser detectados por avaliações seriais e métodos estatísticos. Métodos de avaliação removem valores que estão fora de limites admissíveis para a variável do processo, sendo um procedimento bastante usado em sistemas industriais. Todavia, esses testes não são capazes de detectar erros sistemáticos, como *bias* de instrumentos e vazamentos. Métodos estatísticos, por sua vez, são adequados para detectar esse tipo de comportamento e correspondem ao método mais efetivo para detectar erros grosseiros, embora sejam mais sofisticados, exigindo um modelo mais acurado da planta (CHEN *et al.*, 1998b).

Revisões abrangentes sobre as técnicas envolvidas na retificação de dados podem ser encontradas em diversos trabalhos da literatura (CROWE, 1996; HLAVÁČEK, 1977; HODGE e AUSTIN, 2004; KADLEC *et al.*, 2009; NARASIMHAN e JORDACHE, 1999; PRATA, 2009; ROMAGNOLI e SÁNCHEZ, 2000).

2.4.4 Otimização Econômica

Após a retificação dos dados medidos e a atualização dos parâmetros, a próxima etapa consiste na resolução da otimização econômica, que corresponde ao problema de programação matemática representado pela Equação (2.3). Usando o modelo atualizado da planta, a otimização econômica está sujeita às diversas restrições impostas por equipamentos, especificações de qualidade dos produtos, normas de segurança

e ambientais, bem como restrições econômicas impostas pela administração da produção (ENGELL, 2007). O resultado obtido é uma estimativa de um conjunto de valores dos graus de liberdade disponíveis que promovem o melhor desempenho da planta (DARBY *et al.*, 2011).

Ao otimizar a operação de uma planta, buscam-se benefícios a partir da melhoria de desempenho do processo, o qual pode ser alcançado a partir de: aumento no rendimento de produtos de interesse (ou redução dos rendimentos de produtos secundários e não desejados), redução no consumo de energia, aumento nas taxas de produção e tempos mais longos entre paradas da planta. Outras melhorias de desempenho possíveis podem ser representadas por redução de custos de manutenção, menor desgaste em equipamentos, ou ainda melhor utilização da mão-de-obra e capacidade instalada (EDGAR *et al.*, 2001). Além de melhorias concretas, a otimização pode proporcionar benefícios intangíveis, como maiores velocidade de tomada de decisão e confiabilidade na solução de problemas. A definição da função objetivo econômica varia de acordo com o processo e pode ser definida como qualquer melhoria citada anteriormente, como redução do consumo de energia (LID *et al.*, 2002). Não obstante, é conveniente em muitos processos quantificar as variadas formas de melhoria de desempenho, definindo um objetivo para o problema de otimização por meio de um índice econômico.

Um índice comumente usado como função objetivo em sistemas RTO corresponde ao lucro bruto operacional, *i.e.*, a diferença entre a receita das vendas do produto e o custo dos bens e serviços envolvidos na sua produção (DARBY *et al.*, 2011). Uma típica função objetivo econômica que modela o lucro bruto operacional corresponde a (NAYSMITH e DOUGLAS, 1995)

$$\begin{aligned}
 J^{opt} := \text{lucro} = & \text{[receita de venda dos produtos]} \\
 & - \text{[custos com matéria-prima]} \\
 & - \text{[custos com utilidades]} \\
 & + \text{[outros efeitos econômicos variáveis]} \quad .
 \end{aligned} \tag{2.47}$$

Os valores econômicos usados na otimização irão mudar, embora geralmente a uma taxa mais lenta que a frequência normal do sistema RTO. Tais dados econômicos tipicamente são oriundos de uma variedade de fontes. Alguns são inseridos manualmente, outros são transmitidos pelas camadas de planejamento e programação da produção, e há também aqueles que são determinados pelo mercado. Diante disso, parâmetros necessários ao modelo econômico da operação usado na função objetivo constituem mais uma fonte de incerteza no sistema RTO (MARLIN e HRYMAK, 1997; ZHANG *et al.*, 2002). Todas essas mudanças ocorrem de maneira assíncrona e é necessário garantir que o conjunto usado na otimização seja consistente (WHITE,

1997). Preços de matérias-primas e produtos devem ser escolhidos de maneira apropriada. Uma dificuldade bastante frequente consiste na precificação de correntes intermediárias do processo. Algumas abordagens utilizam como base o preço de mercado e consideram esse valor diretamente ou descontado dos custos associados ao processamento, enquanto outras consistem em adicionar modelos simples ao sistema RTO para abranger um envelope maior do processo e capturar os preços de entrada e/ou saída sem a necessidade de definir preços intermediários (DARBY *et al.*, 2011). Além disso, com alguma frequência pode ser necessário considerar diferentes preços para um mesmo material ou corrente (*e.g.*, quando há um preço de contrato e um preço de mercado), especialmente quando o uso de um preço médio não é aceitável. Esse tipo de situação merece atenção, pois pode causar descontinuidade na função objetivo e mudar a natureza do problema de otimização (NAYSMITH e DOUGLAS, 1995).

BAILEY *et al.* (1993) investigaram a melhoria de desempenho alcançada por um sistema RTO em um estudo simulado com base no processo de hidrocraqueamento da refinaria de Sarnia, descrevendo com detalhes a definição da função objetivo empregada, que incluía consumo de utilidades, preço de produtos químicos, qualidade da gasolina e poder calorífico das correntes destinadas ao combustível. EDGAR *et al.* (2001) discutiram com detalhes o desenvolvimento da função objetivo, bem como métodos para aferição de rendimento e estimativas de custo.

Além da definição da função objetivo, a formulação do problema de otimização econômica exige outros cuidados para uma resolução adequada e confiável. Questões como dimensão e faixa de variação das variáveis, continuidade das funções e limitação do passo a cada iteração podem ser importantes. A limitação da máxima variação das variáveis de decisão previne possíveis problemas ocasionados por mudanças bruscas determinadas pela otimização. Na situação em que novas condições de operação fossem muito diferentes da condição mais recente, uma opção interessante poderia envolver um segundo problema de otimização, para determinar a melhor trajetória na direção da condição ótima (MARLIN e HRYMAK, 1997). Algumas implementações resolvem automaticamente um caso adicional, que avalia os resultados do otimizador sem as limitações de passo (DARBY *et al.*, 2011).

2.4.5 Algoritmos de Otimização

O sistema RTO com base na abordagem em duas etapas é constituído pela resolução de sucessivos problemas de programação matemática, em quase na totalidade das vezes problemas NLP, sujeitos a restrições de igualdade e desigualdade. Por isso, embora não esteja explícito, a existência de um método numérico capaz de encontrar o conjunto correto de valores das variáveis de decisão em ambos os problemas (2.2) e

(2.3), bem como (2.42) e (2.46), assume elevada importância. Diversos algoritmos e ferramentas para resolução de programas não lineares estão disponíveis na literatura (BIEGLER e GROSSMANN, 2004; FLETCHER, 2000; PIBOULEAU *et al.*, 1999).

Métodos determinísticos de otimização, em particular o método SQP, são amplamente empregados em sistemas RTO, tanto na literatura acadêmica quanto em *softwares* comerciais (QUELHAS *et al.*, 2013; ROTAVA e ZANIN, 2005). Algoritmos SQP podem ser facilmente incorporados em simuladores modulares, tendo a vantagem de serem diretamente aplicados em fluxogramas construídos em simuladores comerciais (TRIERWEILER, 2014). O algoritmo SQP converge para as restrições de igualdade e desigualdade juntamente com as condições de otimalidade e requer um número relativamente menor de avaliações da função objetivo. No entanto, o método SQP é um método baseado em gradiente, em que a obtenção das derivadas ocorre geralmente por diferenciação numérica, o que aumenta o esforço computacional e o tempo de convergência. Discussões detalhadas sobre a técnica podem ser encontradas em ALKAYA *et al.* (2001) e BIEGLER (2010).

Embora o método SQP possua limitações bem conhecidas (YANG, 2010), a literatura sobre RTO discute pouco o impacto de problemas numéricos no desempenho de sistemas RTO ou seus componentes. Conforme sugerido por QUELHAS *et al.* (2013), uma causa possível da falta de discussões sobre a questão é a ainda mais ausente discussão sobre implementações mal sucedidas. Poucos trabalhos mencionam abertamente erros e falhas relacionadas aos algoritmos de otimização (CUBILLOS *et al.*, 2007; GOLSHAN *et al.*, 2008). GOLSHAN *et al.* (2008), em especial, mostraram como algoritmos SQP tendem a convergir prematuramente para mínimos locais e a falhar quando etapas intermediárias de cálculo resultam em valores das variáveis de decisão fora da região viável. Além disso, o método SQP é muito sensível à estimativa inicial e não converge se esse valor estiver fora da região viável. De acordo com GOLSHAN *et al.* (2008), que avaliaram o algoritmo SQP para o estudo simulado de um sistema RTO implementado no problema Tennessee Eastman, essa situação ocorreu em aproximadamente 40% das vezes. Analisando uma implementação de sistema RTO em uma unidade industrial de destilação de petróleo pelo *software* Aspen Plus[®] 7.1, que usa o método SQP como algoritmo de otimização padrão (ASPEN PLUS USER GUIDE, 1998), QUELHAS (2013) mostrou que o RTO obteve convergência do algoritmo de estimação/reconciliação 59,7% das vezes. O método Gauss-Newton padrão, que pode ser usado para resolver problemas intermediários do algoritmo SQP, pode igualmente não convergir por mau condicionamento da matriz Hessiana e/ou cálculo impreciso das derivadas, especialmente quando calculadas numericamente (PRATA *et al.*, 2009b). Adicionalmente, é comum admitir que uma boa estimativa inicial seja suficiente para garantir bom desempenho em técnicas baseadas no método de Gauss-Newton, mas isso não é sem-

pre verdade. Avaliando o método Gauss-Newton padrão para a resolução em tempo real do problema de reconciliação de dados dinâmica não linear em um processo de polimerização, PRATA *et al.* (2009b) concluíram que falhas de convergência, mesmo sendo pouco frequentes, podem ser inaceitáveis nesse tipo de aplicação. Quaisquer falhas que ocorram com o algoritmo de otimização, seja durante a atualização dos parâmetros, seja durante a otimização econômica, irá afetar o desempenho do sistema RTO.

Embora a informação contida nas derivadas seja o grande responsável pelo desempenho de algoritmos de otimização determinísticos, nem sempre as derivadas estão disponíveis. Em outras situações, a obtenção das derivadas impõem limitações, como quando avaliações da função objetivo são muito custosas e dificultam a obtenção do gradiente e Hessiana. É importante considerar também a precisão com a qual as derivadas são obtidas; nos casos em que valores da função objetivo estejam contaminados com ruído, a aproximação das derivadas por diferenças finitas pode conter pouca informação. Adicionalmente, o cálculo das derivadas em funções descontínuas pode ocasionar problemas de estabilidade numérica, notadamente ao empregar estimadores robustos na DR (PRATA *et al.*, 2009a). Nesses casos, métodos NLP baseados em gradientes podem não ser adequados e algoritmos sem derivadas devem ser considerados (CONN *et al.*, 2009).

Outra classe de métodos que podem ser considerados são métodos de busca aleatória (métodos de otimização heurísticos ou métodos não determinísticos). A essa classe pertencem métodos como algoritmo genético, *simulated annealing* e métodos baseados em populações. Embora algoritmos não determinísticos tendam a ser computacionalmente custosos, seu caráter global assintótico pode ser garantido. Devido à maior capacidade de exploração da região viável, os resultados produzidos tendem a desviar menos do ponto ótimo e, geralmente, esses algoritmos não apresentam as desvantagens possíveis em técnicas baseadas derivadas, sendo menos sensíveis à estimativa inicial e à existência de mínimos locais e menos suscetíveis à divergência (SCHWAAB *et al.*, 2008; YIQING *et al.*, 2007). Dentre eles, a técnica enxame de partículas (*particle swarm optimization* - PSO) (KENNEDY e EBERHART, 1995) tem sido aplicada com sucesso em casos industriais.

PRATA *et al.* (2009b) aplicaram o algoritmo PSO em uma implementação em tempo real do problema de DR dinâmica não linear em processo industrial de polimerização de propileno. Com base no algoritmo PSO, PRATA *et al.* (2010) foram os primeiros a relatar uma aplicação industrial do procedimento de RDESP não linear utilizando o estimador de Welsch para simultânea detecção de erros grosseiros. O procedimento incluiu um modelo fenomenológico do processo baseado em balanços detalhados de massa e energia, constituindo um sistema de equações algébrico-diferenciais, e foi implementado em tempo real em um processo industrial

de polimerização de propileno. JESUS (2011) implementou um sistema RTO em um processo industrial de produção de etileno, incluindo procedimentos de reconciliação e otimização econômica, cujo objetivo foi definido como maximização do lucro bruto. Tanto o problema de DR, quanto o problema de otimização econômica, foram resolvidos por meio do algoritmo PSO, cujo tempo necessário de resolução do problema permaneceu dentro dos limites que permitem a execução do sistema RTO no processo. De acordo com os trabalhos mencionados, parte do sucesso do uso da técnica PSO em aplicações em tempo real se deve à significativa ausência de falhas do algoritmo, uma vez que não depende do cálculo de derivadas, além da simplicidade de implementação e confiabilidade dos resultados produzidos. Métodos híbridos que combinam o PSO a diferentes abordagens podem ser encontrados na literatura (VAZ e VICENTE, 2007; YIQING *et al.*, 2007). Uma revisão sobre os desenvolvimentos da técnica PSO pode ser encontrada em POLI *et al.* (2007).

Além de eventuais falhas que possam ocorrer com o algoritmo de otimização, a abordagem em duas etapas está susceptível a produzir valores equivocados para o conjunto de parâmetros atualizados e variáveis de decisão, afetando o desempenho do processo. Com efeito, a influência do algoritmo de otimização sobre os resultados é outro tópico que tem recebido pouca atenção na literatura. Por meio de um exemplo simples que emprega as técnicas determinísticas SQP e o algoritmo de busca Nelder-Mead (NELDER e MEAD, 1965), QUELHAS *et al.* (2013) mostraram como a perda de desempenho pode estar relacionada ao algoritmo ou técnica de otimização, bem como à estimativa inicial. Apesar dos argumentos mencionados, que favorecem algoritmos não determinísticos, sua influência no sistema RTO pode ser ainda mais expressiva. Por meio do mesmo exemplo usado para as técnicas determinísticas, QUELHAS *et al.* (2013) avaliaram os métodos heurísticos PSO e algoritmo genético. Além da já observada dependência dos resultados com a estimativa inicial e a incapacidade de eleger qual método foi superior, foram ressaltadas a influência significativa da parametrização dos métodos nos resultados, bem como a natureza estocástica transmitida aos resultados. O exemplo é muito útil para mostrar que o algoritmo de otimização usado também é um elemento do sistema RTO que afeta o desempenho do processo e, portanto, também deve ser considerado ao avaliar estruturas candidatas de sistemas RTO. Nesse sentido, além de não ser possível determinar *a priori* o desvio existente entre os resultados obtidos por diferentes algoritmos, não é possível garantir que um método de otimização terá melhor desempenho que outro (KÖPPEN, 2004; WOLPERT e MACREADY, 1997).

2.4.6 Aspectos das Incertezas e seu Efeito na Solução do Sistema RTO

Tendo resolvido o problema de otimização econômica, o novo conjunto de *setpoints* está disponível para ser transmitido à camada de controle supervisão, que usará esses valores como *setpoints* para determinar movimentos apropriados das variáveis manipuladas (ENGELL, 2007). Uma discussão sobre a implementação de valores ótimos estacionários por controles preditivos lineares pode ser encontrada em RAO e RAWLINGS (1999). Todavia, é prudente que os resultados sejam avaliados antes da sua implementação, mesmo que tal avaliação seja deixada a cargo da subjetividade e experiência dos operadores da planta ou de um supervisor do processo (VAN WIJK e POPE, 1992).

Visto que o processo consiste em uma fonte de incertezas, sendo a causa de perturbações desconhecidas e não controladas, existe a possibilidade de que o estado da planta tenha mudado durante o cálculo do novo ponto ótimo de operação. Assim, testes pertinentes avaliam se mudanças ocorreram na planta durante a execução do sistema RTO, determinando: i) se a planta permanece em estado estacionário e se essa condição equivale à mesma determinada no início da execução do sistema, ii) se os graus de liberdade definidos pela otimização econômica continuam disponíveis para manipulação, e iii) se os limites das variáveis de otimização permanecem os mesmos (MARLIN e HRYMAK, 1997). Se uma mudança significativa ocorre, a execução do RTO deve ser reiniciada.

Além das incertezas de processo, a variabilidade introduzida por meio dos erros de medição, erros de modelagem e incertezas dos parâmetros econômicos sugerem a necessidade de uma metodologia que permita avaliar se os resultados calculados são significativos, seguros e coerentes com a realidade da planta, permitindo uma análise crítica dos valores propostos pelo otimizador. A formulação determinística, tipicamente apresentada para o problema de otimização econômica, dificulta a percepção de que todos os resultados produzidos pelo sistema de RTO possuem natureza estocástica, sendo estimativas para os verdadeiros valores das variáveis. Isso ocorre em virtude da propagação das incertezas de natureza aleatória contidas nas medições dos resultados de cada componente do sistema RTO. Reconhecer essa característica do sistema direciona a atenção para a validade estatística dos resultados \mathbf{u}_{k+1}^* propostos no instante de tempo k . Em outras palavras, mesmo que o resultado \mathbf{u}_{k+1}^* calculado pelo sistema RTO seja numericamente diferente dos valores em vigor das variáveis manipuladas \mathbf{u}_k , ambas podem ser realizações diferentes de um mesmo evento aleatório.

Outra questão pertinente diz respeito à garantia de viabilidade das soluções, *i.e.*, avaliar se a solução \mathbf{u}_{k+1}^* viola uma ou mais restrições, cujo efeito sobre o índice de

desempenho é certo, mas nem sempre explícito. Sabendo que as restrições podem envolver limites críticos de segurança, a violação das restrições pode promover condições de risco, tanto em relação à vida dos operadores, quanto ao dano irreversível de equipamentos. Em todos os casos, a incerteza paramétrica do modelo é admitida e quantificada. Considerando uma única restrição $g_{p,j}$ (incerta) devido ao vetor de parâmetros incertos θ , admitido dentro do espaço $\Theta \subset \mathbb{R}^{n_\theta}$, a viabilidade da solução \mathbf{u}_{k+1}^* corresponde a satisfazer (BUNIN *et al.*, 2013b):

$$g_{p,j}(\mathbf{u}_{k+1}^*) \leq 0 \quad , \quad (2.48)$$

garantindo:

$$\sup_{\theta \in \Theta} g_j(\mathbf{u}_{k+1}^*, \theta) \leq 0 \quad , \quad (2.49)$$

em que $g_j : \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ é um modelo paramétrico e $\sup g_j$ corresponde ao valor superior (ou supremo) da função em Θ .

Os três métodos conhecidos para lidar com as incertezas no sistema RTO incluem: i) análise dos resultados, voltada à utilidade da solução ao objetivo de melhoria de desempenho econômico; ii) estratégias de imposição das restrições, voltados para a viabilidade da solução; e iii) otimização sob incertezas, *i.e.*, programação estocástica e otimização robusta, que contemplam tanto a viabilidade da solução quanto sua utilidade para a melhoria de desempenho econômico. As duas últimas abordagens envolvem formas específicas de lidar com as incertezas no problema de otimização, em que algum nível de conservadorismo é introduzido para garantir a viabilidade dos resultados, em lugar de analisar os resultados após a otimização (ELGSÆTER *et al.*, 2010).

2.4.6.1 Análise dos Resultados

Apesar da validade estatística dos resultados já ter sido abordada anteriormente (BOX e DRAPER, 1969) e considerada em sistemas RTO (KONINCKX, 1988), a questão ganhou notoriedade apenas mais recentemente. A estratégia de análise de resultados em sistemas RTO foi proposta por MILETIC e MARLIN (1998, 1996) como um procedimento baseado em técnicas de controle estatístico de processos e análise de sensibilidade, para avaliar se os resultados do sistema RTO representam mudanças estatisticamente significativas com relação à condição de operação mais recente ou se tal resultado é produto da transmissão de variabilidade aleatória ao longo dos componentes do sistema. Nessa metodologia, a análise é direcionada à matriz de covariância das estimativas das variáveis de decisão, que é resultado da propagação das incertezas de medição (*i.e.*, matriz de covariâncias dos dados medidos) ao longo dos procedimentos de estimação de parâmetros e de otimização

econômica.

Considerando problemas de otimização econômica sujeitos apenas a restrições de igualdade, a estratégia efetua a análise no espaço reduzido das variáveis independentes, aplicando o teste estatístico multivariável T^2 de Hotelling para verificar se os novos valores das variáveis de decisão diferem significativamente dos mais recentes. Esse teste é então usado para aceitar ou rejeitar o novo *setpoint* calculado. A metodologia foi posteriormente estendida por ZHANG *et al.* (2001b) para considerar também o problema com restrições de desigualdade, aplicando a estatística T^2 em um segundo estágio aos multiplicadores de Lagrange associados com as regiões ativas da região de confiança, para determinar se essas restrições estão realmente restringindo a melhoria de rendimento da operação. ZAFIRIOU e CHENG (2004) implementaram o componente de análise dos resultados em um sistema RTO baseado na estratégia alternativa à abordagem em duas etapas proposta por CHENG e ZAFIRIOU (2000).

Os métodos de análise dos resultados se mostraram efetivos para redução da influência do ruído de medição na solução do sistema RTO. Há, no entanto, diversas hipóteses importantes admitidas na dedução do método e que restringem sua aplicação, ou ao menos restringem a validade das análises aos problemas que satisfazem tais hipóteses. Admite-se no método que os erros de medição seguem distribuição normal e que as transformações matemáticas que expressam a propagação dos erros sejam lineares, de modo que as incertezas das variáveis de decisão também serão descritas pela distribuição normal. Adicionalmente, a abordagem de análise de resultados é um procedimento reativo, que não incorpora ao otimizador a capacidade de lidar com as incertezas, não considerando explicitamente as incertezas. Além disso, apesar do interesse legítimo em transmitir à camada de controle apenas as soluções estatisticamente significativas, a metodologia não aborda a influência das incertezas na viabilidade dos resultados. Dados os possíveis efeitos das incertezas de medição e de erros de modelagem, o risco de que a solução estimada viole restrições aumenta, na medida em que mais restrições se tornem ativas, isto é, $g_{p,i}(\mathbf{u}, \boldsymbol{\theta}) \leq 0$. Assim, considerando que as soluções do problema de otimização costumam situar-se próximas ou sobre as restrições quanto mais linear for a estrutura do problema de otimização, esse risco pode apresentar-se sobremodo grande para ser desprezado.

2.4.6.2 Estratégias de Imposição das Restrições

A outra abordagem de lidar com as incertezas é voltada unicamente à garantia da viabilidade das soluções, consistindo em estratégias para impor a satisfação das restrições.

Uma estratégia simples para considerar o problema de viabilidade das soluções na presença de incertezas consiste no uso de margens de segurança na definição das res-

trições (GOVATSMARK e SKOGESTAD, 2005; NARRAWAY e PERKINS, 1991). Nessa metodologia, constantes β são adicionadas às restrições do problema de otimização econômica, de forma a promover um afastamento (ou margem de segurança) que auxilie o algoritmo de otimização a produzir resultados viáveis (DE HENNIN, 1994; LOEBLEIN e PERKINS, 1998; LOEBLEIN *et al.*, 1999). Uma restrição modificada por uma constante de afastamento β é escrita como:

$$g_j(\mathbf{u}, \boldsymbol{\theta}) + \beta \leq 0, \quad \beta > 0 \quad . \quad (2.50)$$

Na prática, essa estratégia equivale a definir restrições mais rígidas que as necessárias para uma operação segura e rentável e, por isso, a viabilidade imposta dessa forma vem ao custo de uma provável redução do índice de desempenho por soluções subótimas (QUELHAS *et al.*, 2013). A determinação do valor das constantes de afastamento envolve a avaliação da variabilidade das restrições em função dos *setpoints* calculados, das incertezas de medição e paramétricas e dos erros de modelagem, sendo calculados a partir da otimização do modelo linearizado, assegurando, para dado nível de confiança, que os limites das restrições não serão violados (LOEBLEIN e PERKINS, 1998). Apesar das constantes de afastamento representarem um mecanismo ativo de garantia de resultados viáveis incorporado ao procedimento de otimização, sua principal desvantagem consiste na falta de acurácia dos resultados, promovida pela baixa frequência de atualização das constantes de afastamento, que, além de serem determinadas *off-line*, admitem que o conjunto de restrições ativas permaneça constante (ZHANG *et al.*, 2002).

Outra estratégia desse tipo consiste em limitar as variáveis de decisão em cada iteração, como restrições no tamanho do passo (CHENG e ZAFIRIOU, 2000; GAO e ENGELL, 2005b):

$$\mathbf{u}_k - \Delta \mathbf{u}_{max} \leq \mathbf{u}_{k+1} \leq \mathbf{u}_k + \Delta \mathbf{u}_{max} \quad . \quad (2.51)$$

Ambas podem ser abordagens conservadoras, sendo fundamentalmente abordagens *ad hoc* que não preenchem as garantias de viabilidade desejadas. Em um estudo simulado usado para o desenvolvimento de uma metodologia para otimização da produção de óleo e gás em plataformas *offshore*, ELGSÆTER *et al.* (2010) combinaram as metodologias de análise dos resultados e restrição do tamanho do passo no sistema RTO para lidar com as incertezas. A análise de resultados desenvolvida julgou a implementação do novo *setpoint* sugerido pela otimização por meio de um teste estatístico que não admite hipóteses sobre a distribuição de probabilidades da função objetivo, mas estimava tal distribuição usando a técnica de Monte Carlo.

Uma terceira alternativa de impor a viabilidade das restrições foi discutida por BUNIN *et al.* (2011), que propuseram uma abordagem simples, mas robusta, usando

um filtro de ganho variável na lei padrão de filtragem das variáveis de decisão:

$$\mathbf{u}_{k+1} = \mathbf{u}_k + K_k (\mathbf{u}_{k+1}^* - \mathbf{u}_k) \quad , \quad (2.52)$$

em que o ganho K_k é definido na iteração k de modo a garantir $g_{p,j}(\mathbf{u}_{k+1}) < 0$ dado que $g_{p,j}(\mathbf{u}_k) < 0$. Assim, limites adaptativos para o ganho K_k do filtro foram determinados a cada iteração, de modo a garantir a viabilidade para qualquer método RTO (BUNIN *et al.*, 2013b).

2.4.6.3 Otimização sob Incertezas

A terceira abordagem para lidar com as incertezas no sistema RTO contempla tanto a viabilidade da solução quanto a sua utilidade para a melhoria de desempenho econômico, correspondendo à única abordagem que considera explicitamente as incertezas para a solução do problema de otimização.

Lembrando que tanto os parâmetros $\boldsymbol{\theta}$ quanto as variáveis de decisão \mathbf{u} são inferências dos valores verdadeiros e desconhecidos dessas variáveis com base nos dados medidos, fica evidente que esses elementos também são variáveis aleatórias. Assim, restrições e função objetivo do problema de otimização econômica correspondem a transformações aplicadas sobre a média dessas variáveis aleatórias, *i.e.*,

$$\begin{aligned} J^{opt} &:= \phi(E[\mathbf{u}], E[\boldsymbol{\theta}]) \\ g_j(E[\mathbf{u}], E[\boldsymbol{\theta}]) &\leq 0 \quad , \end{aligned} \quad (2.53)$$

assumindo, portanto, um caráter igualmente aleatório. Se as funções $\phi(\cdot)$ e $\mathbf{g}(\cdot)$ consistirem apenas de transformações lineares, então a seguinte expressão é válida:

$$\begin{aligned} J^{opt} &:= E[\phi(\mathbf{u}, \boldsymbol{\theta})] \\ E[g_j(\mathbf{u}, \boldsymbol{\theta})] &\leq 0 \quad . \end{aligned} \quad (2.54)$$

Assim, resolver o problema (2.3) com base nas Equações (2.54) por meio de um algoritmo determinístico (*e.g.*, SQP) garante apenas que cada restrição seja atendida em metade dos casos, o que geralmente não é considerado satisfatório. Além disso, a formulação resultante despreza a distribuição de probabilidades das funções, bem como a faixa de valores possíveis, e não limita a quantidade esperada de violações. Adicionalmente, caso as funções $\phi(\cdot)$ e $g_j(\cdot)$ correspondam a transformações não lineares, não há formulação analítica prévia que possa indicar as consequências da resolução determinística na otimalidade e viabilidade da solução, comprometendo a confiabilidade dos resultados sugeridos pelo sistema RTO (QUELHAS, 2013).

A abordagem mais rigorosa para considerar explicitamente as incertezas no problema de otimização e impor a viabilidade do problema descrito pela Equação (2.1)

corresponde à otimização sob incertezas, representada pela programação estocástica e pela otimização robusta, que requerem um modelo das restrições e são geralmente aplicados nas fases iniciais de projeto (BEYER e SENDHOFF, 2007). A abordagem padrão é a programação estocástica, que supõe uma distribuição de probabilidade para os possíveis valores de $\boldsymbol{\theta}$ e tentam satisfazer a Equação (2.49) com um nível de confiança, enquanto a otimização robusta, seguindo a abordagem do pior caso, busca essencialmente o mesmo objetivo, mas sem admitir hipóteses sobre a distribuição de $\boldsymbol{\theta}$ e considerando o nível de confiança 100% (BEN-TAL *et al.*, 2009). A otimização robusta é geralmente considerada uma abordagem muito conservadora, uma vez que ela considera todos os possíveis desvios, enquanto a programação estocástica é afetada pelas hipóteses necessárias á formulação do problema.

Reformular o problema de otimização econômica como um problema de otimização sob incertezas implica escrever as restrições levando em conta, explicitamente, sua natureza probabilística e a probabilidade de violação das restrições sob dado nível de confiança α . Dessa maneira, as restrições podem ser definidas como:

$$\text{JPC: } P [g_j (\mathbf{u}, \boldsymbol{\theta}) \leq 0, j = 1, \dots, n_g] \leq \alpha \quad , \quad (2.55)$$

ou

$$\text{IPC: } P [g_j (\mathbf{u}, \boldsymbol{\theta}) \leq 0] \leq \alpha_j, j = 1, \dots, n_g \quad . \quad (2.56)$$

A função objetivo também deve ser redefinida, sendo que a formulação comumente adotada consiste em uma combinação linear entre a esperança matemática e a variância da função objetivo (DARLINGTON *et al.*, 1999),

$$J^{opt} := w_1 E [\phi (\mathbf{u}, \boldsymbol{\theta})] + w_2 Var [\phi (\mathbf{u}, \boldsymbol{\theta})] \quad . \quad (2.57)$$

Assim, o problema de otimização econômica pode ser rescrito como um problema de otimização sob incertezas:

$$\begin{aligned} \min_{\mathbf{u}} \quad & J^{opt} := w_1 E [\phi (\mathbf{u}, \boldsymbol{\theta})] + w_2 Var [\phi (\mathbf{u}, \boldsymbol{\theta})] \\ \text{sujeito a} \quad & \mathbf{g} (\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}) \leq \mathbf{0} \\ & \text{ou,} \\ & P [g_j (\mathbf{u}, \boldsymbol{\theta}) \leq 0, j = 1, \dots, n_g] \leq \alpha \quad , \\ & \text{ou ainda,} \\ & P [g_j (\mathbf{u}, \boldsymbol{\theta}) \leq 0] \leq \alpha_j, j = 1, \dots, n_g \quad . \end{aligned} \quad (2.58)$$

A formulação apresentada na Equação (2.58) é genérica, sendo tradicionalmente classificada como programação estocástica ou otimização robusta, a depender das características da função objetivo e restrições, e varia de acordo com a estratégia exigida pelo problema para contemplar os efeitos das incertezas. Dentre as caracte-

rísticas mais relevantes, destacam-se:

- A função objetivo apresentada na Equação (2.57) é geralmente usada em problemas de otimização robusta (de acordo com a Equação A.30) e estabelece um critério multiobjetivo entre otimalidade e viabilidade, ponderando o máximo desempenho (valor médio) e a robustez (variância) (MANDUR e BUDMAN, 2014). No entanto, com frequência admite-se que o termo da variância da função objetivo não muda significativamente no ponto ótimo e a expressão é simplificada, contendo apenas o termo da média (KOOKOS, 2003; OSTROVSKY *et al.*, 2010; ZHANG *et al.*, 2002). Adicionalmente, o termo da variância na Equação (2.57) pode não ser uma boa medida de risco, pois exige que a distribuição de probabilidade das variáveis aleatórias seja simétrica em torno do respectivo valor médio (MULVEY *et al.*, 1995). Além disso, no caso específico da otimização econômica, deve-se considerar que a mesma penalização é dada tanto à contribuição da variabilidade, que torna o lucro superior, quanto àquela que torna o lucro inferior.
- O problema de otimização econômica poderia ser reformulado como um problema de programação estocástica, de acordo com a Equação (A.28), para um problema de otimização sujeito a restrições de probabilidade conjunta (JPC); ou, de acordo com a Equação (A.29), para um problema de otimização sujeito a restrições de probabilidade individual (IPC) (KALL e WALLACE, 1994). A formulação das restrições, conforme a Equação (2.55), assegura que os valores buscados pelo algoritmo de otimização para as variáveis de decisão sejam capazes de garantir que todas as restrições sejam simultaneamente satisfeitas ao nível de confiança α . Embora mais rigorosa, as dificuldades impostas pela restrição (2.55) à resolução do problema (A.28) fazem com que o problema IPC (Equação A.29) seja mais abordado na literatura (ZHANG *et al.*, 2002). No caso das restrições IPC, representadas pela Equação (2.56), a correlação entre as restrições é ignorada e procura-se garantir a probabilidade de que cada restrição seja atingida isoladamente.

No contexto da otimização sob incertezas, o custo computacional requerido pelos algoritmos de otimização é, em alguns casos, fator limitante para a resolução do problema. Desse modo, MANDUR e BUDMAN (2014) apresentaram um algoritmo eficiente para solução do problema de estimação de parâmetros, formulado como um problema de otimização robusta, quando incertezas paramétricas são descritas usando o Teorema de Bayes. A dificuldade de resolução do problema consiste na determinação da distribuição de probabilidades que resulta da aplicação do Teorema de Bayes, que exige elevado número de simulações. Assim, uma proposta de metodologia foi baseada na análise de multiresolução, para aproximar o modelo com

elevada acurácia nas regiões do espaço paramétrico em que a probabilidade é relativamente alta. A abordagem define diferentes níveis de resolução, em que cada nível é avaliado com o critério de divergência de Kullback-Leibler para selecionar regiões dos parâmetros onde a mudança na distribuição de probabilidades é maior que uma dada tolerância. Assim, no próximo nível de resolução, bases funcionais são adicionadas apenas nessas regiões de elevada probabilidade, resultando em um refinamento adaptativo. Tendo obtido a descrição das incertezas, uma expansão é usada para propagar a incerteza paramétrica estimada na função objetivo. Uma vez que a expansão permite a determinação da média e variância analiticamente, uma redução significativa no tempo computacional é alcançada quando comparada ao método de Monte Carlo.

Na literatura, os estudos de caso apresentados lidam com funções de restrição nas quais a incerteza aparece linearmente (LI *et al.*, 2008; MESFIN e SHUHAIMI, 2010; ZHANG *et al.*, 2002), com exceção do trabalho de KOOKOS (2003). No contexto da otimização em tempo real, ZHANG *et al.* (2002) desenvolveram uma abordagem para um sistema RTO robusto, que lida efetivamente com incertezas paramétricas. A principal contribuição foi a incorporação da incerteza no problema de otimização econômica, usando métodos de programação estocástica. Por meio de um estudo simulado para o problema de *blending* da gasolina, o sistema RTO foi implementado para determinar o fluxo de matérias-primas que maximizam a rentabilidade, mas sem violar especificações de qualidade. Ambas as formulações JPC e IPC para as restrições foram avaliadas, mostrando que o uso da formulação IPC para casos com correlação entre as restrições probabilísticas reduz a robustez do sistema RTO. Embora a satisfação das restrições de probabilidade individual na formulação IPC pode ser ajustada para produzir robustez similar à formulação JPC, tal ajuste é arbitrário e pode ser inviável para problemas de elevada dimensão. Além disso, os autores ressaltaram o alto custo computacional requerido para resolver o problema pela formulação JPC, causado principalmente pelas simulações de Monte Carlo necessárias para caracterizar as PDFs conjuntas dos parâmetros e pelo algoritmo de otimização usado para resolver o problema JPC. Nesse caso, algoritmos específicos ao problema, computadores de processamento paralelo e métodos de aproximação das distribuições de probabilidade podem reduzir substancialmente o custo computacional.

Dentro do contexto de otimização da produção, mas sem implementar uma estratégia *on-line*, HANSSEN e FOSS (2015) formularam um problema de otimização para estudos simulados de produção de petróleo, tendo como objetivo maximizar o potencial de óleo para um conjunto de poços entre alguns cenários. O modelo de otimização foi formulado como um problema de programação estocástica em dois estágios linear, cuja solução consiste em uma estratégia de operação dos poços em

lugar de um único *setpoint*. Resultando em problema de programação linear inteira mista (MILP), o estudo simulado indicou que a abordagem proposta melhorou a expectativa de produção de óleo em 1,5%.

Embora seja o modo que responde de forma mais completa aos efeitos das incertezas em um problema de otimização, problemas de otimização estocástica aplicados em sistemas industriais não foram encontrados, provavelmente devido à complexidade de implementação e resolução computacional envolvidos (QUELHAS, 2013). Apesar disso, KASSMANN *et al.* (2000) propuseram um algoritmo MPC robusto em duas camadas com restrições de probabilidade conjunta (JPC), composto por cálculo de *setpoint* estacionário seguido por otimização dinâmica, que posteriormente originou uma patente assinada pela Aspen Technology, Inc. (Cambridge, MA, EUA) (KASSMANN *et al.*, 2002). O cálculo do *setpoint* foi formulado como um problema de programação estocástico linear, considerando incertezas paramétricas na matriz de ganho estático do ponto de operação nominal. Entretanto, tais modelos de ganho estacionário podem não representar o processo corretamente quando o ponto de operação nominal muda significativamente. Apesar do controlador permanecer estável para grandes mudanças na operação, os resultados da otimização econômica podem deixar ser ótimos e viáveis (ZHANG *et al.*, 2002).

Enquanto os métodos de otimização sob incertezas são teoricamente plausíveis, existem diversas dificuldades para sua aplicação em sistemas RTO (QUELHAS *et al.*, 2013). A questão imediata que deveria ser considerada consiste na suposição inicial de que a incerteza é unicamente paramétrica, o que não é válido no caso geral. Quando incertezas estruturais de modelagem estão presentes, ambas as abordagens perdem o rigor teórico, em que o grau de perda varia naturalmente com a quantidade de incerteza estrutural (BUNIN *et al.*, 2013b). No entanto, é muito difícil não admitir *a priori* a hipótese de que o modelo seja perfeito, sobretudo para o trabalho de estimação de parâmetros, pois se erros de modelagem fossem conhecidos e o modelo pudesse ser melhorado, não haveria razão, a princípio, para utilizar um modelo errado (SCHWAAB e PINTO, 2007). Adicionalmente, mesmo quando a suposição de incerteza paramétrica é válida, obter a distribuição de probabilidade ou os limites superior e inferior dos parâmetros para o cenário do pior caso pode ser muito desafiador.

Por fim, o problema pode se tornar computacionalmente intratável, dependendo da distribuição de probabilidades e da maneira com que os parâmetros participam das restrições, o que se torna particularmente problemático quando o número de parâmetros incertos e restrições aumenta. Embora algumas propostas sugiram a solução do problema por linearização das restrições com relação aos parâmetros (ZHANG *et al.*, 2002), podem surgir erros significativos quando o processo é não linear, mesmo se o modelo for perfeito (*i.e.*, não houver incerteza estrutural) (BUNIN

et al., 2013b).

2.4.7 Projeto e Avaliação de Desempenho de Sistemas RTO

Conforme discutido, além da atualização do modelo e da otimização econômica, sistemas RTO também incorporam outros subsistemas para validação dos dados e, eventualmente, para análise de resultados. O desempenho global do sistema RTO depende não apenas do método RTO utilizado, mas também do projeto de cada subsistema (FORBES e MARLIN, 1996; MARLIN e HRYMAK, 1997; ROBERTS e WILLIAMS, 1981).

O projeto de um sistema de otimização estacionária em tempo real requer um conjunto razoável de decisões estruturais, especialmente (PERKINS, 1998):

- seleção do modelo estacionário usado no sistema e seu nível de complexidade;
- seleção das medidas da planta usadas no sistema, em especial no procedimento de atualização dos parâmetros;
- seleção do conjunto de parâmetros do modelo que serão atualizados a partir dos dados da planta, bem como o conjunto que irá permanecer fixo em valores nominais;
- seleção dos graus de liberdade usados para transmitir os resultados do sistema RTO para a planta;
- definição da estrutura do controle supervisão em termos de variáveis controladas e manipuladas.

Desse modo, o objetivo que deve guiar decisões em cada uma dessas áreas deve ser alcançar o melhor desempenho possível pela mitigação dos efeitos dos inevitáveis erros de medição e erros estruturais de modelagem no desempenho do processo.

Infelizmente, a modularidade funcional dos sistemas de otimização em tempo real contribuiu para uma aparente independência entre muitas áreas de pesquisa, refletindo em baixa atenção ao comportamento estrutural do sistema como um todo (SEQUEIRA, 2003). Com efeito, a seleção da estratégia de otimização em tempo real faz parte do rol de decisões incluídas no projeto de um sistema RTO e, conforme foi apresentado, a maioria das pesquisas relacionadas não avalia o efeito da estratégia no contexto do sistema, mas faz uma análise isolada.

Com relação à estratégia de otimização em tempo real, considerável atenção foi dedicada ao modelo estacionário empregado, dada a percepção de sua importância crítica ao sucesso da implementação. DURBECK (1965) e BIEGLER *et al.* (1985) apresentaram condições necessárias baseadas apenas nas características do modelo.

Reconhecendo que o modelo e o procedimento de atualização devem ser considerados em conjunto na determinação da adequação do modelo em um sistema RTO, FORBES *et al.* (1994) formularam o critério de adequação de modelos (conforme apresentado na Seção 2.3.1.2, p. 24). Um modelo é considerado adequado para uso em um sistema de otimização estático em tempo real se for capaz de determinar o ponto ótimo verdadeiro da planta \mathbf{u}_p^* , o que implica que o modelo deve ser capaz de estimar os gradientes do processo no ponto ótimo. Embora o critério verifique se a otimização econômica terá condições de determinar o ponto ótimo verdadeiro com determinado modelo e conjunto de parâmetros estimados, \mathbf{u}_p^* é desconhecido. Portanto, o critério de adequabilidade de modelo não pode ser usado para selecionar modelos, tampouco para apontar quais parâmetros estimar (QUELHAS, 2013). Dificuldades no cálculo do gradiente e da Hessiana podem impossibilitar seu uso. Além disso, sua validade é local em torno do ponto ótimo verdadeiro \mathbf{u}_p^* , não havendo garantia de que pontos ótimos de diferentes condições operacionais possam ser encontrados. Vale destacar que, a partir do trabalho de ROBERTS (1979), que mostrou a necessidade de considerar a interação entre a estimação de parâmetros, estratégias de otimização e o modelo do processo, muitas estratégias de otimização em tempo real foram desenvolvidas.

No âmbito da seleção dos parâmetros atualizados e medidas usadas no sistema RTO, KRISHNAN *et al.* (1992, 1993) propuseram uma abordagem para projeto do procedimento de atualização do modelo, selecionando tanto os parâmetros a serem estimados, quanto o conjunto mais adequado de variáveis medidas para estimar tais parâmetros. Sendo uma estratégia de caráter eminentemente prático, embora simplista, a abordagem consiste em realizar a análise de sensibilidade da função objetivo econômica em relação a variações dos parâmetros de até 5%. Embora os efeitos dos parâmetros sobre a função objetivo econômica sejam contabilizados, não é fornecido um critério objetivo que sirva como referência para determinar a inclusão ou não de determinado parâmetro no conjunto daqueles atualizáveis. Além disso, a abordagem apresenta caráter local, não avaliando efeitos combinados de diferentes agrupamentos de parâmetros, constituindo um método monovariável e avaliando isoladamente o efeito de cada mudança, não sendo capaz de prever a interação de cada arranjo com os demais componentes em um sistema. FRALEIGH *et al.* (2003) propuseram uma abordagem sistemática para seleção das medidas no projeto de sistemas RTO com base em uma métrica de seleção que pondera diferenças entre duas abordagens de projeto de experimentos, mostrando-se consistente com o método de avaliação de desempenho de FORBES e MARLIN (1996). Na prática, tentativa e erro consiste no procedimento geralmente usado para selecionar as variáveis medidas que serão reconciliadas, visto que um diferente subconjunto de parâmetros podem levar a um melhor ajuste entre modelo e planta (DARBY *et al.*, 2011).

Em virtude do funcionamento integrado de todos os componentes do sistema RTO, qualquer decisão de projeto, mesmo envolvendo um único elemento em determinado componente, exerce impacto direto no desempenho global do sistema. Por conta disto, qualquer análise sobre alternativas na estrutura de um sistema RTO deve levar em conta o comportamento do sistema como um todo. Assim, para poder avaliar decisões de projeto e discriminar estruturas alternativas; é necessário que o método de análise contemple todos os elementos do sistema, quantificando os efeitos em uma métrica de desempenho vinculada às políticas operacionais da planta (QUELHAS, 2013). KOOKOS (2005) propôs uma formulação que determina o conjunto de variáveis manipuladas e controladas, bem como a operação estacionária que minimiza os efeitos de perturbações na economia do processo, resultando em uma diminuição da sensibilidade do processo às perturbações e uma operação aproximadamente ótima, com execuções menos frequentes do sistema RTO. Embora aborde questões estruturais, a metodologia apresentada incorpora decisões pertinentes à estrutura e funcionalidade do sistema de controle supervisorio no nível de otimização em tempo real, aproximando-se mais da estratégia SOC e não correspondendo ao projeto do sistema RTO propriamente dito.

A avaliação das decisões de projeto em termos do desempenho global do sistema RTO recebeu pouca atenção na literatura. Existem, possivelmente, apenas duas propostas disponíveis na literatura que atendem a este requisito e propuseram-se a lidar de forma sistemática com o problema de projeto (*i.e.*, discriminação de estruturas alternativas) e avaliação de desempenho de sistemas RTO, a saber:

- Abordagem de LOEBLEIN e PERKINS (1996, 1998), desenvolvida com base no trabalho de DE HENNIN *et al.* (1994), denominada desvio médio do ótimo (DMO). A abordagem consiste na caracterização do desempenho de sistemas RTO com base no desvio médio do ótimo e no afastamento requerido a partir das restrições de desigualdade ativas, as quais são introduzidas nas restrições para garantir que apenas pontos de operação viáveis sejam produzidos pelo sistema RTO. Nessa abordagem, melhores projetos RTO resultam em valores menores das constantes de afastamento e do desvio médio do ótimo. Essa proposta foi posteriormente estendida para abranger a camada de controle supervisorio (implementada por meio de um MPC) e quantificar também efeitos dinâmicos do processo (LOEBLEIN e PERKINS, 1999).
- Paralelamente, FORBES e MARLIN (1996) propuseram um índice de desempenho conhecido como custo de projeto (CP), definido como a perda total de desempenho ocasionada pelo sistema RTO em relação à otimização perfeita da planta ao longo de um período de tempo. Representado como a média dos desvios entre o resultado ótimo verdadeiro e o resultado produzido pelo

sistema RTO, o índice quantifica os efeitos das diversas fontes de incerteza na determinação do ponto ótimo verdadeiro pelo sistema RTO. Sendo capaz de atribuir valor ao desempenho do sistema RTO, diferentes decisões de projeto podem ser comparadas de acordo com esse critério, em que as melhores decisões produzem menores custos de projeto. ZHANG e FORBES (2000) estenderam a abrangência do índice para incluir efeitos transientes e o efeito do número de execuções sucessivas da otimização necessárias até a estabilização dos valores propostos. Nessa extensão, o termo transiente refere-se a soluções intermediárias limitadas que são fornecidas pelo sistema RTO, uma vez que o verdadeiro efeito dinâmico do processo é ignorado (NARRAWAY e PERKINS, 1991). Muitas formulações apresentadas por FORBES e MARLIN (1996) foram idênticas às apresentadas por PINTO (1998, 2001), que lidava com os custos de incertezas paramétricas em problemas de otimização, ainda que este não abordasse especificamente um problema de RTO.

Ambas as propostas procuram criar um índice de desempenho para o sistema RTO com base no afastamento, em relação a uma determinada referência, do valor da função objetivo econômica J^{opt} correspondente à estimativa do ponto ótimo produzida pelo sistema RTO, \mathbf{u}^* . Diferentes estruturas conduzem a diferentes estimativas da operação ótima e, conseqüentemente, a diferentes resultados econômicos, refletidos pelo índice de desempenho do processo; uma vez que as diferenças sejam quantificadas em relação a uma referência comum, o índice de desempenho do sistema (DMO ou CP) possibilita atribuir valor a diferentes estruturas e, portanto, fazer a discriminação.

Apesar da aparente diferença percebida na representação matemática dos índices definidos nos trabalhos originais, QUELHAS (2013) propôs a seguinte apresentação para facilitar a análise comparativa (admitindo $J^{opt} := \phi(\mathbf{u})$):

$$\ell_{\text{DMO}} = E [\phi(\mathbf{u}_p^*)] - E [\phi(\mathbf{u}^*)] \quad (2.59)$$

$$\ell_{\text{CP}} = \phi(\mathbf{u}_p^*) - E [\phi(\mathbf{u}^*)] \quad , \quad (2.60)$$

em que ℓ_{DMO} e ℓ_{CP} são números reais que representam os índices de desempenho das abordagens DMO e CP, respectivamente.

Basicamente, a diferença entre as abordagens reside na definição da referência adotada. Na Equação (2.60), admite-se a existência e o conhecimento do ponto ótimo verdadeiro da planta \mathbf{u}_p^* , e do conseqüente valor máximo do índice de desempenho do processo. Por outro lado, a condição de referência na Equação (2.59) é considerada em um grau inferior de idealização. Ainda que os modelos do processo e as medições sejam perfeitas, o conhecimento sobre os parâmetros do modelo não o é, causando variabilidade na informação obtida a respeito do desempenho máximo do processo.

Contudo, a incerteza a respeito dos parâmetros, que consiste geralmente em limites inferior e superior que representam seu intervalo de confiança, é constante e dada *a priori*, sendo inalterada por alterações estruturais do RTO, o que torna inócua o trabalho de calcular $E[\phi(\mathbf{u}^*)]$ para o objetivo de discriminar o desempenho de diferentes sistemas. Dadas as similaridades e a forma de apresentação, ambas as formulações (Equações 2.59 e 2.60) poderiam ser chamadas de desvio médio do ótimo (QUELHAS, 2013). Adicionalmente, os dois métodos deduzem uma expressão analítica para os critérios de desempenho, representando, em ambas as abordagens, a função objetivo econômica por meio de uma expansão em série de Taylor de segunda ordem.

As diferenças entre as abordagens tornam-se pronunciadas a partir da dedução dos elementos que compõem o índice de desempenho do sistema. O método CP propõe expressar o critério de desempenho ℓ_{CP} em função da condição operacional estimada, \mathbf{u}^* , deixando implícitas as fontes de variabilidade e as especificidades do modelo. Desse modo, um limite superior esperado para ℓ_{CP} é definido como (FORBES e MARLIN, 1996):

$$\ell_{\text{CP}} \leq \ell_{\text{CP}}^U := \left\| \left. \frac{d^2 \phi}{d \mathbf{u}^2} \right|_{\mathbf{u}_p^*} \right\|_2 \left(\|\mathbf{u}_p^* - \mathbf{u}^*\|_2^2 + n_u \|\mathbf{V}_u\|_2 \right) \quad , \quad (2.61)$$

em que \mathbf{V}_u é a matriz de covariância das estimativas \mathbf{u}^* , determinada por meio de aproximação linear da propagação das incertezas de medida através dos componentes do sistema RTO (estimação de parâmetros e otimização econômica), sob a suposição de erros gaussianos aditivos não correlacionados no tempo. Os autores da abordagem DMO, por sua vez, valorizaram a apresentação de uma expressão analítica para o critério de desempenho diretamente como função das características primárias de variabilidade, como os erros de medição e a variabilidade estocástica e determinística dos parâmetros, assim como em função do vetor de afastamento, que adiciona uma margem de segurança às restrições. Dessa maneira, hipóteses são admitidas para a obtenção dos termos que compõem o índice, pré-especificando algumas decisões do sistema RTO, entre elas: i) natureza das incertezas de medida (aditivas, gaussianas), ii) procedimento de estimação, definido com o estimador de mínimos quadrados, sem reconciliação de dados, iii) restrições do problema (funções linearizadas), e iv) inclusão de constantes de afastamento para impor a satisfação das restrições. Estas definições prévias restringem a aplicação do índice e o conjunto de variações estruturais que podem ser avaliadas para um sistema RTO. Assim, apesar de gerar como solução uma equação linear pronta para o uso, sem necessidade de procedimentos adicionais de otimização (embora demasiado extensa), a abordagem DMO representa um caso menos genérico que CP. A Tabela 2.1 apresenta uma

comparação resumida entre as principais características das métricas CP e DMO.

Tabela 2.1: Comparação das métricas de desempenho de sistemas RTO, adaptado de QUELHAS (2013).

Características	CP	DMO
Aproximações	<ul style="list-style-type: none"> • função objetivo econômica aproximada por série de Taylor de 2ª ordem; curvatura constante ao longo de qualquer direção; • \mathbf{V}_u calculado por aproximação linear; • restrições ativas não mudam devido à variabilidade. 	<ul style="list-style-type: none"> • função objetivo econômica aproximada por série de Taylor de 2ª ordem; • linearização do modelo e das restrições do processo; • restrições ativas não mudam devido à variabilidade.
Erros de Medição	Erros aditivos gaussianos e não correlacionados.	Erros aditivos gaussianos, não correlacionados; Possibilita avaliar variação dos parâmetros com o tempo.
Elementos já incluídos na estrutura do RTO	Não apresenta hipóteses que pré-especifiquem elementos estruturais do sistema RTO.	<ul style="list-style-type: none"> • Estratégia em duas etapas com função objetivo pré-definida; • Uso de constantes de afastamento β às restrições do problema de otimização econômica.

Prezando por uma formulação analítica do índice de desempenho, ambas as abordagens fazem diversas aproximações. Além disto, alguns cálculos (*e.g.*, determinação dos gradientes envolvidos no cálculo de \mathbf{V}_u) podem impor dificuldades para serem executados em problemas de maior dimensão. Este tipo de formulação, todavia, está muito vinculado aos casos estudados na literatura. Por serem pouco complexos, oferecem condições para o desenvolvimento de uma formulação explícita em termos das equações do modelo e das restrições, cujo número total de equações está algumas ordens de grandeza abaixo dos modelos de processo encontrados em uso industrial. Em casos reais, adicionalmente, a formulação explícita das equações não está disponível na maior parte das vezes, pois o modelo é construído em softwares comerciais de simulação de processos, produzindo mais impedimentos ao uso de propostas que prevejam manipulações algébricas das equações do modelo, como requerido para o cálculo de CP e DMO (QUELHAS, 2013).

Motivado pelo fato de que a natureza aberta do problema de projeto de um sistema RTO exige uma metodologia aplicável a qualquer processo e que não defina de antemão uma expressão única e definitiva para um índice de desempenho do

sistema, QUELHAS (2013) propôs uma metodologia de avaliação exaustiva em todas as estruturas possíveis em um projeto de sistema RTO. Reunindo alternativas para os elementos envolvidos no projeto do RTO, a metodologia prevê a construção do rol de configurações resultantes da combinação entre as alternativas e a simulação do sistema RTO para cada estrutura. Com base em características de notável relevância durante a operação de uma planta sob ação do sistema RTO, os termos escolhidos para compor o índice de desempenho do sistema RTO usado na metodologia são:

1. Subotimalidade do índice de desempenho do processo, definido como (admitindo $J^{opt} := \phi(\mathbf{u})$):

$$\Delta\phi_k = \frac{\phi(\mathbf{u}_{k+1}^*) - \phi(\mathbf{u}_{id,k+1}^*)}{\phi(\mathbf{u}_{id,k+1}^*)} \quad (2.62)$$

em que $\Delta\phi_k$ representa a subotimalidade do índice de desempenho do processo na interação k do sistema RTO; e o subscrito $(\cdot)_{id}$ representa a solução idealizada do sistema RTO em malha fechada com modelo e otimização perfeitos.

2. Variabilidade das variáveis de decisão, representada por:

$$\Delta u_k = \frac{(\mathbf{u}_{k+1}^* - \mathbf{u}_k) - (\mathbf{u}_{id,k+1}^* - \mathbf{u}_{id,k})}{\mathbf{u}_k} \quad (2.63)$$

em que Δu_k representa a variabilidade produzida pelo sistema RTO nos resultados da interação k . A Equação (2.63) procura capturar o efeito da variabilidade desnecessária associada à existência do sistema RTO, expressando de forma relativa a variação das variáveis de decisão entre as iterações, descontada a variabilidade necessária $(\mathbf{u}_{id,k+1}^* - \mathbf{u}_{id,k})$ relacionada a intervenções que idealmente deveriam ser realizadas no processo.

3. Risco de inviabilidade das soluções, cuja quantificação se dá a partir do número de restrições violadas pela resposta do RTO na interação k , *i.e.*,

$$\Delta g_k = \sum_{i=1}^{n_g} v_i : v_i = \begin{cases} 0, & \text{se } G_i(\mathbf{u}_{k+1}^*) \leq 0 \\ 1, & \text{se } G_i(\mathbf{u}_{k+1}^*) > 0 \end{cases} \quad (2.64)$$

Desse modo, diretrizes financeiras, operacionais e de segurança são consideradas para atribuir valor às especificidades contidas em cada decisão de projeto, permitindo uma comparação conveniente entre cada configuração.

Utilizando a abordagem em duas etapas como estratégia de otimização em tempo real, a metodologia permite avaliar as seguintes decisões estruturais: modelo do processo; conjunto de variáveis medidas (*i.e.*, malha de instrumentação); conjunto de

parâmetros e variáveis candidatos a serem estimados e reconciliados, respectivamente, número de parâmetros atualizados e número de variáveis medidas usadas na função objetivo do problema de RDESP. Com o objetivo de discriminar as estruturas dentro de um contexto que leve em conta todas as influências sobre o desempenho global, a metodologia considera também a natureza dos sinais medidos, permitindo avaliar o comportamento do sistema na presença de autocorrelação e *bias* nas incertezas de medida. Buscando abranger as incertezas de processo, a definição de diferentes cenários de operação permite avaliar efeitos da variabilidade das condições impostas pelas variáveis ao processo. Por fim, opções com relação aos graus de liberdade selecionados como variáveis de decisão e restrições operacionais são usadas para compor diferentes versões do problema de otimização. As diferentes estruturas são avaliadas para um número fixo de iterações (*i.e.*, ciclos de execução do sistema), sendo função das respostas produzidas ao longo do número definido de cenários. Assim, os termos do índice de desempenho do sistema RTO são calculados para cada iteração em todas as condições construídas que envolvem as decisões de projeto, avaliando o comportamento dos resultados em função da natureza da corrupção das medidas e dos cenários de operação.

Uma vez que os índices definidos nas Equações (2.62), (2.63) e (2.64) são calculados a cada iteração de uma estrutura específica, a criação de outros índices se faz necessária, representando o desempenho global da estrutura. Com intuito de considerar a informação parametrizada da PDF dos termos de desempenho em cada iteração e ao longo das réplicas, QUELHAS (2013) definiu termos de desempenho das estruturas como a média ponderada da mediana e do 25^o percentil normalizados, sob o argumento de incorporar elementos da tendência central e da dispersão no sentido de aumento do índice, uma vez que o sentido de diminuição do índice não é relevante para a análise. Embora os termos possam ser analisados separadamente, também foi sugerido um índice unificado que supõe a equivalência entre os índices de desempenho globais para cada estrutura.

QUELHAS (2013) aplicou sua metodologia de avaliação de estruturas de sistemas RTO ao processo químico de WILLIAMS e OTTO (1960), composto por reator, trocador de calor, decantador e coluna de destilação, cujo modelo contém 11 variáveis de entrada e 11 variáveis de saída. Fixando cinco cenários de operação, seis iterações do método RTO ($k = 6$, Equações 2.2 e 2.3) e três diferentes configurações de corrupção dos sinais medidos, entre outros fatores, um conjunto total formado de 7228 estruturas possíveis foi construído. De imediato, a elevada dimensão do problema evidencia o caráter intratável da escolha da estrutura do RTO, quando submetido apenas aos critérios subjetivos baseados na experiência. Os resultados obtidos mostraram que não há qualquer garantia de correlação entre os índices globais (desempenho do processo, variáveis de decisão e violação das restrições) e que

a falta de correlação entre as grandezas quantificadas pelos índices é dependente da versão do problema considerada. Além disso, para a abordagem em duas etapas, se o conjunto das variáveis medidas é menor que o conjunto de variáveis estimadas, os parâmetros atualizados não corresponderão à realidade física. Tal perda de equivalência física é incontornável e faz parte das limitações mais fundamentais deste tipo de abordagem, ocorrendo ainda que todas as premissas do método de máxima verossimilhança sejam atendidas e que os métodos numéricos de otimização sejam perfeitos.

A natureza puramente matemática dos valores adaptados compromete a utilidade do sistema RTO como ferramenta de auxílio na tarefa de diagnóstico e manutenção preditiva de equipamentos e processos e revela a baixa efetividade de procedimentos intuitivos de escolha dos parâmetros a serem estimados. As variáveis devem ser escolhidas por meio de um critério vinculado à função objetivo da camada principal de otimização e esta escolha deve ser estritamente pragmática, no sentido de auxiliar o desempenho global do sistema RTO. Assim, na prática, o problema real consiste em escolher os parâmetros estimados que sejam matematicamente úteis para conferir robustez e melhoria dos índices de desempenho. Outro resultado relevante consiste na observação de que, para uma determinada versão do problema de otimização, um conjunto razoável de valores fixos das variáveis de decisão produzem resultados melhores do índice de desempenho unificado que os resultados obtidos com a atuação da melhor estrutura de sistema RTO (QUELHAS, 2013).

Corroborando esse resultado, SCHULTZ e FARENZENA (2015) comentaram que a análise da frequência do distúrbio e do comportamento dinâmico do sistema é determinante na escolha da solução a ser adotada, em que, para cenários específicos, o uso de qualquer ferramenta de otimização acarretará em perdas, quando comparada ao sistema sem otimização. Essa abordagem apenas confirma a visão de TAKAMATSU (1983), quando comentou que metodologias para avaliação de decisões de projeto e mudanças na operação devem ser unificadas.

2.4.8 Aplicações

De acordo com DARBY *et al.* (2011), soluções RTO são desenvolvidas desde o início da década de 1980, estando mais presentes na indústria química e petroquímica, especialmente na produção de etileno e propileno em unidades de craqueamento catalítico em leito fluidizado. Adicionalmente, KONINCKX (1988) relatou aplicações de sistemas de otimização em tempo real em caldeiras, sistemas de utilidades, unidades de produção e separação de olefinas, seções de refinarias, plantas de papel e celulose, processos de produção de amônia, acetaldeído, entre outros. Uma quantidade substancial de aplicações industriais pode ser encontrada na literatura

(ALKAYA *et al.*, 2001; BAILEY *et al.*, 1993; BASAK *et al.*, 2002; BAUSA e DÜNNEBIER, 2006; CHEN *et al.*, 1998b; CUTLER e PERRY, 1983; DARBY *et al.*, 2011; HANMANDLU *et al.*, 1985; MARLIN e HRYMAK, 1997; MERCANGÖZ e DOYLE, 2008; NAYSMITH e DOUGLAS, 1995; SEBORG *et al.*, 2010; VAN WIJK e POPE, 1992; WHITE, 1997, 2005), mencionando a obtenção de melhorias na operação da planta e na rentabilidade do processo da ordem de 3% em média (LAUKS *et al.*, 1992). Entre as empresas envolvidas, é possível identificar Statoil, Sunoco, ExxonMobil, BASF, Bayer, Shell, Hyundai Petrochemical, Suncor, Rhodia, *etc.* Diante da quantidade expressiva de aplicações, especial atenção foi dada às aplicações nacionais.

DE GOUVÊA e ODLOAK (1998) apresentaram os resultados de um sistema RTO aplicado em refinarias da Petrobras. O objetivo do sistema foi maximizar a produção de gás liquefeito de petróleo em reatores FCC, aplicando a estratégia RTO em uma camada, representada pelo MPC econômico, fornecendo também uma comparação com a abordagem em duas etapas. O MPC econômico se mostrou mais vantajoso, por assimilar mais rapidamente as mudanças nos objetivos econômicos e resultando em resposta dinâmica mais suave, embora apresentando respostas lentas para perturbações severas.

ZANIN *et al.* (2000, 2002) descreveram a implementação de um MPC econômico aplicado à maximização a produção de gás liquefeito de petróleo em conversor FCC na refinaria de São José dos Campos - SP. Os primeiros testes de comissionamento mostraram que os benefícios da otimização não se restringiram aos valores econômicos. Além disso, a estratégia foi capaz de manter o controle da planta, mesmo com a perda de variáveis manipuladas pela presença de perturbações severas.

TRESMONDI *et al.* (2001) relataram o desenvolvimento de um sistema RTO em reatores de oxidação de cumeno na planta de fenol da Rhodia Brasil, operando desde 1999. O sistema RTO foi integrado a um controlador multivariável, bem como a analisadores de composição *on-line*. O objetivo definido para o sistema foi a redução da formação de subprodutos durante a reação. Um algoritmo SQP foi usado para determinar os valores de composição dos reatores, que foram transmitidos à camada de controle como *setpoints*. A redução na formação de subprodutos obtida com a implementação do sistema foi de 4,7% (m/m) em média.

PORFÍRIO e ODLOAK (2011) descreveram os resultados de um MPC econômico implementado da coluna de destilação de tolueno na refinaria da Petrobras em Cubatão - SP. O controlador consistia em MPC de horizonte infinito, que apresentava estabilidade nominal e resolvia simultaneamente o problema de controle e otimização por inclusão de termos econômicos na função objetivo. O uso do sistema se mostrou eficiente e robusto para diversos objetivos econômicos do sistema de destilação, mesmo sendo limitado aos casos em que a função objetivo econômica era

côncava.

JESUS (2011) descreveu as etapas de implementação de um sistema RTO para a planta de produção de etileno de uma unidade industrial da Braskem em Duque de Caxias - RJ. O sistema RTO desenvolvido foi baseado na abordagem em duas etapas e, portanto, incluiu os procedimentos de estimação de parâmetros, reconciliação de dados e otimização econômica. O objetivo foi definido como o de maximização do lucro bruto. Ambos os procedimentos de RDESP e otimização econômica foram resolvidos por meio do algoritmo PSO, cujo tempo necessário de resolução do problema permaneceu dentro dos limites que permitiam a execução do sistema RTO no processo. Nas condições sugeridas, o ganho proporcionado pelos procedimentos de otimização foi de 1,8%.

Embora não se trate de uma aplicação propriamente dita, o trabalho de EBRAHIMPOUR *et al.* (2015) apresentou a arquitetura de um protótipo de sistema RTO que permite a avaliação de diferentes metodologias para os componentes do RTO, tendo a expectativa de ser aplicado em uma unidade de separação da Petrobras. Os componentes considerados no sistema RTO foram: detecção de estado estacionário, detecção de erros grosseiros, reconciliação de dados, estimação de parâmetros, otimização do processo e análise econômica.

Assim, é possível perceber, em conjunto com a diversidade de temas, estratégias e metodologias envolvidas e desenvolvidas, o alcance da otimização estática em tempo real em plantas nacionais. O nível de automação é um aspecto importante nessa realidade, que não oferece garantias de sucesso mas diferencia os desafios a serem superados. Considerando a quantidade de sistemas RTO implementados no Brasil, observa-se que os elementos básicos para o desenvolvimento desses sistemas, como a disponibilidade de medidas e de um modelo confiável do processo, continuam a desafiar a expansão das tecnologias de otimização em tempo real a diferentes domínios. Diante disso, esforços que contribuam para aumentar o alcance das técnicas e ferramentas da otimização em tempo real nas indústrias brasileiras possuem especial relevância.

Capítulo 3

Processos de Elevada Dimensão

Em processos químicos e petroquímicos, em especial aqueles relacionados à produção de *commodities*, a dimensão do número de variáveis medidas pode facilmente alcançar a ordem dos milhares. A quantidade de dados coletados da planta atinge dimensões tão elevadas, que a relativa simplicidade de transmitir, organizar e armazenar essa informação exige soluções dedicadas. Dificuldades com processos de dimensão elevada constituem problemas antigos e estratégias envolvendo a identificação recursiva de modelos dinâmicos empíricos *on-line* (BAMBERGER e ISERMANN, 1978; BHATTACHARYA e JOSEPH, 1982; GARCIA e MORARI, 1981; MCFARLANE e BACON, 1989), bem como a incorporação de informações fenomenológicas no esquema adaptativo, já foram propostas em sistemas RTO (GOLDEN e YDS-TIE, 1989), mas apresentaram problemas com a sintonia de parâmetros presentes na implementação e com imprecisões na etapa de identificação.

Assim, a elevada dimensão do número de variáveis, equipamentos, cenários de operação, tipos de perturbações possíveis, dentre outros constitui um importante desafio existente na etapa de identificação (modelagem) de estratégias de otimização em tempo real (BHATTACHARYA e JOSEPH, 1982), conforme fluxograma da Figura 1.2. Procedimentos em tempo real para aplicação em sistemas industriais de grande porte impõem exigências de desempenho computacional e precisão de técnicas numéricas que amplificam ainda mais essa dificuldade. Se o processo possuir um moderno sistema de automação e controle, isso certamente influenciará o desempenho final do sistema RTO.

Diante desse cenário, foi desenvolvida uma estratégia para monitoramento de processos industriais de grande escala com base no uso de modelos empíricos por meio de métodos de controle estatístico de processos (*Statistical Process Control* - SPC). Técnicas de modelagem baseadas em dados, como técnicas de *machine learning*, têm especial capacidade de lidar com a elevada dimensão dos conjuntos de dados com velocidade e eficiência, dispondo de pouco ou nenhum conhecimento do domínio (KADLEC *et al.*, 2009; REIS e GINS, 2017). Além disso, podem ser

usadas para identificar e selecionar as variáveis que mais contribuam na predição de uma variável de interesse. Com base em métodos estatísticos, foram definidos vários indicadores (ou alarmes) com o objetivo de detectar desvios significativos no desempenho do modelo que pudessem indicar perturbações ou falhas indesejáveis no processo real. Os alarmes de adequação propostos consistiram nos seguintes tipos: detecção de *outliers*, detecção de tendências dinâmicas e detecção de *bias*.

Propõe-se também a extensão do conjunto de dados medidos disponível por meio de transformações matemáticas lineares e não lineares das medidas disponíveis com base em conhecimento de domínio. O conhecimento do domínio permite o cálculo de variáveis não medidas a partir de balanços de massa, energia e momento aplicadas em determinadas seções do processo. Este tipo de redundância analítica pode revelar importantes relações entre as variáveis, que podem não ser capturadas por procedimentos estatísticos. Adicionalmente, os dados brutos disponíveis podem ser filtrados e reconciliados para reduzir o ruído de medição e satisfazer as equações de balanço.

A metodologia foi avaliada em um processo de tratamento primário de petróleo contendo aproximadamente 6000 variáveis medidas com valores disponíveis de minuto em minuto. O processo foi descrito por modelos dinâmicos lineares gerados a partir de técnica baseada em dados, cuja variável de interesse foi a produção de óleo da unidade. Simulações com dados reais de planta mostraram a satisfatória capacidade de ajuste do modelo. A aplicação dos alarmes estatísticos aos sinais do processo e do modelo demonstraram que perturbações podem ser detectadas com antecedência, permitindo detectar a ocorrência de eventos indesejados de forma antecipada ou a necessidade de readequação do modelo às novas condições de operação. Observou-se também que a inclusão de sinais construídos com base em conhecimento do domínio favorece a construção de modelos mais precisos, refletido pela melhor capacidade preditiva. Isso resultou em modelos mais confiáveis e, conseqüentemente, em um monitoramento aprimorado do comportamento do processo, levando a uma identificação mais rápida e eficiente das falhas. O modelo resultante pode ser aplicado a diferentes finalidades, embora esse estudo tenha focado no monitoramento da operação da planta ao longo do tempo.

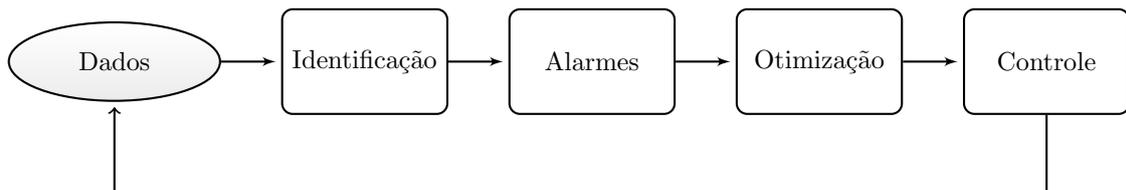


Figura 1.2: Fluxograma alternativo com as etapas envolvidas em uma implementação real de sistema RTO (apresentado na p.8).

O presente trabalho foi realizado em parceria com uma empresa e o arranjo proposto inova as etapas de *Identificação* e *Alarmes* da Figura 1.2. Em virtude dos termos de confidencialidade assinados e dos processos de submissão de patentes que estão em andamento, registrado no United States Patent and Trademark Office (USPTO) nº 15141981 em abril de 2016 sob título *Process Monitoring Based on Large-Scale Combination of Time Series Data*, apresentam-se a seguir descrições resumidas do trabalho desenvolvido, sem que se comprometa o sigilo necessário para a tramitação adequada das patentes e colaborações realizadas. Assim, os resultados apresentados não têm como objetivo caracterizar de forma detalhada os exemplos tratados, mas ilustrar a aplicação do procedimento em uma aplicação real.

3.1 Introdução

Em processos industriais com algum nível de automação existem múltiplos sensores, geralmente com objetivos bem definidos, fornecendo medidas de diferentes grandezas. Essas medidas são continuamente coletadas e armazenadas sob a forma de séries temporais, gerando uma grande quantidade de dados. Embora exista abundância na quantidade de dados, a falta de ferramentas analíticas efetivas e tecnologias computacionais robustas para extrair informação útil desses dados faz com que, na maior parte dos casos, esses dados sejam compactados e arquivados, sendo apenas recuperados para uso em análises após alguma ocorrência crítica, em vez de um uso rotineiro no processo de tomada de decisão (QIN, 2014). Essa realidade de indústrias químicas de manufatura, no entanto, não reflete a postura de empresas de outros setores que, considerando os dados como ativos, coletam e utilizam toda informação disponível (GRAEFF e HARMON, 2002; TUCKER, 2014).

Com o desenvolvimento recente de diversas aplicações baseadas no uso de internet, sensores inteligentes, comunicação sem fio, dispositivos móveis *etc.* (DAVIS *et al.*, 2012), a grande quantidade de dados disponíveis levou muitas áreas do conhecimento e indústrias a reexaminar o valor das informações contidas nos dados. Como resultado, surgiu uma nova disciplina conhecida como ciência de dados (do inglês *data science*) e a análise de grandes conjuntos de dados se tornou a nova base para o aumento de competitividade, produtividade e inovação (MANYIKA *et al.*, 2011). Esta atividade acabou tornando-se conhecida pelo termo *big data*, fazendo referência ao tamanho (ou volume) e variedade dos conjuntos de dados que desafiam a capacidade e velocidade das técnicas e *softwares* tradicionais para captura, armazenamento, gerenciamento e análise. Outro resultado foi o surgimento de novas frentes de pesquisa, desenvolvimento e inovação para iniciativas que buscam oportunidades em *big data*, com incentivos tanto no setor público quanto privado (QIN, 2014).

No contexto da engenharia de processos, algumas ferramentas analíticas têm sido adotadas em aplicações reais de monitoramento de processos, bem como detecção e diagnóstico de falhas, com diferentes graus de sucesso, como estatística multivariada e redes neuronais (REIS e GINS, 2017). Essas técnicas muitas vezes são capazes de detectar e diagnosticar a causa raiz de falhas operacionais relacionadas à qualidade, à degradação do processo e mudanças que podem levar a falhas graves. No entanto, métodos e aplicações de análise multivariada nessa área, aparentemente, não têm feito uso ou conexão com o desenvolvimento recente em aprendizado de máquina (*machine learning*), mineração de dados (*data mining*) e análise de dados de grande dimensão (*big data analytics*) (QIN, 2014). Essas técnicas diferem não apenas em termos de tamanho, mas também em como e quais dados devem ser usados para a solução dos problemas reais de operação. Algumas técnicas desenvolvidas na área de aprendizado da máquina, considerada uma ponte entre estatística e computação, combinando conhecimentos de teoria da informação, processamento de sinais, algoritmos, teoria de controle e teoria de otimização (NATIONAL RESEARCH COUNCIL, 2013), incluem máquinas de vetores de suporte e PCA por funções núcleo (*kernel PCA* - KPCA). Além desses métodos, os chamados modelos de dados latentes constituem ferramentas efetivas para analisar dados multidimensionais, com técnicas de projeção em estruturas latentes e de análise de componentes principais, e têm sido aplicados em quimiometria e monitoramento de processos (MACGREGOR *et al.*, 2015). A maioria dos dados de processo é formada por séries temporais com períodos de amostragem regulares, tornando-os apropriados para aplicação de métodos de mineração de dados e de representação de séries temporais, que têm sido amplamente estudados em outras áreas (CHUNG FU, 2011), embora pouca atenção tenha sido dada a esses avanços no contexto da engenharia química. QIN (2012) revisou as técnicas de *data analytics* disponíveis e suas aplicações para monitoramento de processos, enquanto uma revisão sobre aprendizagem de máquina pode ser encontrada em BENGIO *et al.* (2013).

A identificação do sistema e a modelagem de séries temporais são ferramentas necessárias para aplicar técnicas de controle baseadas em modelo e a avaliação de desempenho do controle. Além disso, estratégias envolvendo a identificação recursiva de modelos dinâmicos empíricos *on-line* já foram propostas em sistemas RTO (BAMBERGER e ISERMANN, 1978; BHATTACHARYA e JOSEPH, 1982; GARCIA e MORARI, 1981; GOLDEN e YDSTIE, 1989; MCFARLANE e BACON, 1989), mas apresentaram problemas com a elevada dimensão do sistema, com a sintonia de parâmetros presentes na implementação e com imprecisões na etapa de identificação. Logo, existe grande benefício quando técnicas baseadas em dados são combinadas com conhecimento do domínio e com procedimentos de detecção de falhas, potencializando o desempenho e robustez da solução. No contexto da “Indústria 4.0”, vários

novos conceitos, abordagens e ferramentas têm sido propostos em todos os setores industriais, promovendo a chamada “Quarta Revolução Industrial” (KLAUS, 2016). No entanto, algumas questões práticas devem ser abordadas por plantas industriais modernas antes da implementação de um estágio tão avançado.

Assim, propõe-se nesta parte do trabalho um método para monitoramento estatístico de processos para sistemas dinâmicos multivariados de elevada dimensão apoiado em *big data*. O uso de técnicas *big data* se faz necessário para que o método possa ser aplicado a processos de elevada dimensão e seja facilmente escalável, se necessário, fazendo uso de todos os tipos de dados disponíveis no processo. Com base em um modelo empírico do processo, testes estatísticos são usados para detectar a existência de falhas ou condições anormais de operação, bem como para o isolamento das fontes de perturbação. Além disso, a estacionariedade do processo também pode ser inferida a partir dessas informações. O método proposto consiste na definição de índices de detecção capazes de indicar perturbações ou falhas indesejáveis no processo real e na visualização desses índices por meio de alarmes, permitindo um monitoramento contínuo do processo. Adicionalmente, propõe-se a construção de novos sinais com base em conhecimento do domínio, como relações originadas de balanços de massa, para melhorar o desempenho do procedimento de monitoramento. A metodologia foi avaliada por meio de simulações com dados reais de um processo de tratamento primário de petróleo contendo aproximadamente 6000 variáveis medidas com valores disponíveis de minuto em minuto.

3.2 Detecção e Diagnóstico de Falhas com Métodos Baseados em Dados

O esquema de detecção e diagnóstico de falhas (*fault detection and diagnosis* - FDD) pode ser dividido em três atividades (ZHANG e JIANG, 2008):

1. Detecção de falhas: indica que algo está errado no sistema; isto é, a ocorrência de uma falha e o instante de ocorrência da falha;
2. Isolamento de falha: determina a local e o tipo de falha (qual componente falhou);
3. Identificação da falha: determina a magnitude (tamanho) da falha.

Na literatura, é comum que as atividades de isolamento e identificação de falhas sejam conjuntamente chamadas de diagnóstico de falhas (ISERMANN, 1997).

As abordagens FDD existentes podem ser classificadas como quantitativas e qualitativas. Essas duas categorias podem ser adicionalmente divididas em abordagens

baseadas em modelos e baseadas em dados (*data-driven* ou *model-free*). Essencialmente, uma abordagem de FDD quantitativa utiliza um modelo matemático (muitas vezes chamado de redundância analítica) para descrever o comportamento nominal, sem falhas, do processo sob análise e realizar a FDD em tempo real. Dentro do grupo de abordagens FDD quantitativas baseadas em modelo, as principais técnicas utilizadas são baseadas em: i) estimação de estados, ii) estimação de parâmetros, iii) espaço de paridade e iv) combinação entre as anteriores. Por outro lado, em comparação com o custo e esforço da modelagem rigorosa das abordagens baseadas em modelo, os métodos de modelagem baseados em dados podem ser usados para detectar mudanças anormais na operação do processo a partir de dados em tempo real devido à capacidade de redução da dimensão, facilidade de visualização e facilidade de implementação em aplicações reais. Estão nesse grupo técnicas como análise em componentes principais (PCA), projeção em variáveis latentes (*partial least squares* - PLS) e análise de correlação canônica (*canonical correlation analysis* - CCA). A literatura apresenta diversos trabalhos com revisões sobre o tema e outras classificações possíveis dos métodos (VENKATASUBRAMANIAN *et al.*, 2003; ZHANG e JIANG, 2008).

A aplicação de abordagens quantitativas segue o procedimento típico da análise multivariada de dados de processo, envolvendo (QIN, 2014):

1. Coleta de dados “normais” com boa abrangência das regiões operacionais; dados de falhas podem ser úteis, mas não obrigatórios;
2. Aplicação de métodos de regressão, geralmente métodos de variáveis latentes, para modelar os dados (GELADI e KOWALSKI, 1986);
3. Definição dos índices de detecção de falhas e limites de controle;
4. Diagnóstico de falhas e solução de problemas.

Assim, o procedimento FDD pode ser considerado como um algoritmo de três etapas (KÖPPEN-SELIGER *et al.*, 1995; SCHUBERT *et al.*, 2011). Após a geração do modelo identificado, a primeira etapa consiste em gerar um ou vários sinais que refletem falhas no comportamento do processo. Esses sinais, geralmente resíduos da reconstrução das variáveis medidas, são compostos por medidas existentes e sinais de referência correspondentes. Também é possível definir índices multivariados, limites e outras referências com base nos sinais do processo. Na segunda etapa, os resíduos e índices são avaliados, determinando-se o instante e a localização das possíveis falhas indicadas. Por fim, a natureza e a causa da falha são estudadas por meio da análise das relações entre os eventos observados nos sinais e suas causas físicas.

Nos processos industriais, devido às frequentes mudanças operacionais, as condições normais de operação assumem distribuições multimodais, tornando inadequa-

das as abordagens unimodais. Nos processos em que a condição de operação muda de um regime para outro, são necessários métodos multimodais. As abordagens para esse caso incluem, por exemplo, métodos de modelagem multimodal, como a aprendizagem local e abordagens de múltiplos modelos. O trabalho de FEITAL *et al.* (2013) fornece uma revisão das abordagens para modelagem e monitoramento de sistemas multimodais baseadas em técnicas estatísticas.

3.3 Procedimento Proposto

Para aplicar o procedimento de FDD, os resíduos e outros sinais devem ser gerados e os índices de detecção de falhas e limites de controle devem ser definidos. Utilizando os dados medidos das variáveis de entrada e saída do sistema de interesse e algum tipo de modelo, são gerados sinais que refletem de alguma maneira a(s) falha(s) que se deseja detectar. Assim, o modelo é usado *on-line* para análise contínua da condição do processo.

3.3.1 Modelagem

Antes de aplicar o procedimento para a avaliação dos resíduos, o modelo deve ser construído (ou treinado) para esta tarefa. Assim, o uso do modelo baseado em dados compreende duas etapas distintas:

- Construção do modelo, em que dados históricos são usados para obter uma relação da forma:

$$\hat{y}_k = f(\mathbf{x}_k) \quad (3.1)$$

em que k indica um instante de tempo e $\mathbf{x} \in \mathbb{R}^{n_x}$ tem dimensão finita previamente estabelecida, cujos elementos são variáveis medidas do processo, defasadas no tempo ou não (*i.e.*, com ou sem *lag*). A defasagem temporal é uma tática para introduzir dinâmica na modelagem do processo. Quando as variáveis do conjunto de dados usado para construção do modelo são consideradas com defasagem temporal, é possível construir modelos dinâmicos, de forma equivalente à técnica de PCA dinâmico (CHEN e LIU, 2002; KU *et al.*, 1995; LEONTARITIS e BILLINGS, 1985; VANHATALO *et al.*, 2017). O número de defasagens no tempo a serem consideradas para cada variável deve ser previamente especificado;

- Uso do modelo para predição, em que o conjunto mais recente de medições do processo é usado para fornecer uma estimativa de \hat{y} .

Nesse trabalho, propôs-se a seguinte relação linear para descrever a variável de

interesse:

$$\hat{y} = \sum_{i=1}^l b_i x_i = \mathbf{b}^T \mathbf{x}_l \quad , \quad (3.2)$$

em que y e $\mathbf{x}_l \subseteq \mathbf{x}$ são variáveis normalizadas, $\mathbf{b} \in \mathbb{R}^l$ são coeficientes a serem estimados e $l \leq n_x$ é a ordem do modelo. Os coeficientes lineares do modelo, que compõem o vetor \mathbf{b} , foram determinados utilizando a técnica de decomposição ortogonal de sinais (OMP) (PATI *et al.*, 1993). Assim, para dada dimensão l , essa técnica seleciona as variáveis de entrada que irão compor \mathbf{x}_l e estima os coeficientes, capturando o conjunto ortogonal no espaço original das variáveis utilizadas, fornecendo os componentes mais relevantes para a predição da variável de interesse, em que o produto $b_i x_i$ representa o efeito (ou contribuição) da variável i na resposta do modelo.

Após a construção do modelo, os resíduos gerados pelo modelo e os sinais obtidos do processo são usados para calcular os índices de detecção, definidos por meio de testes estatísticos. Na sequência, a resposta dos índices é convertida em alarmes, que podem indicar desvios do processo e evitar falhas indesejáveis. Os índices e alarmes são definidos a seguir. Considerando que a modelagem segue uma abordagem unimodal, adota-se a estratégia de múltiplos modelos, admitindo que os modelos possam ser substituídos sempre que necessário.

3.3.2 Índices de Detecção e Limites de Controle

A identificação de falhas está baseada na trajetória de determinadas variáveis ao longo do tempo, em que a avaliação de resíduos desempenha papel essencial na detecção e diagnóstico de falhas. Essa avaliação, e a correspondente definição dos índices de detecção e seus limites, deve considerar o compromisso entre uma detecção rápida e confiável. Na maioria dos trabalhos que utilizam observadores, limites simples são utilizados para a avaliação de resíduos. Classificadores estatísticos também podem ser usados nessa avaliação, como por exemplo uma rede neuronal (KÖPPEN-SELIGER *et al.*, 1995). Não obstante, a análise do comportamento das variáveis de entrada também é importante, sendo bastante útil para fins de diagnóstico da falha.

No contexto da modelagem descrita anteriormente, em que um modelo e uma lista das variáveis medidas mais relevantes estão disponíveis, foram definidos índices de detecção para o erro de predição ($z = y - \hat{y}$), para a variável de interesse (y) e para as variáveis independentes selecionadas (\mathbf{x}_l). Essas variáveis foram incluídas no monitoramento com o objetivo de melhorar a assertividade na distinção entre eventos que alteram o estado do processo, acompanhado ou não da degradação do desempenho do modelo, de flutuações aleatórias presentes no sistema. Além disso, também visou a distinguir eventos ligados à aquisição dos dados, como perda de

calibração dos instrumentos e falhas em sensores, daqueles associados a falhas de processo.

A definição dos índices de detecção esteve associada ao objetivo de auxiliar as seguintes atividades: i) avaliar a qualidade dos dados e apoiar a identificação de problemas associados com a aquisição dos sinais, como valores espúrios; ii) avaliar a pertinência dos dados e apoiar a identificação de mudanças na morfologia dos sinais; iii) avaliar a qualidade do modelo; e iv) avaliar o estado do processo e apoiar a identificação e detecção de falhas. Nesse contexto, a definição dos índices foi baseada na identificação de valores espúrios (ou *outliers*), desvios sistemáticos e tendências no conjunto de dados.

Índice de detecção de valores espúrios (I_o): esse índice tem como objetivo identificar se o sinal é muito distinto da média observada no conjunto de dados usado na construção do modelo, ou seja, avaliar se os sinais do processo tem correspondência com o conjunto de dados normal. O índice é calculado em cada instante de amostragem k , sendo definido de acordo com testes estatísticos aplicados a cada um dos grupos de variáveis monitoradas como segue:

1. Índice para o erro de predição da variável de interesse, mede o desempenho do modelo para representar os dados:

$$I_{o,z} = |z| < c_{o,z} \quad (3.3)$$

2. Índice para a variável de interesse, mede a pertinência da variável resposta frente ao dados de construção do modelo:

$$I_{o,y} = |y - \bar{y}| < c_{o,y} \quad (3.4)$$

3. Índice das variáveis de entrada, mede a pertinência das variáveis de entrada selecionadas frente aos dados de construção do modelo:

$$I_{o,x_i} = |x_i - \bar{x}_i| < c_{o,x_i} \quad (3.5)$$

O valor \bar{y} é o valor médio para as observações da variável de interesse no conjunto de treino, enquanto z é a diferença entre o valor da variável de interesse e sua predição no tempo mais recente; \bar{x}_i é a média da variável selecionada i no conjunto de treino.

Os limites de controle $c_{o,z}$, c_{o,x_i} e $c_{o,y}$ são limites de confiança para cada índice, geralmente determinados por meio de propriedades estatísticas dos dados históricos usados na construção do modelo. Por exemplo, $c_{o,x_i} = 3\sigma_{x_i}$ representa um intervalo

de confiança de três vezes o desvio padrão de x_i observado no conjunto de treino e corresponde ao nível de confiança de 99,56% da distribuição normal.

No caso em que um dos índices viola o limite de tolerância, *i.e.*, quando alguma das desigualdades não é satisfeita, há indícios de um valor espúrio na variável monitorada pelo índice, que pode estar associado a flutuações aleatórias do sistema de medição. Os índices são calculados de maneira independente, de modo que várias variáveis podem violar os respectivos limites de controle ao mesmo tempo.

Índice de detecção de desvio sistemático (I_b): o segundo índice tem como objetivo detectar a existência de desvio sistemático ao longo de uma janela de tempo pré-definida. O índice é calculado para uma janela de dados contendo N pontos, como segue:

1. Índice para desvio sistemático do erro de predição da variável de interesse, indica o nível de degradação do ajuste do modelo para os dados da janela:

$$I_{b,z} = |\bar{z}_N| < c_{b,z} \quad (3.6)$$

2. Índice para desvio sistemático da variável de interesse, indica a mudança de comportamento da variável resposta para os dados da janela:

$$I_{b,y} = |\bar{y}_N - \bar{y}| < c_{b,y} \quad (3.7)$$

3. Índice para desvio sistemático das variáveis selecionadas, indica a mudança de comportamento das variáveis de entrada para os dados da janela:

$$I_{b,x_i} = |\bar{x}_{i,N} - \bar{x}_i| < c_{b,x_i} \quad (3.8)$$

em que \bar{y}_N e \bar{z}_N correspondem à média dos valores da variável de interesse e dos erros de predição, respectivamente, na janela contendo os últimos N pontos; e $\bar{x}_{i,N}$ é a média da i -ésima variável selecionada na construção do modelo na janela contendo os últimos N pontos.

Como descrito para o índice anterior, $c_{b,z}$, $c_{b,y}$ e c_{b,x_i} são limites de confiança para cada variável monitorada, geralmente determinados por meio de propriedades estatísticas. Para o caso da janela de dados com N pontos, $c_{b,x_i} = 3\sigma_{x_i}/\sqrt{N}$ representa um intervalo de confiança de três vezes o desvio padrão de x_i observado na janela de dados e corresponde ao nível de confiança de 99,56% da distribuição normal.

O tamanho da janela de dados, N , é um parâmetro de sintonia, que tem relação com a dinâmica do processo. Assim, o valor de N deve ser avaliado para cada sistema. Para fins de avaliação, foram testadas janelas de tamanho 10 e 60 minutos.

Nesse cenário, quando um índice viola o limite de tolerância, existe evidência de desvio sistemático na variável monitorada para a janela de tempo considerada, que pode estar associado a uma falha de medição, falha de processo ou degradação do desempenho do modelo. Novamente, os índices são independentes e podem violar os respectivos limites de controle simultaneamente.

Índice de detecção de tendência (I_d): o terceiro índice tem como objetivo detectar tendência para as medidas de uma variável monitorada em um intervalo de tempo. A tendência avaliada pelo índice é apontada pela significância estatística do coeficiente angular obtido para uma regressão linear dos dados normalizados no interior de uma janela de N pontos, conforme ilustrado na Figura 3.2.

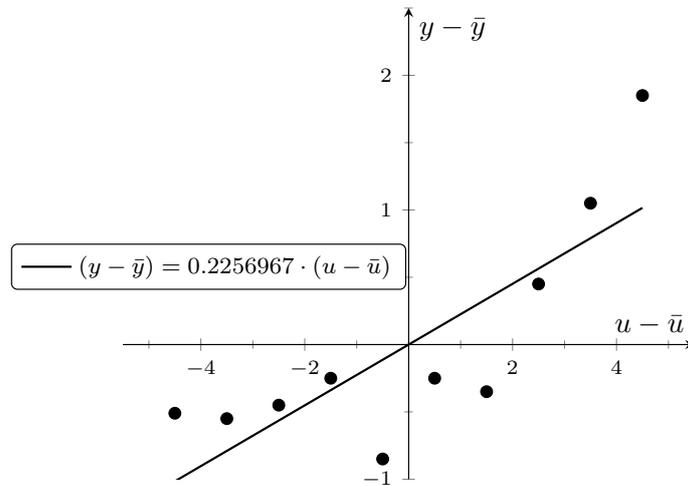


Figura 3.2: Tendência em uma série temporal com janela de tamanho $N = 10$.

Para a variável de interesse (y), a regressão linear com os dados normalizados pode ser representada pela seguinte equação linear:

$$(y_u - \bar{y}) = a (u - \bar{u}) \quad ,$$

em que $u = 0, \dots, N - 1$ é o instante de tempo correspondente a y_u no interior da janela e \bar{u} é o valor médio de u . Assim, com base em uma regressão de mínimos quadrados, o coeficiente a da regressão linear pode ser determinado como (HIMMELBLAU, 1970):

$$\hat{a} = \frac{\sum_{u=0}^{N-1} (y_u - \bar{y}) (u - \bar{u})}{\sum_{u=0}^{N-1} (u - \bar{u})^2} \quad . \quad (3.9)$$

Assim, o índice consiste em avaliar o grau de significância do valor estimado do coeficiente linear de acordo com limites de controle:

1. Índice para tendência do erro de predição da variável de interesse, indica tendência na qualidade do ajuste do modelo aos dados da janela:

$$I_{d,z} = |\hat{a}_z| < c_{d,z} \quad (3.10)$$

2. Índice para tendência da variável de interesse, indica o comportamento dinâmico da variável resposta para os dados da janela:

$$I_{d,y} = |\hat{a}_y| < c_{d,y} \quad (3.11)$$

3. Índice para tendência das variáveis selecionadas, indica o comportamento dinâmico nas variáveis de entrada para os dados da janela:

$$I_{d,x_i} = |\hat{a}_{x,i}| < c_{d,x_i} \quad (3.12)$$

em que \hat{a}_z , \hat{a}_y e $\hat{a}_{x,i}$ são coeficientes estimados com dados para o erro de predição, para a variável de interesse e para a variável selecionada i , respectivamente.

Os limites c_{d,x_i} , $c_{d,y}$ e $c_{d,z}$ podem ser definidos em termos da variância do coeficiente angular estimado (σ_a^2), de acordo com a equação:

$$\sigma_a^2 = \frac{\sigma^2}{\sum_{u=0}^{N-1} (u - \bar{u})^2} \quad , \quad (3.13)$$

em que

$$\sigma^2 = \sum_{u=0}^{N-1} \frac{[(y_u - \bar{y}) - a u]^2}{N - 1} \quad .$$

Se o valor absoluto do coeficiente linear viola os limites de tolerância, existe evidência de tendência no conjunto de dados para a janela considerada, que pode estar associada a uma falha de processo ou degradação do desempenho do modelo.

De modo similar ao caso do índice anterior, o tamanho da janela de dados, N , é um parâmetro de sintonia que tem relação com a dinâmica do processo e deve ser avaliado. Para fins de avaliação, foram testadas janelas de tamanho 10 e 60 minutos.

3.3.3 Função Alarme e Visualização por Diagrama de Cores

Para facilitar a visualização e interpretação dos diversos índices apresentados, propôs-se o uso de alarmes, calculados por meio da função alarme (f_{al}).

Partindo da definição da função sinal (ROBERTS, 1980; SAYED, 2008):

$$\text{sgn}(x) := \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}, \quad (3.14)$$

buscou-se definir uma função que, quando aplicada a determinado índice, retornasse zero quando o índice não violasse o limite de controle e +1 e -1 quando o índice violasse o limite superior e inferior, respectivamente. Assim, essa função, denominada de função alarme, foi definida como:

$$f_{al}(x, c) := \frac{1}{2} [\text{sgn}(x - c) + \text{sgn}(x + c)] \quad , \quad (3.15)$$

para os argumentos $x \in \mathbb{R}$ e $c \in \mathbb{R}$, em que c e $-c$ representam os pontos de descontinuidade. O comportamento da função alarme para um valor arbitrário de c é apresentado na Figura 3.3A.

Dessa maneira, o cálculo dos alarmes consistiu em aplicar a Equação (3.15) a cada um dos índices definidos nas Equações (3.3)-(3.12) com os respectivos limites de controle. Por exemplo, o alarme para o índice de detecção de valores espúrios do resíduo de predição é determinado por

$$f_{al}(I_{o,z}, c_{o,z}) \quad .$$

Diz-se que um alarme está acionado quando apresenta valor diferente de zero.

Além da definição da função alarme, propôs-se a apresentação dos alarmes (ou seja, dos resultados da função alarme aplicada aos índices de detecção) na forma de diagrama de cores, conforme ilustra a Figura 3.3B. A apresentação na forma de diagrama de cores tem por objetivo promover uma rápida interpretação visual dos índices de detecção, que torna-se complexo à medida que mais índices e variáveis passam a ser monitorados.

3.3.4 Extensão do Conjunto de Dados com Variáveis Calculadas

A técnica OMP, e outros métodos lineares de construção de modelos, como o PCA, não é capaz de reproduzir as relações não lineares que eventualmente existam no processo apenas a partir dos dados. De maneira similar, esse método também não é capaz de capturar a dependência temporal existente nos dados, *i.e.*, os efeitos dinâmicos.

A abordagem mais simples para lidar com essas limitações consiste no aumento da dimensão do problema, em que o conjunto de dados é estendido com variáveis sem significado físico, que representam os efeitos não lineares de mistura dos sinais.

No caso da dependência temporal, por exemplo, variáveis defasadas no tempo foram incluídas no conjunto de dados, como descrito anteriormente. Somente após esse aumento da dimensão dos dados é que a dependência entre eles pode ser capturada com técnicas lineares. Duas técnicas que usam este tipo de abordagem são PCA generalizado (*Generalized PCA - GPCA*) (GNANADESIKAN, 1997) e KPCA (SCHÖLKOPF *et al.*, 1998).

Uma vez que um conjunto infinito de efeitos pode ser construído, o conhecimento do domínio da aplicação é útil para identificar os efeitos mais relevantes para a modelagem. No contexto de processos químicos transientes, algumas relações relevantes podem ser destacadas.

A derivada representa a taxa de variação, sendo adequada para identificar dinâmica no processo, e pode ser determinada por operações simples, como

$$\frac{dx_i}{dt} = \frac{x_i(t_k) - x_i(t_{k-1})}{t_k - t_{k-1}}$$

Outros procedimentos numéricos também podem ser usados para diferenciação sem qualquer perda de generalidade. Além disso, a diferenciação pode ser usada para calcular variáveis não medidas, como ocorre na determinação de dados de vazão

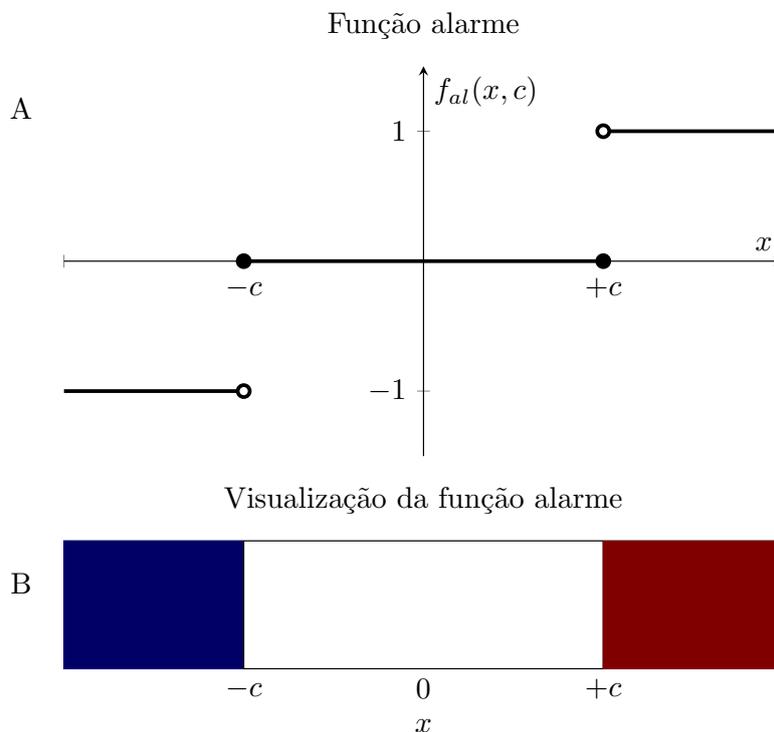


Figura 3.3: (A) Ilustração da função alarme definida na Equação 3.15; (B) estratégia de visualização do resultado da função alarme na forma de diagrama de cores: regiões em vermelho representam valores de x acima do limite $+c$, enquanto regiões em azul representam valores de x abaixo do limite $-c$.

volumétrica a partir das séries temporais provenientes de sensores de nível. Outra aplicação da diferenciação consiste em transformar sinais integrados no tempo, como dados medidos por totalizadores, instrumentos bastante comuns em processos reais para medir vazões. Em resumo, casos típicos em conjuntos de dados reais são a diferenciação de sinais fornecidos por totalizadores de massa (vazões), indicadores de nível (vazões e taxas de acúmulo de massa) e instrumentos de temperatura (taxas de calor).

Importantes combinações lineares podem ser identificadas a partir das equações de balanço de massa, energia e momento aplicadas em determinadas seções do processo, permitindo o cálculo de variáveis não medidas. Por exemplo, em um vaso de mistura de volume $V = AL$, em que A é a área da seção transversal do vaso (uma variável de projeto) e L é o nível (geralmente medido), com correntes únicas de alimentação (F_{in} , geralmente medido) e saída (F_{out} , geralmente medido), pode-se escrever:

$$A \frac{dL}{dt} - F_{in} + F_{out} = 0 \quad . \quad (3.16)$$

A combinação linear de sinais de processo medidos e transformados de acordo com a Equação 3.16 representa o balanço de massa, que pode ser usado para fins de monitoramento e detecção de falhas.

Continuando o exemplo anterior, se dados de analisadores de composição estiverem disponíveis, relações bilineares entre dados de vazão (F) e de composição de um componente i qualquer (x_i), podem ser usadas para calcular dados de vazão mássica (\dot{m}_i) a partir da equação:

$$\dot{m}_i = x_i F \quad .$$

Assim, diversos efeitos importantes podem ser construídos a partir do conjunto de dados original, permitindo que a técnica linear capture outras relações existentes no processo.

3.3.5 Algoritmo de Monitoramento do Processo

Admitindo que dados de processo, categorizados como dados de entrada e saída, estejam disponíveis (e que o conjunto de dados foi adequadamente estendido com variáveis calculadas), a primeira etapa consiste na geração do modelo, usando o método de modelagem multivariável. O modelo gerado é uma combinação linear das variáveis de entrada, em que a seleção das variáveis e o cálculo dos coeficientes são realizados pela técnica OMP. Após esta etapa, um procedimento iterativo é iniciado, envolvendo a aquisição dos dados de processo, a predição da variável de interesse e o cálculo do resíduo de predição, dos diversos índices de detecção e dos alarmes. Caso a avaliação dos alarmes aponte algum indício de falha, os termos de

contribuição das variáveis do modelo para a magnitude do resíduo podem ser usados para diagnóstico de falha. A análise dos alarmes não é rígida e diferentes hierarquias podem ser definidas.

Um possível fluxograma para detecção e diagnóstico de falhas usando o conjunto de alarmes proposto é apresentado na Figura 3.4. A primeira análise avalia os alarmes para o erro de predição da variável de interesse. Se nenhum desvio for detectado, os processos da planta são considerados em condição normal e o modelo é considerado apropriado para o monitoramento do processo e execução de RTO. Caso contrário, a detecção indica que o desvio precisa ser investigado. A segunda análise avalia os alarmes das variáveis selecionadas. Se uma ou mais variáveis diferirem de sua média histórica no período de treinamento, essas variáveis são apontadas como em condição de falha. Após o acionamento de um ou mais alarmes das variáveis selecionadas, uma terceira análise é realizada com o objetivo de julgar a necessidade de atualizar o modelo. Esta decisão pode ser baseada na intervenção de um usuário ou em um procedimento automático com base em critérios adicionais. Caso o desvio de predição não apresente efeitos nos alarmes das variáveis, pode indicar a necessidade de atualização do modelo. Em outras palavras, entende-se que se nenhuma variável foi apontada em condição de falha, os desvios de predição são causados por perda da validade do modelo. Assim, conclui-se que o processo encontra-se em novo regime de operação e o modelo precisa ser substituído.

O algoritmo de monitoramento também pode ser aplicado a subconjuntos das variáveis medidas. Assim, a construção de modelos e o monitoramento de índices pode envolver dados obtidos a partir de sensores de um ou mais tipos (como transmissores de pressão), de uma ou mais áreas da planta (como utilidades, tratamento de água *etc.*), ou de intervalos de tempo diferentes para cada grupo de sensores. O resultado é que várias instâncias do algoritmo de monitoramento podem ser usadas ao mesmo tempo, fornecendo diferentes análises para o mesmo processo industrial.

A inovação fundamental dessa aplicação diz respeito à visão integrada do monitoramento de processos, envolvendo a combinação das atividades de modelagem baseada em dados, das técnicas de controle estatístico de processos e FDD, em um contexto de *big data* com técnicas massivamente paralelas, e da forma de apresentação dos resultados dos alarmes por diagramas de cores, para caracterização do desempenho de processos reais e identificação de anormalidades que impeçam o uso de modelos fenomenológicos em um contexto de predição e otimização. Até onde se pode determinar, essa é a primeira implementação integrada desses métodos para a interpretação *on-line* de dados de processo em aplicação envolvendo milhares de variáveis.

3.4 Aplicação em Dados Industriais Reais

A estratégia de monitoramento para detecção e diagnóstico de falhas foi avaliada por meio de simulações aplicadas a um conjunto de dados reais. Os dados são provenientes de uma plataforma de produção de petróleo. A falha de operação considerada nesse sistema consiste na parada não programada de produção, comumente chamada de *shutdown* da unidade. Nesse contexto, admite-se que detectar e diagnosticar esse tipo de falha com a máxima antecedência possível de modo a evitá-la é uma otimização do processo, pois minimiza as perdas de produção. Além disso, o uso de um modelo incapaz de representar os dados é temerário em um sistema RTO.

O processo foi modelado de acordo com a Equação 3.2, em que a variável de interesse y no instante de tempo k ($y(t_k) = y_k$) é representada como a combinação

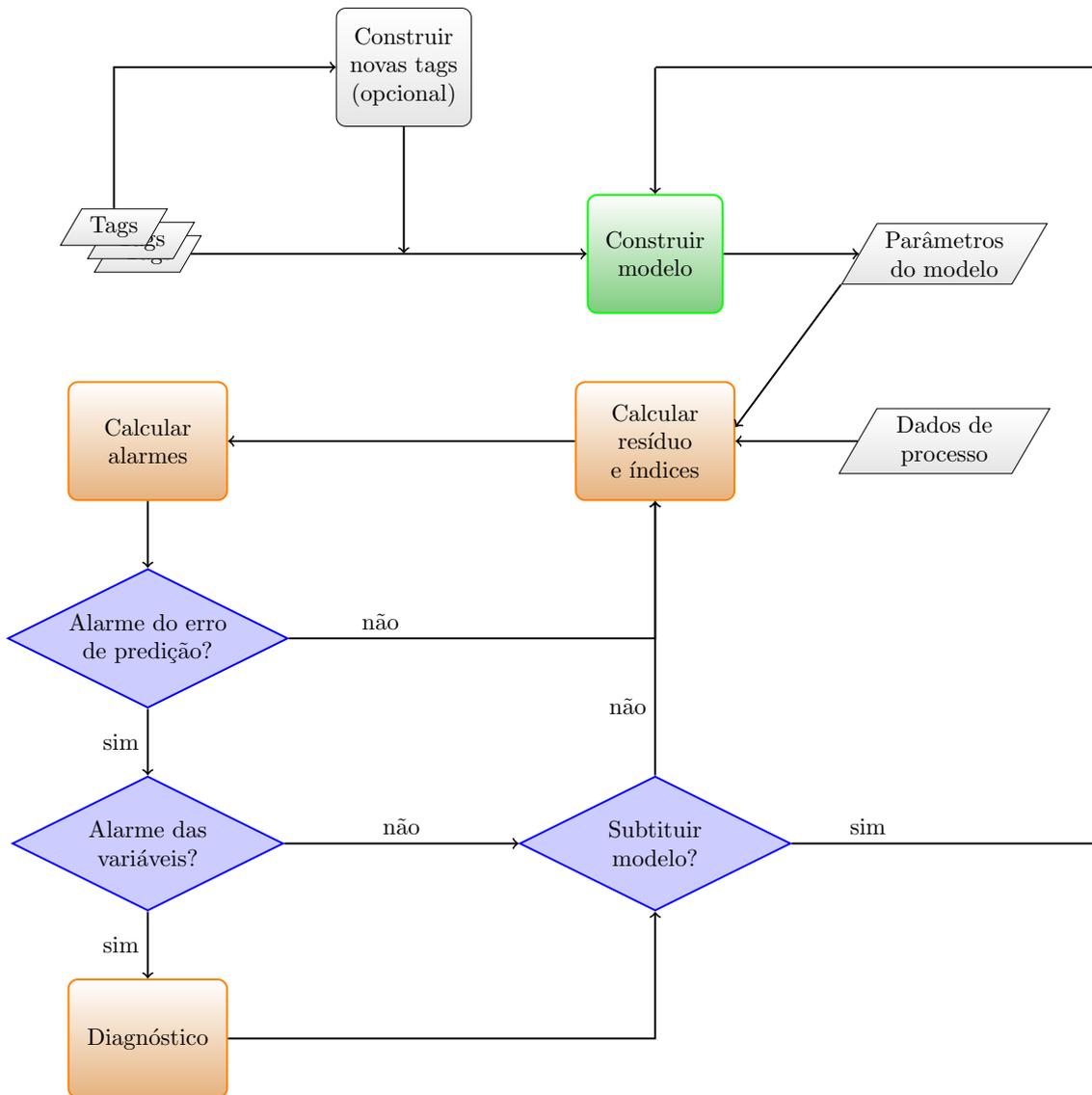


Figura 3.4: Proposta de fluxograma para detecção e diagnóstico de falhas usando o conjunto de alarmes proposto.

linear das outras variáveis medidas do processo, representadas por $\mathbf{x}_{l,k}$. A variável de interesse selecionada foi a vazão total de óleo produzida, mas outros índices de produção, ou mesmo um índice econômico, como a rentabilidade da operação, poderiam ser utilizados, simplificando a eventual implementação do RTO.

A determinação dos coeficientes \mathbf{b} e a seleção das variáveis \mathbf{x}_l do modelo foram realizadas com a técnica OMP (PATI *et al.*, 1993). Um algoritmo proprietário de uma versão paralelizada da técnica OMP foi utilizado, viabilizando o processo de identificação em tempo compatível com o necessário para o seu uso no monitoramento de processos, quando aplicado em sistemas constituídos por milhares de variáveis.

O modelo resultante foi então usado para monitorar a trajetória temporal de y , \mathbf{x}_l e do erro de predição de y , $z_k = y_k - \hat{y}_k$, em que \hat{y}_k é a predição do modelo para a variável de interesse y no instante de tempo k . O modelo foi construído usando um conjunto de dados de treinamento \mathbf{X} , que corresponde a um período histórico de valores medidos. Para a realização dos cálculos, todas as variáveis de processo \mathbf{x} foram normalizadas de acordo com a média e variância do conjunto de dados de treinamento.

O período de operação usado no estudo, para ambos os conjuntos de treinamento e validação, foi escolhido de modo a conter um evento da falha de interesse e ter duração suficientemente longa, para que a simulação do monitoramento contínuo fosse não tendenciosa e representasse o resultado obtido durante uma operação real, permitindo avaliar a robustez e o desempenho do monitoramento. O número de variáveis selecionadas para a construção do modelo também esteve baseado na visão de proporcionar simulações equivalentes ao de um sistema implementado em um processo real. Não obstante, a definição do número de variáveis selecionadas também levou em conta a predição da variável de interesse, para que o modelo não estivesse sobreparametrizado, ou seja, para que o modelo não incorporasse flutuações aleatórias na predição da variável resposta.

O conjunto de simulações foi realizado utilizando a plataforma Hadoop na infraestrutura computacional da empresa parceira EMC. Hadoop é a principal plataforma de processamento *big data* devido às altas confiabilidade, eficiência e escalabilidade (VANACE, 2009), baseando-se no sistema de arquivos do Google (GFS) e no modelo de programação *MapReduce* (DEAN e GHEMAWAT, 2008; LÄMMEL, 2008). O Hadoop é composto por duas partes principais: um sistema de arquivos distribuídos para suportar o armazenamento massivo de dados e uma estrutura de computação distribuída para suportar o processamento de dados usando o modelo MapReduce. A plataforma Hadoop, e o conseqüente uso do modelo MapReduce, foi aplicada no armazenamento, acesso e processamento distribuído dos dados de processos. Maiores detalhes sobre a plataforma Hadoop podem ser encontrados em VANACE (2009).

3.4.1 Avaliação do Monitoramento com o Conjunto de Dados Original

Dados das séries temporais ao longo de 26 dias de operação normal foram usados como dados de treinamento, selecionando as variáveis mais relevantes e construindo o modelo para predição da produção total de óleo (variável de interesse). Definiu-se em 100 o número de variáveis selecionadas para a construção do modelo. Em virtude das diferentes características do conjunto de dados exportados do sistema da planta industrial, as séries temporais foram tratadas com algoritmos de detecção de valores não válidos, interpolação, entre outros. Juntamente com as séries temporais amostradas de 1 em 1 minuto, novas séries atrasadas no tempo foram geradas para cada variável. As defasagens aplicadas foram os valores de 1, 2, \dots , 20 instantes de amostragem, correspondendo ao intervalo $[1, 20] \in \mathbb{N}$. Desse modo, o conjunto de dados utilizado, denominado conjunto de dados original, compreendeu as variáveis armazenadas no banco de dados da plataforma e essas mesmas variáveis defasadas no tempo.

A Figura 3.5A mostra os resultados de predição do modelo gerado frente aos dados medidos do conjunto de treino, apresentando um coeficiente de correlação (R^2) de 75,6% e raiz do desvio médio quadrático de predição (RMSE) de 45,3 unidades de produção, que corresponde a 11,2% da produção total média no período. Tratando-se de dados industriais reais, esses resultados são considerados bons para o contexto de unidades de produção de óleo *offshore*.

Resultados do alarme para o índice de detecção de valores espúrios do resíduo de predição ($I_{o,z}$) ao nível de confiança de 99,56% são apresentados na Figura 3.5B na forma de diagrama de cores, conforme definido na Figura 3.3B. Admitindo que o resíduo seja uma variável aleatória normalmente distribuída, espera-se que o alarme seja acionado em 164 instantes de amostragem ao longo do período de treino de acordo com o nível de confiança utilizado. Nesse sentido, observa-se que o alarme do índice $I_{o,z}$ se comporta dentro do esperado para uma operação normal para o conjunto de treino.

As variáveis selecionadas e o modelo foram usados em um experimento simulado de monitoramento do processo durante um período diferente daquele usado para a geração do modelo, denominado conjunto de validação, correspondendo aos 29 dias imediatamente subsequentes ao período dos dados de treino. No conjunto de validação, há dois eventos de falha de parada não programada da produção, sendo um durante o dia 3 e outro durante os dias 24, 25 e 26 do período. A Figura 3.6 mostra os resultados de predição do modelo frente aos dados medidos do conjunto de validação, bem como a trajetória do RMSE.

A Figura 3.7 mostra os resultados do alarme para o índice de detecção de va-

lores espúrios aplicado às variáveis selecionadas pelo modelo ($I_{o,xi}$) e ao resíduo de predição ao nível de confiança de 99,56%. Vale destacar que a estratégia de visualização por meio de diagrama de cores possibilita o empilhamento dos diagramas de diferentes alarmes, combinando-os em um único gráfico, como feito na Figura 3.7A. Isso permite a visualização imediata de como as variáveis, em conjunto, variam no tempo, por meio da evolução temporal dos alarmes, de quais alarmes foram acionados e com que frequência o acionamento ocorreu, gerando um padrão de desempenho. Esses padrões podem ser adicionalmente explorados por algoritmos específicos de identificação de padrões para avaliar o sistema. A sequência de empilhamento dos diagramas pode ser livre ou definida de acordo com relações de dependência entre as variáveis monitoradas segundo algum critério (por exemplo, variáveis de uma mesma seção do processo).

É interessante notar o comportamento do alarme do índice $I_{o,z}$ durante o período compreendido entre os dias 1 e 2 do conjunto de validação (Figura 3.7B). Nesse período, que antecede a ocorrência da falha no dia 3, a frequência de acionamento do alarme aumenta. Observa-se também o acionamento do alarme do índice de detecção de valores espúrios para um grupo de variáveis selecionadas pelo modelo. Comportamento bastante similar ocorre no período que antecede a segunda falha do período de validação, que inicia no dia 24. Além disso, há um intervalo entre os dias 10 e 17 em que os alarmes da maior parte das variáveis selecionadas não está ativo, enquanto o alarme do resíduo de predição permanece ativo em quase a totalidade desse período. Esse resultado indica uma sensível diferença no comportamento do processo com o restabelecimento da produção após a primeira falha, evidenciado

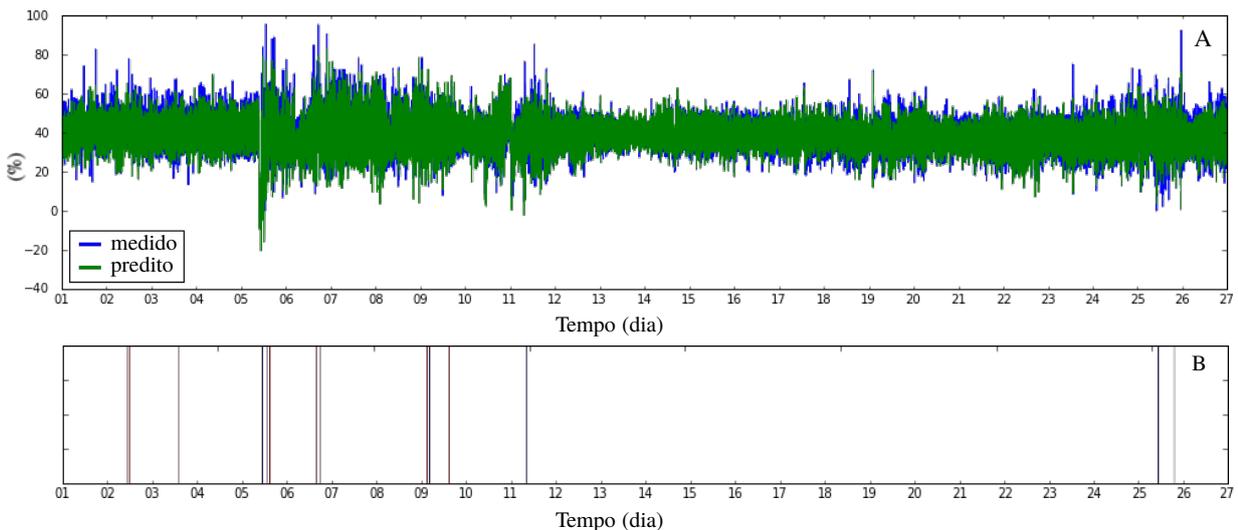


Figura 3.5: Resultados para o período de treino: A) valores medidos e preditos pelo modelo para a produção total de óleo; B) alarme para o índice de detecção de valores espúrios do resíduo de predição.

com mais intensidade pela informação estatística do resíduo de predição. Nesse caso, a ausência de alarmes nas variáveis selecionadas e a presença de alarmes nos desvios de predição podem indicar perda da validade do modelo e necessidade de substituição.

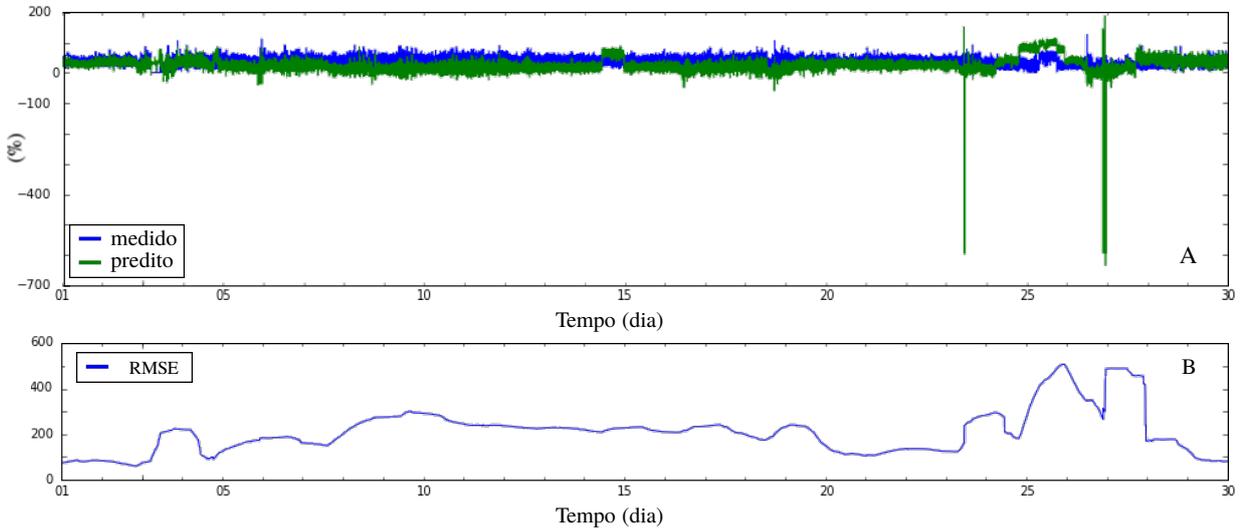


Figura 3.6: Resultados para o período de validação: A) valores medidos e preditos pelo modelo para a produção total de óleo; B) raiz do desvio médio quadrático de predição (RMSE).

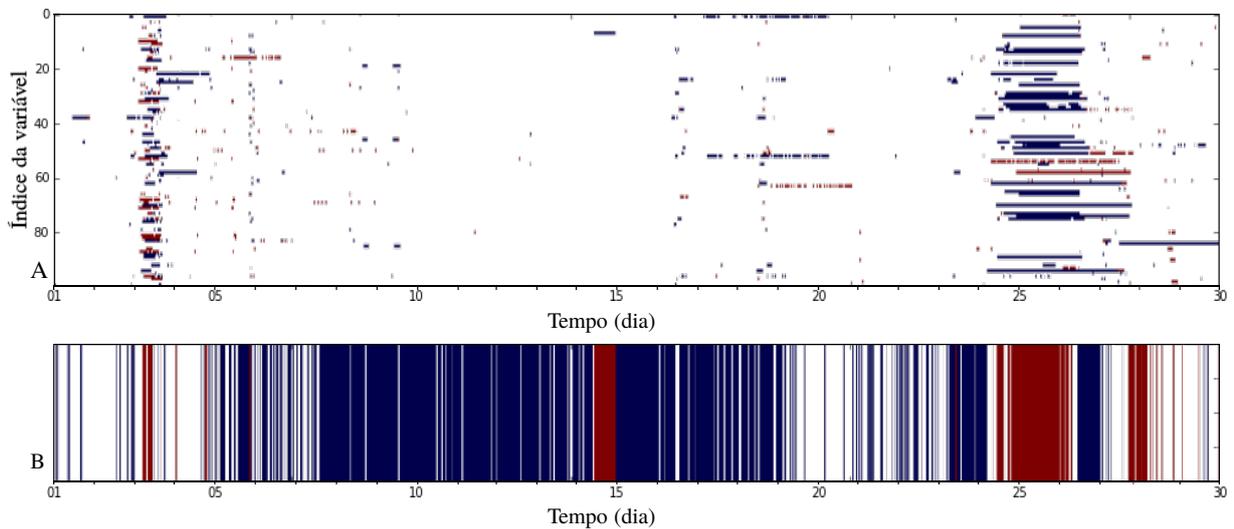


Figura 3.7: Diagrama de cores apresentando os resultados do alarme do índice de detecção de valores espúrios no período de validação: A) variáveis selecionadas; B) desvio de predição. Pontos vermelhos representam violação do limite máximo do teste, enquanto pontos azuis representam violação do limite mínimo.

Embora a análise qualitativa dos alarmes do índice I_o para as variáveis selecio-

nadas possa não indicar a ocorrência da falha de maneira inequívoca, a investigação das variáveis alarmadas mostrou-se bastante informativa para a percepção de anormalidades na operação. Isso foi concluído a partir das informações do relatório de operação, que relatou problemas relacionados a diversas variáveis alarmadas. Dessa forma, quando a utilização dos alarmes é comparada com o uso isolado do RMSE do modelo, nota-se que os sinais dos alarmes contribuem com informação útil. Quando não há falha, essa informação pode contribuir para a análise do desempenho do modelo, e quando há indício de falha, os sinais dos alarmes são utilizados para diagnóstico da falha. Assim, o uso combinado desses indicadores contribui para melhorar a avaliação do desempenho do modelo, a confiabilidade da análise da operação e o diagnóstico de um evento anormal.

3.4.2 Avaliação do Monitoramento com o Conjunto de Dados Estendido

Conforme discutido anteriormente, a extensão do conjunto de dados com variáveis calculadas pode permitir uma melhor modelagem do sistema a partir de informações extras que não são medidas pelos sensores e que não são capturadas pelas técnicas lineares com o conjunto de dados original.

Para compor o conjunto de dados estendido, um total de 55 novas séries temporais, originadas por diferenciação numérica e outras operações provenientes de balanços de massa, foram produzidas e incluídas no conjunto de dados, conforme discutido na Seção 3.3.4. As novas variáveis consistiram de séries temporais provenientes de sensores de nível, diferenciadas numericamente para produzir dados de vazão volumétrica, e séries temporais de dados de vazão volumétrica e de analisadores de composição, que foram multiplicadas para produzir dados de vazão mássica específica de um componente. Produto das séries originais, cada uma das novas séries também foi amostrada de 1 em 1 minuto e usada para produzir variáveis atrasadas no tempo com as mesmas defasagens aplicadas anteriormente, correspondendo, assim, a um total de 1155 novas séries temporais no conjunto de dados. As características do conjunto de dados, à exceção da inclusão das novas séries, os procedimentos de treinamento e teste do modelo e as análises dos alarmes foram idênticos aos descritos anteriormente.

Em virtude de um componente estocástico presente no algoritmo usado da implementação paralela da técnica OMP, o efeito da inclusão das novas séries temporais foi investigado não para um modelo, mas para um conjunto de modelos. Os modelos pertencentes a esse conjunto foram gerados de maneira independente, para quatro períodos de operação distintos, com duração média de 30 dias, duração essa equivalente àquela usada no monitoramento com o conjunto de dados original. O número

de modelos presentes nesse conjunto amostral correspondeu a 100, definido com base na convergência dos resultados estatísticos do conjunto em função do número de modelos.

Conforme mostram as Tabelas 3.1 e 3.2, a melhoria observada na geração do modelo com o conjunto de dados estendido é sistemática. Ambas as tabelas mostram que a adição das novas séries temporais melhora a qualidade do ajuste do modelo aos dados usados no treinamento, de acordo com os valores de R^2 e RMSE. Em média, 20% das variáveis selecionadas pelo algoritmo paralelizado da técnica OMP foram provenientes do conjunto de novas séries temporais, conforme mostra a Tabela 3.3. Isso constitui importante evidência de que a informação presente nas novas séries, calculadas a partir de transformações simples baseadas no conhecimento do sistema, corresponde a efeitos não lineares que a técnica OMP não foi capaz de modelar usando o conjunto de dados original.

Tabela 3.1: Valores do coeficiente de correlação (R^2) do modelo avaliado com os dados de treinamento de diferentes períodos para o conjunto de dados original e conjunto estendido. Estatísticas calculadas para 100 modelos.

Período	1		2		3		4	
	Original	Estendido	Original	Estendido	Original	Estendido	Original	Estendido
média	0.832	0.913	0.716	0.801	0.796	0.838	0.828	0.885
desvio padrão	0.014	0.014	0.020	0.026	0.020	0.022	0.034	0.021
mín	0.796	0.845	0.664	0.695	0.755	0.760	0.752	0.821
$P_{25\%}$	0.823	0.909	0.703	0.790	0.780	0.830	0.797	0.876
$P_{50\%}$	0.835	0.915	0.717	0.811	0.799	0.845	0.841	0.891
$P_{75\%}$	0.844	0.921	0.731	0.819	0.812	0.852	0.858	0.901
máx	0.857	0.930	0.757	0.832	0.826	0.871	0.877	0.913

Tabela 3.2: Valores de RMSE do modelo avaliado nos dados de treinamento de diferentes períodos para o conjunto de dados original e conjunto estendido. Estatísticas calculadas para 100 modelos.

Período	1		2		3		4	
	Original	Estendido	Original	Estendido	Original	Estendido	Original	Estendido
média	47.88	34.52	48.86	40.82	38.15	33.94	56.43	46.16
desvio padrão	2.00	2.61	1.70	2.59	1.82	2.20	5.49	4.03
mín	44.22	30.96	45.18	37.57	35.18	30.37	48.01	40.26
$P_{25\%}$	46.19	32.93	47.58	39.07	36.63	32.54	51.50	43.04
$P_{50\%}$	47.56	34.10	48.82	39.84	37.84	33.28	54.58	45.20
$P_{75\%}$	49.26	35.30	50.01	42.02	39.58	34.83	61.64	48.16
máx	52.89	46.02	53.18	50.68	41.78	41.40	68.06	57.89

Tabela 3.3: Número de variáveis provenientes do conjunto de novas séries temporais selecionadas na construção do modelo. Estatísticas calculadas para 100 modelos.

Período	1	2	3	4	Média
média	24.2	18.7	17.4	19.5	19.9
desvio padrão	3.9	4.1	3.5	3.7	3.8
mín	15.0	10.0	6.0	12.0	10.8
$P_{25\%}$	21.0	16.0	15.0	17.0	17.3
$P_{50\%}$	24.0	19.0	18.0	19.5	20.1
$P_{75\%}$	27.0	22.0	19.3	22.0	22.6
máx	34.0	28.0	27.0	30.0	29.8

Utilizando um dos modelos gerados para o mesmo período estudado na seção anterior, referente ao conjunto de dados original, efetuou-se o estudo do uso do modelo para FDD. A Figura 3.8 mostra os resultados de predição do modelo gerado frente aos dados medidos, apresentando R^2 de 82,5% e RMSE de 38,3 unidades de produção, que corresponde a 9,5% da produção total média no período. O modelo apresentou maior acurácia na predição, conforme refletem os valores de R^2 e RMSE. Novamente, esses resultados são considerados bons, considerando o contexto de unidades de produção de óleo *offshore*, mas principalmente pela comparação com os resultados obtidos anteriormente para o conjunto de dados original. Essa melhoria foi refletida também no alarme do índice $I_{o,z}$, Figura 3.8B, em que o alarme foi acionado com frequência menor, quando comparado aos resultados obtidos para o conjunto de dados original no período de treino.

Particularmente, algumas das novas séries temporais, como vazões, estão mais diretamente relacionadas à variável de interesse (produção de óleo). Além disso, os inventários em processo dependem das taxas de acúmulo nos vasos, relacionadas às derivadas temporais das medidas de nível. Portanto, o cálculo de derivadas temporais (como das medidas de nível) e vazões por balanço de massa é útil na descrição mais acurada da produção, permitindo também a redução do tamanho do modelo com o uso de derivadas, em vez de múltiplas variáveis medidas defasadas no tempo.

Conforme feito anteriormente, o modelo foi usado para simular o monitoramento *on-line* do processo durante o período de validação. A Figura 3.9 mostra os resultados de predição do modelo frente aos dados medidos para o conjunto de validação, bem como a trajetória dos valores de RMSE. A Figura 3.10 mostra os resultados do alarme do índice de detecção de valores espúrios aplicado às variáveis selecionadas pelo modelo e para o resíduo de predição.

O comportamento geral observado é semelhante àqueles obtidos para o conjunto de dados original. Todavia, o aumento da acurácia do modelo foi também acompa-

nhado por um aumento na sensibilidade à condição da falha, conforme comparado na Figura 3.9B entre os dias 3 e 4 do período. Isso permite detectar a falha com maior clareza, aumentando a confiabilidade do procedimento. Outro fator observado é que a falha afeta menos a acurácia do modelo no período posterior à sua ocorrência, quando comparado aos resultados observados para o conjunto de dados original, revelando uma maior robustez do modelo e maior capacidade de permane-

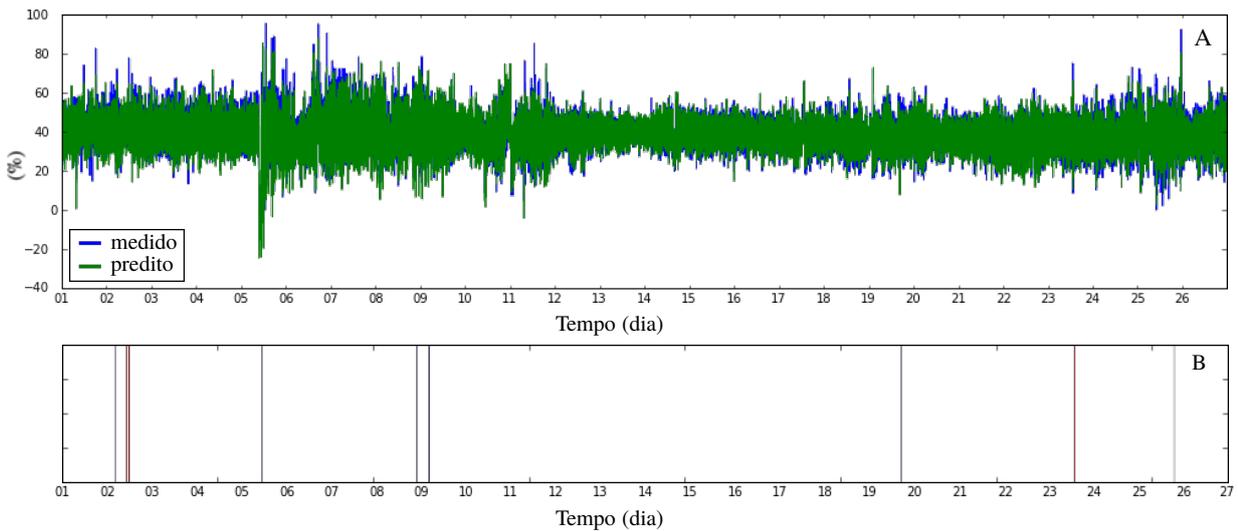


Figura 3.8: Resultados para o período de treino com o conjunto estendido de variáveis: A) valores medidos e preditos pelo modelo para a produção total de óleo; B) alarme do índice de detecção de valores espúrios para o resíduo de predição.

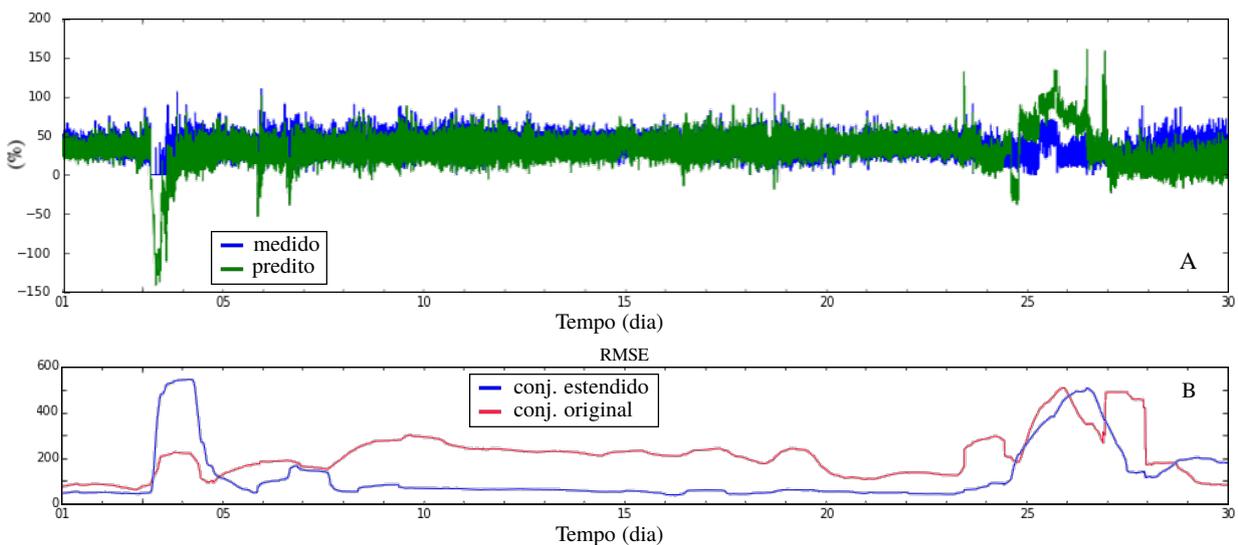


Figura 3.9: Resultados para o período de validação com o conjunto estendido de variáveis: A) valores medidos e preditos pelo modelo para a produção total de óleo; B) raiz do desvio médio quadrático de predição (RMSE).

cer válido após condições anormais. Assim, quando o conjunto de dados é estendido para incluir variáveis de processo calculadas com a ajuda de equações de balanço de massa, novas variáveis observadas (embora não medidas) são usadas e permitem a criação de modelos empíricos mais precisos.

Apesar da criação das novas séries temporais ser avaliada a partir da comparação entre os modelos gerados, o valor do modelo está vinculado à sua utilidade na aplicação. Dessa forma, não defende-se o uso do modelo obtido a partir do conjunto estendido, mas sim do modelo que melhor atende aos objetivos pretendidos. Nesse sentido, o uso de múltiplos modelos seria também uma possibilidade válida (e viável). Características obtidas com o uso de múltiplos modelos (inclusive modelos para diferentes variáveis de interesse) poderiam ser investigadas para produzir resultados úteis à atividade de FDD. Além disso, também é possível fornecer uma avaliação de desempenho do modelo que indica a necessidade de atualização ou substituição do modelo devido à ocorrência de falhas ou novas condições de operação.

Embora esse estudo tenha focado no monitoramento da operação da planta ao longo do tempo, o modelo resultante pode ser aplicado a diferentes finalidades. Como o modelo é usado *on-line* e dados podem ser usados em tempo real para identificar um ou mais modelos, o método pode ser aplicado não apenas para a análise contínua da condição do processo, mas também para outras aplicações, como controle de processos (EATON *et al.*, 2017). O procedimento de identificação do modelo em tempo real não interrompe a operação e a inclusão de novas variáveis o torna

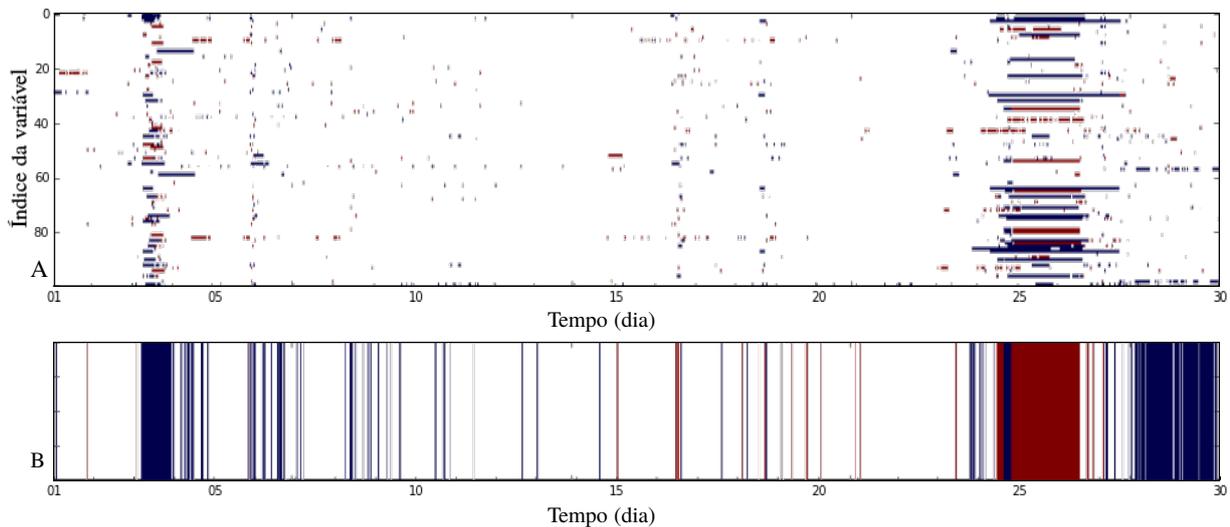


Figura 3.10: Diagrama de cores representando os resultados dos alarmes do índice de detecção de valores espúrios no período de validação com o conjunto estendido de variáveis: A) variáveis selecionadas; B) desvio de predição. Pontos vermelhos representam violação do limite máximo do teste, enquanto pontos azuis representam violação do limite mínimo.

adequado para processos não lineares, bem como aos casos em que a identificação do modelo *off-line* é difícil.

Conforme observado anteriormente, a investigação das variáveis alarmadas mostrou-se informativa para a percepção da falha analisada, considerando as informações do relatório de operação. No entanto, a presença das variáveis calculadas reforçou a interpretação da descrição da falha com base nas variáveis alarmadas.

3.5 Conclusões

O desafio de tratar e explorar grandes conjuntos de dados não é recente, mas ganhou novas atenções nos últimos anos. Muitos incentivos surgiram nessa área com o amplo acesso e maior disponibilidade de dispositivos e sensores sem fio, acompanhado por novos modelos de negócio na área de sistemas de informação. Nas indústrias, esse movimento tem recebido o nome de Indústria 4.0 e materializa-se, por exemplo, com a aplicação de plataformas de *big data* para processamento distribuído e técnicas de aprendizado de máquina aos dados de processo, para extrair informações que contribuam com o aumento de eficiência da operação.

Foi proposto nesse trabalho um procedimento de detecção e diagnóstico de falhas com base em técnicas de controle estatístico de processos apoiado em *big data*. O procedimento fez uso da técnica OMP para a construção de um modelo dinâmico linear do processo e para a análise dos sinais e do resíduo de predição. Testes de hipóteses, definidos para detecção de *outliers*, tendências dinâmicas e *bias*, foram aplicados a diferentes sinais, refletindo o estado do sistema, em relação a um conjunto de dados de referência, e a qualidade de ajuste do modelo. Propôs-se também a conversão do resultado dos testes de hipótese, chamados de índices de detecção, em alarmes e a visualização desses alarmes por diagrama de cores, que mostraram-se úteis para uma rápida interpretação do estado do sistema. Também foi proposto um conjunto de operações para estender o conjunto de dados medidos do processo com base em conhecimento do domínio, como balanços de massa, permitindo a construção de modelos não lineares. A plataforma Hadoop de processamento *big data* foi empregada para viabilizar a construção rápida de modelos e o cálculo contínuo dos alarmes para aplicações *on-line* em processos com milhares de variáveis medidas. O procedimento mostrou-se capaz de detectar o comportamento anormal e de fornecer informações relevantes para o diagnóstico da causa provável, tornando possível que problemas possam ser detectados e resolvidos antes de sua ocorrência. O procedimento é capaz de auxiliar processos de tomada de decisão e análises de causalidade, complementando o sistema de automação e controle.

Embora diversas características da proposta possam ser encontradas na literatura, sua utilização ocorre de forma isolada. O procedimento proposto é único ao

integrar diversos aspectos na composição do sistema de monitoramento *on-line* de processos, envolvendo a combinação de: i) modelagem dinâmica baseada em dados, valorizando transformações não lineares específicas ao domínio da aplicação; ii) técnicas de controle estatístico de processos e FDD; iii) estrutura computacional *big data* com técnicas massivamente paralelas; iv) forma de apresentação dos resultados dos alarmes por diagramas de cores; e v) aplicação para a caracterização do desempenho de processos reais e identificação de anormalidades, em que a elevada dimensão do sistema impõe grandes desafios quanto ao uso de modelos fenomenológicos para predição e otimização.

Além dos desafios relativos ao desenvolvimento e uso de modelos fenomenológicos, a aplicação de sistemas RTO em processos de elevada dimensão ainda é um desafio. Embora existam exceções, o escopo típico de sistemas RTO é, em geral, o mesmo de controladores MPC e desafios adicionais surgem com a presença de muitos produtos intermediários e restrições logísticas (caso de refinarias de petróleo), e com o elevado número de modos de operação ou opções de equipamentos em uma determinada unidade (DARBY *et al.*, 2011). Ademais, permanece a necessidade de garantir a estacionariedade do processo em virtude do uso de modelos estáticos. Diante disso, o procedimento proposto contribui nessas duas frentes, tanto para o desafio de aplicações em processos de elevada dimensão, quanto para a detecção de estado estacionário. Uma vez que o sistema identifica o estado do processo, é possível informar se o sistema está em regime transiente, ou se está estacionário sob ação de alguma anomalia, situações em que a execução do RTO não produzirá resultados confiáveis. De outro modo, informações sobre a degradação do modelo do sistema de monitoramento podem ser combinadas ao sistema RTO para identificar regiões operacionais em que o modelo do RTO não apresenta bom desempenho. Até onde se pode determinar, aplicações desse tipo, conectando sistemas de monitoramento e FDD a sistemas RTO, não existem relatadas na literatura.

Capítulo 4

Processos com Baixa Disponibilidade de Informações

Conforme discutido no capítulo anterior, a dimensão do problema impacta significativamente a complexidade da solução, influenciando a estratégia e a abordagem usadas na implementação de procedimentos em tempo real. Todavia, as dificuldades de modelagem e implementação de procedimentos em tempo real não é eliminada pela redução da dimensão do problema.

Diversos processos industriais, como aqueles relacionados à tecnologia de polimerização e biotecnologia, são regidos por fenômenos que nem sempre dispõem de tecnologias de sensoriamento e medição maduras e adequadas ao ambiente fabril, ao contrário da área petroquímica. Na grande parte das situações, tais processos contam com o indispensável auxílio de um laboratório analítico para acompanhamento das rotinas de produção, mas cuja estrutura está dimensionada para atender pequenas demandas; ademais, alguns procedimentos analíticos podem ser caros, laboriosos e exigir longos tempos de análise. Por detrás da dificuldade de medição de diversas grandezas importantes para esses processos, reside o desafio de construir e atualizar modelos fenomenológicos para uso em procedimentos em tempo real, que corresponde à etapa de identificação, conforme fluxograma da Figura 1.2.

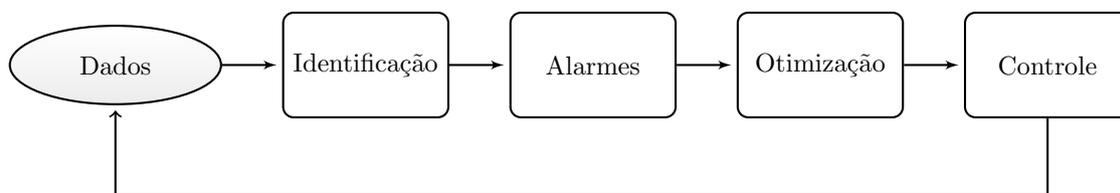


Figura 1.2: Fluxograma alternativo com as etapas envolvidas em uma implementação real de sistema RTO (apresentado na p.8).

Nesse contexto, em que a quantidade de informação é baixa, o uso de modelos empíricos é inviável para a identificação de processo e, dentro da perspectiva de mo-

delagem caixa cinza, o uso de modelos semimecanísticos torna-se uma alternativa promissora. Modelos semimecanísticos são modelos cuja estrutura matemática está baseada em modelos fenomenológicos, mas que não têm a pretensão de representar detalhadamente todos os mecanismos envolvidos. Assim, esse tipo de modelo apresenta caráter mais pragmático, que pode ser explorado em procedimentos em tempo real.

Tendo estas questões em vista, foi desenvolvido um procedimento de identificação *on-line* para monitoramento de processos. A estratégia abrangeu o desenvolvimento e uso de um modelo fenomenológico, semimecanístico e não estruturado com base em reações cinéticas e balanços de massa e energia. O modelo foi utilizado como um modelo de referência no procedimento *on-line* de identificação. O procedimento de identificação *on-line* consiste em atualizar o modelo de referência por meio da estimação de parâmetros corretivos para determinadas taxas de reação, usando as medidas de processo mais recentes. Abordagens similares têm sido usadas de diferentes maneiras na literatura (ASSIS e MACIEL FILHO, 2000; MAIA *et al.*, 2018).

O objeto desse estudo foi um processo industrial de fermentação para produção de etanol, mais especificamente a sacarificação e fermentação simultâneas (SSF) de hidrolisado de milho. A estratégia fez uso de medidas padrão de laboratório para quantificação de açúcares e dados *in situ* de temperatura e nível de líquido. O modelo, juntamente com o procedimento de identificação *on-line*, foi testado em dados industriais reais pela primeira vez e conseguiu prever com precisão as principais variáveis de interesse operacional: variáveis de estado com sua dinâmica e indicadores do processo.

O estudo mostra que modelos fenomenológicos simples podem ser usados para descrever a operação de processos, se parâmetros de ajuste do modelo são estimados com os dados disponíveis de nível e temperatura do fermentador e brix do mosto, uma vez conhecida a carga inicial da unidade. Estas medidas estão usualmente disponíveis em plantas reais. Os resultados mostram que a estratégia foi capaz de monitorar em tempo real esta complexa fermentação industrial de biomassa. A atividade é relevante, pois processos que usam enzimas e/ou micro-organismos estão sujeitos a elevada variabilidade, por conta das inevitáveis flutuações de carga, qualidade dos insumos naturais e contaminação de linhas e vasos por micro-organismos, de maneira que os parâmetros estão sujeitos a flutuações pouco previsíveis durante a operação real. Esta nova estratégia oferece uma excelente ferramenta de suporte à decisão e abre uma nova gama de oportunidades para otimização industrial nesses processos, em ambiente em que as atualizações de processo são fundamentais.

O presente trabalho foi realizado em colaboração com a empresa Global Yeast do Brasil Ltda. e publicado no periódico *Bioprocess and Biosystems Engineering* (CÂMARA *et al.*, 2017a). O arranjo proposto inova as etapas de *Identificação*,

Alarmes e Otimização da Figura 1.2, em particular como estratégia de operação de vasos industriais de fermentação.

4.1 Introdução

O aumento das preocupações com a segurança energética e as questões ambientais aceleraram a pesquisa e o desenvolvimento de fontes de energia renováveis, especialmente os biocombustíveis (IPCC, 2014). Ademais, ações que favorecem a produção e o comércio de biocombustíveis foram tomadas por governos em todo o mundo, por meio de novas políticas e incentivos para criar e apoiar a crescente demanda, sendo que os principais instrumentos comumente empregados são metas obrigatórias de mistura, isenções fiscais e subsídios (ANG *et al.*, 2016; CHANTHAWONG e DHAKAL, 2016; DUFFIELD *et al.*, 2015; PAIANO e LAGIOIA, 2016; PEREIRA *et al.*, 2012).

O etanol é um dos biocombustíveis renováveis mais importantes, sendo comercialmente produzido a partir de amido e culturas à base de açúcar, incluindo cana-de-açúcar, beterraba, sorgo, milho, trigo, cevada, batata, inhame e mandioca. Além desse etanol de primeira geração, houve também o lançamento de plantas de escala comercial capazes de produzir etanol a partir de resíduos de biomassa lignocelulósica, especialmente nos Estados Unidos, Brasil, Europa e Canadá (BACOVSKY *et al.*, 2013). Em 2015, a produção comercial global de bioetanol atingiu quase 25,5 bilhões de galões, o que corresponde a aproximadamente 57%, 28%, 5%, 3% e 2% da produção para Estados Unidos, Brasil, União Européia, China e Canadá, respectivamente (RFA, 2016). Nos Estados Unidos, o maior produtor mundial de etanol, quase toda a produção de etanol (ou >98%) usa o milho como matéria-prima, conferindo à indústria de etanol de milho o título de maior produtor de biocombustíveis (DUFFIELD *et al.*, 2015).

A principal rota para a produção de etanol a partir do milho envolve basicamente quebrar o amido presente no milho em açúcares fermentescíveis (glicose), formar um caldo, inocular este caldo com levedura e recuperar o produto principal (etanol) e subprodutos (por exemplo, alimentação animal). O processo pode ser representado em cinco etapas gerais: i) moagem, ii) liquefação, iii) sacarificação, iv) fermentação e v) destilação e recuperação. Comercialmente, duas tecnologias principais são utilizadas para a produção de etanol de milho, que diferem em termos do processo de moagem usado: moagem úmida e moagem seca (ou moagem a seco) (MOSIER e ILELEJI, 2014). Da produção de 14,7 bilhões de galões de etanol em 2015 nos EUA, 10% vieram de usinas de moagem úmida, enquanto a grande maioria de 90% deveu-se a processos de usinas de moagem seca, que somam 199 plantas de processo (RFA, 2016).

Consideradas verdadeiras biorrefinarias, os processos de moagem úmida permitem a produção de muitos produtos de alto valor agregado, além do etanol combustível (BOTHAST e SCHLICHER, 2005; CARDONA e SÁNCHEZ, 2007). O termo “moagem úmida” é devido à primeira etapa no processo, que consiste em embeber o grão em água. Durante a etapa de moagem, o grão é separado em seus componentes, incluindo amido, fibra, glúten e gérmen. Uma solução de amido é separada dos sólidos e utilizada para produzir etanol. Os demais componentes, fibra, glúten e gérmen, são utilizados para produzir dois produtos para alimentação animal, farelo de milho (alta proteína, 40%) e glúten de milho (baixa proteína, 28%), e óleo de milho (extraído do gérmen de milho), que são os principais coprodutos deste método. No entanto, a capacidade de separar e produzir outros componentes valiosos torna o processo de moagem úmida mais intensivo em capital e energia. Por esta razão, todas as fábricas de etanol recentes empregam alguma variação no processo básico de moagem a seco, uma vez que essas plantas podem ser construídas com investimento menor, com foco na maximização do retorno de capital por litro de etanol (BOTHAST e SCHLICHER, 2005; MOSIER e ILELEJI, 2014).

Em plantas de moagem a seco, o grão inteiro é processado, de modo a fermentar o máximo possível, e os componentes residuais são separados no final do processo. Primeiro, o milho é moído a uma farinha, que é então embebido com água e aquecido a temperaturas de 110–120 °C (etapa chamada de cozimento). Na sequência, a etapa de liquefação (80–90 °C) ocorre na presença de enzimas amilases, para hidrolisar o amido presente no endosperma da semente em moléculas de cadeia curta (dextrinas) (GRAY *et al.*, 2013; GUO *et al.*, 2015). O hidrolisado de milho segue para a etapa de sacarificação (30 °C), que completa a quebra de amido e dextrinas em glicose, com a ajuda de enzimas α -amilase and glucoamilase. Adiciona-se então a levedura ao hidrolisado de milho sacarificado a aproximadamente 32 °C para converter a glicose em etanol. A fermentação dura cerca de 50–60 h, produzindo cerca de 8–12% em massa de etanol (GUO *et al.*, 2015). Finalmente, o etanol é recuperado por destilação, passando por processos de peneiras moleculares para produção de etanol anidro. A água residual e os sólidos de milho que permanecem após a destilação, também chamado de “stillage”, são então centrifugados e secos para produzir outro produto conhecido como *distillers’ dried grains with solubles* (DDGS), que é vendido como suplemento para alimentação animal (MOSIER e ILELEJI, 2014).

Plantas modernas de etanol de moagem seca nos EUA mostraram rendimentos de 2,8–2,9 galões de etanol (RFA, 2016) e cerca de 17,5 libras de DDGS (DUFFIELD *et al.*, 2015) por *bushel* de milho (25.4 kg, ou 56 lb). A quantidade de enzimas pode ser reduzida se a sacarificação do hidrolisado de milho for feita antes da fermentação a temperaturas de aproximadamente 65 °C. No entanto, a maior parte do etanol combustível dos EUA é produzido por sacarificação e fermentação simul-

tâneas (SSF), pois previne a contaminação microbiana, diminui o estresse osmótico inicial da levedura, evitando uma solução de glicose concentrada, e geralmente é mais eficiente em termos de energia (BOTHAST e SCHLICHER, 2005; DEVANTIER *et al.*, 2005; GRAY *et al.*, 2013). Durante a SSF, as dextrinas presentes no hidrolisado de milho, e não o amido, são sacarificadas pelas enzimas para a produção de glicose. Na prática, a sacarificação ocorre geralmente enquanto o hidrolisado está sendo alimentado no vaso de fermentação e continua durante todo o período da fermentação. O SSF é considerado a etapa mais crítica do processo, pois determina os rendimentos de etanol e a eficiência de conversão de todo o processo. Portanto, o biorreator em que o hidrolisado de milho é convertido em etanol é um foco natural para buscar melhorias no processo.

Geralmente, a melhoria da operação e o aumento de rentabilidade de bioprocessos são alcançadas por meio do aumento da qualidade da matéria-prima (SINGH, 2012), melhoria do fenótipo das linhagens celulares (MUKHERJEE *et al.*, 2014; PAIS *et al.*, 2013; WANG *et al.*, 2009) e uso de ferramentas de engenharia de processos, como técnicas de monitoramento, controle e otimização (STEPHANOPOULOS e REKLAITIS, 2011). Apesar da presença e aplicação dos dois primeiros, o último está sempre presente até certo ponto. De fato, o monitoramento do processo, isto é, a atividade contínua de determinar o estado atual de um sistema (físico, químico ou biológico) (ISERMANN e BALLÉ, 1997), é essencial para o controle de processos e de qualidade. Além disso, é fundamental para diminuir os riscos para a saúde e outros riscos relacionados aos processos de biotecnologia. Nesse contexto, os métodos de monitoramento *on-line* fornecem informações valiosas sobre o sistema em tempo real, permitindo a identificação antecipada de falhas, o diagnóstico das fontes de distúrbios do processo e que operadores de processos e sistemas automatizados implementem decisões fundamentadas no conhecimento do sistema. Assim, o monitoramento *on-line* constitui um passo importante na redução da variabilidade do processo e dos custos de produção.

Apesar disso, muitas plantas de etanol de milho ainda dependem de métodos de monitoramento *off-line*, em que a maioria dos parâmetros de valor real para a fermentação não são monitorados ou controlados em tempo real. Isto é em parte devido à complexidade do processo SSF, que não é completamente monitorado em plantas de moagem seca de milho. Mesmo que existam tecnologias de sensoriamento capazes de fornecer dados em tempo real de alguns parâmetros de interesse, apenas uma pequena fração deles é adequada para uso em aplicações industriais. Vários autores discutiram e analisaram equipamentos de medida (sensores), sistemas analíticos e técnicas de monitoramento *on-line* de bioprocessos (ALFORD, 2006; BEUTEL e HENKEL, 2011; CLEMENTSCHITSCH e BAYER, 2006; HARMS *et al.*, 2002; LOCHER *et al.*, 1992; LOURENÇO *et al.*, 2012; OLSSON *et al.*, 1998; PONS, 1993;

RICH e MYSZKA, 2003; SCHÜGERL, 2001; VOJINOVIĆ *et al.*, 2006).

Os desafios para medir as variáveis críticas de processo e a aquisição de dados *on-line* decorrem de vários fatores. Os meios de cultura geralmente possuem substâncias que interferem nas leituras dos sensores ou podem precipitar na superfície de contato. Além disso, culturas em crescimento podem infiltrar-se nos sensores, invalidando seus resultados (MURTHY *et al.*, 2012b). As técnicas de espectroscopia de infravermelho próximo (*near-infrared spectra*, NIR) (FINN *et al.*, 2006) e infravermelho médio (*mid-infrared spectra*, MIR) (SIVAKESAVA *et al.*, 2001) podem ser usadas para determinar vários componentes, *e.g.*, glicose, etanol e densidade celular, mas o alto custo e a dificuldade de calibração impedem o uso generalizado desses métodos. A espectroscopia Raman também mostrou-se útil para o monitoramento *in situ* de bioprocessos, incluindo SSF de hidrolisado de milho em condições industrialmente relevantes (GRAY *et al.*, 2013). No entanto, como acontece com NIR e MIR, os dados são geralmente difíceis de interpretar devido à sobreposição espectral e interferências complexas, o que pode resultar em menor precisão. Biossensores, sensores químicos em que o sistema de reconhecimento é baseado em mecanismos bioquímicos ou biológicos (BĂNICĂ, 2012), podem apresentar seletividades interessantes para muitas substâncias, mas muitas vezes apresentam problemas de calibração durante bioprocessos longos devido à degradação dos componentes biológicos (PONS, 1993). Como as fermentações de etanol duram vários dias, os sensores devem ser estáveis e capazes de operar ao longo de todo o período, apesar das mudanças na temperatura e na composição do caldo.

Poucos estudos investigaram o uso de uma técnica de monitoramento *on-line* em condições industriais relevantes. A aplicação de alguns dispositivos na produção industrial, como os sistemas de análise de injeção de fluxo (*flow injection analysis*, FIA), são possíveis apenas se forem totalmente automáticos e sua confiabilidade melhorada, exigindo experiência e tempo (SCHÜGERL, 2001). De fato, várias limitações, incluindo instrumentação e manipulação complicadas, equipamentos caros e interferências, fazem com que alguns métodos sejam mais apropriados para pesquisa que para monitoramento de produção (VOJINOVIĆ *et al.*, 2006). Medições típicas disponíveis durante a execução da batelada são temperaturas, volume ou nível de líquido e vazões volumétricas ou mássicas de algumas correntes. Informações relacionadas a restrições de qualidade, tais como concentrações, são realizadas em laboratório, geralmente com elevada precisão, mas são disponibilizadas com atraso. Em alguns casos, elas podem ser bastante escassas, disponíveis apenas ao final de cada batelada ou até mesmo apenas para poucas bateladas. Uma parte da razão para essas medidas serem pouco frequentes está relacionada ao custo do operador, da amostragem e do procedimento de análise, que não consiste em uma razão técnica.

Uma solução interessante para monitorar processos quando não existe uma rede

de sensores adequada ou quando equipamentos confiáveis e precisos não estão disponíveis é estimar variáveis de interesse. Existe um grande interesse tanto na academia quanto na indústria pelo desenvolvimento de sensores virtuais e muitos trabalhos trazem uma revisão sobre o tema (BASTIN e DOCHAIN, 1990; DOCHAIN e PERRIER, 1997; FORTUNA *et al.*, 2007). Esta estratégia emprega modelos matemáticos, conhecidos como sensores virtuais (ou *soft sensors*), para prever quantidades que são não medidas ou não são mensuráveis em tempo real com base no dados de processo disponíveis *on-line* e *off-line*. Além de poder fornecer um sistema de medição alternativo para variáveis de processo e propriedades que não são diretamente mensuráveis, também podem ser construídos sensores virtuais para identificar falhas em instrumentos de medida (LENNOX *et al.*, 2001).

O desenvolvimento de modelos para sensores virtuais pode basear-se apenas em dados medidos, resultando em modelos de natureza empírica orientados por dados (KADLEC *et al.*, 2009; LIN *et al.*, 2007), ou podem basear-se em conhecimentos fenomenológicos sobre o processo (ASSIS e MACIEL FILHO, 2000), em que um modelo de primeiros princípios (FPM) descreve o processo. Os métodos comuns aplicados na abordagem baseada em dados são as técnicas de aprendizagem de máquina, ou seja, análise de componentes principais (PCA) e mínimos quadrados parciais (PLS), redes neurais artificiais (ANN), análise independente de componentes (ICA) e máquinas de vetores suporte (SVM). Apesar da maior popularidade, sensores virtuais baseados em dados são geralmente pouco confiáveis para extrapolação e sua capacidade de generalização não pode ser garantida. Além disso, algumas técnicas baseadas em otimização, como redes neurais, podem convergir para mínimos locais, degradando assim a qualidade das inferências do sensor virtual (YU, 2012). Nesses casos, a falta de informação sobre o fenômeno observado compromete o resultado de procedimentos baseados em dados e a construção do modelo empírico é difícil devido a escassez de dados, ao contrário do cenário apresentado no Capítulo 3 e dos excelentes resultados obtidos.

Vários sensores virtuais foram desenvolvidos para monitorar a fermentação de glicose a etanol (KABBAJ *et al.*, 2001; MELEIRO e MACIEL FILHO, 2000; ÖDMAN *et al.*, 2009), mas muito poucos foram baseados em modelos de primeiros princípios. Quando os dados são escassos, a informação contida em modelos fenomenológicos pode ser crucial o sucesso da aplicação. A característica básica do método fenomenológico é deduzir equações matemáticas baseadas em balanços de massa e energia, bem como no conhecimento sobre os fenômenos que ocorrem no processo. Uma vez que um modelo está disponível, outras técnicas baseadas em modelos podem ser implementadas.

Para ser aplicável no contexto industrial, uma ferramenta de monitoramento SSF deve fornecer medidas rápidas e confiáveis de variáveis chave para a determinação

de critérios de qualidade, como etanol, oligossacarídeos e glicose. No entanto, como a fermentação é um processo exotérmico e os vasos de fermentação industrial são muito grandes, é importante desenvolver uma ferramenta capaz de monitorar não só a composição do caldo, mas também os efeitos térmicos, que raramente são considerados. Poucos estudos avaliaram a aplicação de um sensor virtual FPM para o monitoramento da produção de etanol a partir do milho em tempo real, e ainda menos trabalhos avaliaram essa técnica com dados industriais reais.

Considerando que a disponibilidade de dados medidos e a atividade de identificação são essenciais no contexto de sistemas RTO, este capítulo apresenta o desenvolvimento de um sensor virtual para monitoramento do SSF de hidrolisado de milho. Para isso, propôs-se um modelo fenomenológico e uma estratégia para estimação de parâmetros em tempo real. O modelo é baseado em reações cinéticas semimecanísticas e não estruturadas para hidrólise enzimática e fermentação de glicose a etanol. A produção de glicerol durante a fermentação também é considerada. Além disso, o modelo descreve oligossacarídeos de cadeia curta e incorpora a dependência da temperatura na cinética de reação e nos balanços de energia, de modo que a remoção de calor do fermentador também pode ser monitorada. A estratégia de monitoramento inclui o uso de dados de processo tanto *on-line* quanto *off-line* para ajustar os parâmetros específicos do modelo, fornecendo inferências estimadas pelo modelo em tempo real durante toda a fermentação.

4.2 Estratégia para Monitoramento de SSF

Embora alguns trabalhos tenham sido publicados sobre o desenvolvimento de modelos para SSF de milho e sua aplicação em controle ótimo em ambientes industriais (DAI *et al.*, 2014; MURTHY *et al.*, 2011), esses modelos não conseguem capturar o efeito das incertezas e, portanto, podem precisar de procedimentos para recalibração. Mesmo que um modelo válido esteja disponível para uso, vários fatores influenciam a qualidade das inferências. Esses fatores estão relacionados principalmente com as hipóteses adotadas para a dedução do modelo e distúrbios que afetam o processo SSF. Por exemplo, à medida que a fermentação prossegue, os produtos gasosos escapam da solução e alteram a concentração de cada substância, afetando, conseqüentemente, as propriedades do meio reacional, como a massa específica e a capacidade calorífica (DAI *et al.*, 2014). Os balanços de massa e energia incluem vários parâmetros físicos, cujos valores não são fáceis de determinar e geralmente são obtidos a partir da literatura. Além da dependência com a temperatura, eles também variam de acordo com a composição dos componentes dentro do fermentador. Essas dependências são muitas vezes desprezadas e as propriedades são consideradas como constantes. Além disso, os processos industriais estão sujeitos a mudanças

frequentes. Apesar dos esforços para manter os padrões de qualidade da matéria-prima e as condições do processo operacional, as propriedades da enzima, levedura e hidrolisado podem variar de maneira imprevisível.

A fim de capturar essas muitas fontes de incerteza, mantendo a qualidade das inferências do modelo, desenvolveu-se um método para identificar o sistema *on-line*, detectando mudanças em tempo real. Em relação a soluções para bioprocessos protegidos por direitos de propriedade intelectual, existem patentes dedicadas, por exemplo, a métodos de simulação e predição (BLUCK *et al.*, 2014), dispositivos, sistemas e métodos de sensoriamento para monitoramento (DEYOU, 2014; DISCENZO, 2014; NORKIN e COLOME, 2014), métodos de controle e controle ótimo (BARTEE *et al.*, 2013; MURTHY e SINGH, 2011), e sistemas de controle (SHANRUI, 2016). No entanto, não foi possível encontrar patentes relacionadas a sistemas inteligentes de monitoramento de processos de fermentação para identificar os estados do sistema, como o que se propõe nessa tese.

Para realizar isso, propõe-se uma estratégia de monitoramento baseada em modelo composta por dois elementos principais:

- i) Modelo de processo validado;
- ii) Procedimento de identificação *on-line*.

Nas seções seguintes, cada um desses elementos é discutido em detalhes.

4.2.1 Modelagem da SSF de Hidrolisado de Milho

A SSF de hidrolisado de milho consiste em dois processos acoplados. A sacarificação representa a conversão de oligossacarídeos (produzidos a partir de hidrólise de amido durante a etapa de liquefação) em glicose por meio de reações catalisadas por enzimas. Esta conversão é feita por enzimas do tipo glucoamilase que quebram as ligações $\alpha(1-4)$ das extremidades não redutoras de dextrinas e liberam moléculas de glicose. A etapa de fermentação envolve o uso de leveduras para converter a glicose em etanol por meio do metabolismo das células vivas. A realização da sacarificação e fermentação simultaneamente oferece a possibilidade de combinar a taxa de produção de glicose pela glucoamilase com a taxa de consumo da glicose pela levedura. Isso pode ser explorado para manter baixas concentrações de açúcar no meio reacional, o que contribui para minimizar as contaminações bacterianas e reações secundárias envolvendo açúcares redutores e aminoácidos livres, aumentando assim a eficiência da sacarificação. Além disso, ajuda a manter o equilíbrio osmótico adequado para a levedura.

Vários modelos cinéticos foram propostos para a produção de etanol a partir do milho pelo processo SSF (ALTINTAŞ *et al.*, 2002; DAI *et al.*, 2014; JANG e

CHOU, 2013; KROUMOV *et al.*, 2006; LEE *et al.*, 1992; MURTHY *et al.*, 2012a; OCHOA *et al.*, 2007). Com foco no uso industrial da estratégia de monitoramento, buscou-se desenvolver um modelo para a SSF do hidrolisado de milho que fosse o mais simples possível, mas ainda capaz de descrever a dinâmica das principais espécies químicas envolvidas. Considerando uma aplicação de monitoramento em tempo real, isso ajuda a reduzir os custos computacionais. Além disso, a descrição das taxas de síntese das principais enzimas pelo micro-organismo é uma parte central da modelagem cibernética (KOMPALA *et al.*, 1984; VARNER e RAMKRISHNA, 1999). A dedução de modelos cibernéticos para processos SSF (ALTINTAŞ *et al.*, 2002; MURTHY *et al.*, 2012a; OCHOA *et al.*, 2007) geralmente inclui a secreção das enzimas amilolíticas para o estudo de micro-organismos recombinantes. A indústria de etanol de milho geralmente emprega a levedura *Saccharomyces cerevisiae* na fermentação. Como essa levedura não produz as enzimas necessárias para a decomposição do amido (OSTERGAARD *et al.*, 2000), a modelagem da produção dessas enzimas não fez parte do escopo desse trabalho.

Entre os modelos não segregados disponíveis para descrever a SSF do hidrolisado de milho, alguns deles consideram explicitamente produtos de hidrólise intermediária, como maltodextrinas. O conhecimento dos produtos de hidrólise intermediária desempenha um papel importante na estimativa das taxas de produção e consumo de glicose, que são críticas para o monitoramento do processo SSF. Outra limitação está relacionada com a formação de glicerol como subproduto da fermentação de *S. cerevisiae*. O glicerol também é geralmente desprezado entre esses modelos. Devido a sua relação com o metabolismo e fisiologia das células de levedura, o glicerol consiste em uma espécie importante a ser monitorada durante o bioprocessamento.

Focando nesses fatos, foram feitas tentativas de propor um modelo confiável para o processo em uma única etapa, que possa ser usado para monitorar processos SSF industriais. A filosofia geral adotada na modelagem foi que o modelo deveria basear-se na bioquímica da hidrólise enzimática e do metabolismo das células de levedura, incluindo os efeitos térmicos, mas em que os parâmetros estivessem baseados em fenômenos observáveis na medida do possível. Além disso, a preferência geral foi minimizar o número de parâmetros do modelo, não incorporando todas as complexidades do sistema, de modo a obter um modelo simples e efetivo. O modelo desenvolvido considera um processo de fermentação em regime batelada alimentada em que ocorre a SSF de hidrolisado de milho em etanol. No entanto, o escopo do sistema de equações é genérico e pode ser aplicado a ambos os sistemas de operação batelada (alimentada) e contínuo.

4.2.2 Descrição do Modelo

O sistema é representado por um modelo semimecanístico e não estruturado, composto pelas seguintes variáveis de estado: concentrações de maltotetraose ($DP4$), maltotriose ($DP3$), maltose (M), glicose (G), etanol (Et), glicerol (Gy), água (W), células (X), e enzimas (E). As variáveis de estado foram escolhidas com base nos elementos desejados para monitoramento no processo SSF. Admitiu-se um total de 7 reações que ocorrem dentro do vaso (Tabela 4.1), das quais 3 reações são catalisadas por enzimas, representando a hidrólise, e as 4 restantes para a ação das células de levedura. O processo SSF é feito pela ação das enzimas do tipo glucoamilase, que convertem os oligossacarídeos em glicose, e de células de levedura, como *S. cerevisiae*.

Tabela 4.1: Reações cinéticas. Adaptado de PASCAL *et al.* (1995) e PASCAL *et al.* (1993).

Conversão de maltotetraose a maltose	$DP4 + H_2O \longrightarrow 2M$
Conversão de maltotriose a maltose e glicose	$DP3 + H_2O \longrightarrow M + G$
Conversão de maltose a glicose pela enzima	$M + H_2O \longrightarrow 2G$
Conversão de maltose a glicose pela levedura	$M + H_2O \longrightarrow 2G$
Conversão de glicose a etanol	$G \longrightarrow 2Et + 2CO_2$
Crescimento da levedura	$G + X \longrightarrow 5.80X + 0.20CO_2 + 2.29H_2O$
Produção de Glicerol	$G + \frac{6}{7}H_2O \longrightarrow \frac{12}{7}Gy + \frac{6}{7}CO_2$

4.2.2.1 Modelo de Sacarificação

A hidrólise enzimática relacionada ao processo de sacarificação considerado no modelo está ilustrada na Figura 4.2. O hidrolisado de milho é considerado como uma mistura de (oligo)sacarídeos constituídos por maltotetraose, maltotriose, maltose e glicose.

No modelo de sacarificação, a hidrólise enzimática simplificada do hidrolisado para produzir glicose consiste em três reações distintas, catalisadas pela enzima amiloglucosidase: r_{H1} , r_{H2} , e r_{H3} . A reação r_{H1} representa a conversão de maltotetraose em duas moléculas de maltose; a hidrólise de maltotriose em maltose e glicose é considerada em r_{H2} ; por fim, a reação r_{H3} descreve a conversão de maltose em glicose. Admite-se que essas reações catalisadas por enzimas seguem a cinética clássica de Michaelis-Menten. Este é um ponto relevante, pois mesmo sendo um modelo simples, essa cinética será usada como uma referência no procedimento de identificação para representar um sistema sabidamente complexo. A inibição do substrato e do produto não foram consideradas. Os efeitos da temperatura nas taxas enzimáticas não foram considerados neste modelo, uma vez que a faixa de temperaturas encontradas nos reatores SSF exerce pouco efeito sobre a atividade enzimática

(MURTHY *et al.*, 2012a). Os efeitos do pH na hidrólise enzimática também não foram considerados no modelo. Além disso, admite-se que a atividade enzimática (Ea) é proporcional à concentração de proteína no meio reacional (E), e que os efeitos térmicos e mecânicos na desativação enzimática são insignificantes. A atividade específica da glucoamilase (β , expressa por unidades de atividade por massa de proteína) é utilizada para relacionar a atividade enzimática com a concentração de proteína ($Ea = \beta C_E$), de modo que o modelo é independente de variações nas atividades enzimáticas de diferentes produções de glucoamilase. Assim, as equações cinéticas para a sacarificação do hidrolisado de milho são:

$$r_{H1} = \frac{k_{h1} Ea C_{DP4}}{(K_{H1} + C_{DP4})} = \frac{k_{h1} \beta C_E C_{DP4}}{(K_{H1} + C_{DP4})} = \frac{k_{H1} C_E C_{DP4}}{(K_{H1} + C_{DP4})} \quad (4.1)$$

$$r_{H2} = \frac{k_{H2} C_E C_{DP3}}{(K_{H2} + C_{DP3})} \quad (4.2)$$

$$r_{H3} = \frac{k_{H3} C_E C_M}{(K_{H3} + C_M)} \quad (4.3)$$

em que r_{Hi} é a taxa ($kg m^{-3} h^{-1}$) da reação i ($i = 1, 2, 3$); $k_{Hi} = \beta k_{hi}$ é o parâmetro agrupado da constante específica de reação da reação i ($i = 1, 2, 3$) (h^{-1}); K_{Hi} é a constante de saturação do substrato para a reação i ($i = 1, 2, 3$) ($kg m^{-3}$); C_E é a concentração de glucoamilase na solução ($kg m^{-3}$); e C_{DP4} , C_{DP3} , e C_M são concentrações ($kg m^{-3}$) de maltotetraose, maltotriose, e maltose, respectivamente.

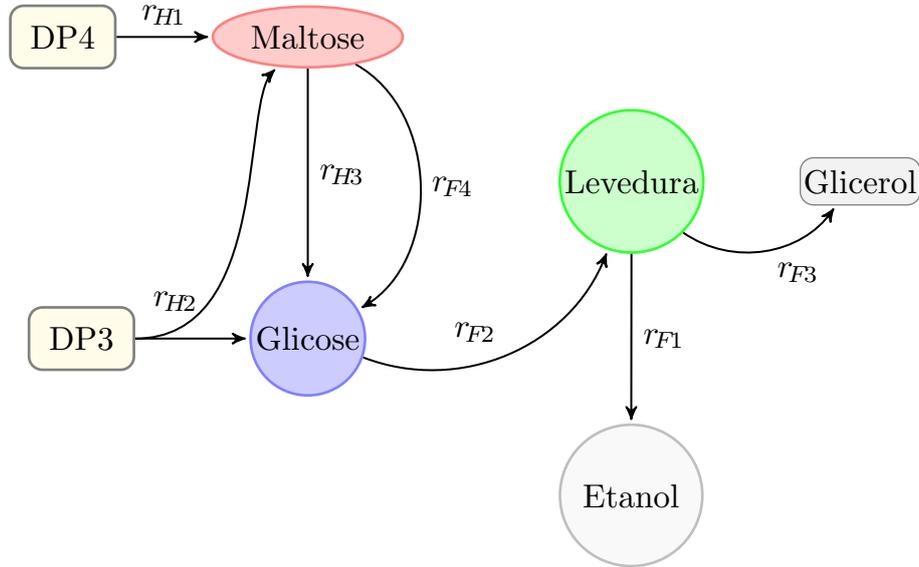


Figura 4.2: Mecanismo simplificado do processo SSF considerado.

4.2.2.2 Modelo de Fermentação

O mecanismo simplificado da fermentação consiste em quatro reações distintas, Equações (4.5)–(4.8), que são realizadas por células de levedura. O mecanismo cinético relacionado à fermentação também é ilustrado na Figura 4.2.

A produção de etanol é realizada pelo consumo de glicose pelas leveduras, r_{F1} , estando baseada na cinética clássica de Monod. Este também é um ponto relevante, pois, novamente, um modelo simples será usado como uma referência no procedimento de identificação para representar um sistema sabidamente complexo. Além disso, considera-se também sua dependência com a temperatura do reator. Esta dependência com a temperatura é expressa por meio da variável k_g , que está relacionada à atividade da levedura. A dependência da atividade da levedura foi proposta com base no conhecimento operacional empírico e na conveniência matemática. Apesar disso, é bem conhecido que as taxas de fermentação geralmente passam por um máximo com o aumento de temperatura, como tem sido observado experimentalmente na literatura por diferentes autores (CLAES e VAN IMPE, 2000; HOFFMANN *et al.*, 2000; VAN IMPE, 1996).

Considera-se que a atividade corresponde a 0,2 em 30 °C, aumentando linearmente até 36 °C, onde apresenta um máximo. Acima desta temperatura, a atividade diminui linearmente, atingindo atividade zero a 38 °C. Assim, admite-se que o intervalo de temperaturas compreendido entre 30 e 38 °C representa a operação típica encontrada em reatores SSF industriais (MURTHY *et al.*, 2011). Além disso, a queda acentuada da atividade após o ponto de máximo foi considerada como uma estratégia para a implementação futura de rotinas de otimização com base na ferramenta de monitoramento, a fim de evitar que o processo se aproximasse de pontos de operação em altas temperaturas. A dependência da atividade da levedura k_g é representada graficamente na Figura 4.3.

A produção de etanol também é dependente da fase *lag* das leveduras, representada pela variável k_{lag} . Com base na investigação dos dados industriais, a seguinte expressão matemática, adaptada da função de Bělehrádek (AHLGREN, 1987), foi proposta para descrever a fase *lag*:

$$k_{lag} = \begin{cases} (t/10)^3 & \text{se } 0 \leq t < 10 \text{ h,} \\ 1 & \text{se } t \geq 10 \text{ h.} \end{cases} \quad (4.4)$$

O crescimento da levedura no modelo (r_{F2}) segue uma cinética de primeira ordem, limitada apenas pela glicose. Além disso, é inibido pela concentração de etanol. Admite-se que as células de levedura continuam produzindo etanol e permanecem viáveis durante toda a fermentação, e que as concentrações de nutrientes e oxigênio dissolvido são sempre adequadas para a produção de etanol. A formação de glicerol

por *S. cerevisiae* durante a fermentação de glicose a etanol é conhecida como resposta à pressão osmótica do meio (MAIORELLA *et al.*, 1983; WANG *et al.*, 2001). Para explicar isso no modelo de fermentação, o efeito da pressão osmótica sobre a levedura foi considerado como associado à concentração total de açúcar. Assim, a reação r_{F3} representa a taxa de produção de glicerol e segue uma cinética de primeira ordem em função de todos os açúcares presentes no meio. Para considerar a regulação da pressão osmótica do meio pelo glicerol, admitiu-se que a descrição matemática desse efeito é equivalente à uma reação com inibição pelo produto.

Além do crescimento da levedura, produção de etanol e glicerol, também considerou-se a glicose como produto do metabolismo da levedura. Esta última reação explica a atividade de enzimas expressas por células de levedura para metabolizar a maltose (VOLPE, 1997). Admite-se que a conversão de maltose a glicose (r_{F4}) está baseada na cinética de Monod.

A reação r_{F4} também apresenta dependência com a temperatura, que é representada como uma função gaussiana centralizada em 34 °C (CLAES e VAN IMPE, 2000; HOFFMANN *et al.*, 2000; MURTHY *et al.*, 2011; VAN IMPE, 1996). Efeitos de pH nas reações de fermentação foram desprezados. Portanto, as taxas cinéticas para a fermentação são:

$$r_{F1} = k_{F1} \frac{C_G}{(K_{F1} + C_G)} k_g(T) k_{lag}(t) C_X \quad (4.5)$$

$$r_{F2} = k_{F2} \frac{C_G}{(K_{F2} + C_{Et})} C_X \quad (4.6)$$

$$r_{F3} = k_{F3} \frac{C_{DP4} + C_{DP3} + C_M + C_G}{(K_{F3} + C_{Gy})} C_X \quad (4.7)$$

$$r_{F4} = k_{F4} \frac{C_M}{(K_{F4} + C_M)} e^{\frac{-(T-34)^2}{0.5}} C_X \quad (4.8)$$

em que r_{Fi} é a taxa de reação ($kg\ m^{-3}\ h^{-1}$) para a reação i ($i = 1, 2, 3, 4$); k_{Fi} é o parâmetro agrupado da velocidade específica máxima para a reação i ($i = 1, 2, 3, 4$) (h^{-1}); K_{Fi} é a constante de saturação da reação i ($i = 1, 2, 3, 4$) ($kg\ m^{-3}$); C_X é a concentração de célula em solução ($kg\ m^{-3}$); T é a temperatura do reator (°C); e C_G , C_{Et} , and C_{Gy} são concentrações instantâneas ($kg\ m^{-3}$) de glicose, etanol, and glicerol, respectivamente.

4.2.2.3 Modelo do Processo

Uma representação do processo SSF presente na maioria das plantas de etanol de milho é mostrada na Figura 4.4. O esquema ilustra um reator com correntes de alimentação para hidrolisado de milho, células de levedura e enzima. A temperatura do reator é controlada pela manipulação da vazão de água em um trocador de calor,

em que uma corrente do meio reacional (também chamado de vinho ou *beer*) perde calor para o fluido refrigerante (água) e retorna ao reator. Com base neste sistema e nas expressões cinéticas para hidrólise enzimática e fermentação, o conjunto de equações diferenciais que descrevem os balanços de massa para o sistema e cada componente (maltotetraose, maltotriose, maltose, glicose, levedura, etanol, água, glicerol e enzima) corresponde a:

$$\frac{dm_{sys}}{dt} = F_{mash}\rho_{mash} + F_{yeast}\rho_{yeast} + F_{enz}\rho_{enz} - F_{CO_2}\rho_{CO_2} - F_{beer}\rho_{beer} \quad (4.9)$$

$$\frac{dDP4}{dt} = F_{mash}C_{DP4}^{mash} - r_{H1}V - F_{beer}DP4/V \quad (4.10)$$

$$\frac{dDP3}{dt} = F_{mash}C_{DP3}^{mash} - r_{H2}V - F_{beer}DP3/V \quad (4.11)$$

$$\frac{dM}{dt} = F_{mash}C_M^{mash} + (n_{H1,M}r_{H1} + n_{H2,M}r_{H2} - r_{H3} - r_{F4})V - F_{beer}M/V \quad (4.12)$$

$$\begin{aligned} \frac{dG}{dt} = F_{mash}C_G^{mash} + (n_{H2,G}r_{H2} + n_{H3,G}r_{H3} \\ - r_{F1} - r_{F2} - r_{F3} + n_{H3,G}r_{F4})V - F_{beer}G/V \end{aligned} \quad (4.13)$$

$$\frac{dX}{dt} = F_{yeast}C_X^{yeast} + n_{F2,X}r_{F2}V - F_{beer}X/V \quad (4.14)$$

$$\frac{dEt}{dt} = F_{mash}C_{Et}^{mash} + n_{F1,Et}r_{F1}V - F_{beer}Et/V \quad (4.15)$$

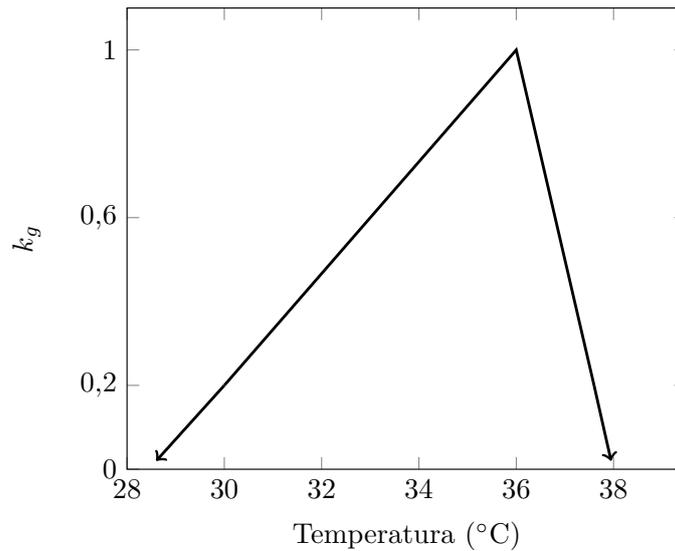


Figura 4.3: Atividade admitida para a produção de etanol em função da temperatura (k_g). Adaptado de CÂMARA *et al.* (2017a).

$$\begin{aligned} \frac{dW}{dt} = & F_{mash} C_W^{mash} + (-n_{H1,W}r_{H1} - n_{H2,W}r_{H2} - n_{H3,W}r_{H3} \\ & + n_{F2,W}r_{F2} - n_{F3,W}r_{F3} - n_{H3,W}r_{F4})V - F_{beer} W/V \end{aligned} \quad (4.16)$$

$$\frac{dGy}{dt} = F_{mash} C_{Gy}^{mash} + n_{F3,Gy}r_{F3}V - F_{beer} Gy/V \quad (4.17)$$

$$\frac{dE}{dt} = F_{enz} C_E^{enz} - F_{beer} E/V \quad (4.18)$$

em que m_{sys} é massa total do sistema (kg); F_{mash} , F_{yeast} , F_{enz} , F_{CO2} , and F_{beer} são as vazões volumétricas de hidrolisado de milho (*corn mash*), levedura (*yeast*), enzima, dióxido de carbono e vinho (*beer*), respectivamente ($m^3 h^{-1}$); ρ_{mash} , ρ_{yeast} , ρ_{enz} e ρ_{CO2} são as densidades das correntes de hidrolisado de milho, levedura, enzima e vapor, respectivamente ($kg m^{-3}$); C_i^{mash} , C_i^{yeast} e C_i^{enz} são as concentrações do componente i nas correntes de alimentação de hidrolisado, levedura e enzima, respectivamente ($i = DP4, DP3, M, G, X, Et, W, Gy, E$) ($kg m^{-3}$); $n_{j,i}$ é o fator mássico de conversão do componente i na reação j ($i = M, G, X, Et, W, Gy$) ($j = H1, H2, H3, F1, F2, F3$); e V é o volume de líquido no fermentador (m^3).

Os parâmetros para o vinho (mais especificamente, F_{beer}) são introduzidos no sistema de equações de tal forma que o modelo pode ser aplicado não apenas à operação batelada (alimentada), mas também para operação contínua. Nesse caso, parâmetros do vinho correspondem à corrente de saída do reator. As propriedades físico-químicas do sistema, como capacidade calorífica e densidade, são calculadas a cada instante com os valores atuais da massa de cada espécie química, considerando que a mistura se comporta como uma solução ideal. Não foram consideradas abordagens mais sofisticadas para determinar propriedades físicoquímicas, uma vez que tais propriedades não afetam os resultados significativamente na prática e eventuais desvios sempre podem ser compensados com ajustes em parâmetros estimados, tais como constantes cinéticas.

A corrente F_{CO2} representa o dióxido de carbono produzido pelas reações r_{F1} , r_{F2} , e r_{F3} . Assim,

$$F_{CO2} \rho_{CO2} = (n_{F1,CO2}r_{F1} + n_{F2,CO2}r_{F2} + n_{F3,CO2}r_{F3})V \quad (4.19)$$

em que $n_{j,CO2}$ é o fator mássico de conversão do dióxido de carbono na reação j ($j = F1, F2, F3$).

Para dedução do balanço de energia, considera-se que o sistema é composto pelo meio reacional dentro do fermentador, que as capacidades caloríficas Cp e entalpias de reação ΔH são constantes e que a perda de calor para o ambiente pode ser desprezada. Dado que a fermentação é um processo exotérmico, o controle de temperatura é obtido por meio de um trocador de calor, que opera continuamente

removendo a energia Q_{HE} da corrente F de vinho que contém etanol, que retorna ao fermentador com uma temperatura T^{out} . Além disso, o balanço de energia inclui a energia das correntes de entrada e saída $H_{streams}$ de hidrolisado, levedura, enzima e vinho. Assim, o balanço de energia corresponde a:

$$\frac{dT}{dt} = \frac{H_{streams} + Hr - Q_{HE}}{m_{sys} Cp} \quad (4.20)$$

$$\begin{aligned} H_{streams} = & (F_{mash}\rho_{mash}Cp_{mash} \\ & + F_{enz}\rho_{enz}Cp_{enz} \\ & + F_{yeast}\rho_{yeast}Cp_W)(T_{feed} - T) \end{aligned} \quad (4.21)$$

$$\frac{-Hr}{V} = \sum_{i=1,2,3} \Delta H_{Hi}r_{Hi} + \sum_{i=1,2,3,4} \Delta H_{Fi}r_{Fi} \quad (4.22)$$

$$Q_{HE} = F \rho Cp (T - T^{out}) \quad (4.23)$$

$$T_{HE}^{out} = T_{HE}^{in} + \frac{Q_{HE}}{F_{HE} \rho_{HE} Cp_{HE}} \quad (4.24)$$

em que $H_{streams}$ é a energia interna total das correntes ($J h^{-1}$); Hr é a energia resultante das reações químicas ($J h^{-1}$); Q_{HE} é a energia removida pelo trocador de calor ($J h^{-1}$); F , ρ e Cp são vazão, massa específica, e capacidade calorífica do vinho, respectivamente; e T_{HE}^{in} , T_{HE}^{out} , F_{HE} , ρ_{HE} e Cp_{HE} são as temperaturas de entrada e saída, vazão, massa específica e capacidade calorífica do fluido de troca térmica, respectivamente.

A ausência de medidas de temperatura para todas as correntes envolvidas no tro-

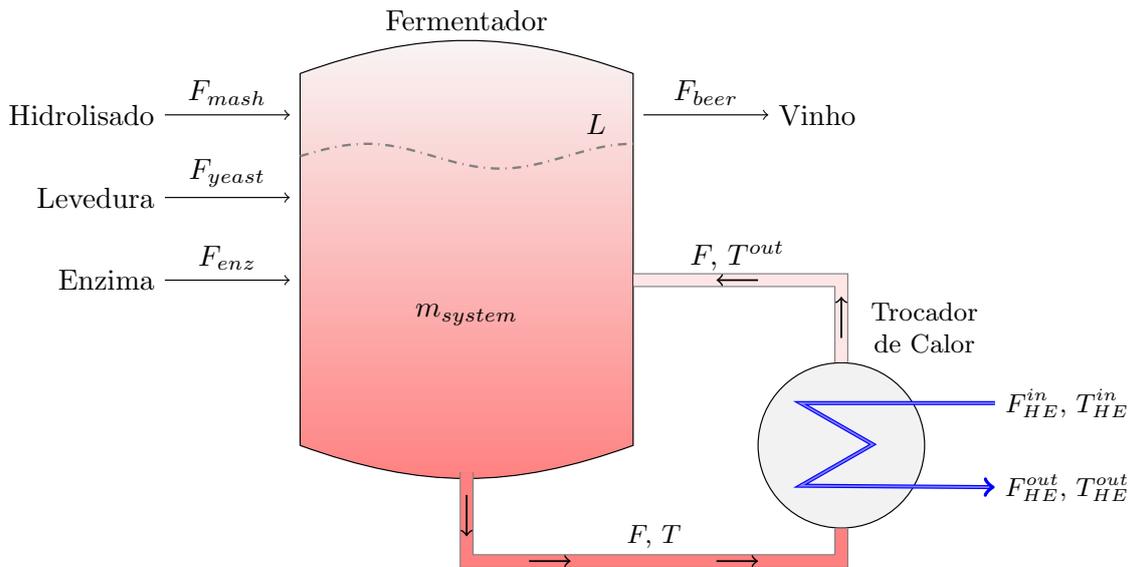


Figura 4.4: Representação esquemática do processo SSF industrial de hidrolisado de milho.

calor não permite a inferência de efeitos dinâmicos associados à incrustação por meio da equação típica $Q_{HE} = U A \Delta_{LMDT}$, que poderia fornecer informações valiosas sobre a condição do equipamento. Essa é uma condição usual na maior parte dos processos industriais de fermentação para produção de etanol.

4.2.3 Estimação de Parâmetros e Implementação do Modelo

Uma vez que um modelo matemático esteja disponível em conjunto com os dados experimentais, parâmetros do modelo podem ser determinados por meio de técnicas de estimação de parâmetros. De acordo com as equações descritas anteriormente, os seguintes vetores são definidos para variáveis de estado (\mathbf{x}), entrada (\mathbf{u}) e saída (\mathbf{y}), e parâmetros ($\boldsymbol{\theta}$). Considera-se que o vetor de parâmetros ($\boldsymbol{\theta}$) possui 14 elementos a serem estimados a partir de medidas experimentais.

$$\begin{aligned}\mathbf{x} &= [m_{sys}, DP4, DP3, M, G, X, Et, W, Gy, E, T]^T \\ \mathbf{u} &= [F_{mash}, F_{yeast}, F_{enz}, T_{feed}, F, T^{out}, F_{HE}, T_{HE}^{in}]^T \\ \mathbf{y} &= [DP4, DP3, M, G, X, Et, W, Gy, E, T, L]^T \\ \boldsymbol{\theta} &= [k_{H1}, k_{H2}, k_{H3}, k_{F1}, k_{F2}, k_{F3}, k_{F4}, \\ &\quad K_{H1}, K_{H2}, K_{H3}, K_{F1}, K_{F2}, K_{F3}, K_{F4}]^T\end{aligned}$$

O problema de estimação de parâmetros é resolvido usando o método de mínimos quadrados ponderados (BARD, 1974), o que resulta no seguinte problema de otimização, como definido no Capítulo 2:

$$\begin{aligned}\min_{\boldsymbol{\theta}} \quad & \sum_{k=0}^{t_{drop}/ta+1} \mathbf{w}_k [\mathbf{y}(t_k) - \mathbf{y}^{ms}(t_k)]^T \mathbf{V}^{-1} [\mathbf{y}(t_k) - \mathbf{y}^{ms}(t_k)] \\ \text{sujeito a} \quad & \dot{\mathbf{x}}(t) = f[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}], \quad \forall t \in [0, t_{drop}] \\ & \mathbf{x}(0) = \mathbf{x}_0 \\ & \mathbf{y} = h(\mathbf{x})\end{aligned}, \quad (4.25)$$

em que t_k e t_{drop} são os tempos no instante de amostragem k e ao fim da fermentação, respectivamente (h); ta é o intervalo de amostragem (h); $\mathbf{y}(t_k)$, $\mathbf{x}(t_k)$ e $\mathbf{u}(t_k)$ são os vetores das variáveis de saída, estado e entrada no tempo t_k , respectivamente; $\mathbf{y}^{ms}(t_k)$ é o vetor de variáveis medidas; $\boldsymbol{\theta}$ é o vetor de parâmetros; \mathbf{V} é a matriz de ponderação entre as variáveis medidas; \mathbf{w}_k é o fator de ponderação para as variáveis medidas no instante de tempo k ; f é o vetor de equações do modelo do processo; e h é o vetor de funções de saída do sistema.

Se os erros de medição forem independentes e normalmente distribuídos com média zero e variância conhecida, o problema de mínimos quadrados ponderados corresponde ao estimador de máxima verossimilhança, em que \mathbf{V} é a matriz de

covariâncias dos erros de medição (BARD, 1974). Admite-se que o vetor de variáveis de entrada \mathbf{u} é perfeitamente conhecido. Geralmente, dados industriais possuem mais abundância de medidas de temperatura e nível de líquido, pois são adquiridas em linha. Por outro lado, medições que dependem da análise laboratorial, como a determinação de concentrações por HPLC, estão disponíveis em frequências mais baixas. Assim, o vetor de ponderação \mathbf{w}_k é introduzido para lidar com a natureza assíncrona das medidas de diferentes variáveis ao longo do tempo de operação do processo. Esse vetor é formado por componentes binários que podem receber valores 1 ou 0 se uma medida está disponível ou não, respectivamente.

Além disso, nos casos em que o conhecimento das condições iniciais das variáveis de estado é incompleto ou ausente, o vetor das variáveis de decisão $\boldsymbol{\theta}$ pode ser aumentado com essas variáveis desconhecidas, na forma:

$$\boldsymbol{\theta}_{aug} = [\boldsymbol{\theta}^T, \mathbf{x}_0^T]^T .$$

O problema de otimização indicado na Equação (4.25) foi resolvido empregando a abordagem do caminho viável (*feasible path approach*) (TJOA e BIEGLER, 1991). Como o modelo de processo constitui uma restrição do problema de otimização, essa abordagem é preferível em relação a outras metodologias, pois os parâmetros estimados sempre satisfazem as restrições do modelo. Em outras palavras, a solução para o problema de otimização sempre satisfaz os balanços de massa e energia. Neste método, também chamado de abordagem sequencial (KIM *et al.*, 1991), a integração numérica das equações do modelo precede a avaliação da função objetivo (e seus gradientes) em cada iteração da técnica de otimização.

A técnica de otimização usada para resolver a Equação (4.25) foi a de otimização por enxame de partículas (PSO), originalmente proposta por EBERHART e KENNEDY (1995) e adaptada por SCHWAAB *et al.* (2008) e PRATA *et al.* (2009b) para a estimação de parâmetros e reconciliação de dados, respectivamente. PSO é um método estocástico que não utiliza cálculo de gradientes, em que uma série de pontos (partículas) são colocados no espaço de busca das variáveis de decisão e a função objetivo é avaliada nesses locais. Para um certo número de iterações (gerações), a localização de cada uma das partículas (estimativa da solução) é atualizada de acordo com uma regra que combina a história de suas atuais e melhores localizações, com a melhor localização global entre todas as partículas e algumas perturbações aleatórias (POLI *et al.*, 2007). Desse modo, a posição das partículas varia de acordo com (PRATA *et al.*, 2009b):

$$v_{p,d}^{k+1} = \omega v_{p,d}^k + c_1 r_1 (x_{p,d}^{ind} - x_{p,d}^k) + c_2 r_2 (x_d^{glo} - x_{p,d}^k) \quad (4.26)$$

$$x_{p,d}^{k+1} = x_{p,d}^k + v_{p,d}^{k+1} \quad , \quad (4.27)$$

em que k é a iteração (geração); p é o número da partícula e d a direção de busca (índice que identifica a variável de decisão); ω , c_1 e c_2 são os parâmetros de controle da velocidade de convergência e da natureza estocástica da busca (SCHWAAB *et al.*, 2008); r_1 e r_2 são valores aleatórios originados da distribuição uniforme na faixa $[0, 1]$; v e x são a velocidade e a posição das partículas, respectivamente; $x_{p,d}^{ind}$ é a melhor solução encontrada pela partícula p e x_d^{glo} é a melhor solução encontrada entre as partículas. As iterações são repetidas até atingir-se uma tolerância numérica pré-especificada para a diferença entre os valores da função objetivo nas duas últimas iterações sucessivas ou até atingir o número de iterações especificado.

Assim, a função objetivo é minimizada por meio da varredura do espaço de busca (definido pelo usuário). Devido à sua capacidade de lidar com funções objetivo multimodais, não-lineares e descontínuas em problemas de alta dimensão, com esforço computacional relativamente pequeno e implementação mais simples, o método PSO constitui uma ferramenta robusta e eficaz para resolver problemas complexos de otimização, já tendo sido utilizado com sucesso em outras aplicações industriais (PRATA *et al.*, 2009b, 2010).

O modelo e o procedimento de estimação de parâmetros foram ambos implementados em MATLAB (MATLAB, 2010). As equações do modelo foram integradas com uma técnica de ordem variável com base nas fórmulas de diferenciação numérica `ode15s`. Os parâmetros do algoritmo PSO estão listados na Tabela 4.2 e foram baseados nos valores usados por PRATA *et al.* (2009b). Os cálculos foram realizados em computador com sistema operacional Windows 64 bits, CPU Intel Core i7 2,0 GHz e 6 GB RAM.

Tabela 4.2: Valores dos parâmetros do algoritmo PSO.

Parâmetro	Valor
Número de gerações	15
Número de partículas	100
Parâmetro de busca local (c_1)	1
Parâmetro de busca global (c_2)	1
Peso inercial (ω)	0,7
Tolerância	$1,0 \cdot 10^{-3}$

4.2.4 Procedimento *On-line* de Identificação

Uma vez que o modelo descrito anteriormente esteja validado com dados de processo, ele pode ser usado para realizar inferências e aplicações de controle e otimização do processo. No entanto, o modelo permanece suscetível à degradação devido às muitas incertezas que afetam o processo.

Para contornar esse problema, o modelo foi modificado para uso na estratégia de monitoramento *on-line*. As Equações. (4.5), (4.6) e (4.8) foram modificadas pela introdução de parâmetros de eficiência α_{F1} , α_{F2} , e α_{F4} multiplicando as taxas de reação, como segue:

$$r_{F1}^{id} = \alpha_{F1} r_{F1} \quad (4.28)$$

$$r_{F2}^{id} = \alpha_{F2} r_{F2} \quad (4.29)$$

$$r_{F4}^{id} = \alpha_{F4} r_{F4} \quad (4.30)$$

em que $(\cdot)^{id}$ corresponde às taxas de reação identificadas. Substituindo as taxas de reação correspondentes no vetor de equações validadas, f , pelas Eqs. (4.28)–(4.30), tem-se um novo vetor de equações, f^{id} , que é uma função de $\boldsymbol{\alpha} = [\alpha_{F1}, \alpha_{F2}, \alpha_{F4}]^T$.

Um ponto relevante desse trabalho é o uso de equações baseadas nas cinéticas de Michaelis-Menten e Monod no modelo, pois cinéticas simples são usadas para representar um sistema complexo. Isso é possível a partir de uma estratégia de adaptação do modelo, que realiza a atualização de parâmetros específicos para promover uma adaptação local do modelo ao processo.

Para atualizar o modelo *on-line* durante o processo SSF, formulou-se um problema de estimação de parâmetros, modificado para ser resolvido em tempo real, em que os parâmetros de eficiência têm como objetivo atualizar localmente o modelo identificado f^{id} . Este problema é implementado seguindo a abordagem de janela (ou horizonte) móvel (KIM *et al.*, 1991; LIEBMAN *et al.*, 1992; ROBERTSON *et al.*, 1996), na qual uma janela de tempo se estende H intervalos de amostragem para trás a partir do tempo atual (ou mais recente) t_c , definindo uma janela de dados de $(t_c - H ta)$ a t_c , em que ta é o intervalo de tempo entre as aquisições de dados medidos. Assim, todas as medidas correspondentes ao tempo representado pela janela de dados $(t_c - H ta \leq t \leq t_c)$ estão disponíveis para resolver o problema de estimação de parâmetros durante o procedimento *on-line*. Em outras palavras, durante o procedimento de identificação, os parâmetros de eficiência são estimados em cada janela de dados. Esse procedimento é diferente da estimação de parâmetros *off-line*, representada na Equação (4.25), para a validação do modelo, em que os parâmetros do modelo $\boldsymbol{\theta}_{aug}$ são estimados com base em dados históricos.

Aplicando a estratégia da janela móvel finita e a função objetivo de mínimos

quadrados ponderados, a Eq. (4.25) pode ser rescrita como:

$$\begin{aligned}
\min_{\boldsymbol{\alpha}} \quad & J_{\text{mon}} := \sum_{k=c-H}^c \mathbf{w}_k [\mathbf{y}(t_k) - \mathbf{y}^{ms}(t_k)]^T \mathbf{V}^{-1} [\mathbf{y}(t_k) - \mathbf{y}^{ms}(t_k)] \\
\text{sujeito a} \quad & t_c \geq t_H \\
& \dot{\mathbf{x}}(t) = f^{id}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\alpha}], \quad \forall t \in [t_{c-H}, t_c] \\
& \mathbf{x}(t_{c-H}) = \mathbf{x}_{c-H} \\
& \mathbf{y}(t_k) = h[\mathbf{x}(t_k)] \\
& (1 - \alpha^{LU})\boldsymbol{\alpha}_0 \leq \boldsymbol{\alpha} \leq (1 + \alpha^{LU})\boldsymbol{\alpha}_0
\end{aligned} \tag{4.31}$$

em que H é o tamanho da janela em número de instantes de tempo, que corresponde a uma janela de tempo com tamanho $H \cdot ta(h)$; $k = c - H$ e $k = c$ são o primeiro e o mais recente instantes de amostragem da janela móvel, respectivamente; t_{c-H} e t_c são os tempos de amostragem no início e no final da janela móvel, respectivamente (h); $x(t_{c-H})$ é a condição inicial para as variáveis de estado na janela móvel; x_{c-H} é o vetor de variáveis de estado no início da janela móvel; $\boldsymbol{\alpha}$ é o vetor de parâmetros de eficiência a ser estimado e α^{LU} é a variação admissível para os parâmetros na janela atual em relação aos valores $\boldsymbol{\alpha}_0$ determinados na janela anterior; f^{id} é o vetor de equações do modelo identificado.

A janela de tempo é atualizada a cada instante de amostragem, à medida que novos dados chegam por meio de sensores e análises de laboratório. Assim, de acordo com a Equação (4.31), o problema não-linear de estimação de parâmetros é resolvido sequencialmente a cada novo tempo de amostragem, re-estimando os parâmetros de eficiência e simulando inferências atualizadas para as variáveis de estado e saída.

A janela móvel reduz a dimensão do problema de otimização e é utilizada para capturar o comportamento mais recente do processo, sendo muito útil quando os parâmetros do processo mudam ao longo do tempo de operação (PRATA *et al.*, 2009b). A definição adequada do tamanho da janela é um parâmetro específico do sistema físico, *i.e.*, o comprimento apropriado da janela está relacionado à dinâmica do sistema sob análise (LIEBMAN *et al.*, 1992). Se H é muito pequeno, pode não haver informação suficiente disponível sobre a dinâmica para a estimação. Se H é muito grande, o problema resultante será muito grande, resultando em elevado custo computacional. Neste contexto, o tamanho apropriado da janela móvel também está fortemente relacionado ao tempo de amostragem em que as medidas estão disponíveis.

A estratégia de monitoramento foi implementada em MATLAB (2010), onde a minimização de J_{mon} também foi feita utilizando a abordagem de caminho viável por meio da técnica PSO, utilizando os parâmetros apresentados na Tabela 4.2, como descrito anteriormente.

4.2.5 Dados Industriais Reais

A avaliação da estratégia de monitoramento baseada em modelo para o processo SSF foi realizada com dados industriais de uma planta de etanol de milho de moagem seca nos EUA, que tem como produto principal o etanol biocombustível. O processo SSF é realizado em fermentadores de 300.000 gal, em que a levedura e a glucoamilase são adicionadas em linha, à medida que o hidrolisado de milho é alimentado ao vaso durante os estágios iniciais da batelada. Devido a esta adição *inline* de leveduras e enzimas, o processo SSF inicia durante o tempo de enchimento. A levedura é obtida em um reator batelada separado, específico para a propagação do micro-organismo. Após a conclusão, o hidrolisado fermentado é bombeado para um tanque de vinho. Foram fornecidos dados de 30 bateladas sequenciais realizadas durante o período de um mês. O conjunto de dados pode ser agrupado em três tipos principais: análises laboratoriais para amostras retiradas durante a batelada, medições em linha disponibilizadas pelos sensores e condições de operação estabelecidas pelos operadores do processo. As informações disponíveis no laboratório são concentrações de oligossacarídeos maiores que quatro unidades de glicose, maltotriose, maltose, glicose, etanol e glicerol em sete tempos de amostragem diferentes ao longo do processo SSF, incluindo o final da batelada. A condição inicial da fermentação não está disponível no conjunto de dados. As medições em linha são muito mais abundantes, sendo registradas a cada 5 minutos para a temperatura do reator, o nível do líquido no interior do vaso e a temperatura do fluido refrigerante na saída do trocador de calor. As condições de operação envolvem o conhecimento de algumas variáveis manipuladas, como perfis de alimentação de hidrolisado de milho, inóculo e enzimas, e vazão de fluido refrigerante. Os perfis de alimentação de enzima foram fornecidos como funções do tempo, admitindo que também são conhecidos F_{enz} (vazão de alimentação de enzima) e E_{enz} (concentração de enzima na corrente de alimentação de enzima), conforme especificado pela Equação (4.18). As constantes cinéticas agrupadas (parâmetros k_{H1} , k_{H2} e k_{H3} nas Equações (4.1)-(4.3)) foram estimados com as respostas disponíveis do processo, conforme discutido anteriormente. Desse modo, do ponto de vista da simulação do modelo, as atividades enzimáticas já estão consideradas nas constantes cinéticas agrupadas e são capturadas pelo procedimento de estimação de parâmetros.

Além dos dados de processo já descritos, a informação relativa aos parâmetros físicos, como dimensões dos vasos, também estava disponível. As propriedades físico-químicas exigidas pelo modelo, como capacidades caloríficas e entalpias de formação, foram obtidas da literatura (FLICKINGER, 2013). Todos os dados foram coletados, organizados e utilizados para estimação de parâmetros, validação do modelo de processo e avaliação do sensor virtual proposto.

4.3 Resultados

Esta seção apresenta os resultados obtidos nas etapas de estimação de parâmetros, validação do modelo e monitoramento *on-line* do processo. Os resultados para essas três atividades foram obtidos usando dados industriais de 30 bateladas sequenciais de fermentação de uma planta de etanol de milho. Os dados cobrem o período de um mês de operação do processo. Informações não presentes no conjunto de dados da planta industrial foram obtidas na literatura.

4.3.1 Estimação de Parâmetros e Validação do Modelo

Conforme descrito anteriormente, os valores dos parâmetros foram determinados pela resolução da Equação (4.25) usando o algoritmo PSO, empregando a abordagem do caminho viável. Deve notar-se que as medições de cromatografia líquida de alta performance (HPLC) para os oligossacarídeos maiores que 4 unidades de glicose são consideradas como maltotetraose no modelo. Conforme discutido anteriormente, a preferência geral foi obter um modelo simples e efetivo. Como uma primeira tentativa de modelagem, o modelo foi simplificado para descrever dados de oligossacarídeos maiores que quatro unidades de glicose como maltotetraose. Uma vez que resultados satisfatórios foram obtidos, mostrando que esta simplificação de modelagem exerceu pouco efeito sobre os resultados, decidiu-se não introduzir no modelo a complexidade de modelar oligossacarídeos com mais de quatro unidades de glicose. O modelo de processo é composto por um sistema de equações diferenciais ordinárias e requer um conjunto de condições iniciais para serem integradas. As condições iniciais para algumas variáveis de estado não estavam disponíveis a partir dos dados da planta, sendo por isso estimadas com o aumento do vetor de variáveis de decisão.

As condições iniciais e os valores dos parâmetros apresentadas nas Tabelas 4.3 e 4.4 foram determinados utilizando dados de processo da Batelada 01. Os perfis simulados correspondentes são mostrados na Figura 4.5. Pode-se ver que o modelo desenvolvido é capaz de capturar os fenômenos mais importantes que ocorrem durante a processo SSF, apesar da sua relativa simplicidade. Os valores simulados para maltotetraose, maltotriose, temperatura e nível do fermentador aproximam-se dos dados experimentais com elevada acurácia. Enquanto os perfis de concentração para glicose e etanol seguem a tendência dos valores experimentais com concordância razoável, a evolução do glicerol no meio foi mais lenta que aquela predita pelo modelo. A trajetória da maltose, por sua vez, foi mais rápida que as previsões do modelo. O vetor aumentado de parâmetros estimados também foi determinado para outras bateladas. Para todas essas outras bateladas (dados não apresentados), os perfis simulados apresentaram boa concordância com os dados da planta, semelhantes aos

apresentados na Figura 4.5.

Para avaliar o modelo, os parâmetros estimados a partir de bateladas diferentes foram usados para construir um conjunto de valores médios e então simulações do modelo foram realizadas para cada uma dessas bateladas usando o conjunto médio. Uma vez que todo o conjunto de dados envolve bateladas espalhadas ao longo de um mês de operação, bateladas distintas podem ter sofrido diferentes distúrbios. Várias condições no início e durante a fermentação impactam o processo de fermentação e podem variar entre as bateladas.

De acordo com os resultados (dados não apresentados), as simulações do modelo para este conjunto médio de valores dos parâmetros também apresentaram boa concordância com os dados. Os perfis simulados para maltotetraose, maltotriose, temperatura e etanol foram muito pouco afetados nas bateladas testadas. Os perfis de glicerol, em especial, mostraram melhor concordância quando o conjunto médio foi usado para algumas bateladas. Os resultados das concentrações de maltose apresentaram uma ligeira melhoria em algumas bateladas. Em termos gerais, o uso do conjunto médio de parâmetros apresentou resultados ligeiramente melhores.

Tabela 4.3: Condição inicial para a Batelada 01.

Parâmetro	Valor	Unidade
X_0	0	kg/m ³
Et_0	0	kg/m ³
W_0	0	kg/m ³
Gy_0	0	kg/m ³
E_0	0	kg/m ³
T_0	35	°C

Tabela 4.4: Valores estimados utilizando os dados de processo da Batelada 01.

Parâmetro	Valor	Unidade	Parâmetro	Valor	Unidade
k_{H1}	816,70	h ⁻¹	K_{H1}	2040,73	kg/m ³
k_{H2}	1778,96	h ⁻¹	K_{H2}	2220,51	kg/m ³
k_{H3}	1590,68	h ⁻¹	K_{H3}	2939,60	kg/m ³
k_{F1}	1334,75	h ⁻¹	K_{F1}	4913,76	kg/m ³
k_{F2}	0,0459	h ⁻¹	K_{F2}	8,10	kg/m ³
k_{F3}	101,35	h ⁻¹	K_{F3}	114,13	kg/m ³
k_{F4}	0,4251	h ⁻¹	K_{F4}	245,21	kg/m ³
$DP4_0$	180,52	kg/m ³	$DP3_0$	40,11	kg/m ³
M_0	24,48	kg/m ³	G_0	31,20	kg/m ³

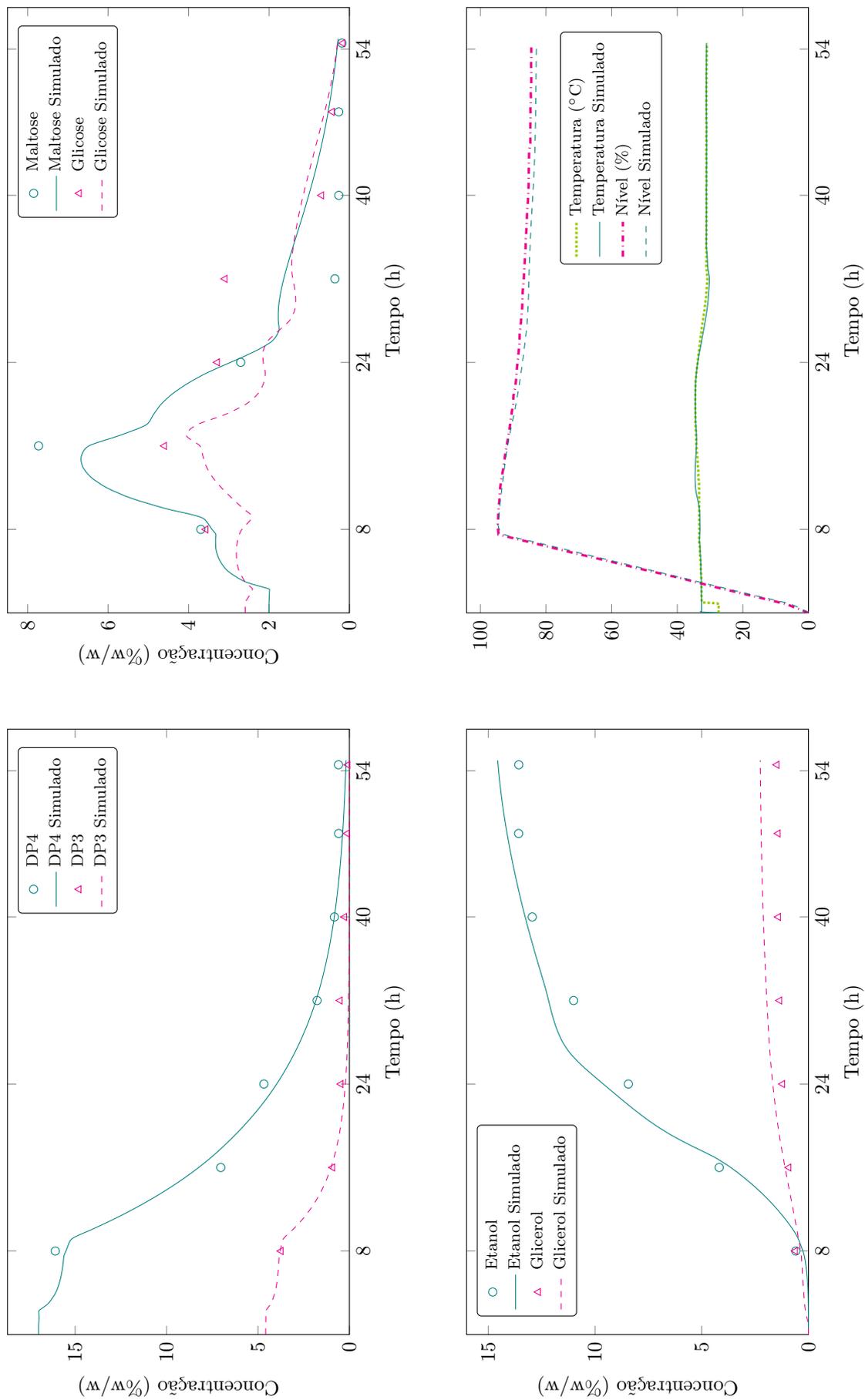


Figura 4.5: Dados experimentais e simulados obtidos com os parâmetros estimados com dados da Batelada 01.

A validação do modelo foi feita comparando-se as inferências de temperatura, nível e concentrações com os dados da planta de cinco bateladas não utilizadas na estimação dos parâmetros. Curiosamente, os resultados foram muito semelhantes aos já apresentados, em que as simulações se encaixam bem aos dados e mostram, assim, que o modelo é capaz de descrever o comportamento do processo. A validação foi feita tanto para parâmetros da Batelada 01 como para do conjunto médio.

A discrepância dos valores calculados pelo modelo em termos quantitativos pode ser parcialmente relacionada à incerteza em relação às propriedades físico-químicas e informações no início da fermentação. Além disso, a simplificação das premissas feitas para a dedução do modelo também pode influenciar esses desvios. Para mencionar algumas dessas simplificações, os efeitos da temperatura e do pH na atividade enzimática não foram considerados. Perturbações no pH podem ocorrer devido a vários fatores, como a infecção bacteriana. Além disso, uma interação complexa entre temperatura, pH e ácidos orgânicos fracos contribuem para a perda de viabilidade celular e não foram contabilizadas no modelo. A produção de ácidos orgânicos pela levedura e seus efeitos sobre a dinâmica de crescimento de células foram desprezadas. No entanto, como já foi visto, o modelo não estruturado é capaz de calcular variáveis-chave, sendo capaz de seguir as tendências gerais dos dados da planta. Com foco no monitoramento do processo, tais discrepâncias não comprometem o uso do modelo para esse objetivo. Como será apresentado na próxima seção, as inferências do modelo podem ser melhoradas pela aplicação de um procedimento de identificação em tempo real, apesar do grau de complexidade do modelo.

4.3.2 Monitoramento do Processo

A etapa de validação do modelo é importante para avaliar a capacidade do modelo capturar o comportamento geral do processo e as principais fontes de variabilidade. Além da imperfeição intrínseca do modelo matemático, outros efeitos contribuem para os desvios observados, como medições corrompidas por ruído e distúrbios do processo, especialmente em ambientes industriais. Assim, não se poderia esperar obter estimativas perfeitas. Neste contexto de incertezas inevitáveis que afetam os resultados do modelo, é possível melhorar ainda mais as inferências do modelo, aplicando um procedimento de identificação *on-line*. Propõe-se um procedimento de identificação baseado na abordagem de janela móvel, que envolve a resolução da Equação (4.31) em cada tempo de amostragem. Quanto à estimação dos parâmetros, o problema de otimização resultante é resolvido usando o algoritmo PSO com os parâmetros da Tabela 4.2. O modelo utilizado no procedimento de identificação foi uma versão modificada do modelo validado. A modificação consiste na introdução dos parâmetros de eficiência nas expressões de taxa de algumas reações, conforme

discutido na Seção 4.2.4.

4.3.2.1 Avaliação dos Parâmetros de Ajuste

É possível listar ao menos três variáveis de projeto para a estratégia de identificação *on-line*: i) tamanho da janela de dados H , ii) parâmetros de eficiência estimados, *i.e.*, quais taxas de reação são identificadas, e iii) tamanho do espaço de busca explorado pelo algoritmo PSO para a estimação dos parâmetros de eficiência. Considerando a Equação (4.31), o tamanho do espaço de busca é definido pelo parâmetro α^{LU} , que estabelece a variação admissível para os parâmetros durante o procedimento de estimação. Adicionalmente, o intervalo de amostragem, ta , também poderia ser incluído como um parâmetro de sintonia, mas não foi considerado no presente estudo. Uma vez que não existe uma solução geral ou padrão que embase essas decisões de projeto, elas devem ser estudadas e avaliadas em cada aplicação. Além disso, elas podem afetar a convergência do esquema proposto.

Para avaliar o desempenho da identificação *on-line*, foram selecionados, com base na análise dos dados e na experiência adquirida durante a etapa de validação do modelo, três níveis para cada variável de projeto estudada na estratégia de monitoramento, conforme apresentado na Tabela 4.5. Os níveis das variáveis foram combinados entre si, resultando em 26 cenários, e experimentos de simulação foram conduzidos utilizando os dados e parâmetros estimados (θ_{aug}) para a Batelada 01 (Tabela 4.4).

Os efeitos dessas três variáveis de projeto sobre a qualidade do ajuste dos perfis de monitoramento foram avaliados e comparados em termos das seguintes métricas estatísticas: soma do erro quadrado (SSE), erro quadrático médio (MSE), raiz do erro quadrático médio (RMSE), erro absoluto médio (MAE) e coeficiente de determinação (R^2) (KVÅLSETH, 1985), conforme definido:

$$SSE = \sum_{j=1}^n (y_j^{ms} - y_j)^2 \quad (4.32)$$

$$MSE = \frac{SSE}{n - p} \quad (4.33)$$

$$RMSE = \sqrt{\frac{SSE}{n}} \quad (4.34)$$

$$MAE = \frac{\sum_{j=1}^n |y_j^{ms} - y_j|}{n} \quad (4.35)$$

$$R^2 = 1 - \frac{\sum_i \sum_{k=1}^m (y_{i,k}^{ms} - y_{i,k})^2}{\sum_i \sum_{k=0}^{k=m} (y_{i,k}^{ms} - \bar{y}_i)^2} \quad (4.36)$$

em que $y_{i,k}$, $y_{i,k}^{ms}$ e \bar{y}_i são valores preditos, medidos e médios dos dados medidos, respectivamente, para variáveis de saída i ($i = DP4, DP3, M, G, Et, Gy, T, L$); R^2 é o coeficiente de determinação do conjunto total de predições. Como o tamanho do conjunto de dados de laboratório é diferente do tamanho do conjunto de dados medidos em linha, os dados usados para o cálculo das métricas estatísticas corresponderam aos tempos em que ambas as medições de laboratório e processo estavam disponíveis, resultando em $m = 7$ valores experimentais para cada variável. Em um experimento de monitoramento simulado, isso correspondeu a um número total de $n = 56$ valores medidos em cada experimento de monitoramento.

Dessa maneira, propôs-se um procedimento de sintonia do problema de identificação, que envolveu: i) selecionar os parâmetros de projeto e níveis de variação, ii) selecionar um conjunto de dados históricos de referência, iii) conduzir experimentos para os cenários possíveis, construídos a partir dos níveis selecionados, e iv) avaliar as realizações obtidas, identificando o cenário com o melhor desempenho. Até onde se pode determinar, essa abordagem de sintonia não é usual na literatura.

Tabela 4.5: Parâmetros avaliados no estudo de sintonia do procedimento de identificação *on-line*.

Parâmetro	Descrição	Níveis
$H \cdot ta$	Tamanho da janela (h)	1, 3, 5
α	Conjunto de reações identificadas	$[\alpha_{F1}]$, $[\alpha_{F1}, \alpha_{F2}, \alpha_{F4}]$, $[\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}]$
α^{LU}	Espaço de busca dos parâmetros estimados	1%, 5%, 10%

Os resultados para os experimentos de monitoramento realizados são apresentados na Tabela 4.6. O melhor resultado obtido foi para o experimento 23, que refere-se a um horizonte de dados de 5 h, modificação das expressões cinéticas das reações r_{F1} , r_{F2} , e r_{F4} e variação admissível de 10% para os parâmetros de eficiência no algoritmo PSO. É interessante notar que, de acordo com os dados das seis melhores simulações, *i.e.*, experimentos 18 – 23, existe consenso sobre o tamanho da janela móvel de 5 h. Entre estes resultados, vale a pena notar que melhores ajustes não foram obtidos à medida que se aumentou o número de expressões de taxa modificadas. Pelo contrário, modificar apenas a reação r_{F1} ou o conjunto r_{F1} , r_{F2} , e r_{F4} forneceu melhores resultados que modificar quatro expressões cinéticas. Embora a faixa de 10% tenha se mostrado como o melhor valor de acordo com as métricas utilizadas, pode-se observar a existência de tendência de sobreajuste. O sobreajuste, ou *overfitting*, corresponde à incorporação do ruído às predições do modelo. Conforme dados da Tabela 4.6, o valor de R^2 para o experimento 23 é superior,

inclusive, ao valor obtido na condição de referência, indicada pelo experimento 0. Assim, de modo a evitar o sobreajuste pela incorporação de ruído, os parâmetros do experimento 21, associado ao intervalo de 1%, foram selecionados.

Tabela 4.6: Métricas estatísticas sobre a qualidade do ajuste para diferentes configurações do procedimento de identificação *on-line*.

Experimento	Parâmetros de ajuste			SSE	MSE	RMSE	MAE	R^2
	$H \cdot ta$	α	Faixa					
0		Validação		29.7629	0.5315	0.7290	0.5142	0.9470
1	1	α_{F1}	1%	47.3026	0.8600	0.9191	0.6797	0.9157
2	1	α_{F1}	5%	50.1551	0.9119	0.9464	0.6780	0.9106
3	1	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	1%	42.9858	0.8111	0.8761	0.6343	0.9234
4	1	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	5%	130.7913	2.4678	1.5283	0.8080	0.7670
5	1	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	10%	117.6690	2.2202	1.4496	0.8180	0.7904
6	1	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	1%	45.8501	0.8817	0.9048	0.6251	0.9183
7	1	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	5%	150.3028	2.8904	1.6383	0.8739	0.7322
8	1	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	10%	114.1601	2.1954	1.4278	0.7848	0.7966
9	3	α_{F1}	1%	36.6597	0.6665	0.8091	0.6026	0.9347
10	3	α_{F1}	5%	39.4607	0.7175	0.8394	0.6440	0.9297
11	3	α_{F1}	10%	35.1565	0.6392	0.7923	0.6041	0.9374
12	3	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	1%	38.2665	0.7220	0.8266	0.6039	0.9318
13	3	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	5%	42.8333	0.8082	0.8746	0.6192	0.9237
14	3	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	10%	44.5507	0.8406	0.8919	0.6360	0.9206
15	3	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	1%	35.7942	0.6884	0.7995	0.5837	0.9362
16	3	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	5%	70.6986	1.3596	1.1236	0.6326	0.8740
17	3	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	10%	58.1220	1.1177	1.0188	0.6680	0.8965
18	5	α_{F1}	1%	33.2743	0.6050	0.7708	0.5808	0.9407
19	5	α_{F1}	5%	33.7425	0.6135	0.7762	0.5815	0.9399
20	5	α_{F1}	10%	33.6353	0.6116	0.7750	0.5800	0.9401
21	5	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	1%	33.5158	0.6324	0.7736	0.5790	0.9403
22	5	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	5%	34.0350	0.6422	0.7796	0.5713	0.9394
23	5	$\alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	10%	28.9275	0.5458	0.7187	0.5211	0.9485
24	5	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	1%	38.0356	0.7315	0.8241	0.5851	0.9322
25	5	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	5%	80.8291	1.5544	1.2014	0.7272	0.8560
26	5	$\alpha_{H3}, \alpha_{F1}, \alpha_{F2}, \alpha_{F4}$	10%	48.5237	0.9331	0.9309	0.5962	0.9136

4.3.2.2 Desempenho do Sensor Virtual

O desempenho da estratégia de monitoramento foi avaliado por meio de simulações. A aplicação do procedimento de identificação *on-line* usando o modelo validado com dados da Batelada 01 (que corresponde ao mesmo conjunto de dados usado na etapa de validação) é apresentada na Figura 4.6, juntamente com as inferências do modelo validado em uma simulação *off-line*. Os resultados do procedimento de identificação *on-line* usando o modelo validado com dados de outras três bateladas não utilizadas na estimação de parâmetros, nem na validação do modelo, são apresentados na Figura 4.7, em que as concentrações de etanol e glicose são mostradas. As simulações

de monitoramento foram realizadas adotando os valores para os parâmetros de ajuste de acordo com o experimento 21 da Tabela 4.6. O procedimento de identificação melhorou significativamente os cálculos de glicose e etanol, dois componentes-chave do processo SSF, para todas essas bateladas.

De acordo com a trajetória dos parâmetros de eficiência estimados, Figura 4.6, é possível observar que os parâmetros do modelo validado sofrem variações locais. Não obstante, a variação dos parâmetros de eficiência, restrita ao intervalo $[0, 35; 2, 41]$, indica que a perturbação da cinética baseada no modelo de Monod nas reações r_{F1} , r_{F2} e r_{F4} foi pequena. Por inspeção dos perfis de α , é interessante notar as diferentes quantidades de informação que cada variável medida possui. Pode-se observar picos que geralmente ocorrem quando uma nova análise de laboratório está disponível. As medidas de concentração contribuem com informação nova, que as medidas de temperatura e nível não abrangem. Assim, as medidas de laboratório contribuem para corrigir as trajetórias inferidas pelo modelo. Este comportamento é refletido nas trajetórias dos parâmetros α ao longo da batelada, em que os coeficientes de eficiência ajustam as expressões de taxa correspondentes.

Esses resultados foram obtidos por meio de simulações computacionais. No entanto, como o procedimento de identificação não é um sistema em malha fechada, as previsões observadas representam os resultados reais que seriam obtidos se o sensor virtual estivesse implementado na planta. No contexto de uma aplicação industrial, os custos computacionais podem mostrar-se proibitivos em determinadas circunstâncias. Por isso, é possível perceber na literatura uma predominância do uso de algoritmos determinísticos para implementações reais em plantas industriais. Deve-se enfatizar, porém, que em todos os experimentos de monitoramento, o algoritmo PSO foi capaz de convergir em menos de 20 segundos a cada instante de amostragem, o que é muito menor que o tempo de amostragem de 5 minutos. Esse desempenho é muito importante, pois, a despeito do fato da otimização ser conduzida por um algoritmo de busca aleatória, demonstra a viabilidade da utilização desse algoritmo em implementações reais.

Para que tenha valor e utilidade em atividades de controle e otimização de processos, uma solução de monitoramento deve ser capaz de prever com segurança as principais características do processo, apesar das incertezas e distúrbios que podem estar presentes. Uma incerteza pronunciável está relacionada à descrição matemática do processo. O monitoramento é uma tarefa desafiadora, mesmo quando se utilizam modelos de processo detalhados, validados em amplas faixas de operação em relação a, *e.g.*, concentração de substrato, carga de enzima e temperatura do reator. Para lidar com essas incertezas, propõe-se a modificação de algumas expressões de taxa pela inclusão de parâmetros de eficiência, conforme descrito anteriormente. Uma abordagem semelhante foi proposta por DAI *et al.* (2014), que modificaram o ba-

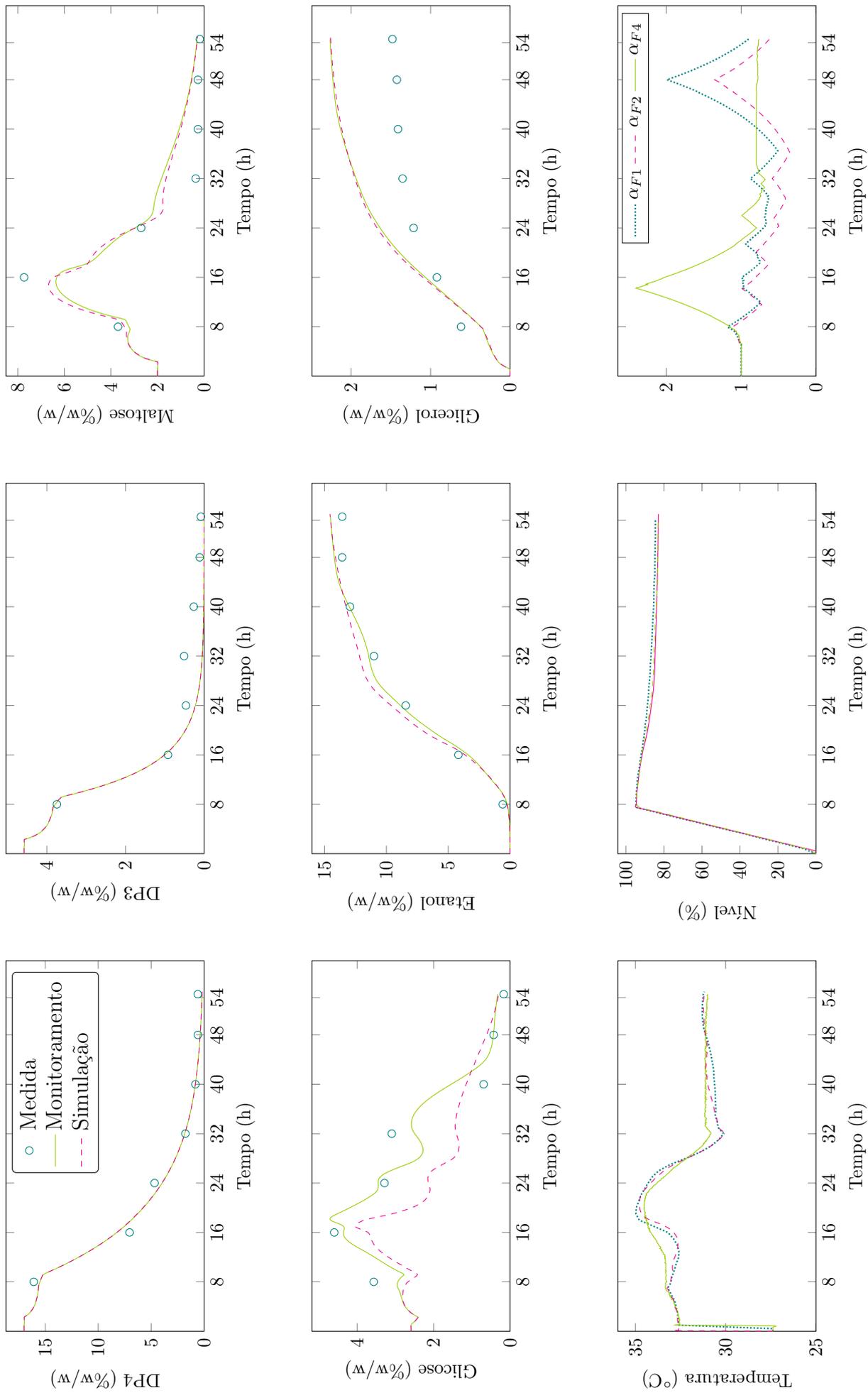


Figura 4.6: Dados medidos da planta e resultados simulados obtidos para o modelo validado e para a estratégia de monitoramento. Resultados baseados nos dados da Batelada 01.

lanço de energia com a introdução de funções lineares e quadráticas da concentração de etanol para capturar incertezas. Eles relatam que, entre outras tentativas, essas expressões proporcionaram resultados significativamente melhores para as medições de temperatura. É importante notar, no entanto, que apesar dos esforços para selecionar uma função apropriada e suas variáveis independentes para a modificação, a extrapolação dessa estratégia para outros processos e condições operacionais não é direta e precisa ser reavaliada.

Outra questão importante em relação ao monitoramento do processo baseado em modelos está relacionada à sensibilidade dos parâmetros a distúrbios, caso em que o modelo pode não ser adequado para o monitoramento do processo. Por exemplo,

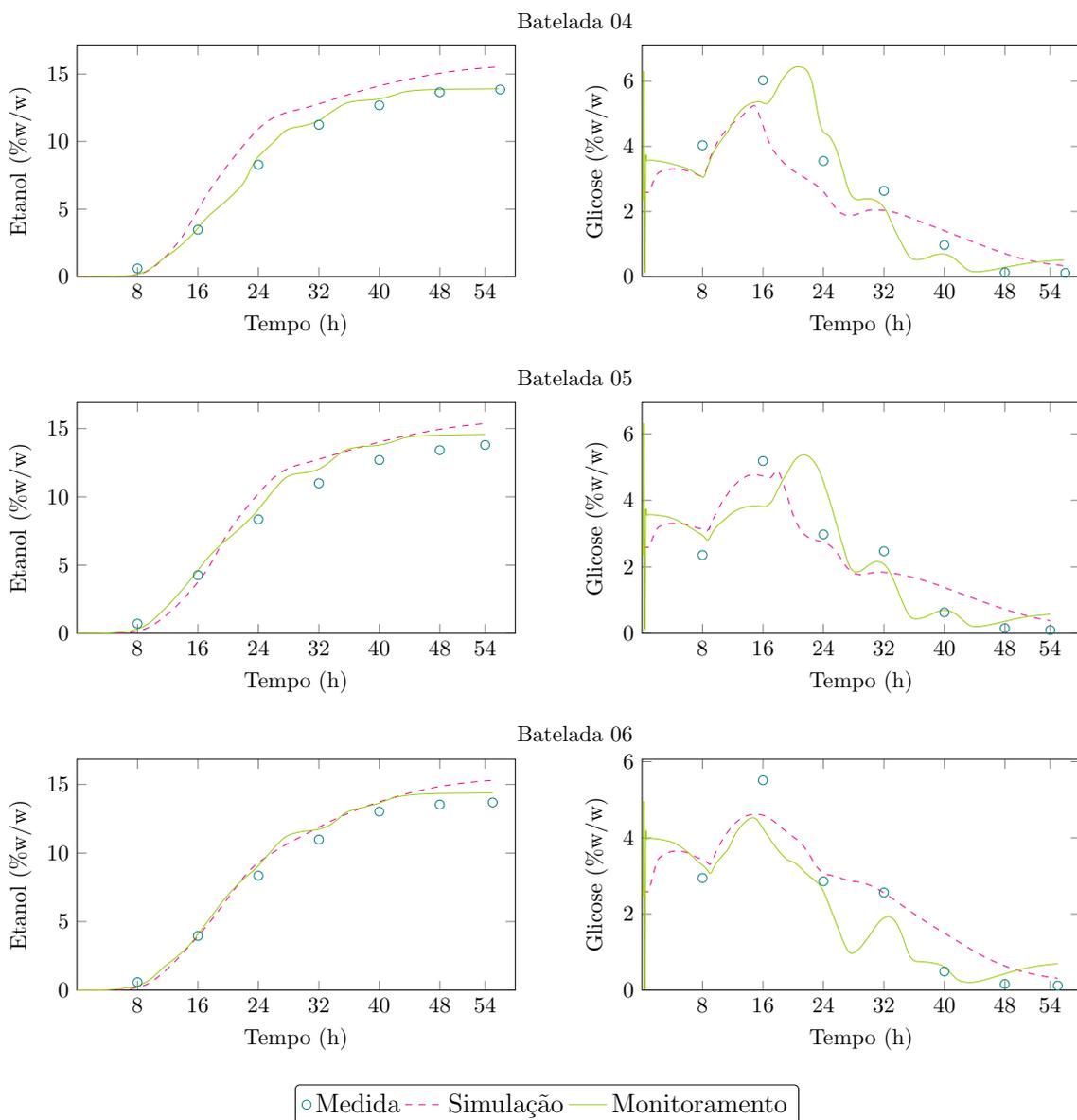


Figura 4.7: Dados de processo, previsões do modelo validado e resultados simulados do sensor virtual para bateladas não utilizadas na estimação de parâmetros ou na validação do modelo.

os parâmetros do modelo podem depender das condições iniciais da batelada, o que é relevante para o monitoramento SSF, uma vez que a adição de substrato afeta o crescimento da levedura e o rendimento do processo. Para lidar com a dependência dos parâmetros do modelo com a concentração inicial de substrato, OCHOA *et al.* (2007) propuseram um método de identificação para distinguir entre parâmetros que podem ser considerados independentes das condições iniciais daqueles que necessariamente mudam com elas e devem ser ajustados sempre que mudanças ocorrem. Para conseguir isso, foram utilizados diferentes conjuntos de dados experimentais em diferentes condições iniciais. No entanto, sempre haverá a necessidade de estimar esses parâmetros, uma vez que as condições iniciais mudem. Embora OCHOA *et al.* (2007) forneçam um método valioso para lidar com a mudança de parâmetros em algumas condições, pode não ser prático para uso em processos reais, pois pode ser muito complexo e requerer tempo e esforço de pessoal especializado para atualizar o modelo. Além disso, o método lida com um distúrbio conhecido, o que pode não ser adequado para uma aplicação industrial. Se algum distúrbio, seja conhecido ou não, não for refletido no conjunto de dados disponíveis, seu efeito nos parâmetros permanecerá desconhecido, afetando as previsões do modelo. Como OCHOA *et al.* (2007) argumentaram, essa estratégia pode reduzir muito o esforço computacional, mas os parâmetros devem ser periodicamente recalculados usando técnicas de adaptação *on-line*.

Conforme apresentado nas Figuras 4.6 e 4.7, o sensor virtual é capaz de prever a operação de diferentes bateladas com razoável precisão. Ao combinar uma abordagem para captura de incertezas com um procedimento de identificação *on-line*, é possível atualizar o modelo localmente com base no comportamento mais recente da planta. Embora o sensor virtual use um modelo de primeiros princípios baseado em cinéticas simples de Michaelis-Menten e Monod, que são modelos não segregados, os resultados mostraram que a estratégia permitiu o acompanhamento de variáveis-chave que permitem o monitoramento do processo. Com efeito, o valor de modelos matemáticos simples para fornecer inferências precisas com base em várias medidas já foi demonstrado em outros sistemas (CLAES e VAN IMPE, 2000; HOFFMANN *et al.*, 2000). A identificação *on-line* proposta combina os conhecimentos incorporados no modelo com todas as informações disponíveis nas medidas da planta, ajustando as inferências do modelo em tempo real. Isso permite não apenas capturar o comportamento dinâmico dos microrganismos vivos, mas também detectar perturbações relacionadas ao processo. Considerando essas características, o procedimento de identificação *on-line* é mais apropriado para o controle e otimização de bioprocessos (VAN IMPE, 1996), em que a proposta de correção multiplicativa das taxas pode ser usada em muitos outros problemas.

4.4 Conclusão

O desenvolvimento e aplicação de um sensor virtual para monitoramento de SSF de hidrolisado de milho foi apresentado. A estrutura é facilmente extensível a outros sistemas e emprega um procedimento para identificar o processo *on-line*, usando um modelo caixa cinza. O modelo foi capaz de prever os perfis dinâmicos de etanol, glicerol, maltotetraose, maltotriose, maltose, glicose, temperatura e nível de líquido dentro do vaso de fermentação. Ao adotar a abordagem de horizonte móvel, o procedimento faz uso de medidas geralmente presentes em plantas típicas de etanol de milho para ajustar o modelo *on-line*, estimando os coeficientes de eficiência. A ferramenta de monitoramento foi avaliada com estudos de simulação, usando dados industriais reais de uma planta de etanol de milho. Os dados considerados neste trabalho consistiram em medidas *on-line* da temperatura de saída da corrente de água do trocador de calor e do nível do vaso, e análises de laboratório para etanol, glicerol e açúcares via HPLC. O procedimento de identificação *on-line* foi resolvido pelo algoritmo metaheurístico PSO. O estudo mostrou que mesmo os modelos fenomenológicos simples, como as cinéticas de Michaelis-Menten e Monod, podem ser usados para inferir o estado operacional da fermentação em batelada e identificar o modelo com as medidas mais recentes. A estratégia também pode ser usada para inferir as taxas reais de fermentação e hidrólise, bem como a energia envolvida no processo, tornando-a uma ferramenta poderosa para examinar mais de perto o que está acontecendo dentro do fermentador.

A SSF é uma etapa crítica em muitos processos industriais. Portanto, informações muito úteis para reduzir o consumo de energia e aumentar a produção podem ser obtidas explorando o processo de fermentação em plantas industriais de etanol de milho. Na indústria, biorreatores são, geralmente, operados manualmente por operadores que usam um conjunto de medidas de processo. Embora esse tipo de operação não exija nenhum modelo, pode levar a uma baixa reprodutibilidade se não forem detectados desvios significativos. Especialmente em bioprocessos, as inevitáveis flutuações na carga, qualidade das matérias-primas naturais e contaminação afetam a ação de enzimas e microorganismos, de tal forma que os distúrbios do processo real são difíceis de prever. Uma informação dinâmica rica e realista fornecida por um sensor virtual deve ajudar o operador de várias maneiras, fornecendo informações adicionais sobre o processo que não estavam disponíveis pelas variáveis medidas. O monitoramento *on-line* do processo é um passo importante para melhorar a operação do processo, permitindo uma maior reprodutibilidade para controle em laço aberto com otimização de batelada em batelada. Em plantas mais automatizadas, ele é útil para dar suporte ao controle e aplicações de otimização em tempo real. O custo de instalação e manutenção do equipamento envolvido no sen-

sensor virtual é vantajoso em relação ao benefício de suas inferências para a estratégia de monitoramento e controle. Portanto, o sensor virtual para SSF de hidrolisado de milho contribui significativamente para a melhoria do processo de produção industrial de etanol combustível. Como exemplos de oportunidades criadas pela aplicação da estratégia de monitoramento proposta, a estratégia de amostragem poderia ser simplificada e as atividades dos operadores repensada; simulações poderiam indicar cenários críticos e antecipar desvios do processo; os dados *on-line* que refletem atividades de enzimas reais poderiam ser usados para redução de custos com enzimas e outros aditivos; e novos conhecimentos sobre a operação do processo podem ajudar a conceber diferentes estratégias de controle, aumentando a produtividade.

Portanto, estratégias de construção de modelos de referência foram desenvolvidas a contento em dois cenários muito distintos: com excesso (Capítulo 3) e com escassez (Capítulo 4) de dados. Isso permite levar as ferramentas de otimização em tempo real a diferentes domínios.

Capítulo 5

Aspectos Numéricos e Abordagens de Resolução

Embora a maior parte da discussão envolvendo estudos de RTO esteja relacionada às dificuldades existentes para obter uma representação matemática precisa do processo com o conjunto de informações disponível, há casos em que isso não constitui o problema principal. No entanto, a disponibilidade de um modelo fenomenológico do processo associado a um conjunto adequado de variáveis medidas não é suficiente para garantir a resolução adequada do problema de programação matemática, seja na etapa de atualização do modelo, seja na etapa de otimização econômica. Nesse contexto, torna-se necessário entender como a estrutura do problema de otimização e o algoritmo numérico, em conjunto com sua parametrização, afetam a identificação do processo. Assim, realizou-se um estudo de caracterização de desempenho de diferentes técnicas numéricas empregadas para resolução do problema de reconciliação de dados, associado às etapas de *Identificação* do processo e de *Otimização* em tempo real, conforme fluxograma da Figura 1.2.

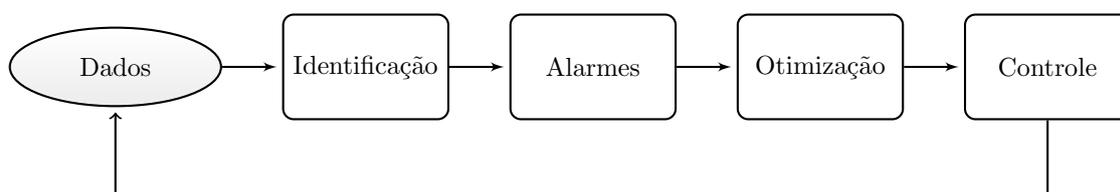


Figura 1.2: Fluxograma alternativo com as etapas envolvidas em uma implementação real de sistema RTO (apresentado na p.8).

A reconciliação de dados é uma técnica baseada em modelo amplamente aplicada em indústrias de processos modernas, estando disponível em *softwares* comerciais. Apesar disso, vários aspectos importantes envolvidos na solução do problema são geralmente ignorados e uma avaliação detalhada de como os procedimentos numéricos afetam a solução do problema não pode ser desprezada. Com base em aplicações

industriais relatadas na literatura, diferentes configurações foram identificadas e testadas, fornecendo uma avaliação numérica sobre o desempenho da reconciliação de dados estacionária e não linear. Os itens discutidos foram a formulação do problema, com relação à presença de parâmetros estimados na função objetivo; a abordagem de solução, quando se aplicam algoritmos de programação não linear; técnicas para estimar gradientes de função objetivo; a estimativa inicial; e o algoritmo de otimização.

O estudo baseou-se nas simulações de um modelo rigoroso e validado de um sistema real de produção de petróleo *offshore*. A avaliação consistiu em comparar a robustez da solução, a violação de restrições e o custo computacional. Além disso, foi proposto o uso de um teste global para detectar inconsistências na formulação e na solução do problema. Os resultados mostram que diferentes configurações exercem um grande impacto sobre o desempenho dos procedimentos de reconciliação, muitas vezes resultando em soluções locais. Na essência, discute-se a questão de como resolver satisfatoriamente o problema de reconciliação de dados para obter melhores estimativas. O presente trabalho foi publicado no periódico *Processes* (CÂMARA *et al.*, 2017b).

5.1 Introdução

Dados confiáveis são essenciais para qualquer processo industrial. Da operação do processo ao planejamento e programação da produção, dados são usados para auxiliar a tomada de decisões que afetam a qualidade do produto, rentabilidade e segurança da planta. Dados também estão envolvidos em questões ambientais e legais, como no monitoramento da poluição pelas agências estatais. Na contabilidade industrial, estão envolvidos procedimentos importantes, como na transferência de custódia e nos inventários de processo. No entanto, especialmente em ambientes industriais, a obtenção de informações confiáveis durante a operação diária não constitui uma tarefa trivial.

As medições das variáveis de processo, tanto *on-line* quanto *off-line*, estão sujeitas a erros aleatórios (ruído), tais como distúrbios naturais, erros de amostragem, leituras de instrumentos não confiáveis e imprecisões de análises laboratoriais, e erros não aleatórios, como a falta de calibração ou falha de sensores (erros grosseiros) (HODOUIN e EVERELL, 1980). O ruído aleatório também pode ser o resultado de outros efeitos, como interferências de alta frequência, baixa resolução e erros na transmissão e conversão (incluindo conversão A/D e D/A) (TAYLOR e DEL PILAR MORENO, 2013), geralmente atribuídos à irreprodutibilidade dos instrumentos de medição (MORAD *et al.*, 2005). Conseqüentemente, os dados coletados geralmente não satisfazem nem mesmo as restrições básicas do processo, como os

balanços de massa e energia. Dados imprecisos de processo podem resultar em conclusões enganosas ou imprecisas, levando a decisões equivocadas que podem afetar adversamente muitas funções da planta. Assim, para reduzir os impactos dos erros de medição nas medições da planta e aumentar o valor dos dados acessíveis através de sistemas de gerenciamento de dados implementados, a retificação de dados deve ser empregada (JOHNSTON e KRAMER, 1995).

A retificação de dados compreende um conjunto de técnicas que visa corrigir dados, sendo composta por diferentes etapas, como classificação de variáveis, detecção de erros grosseiros, estimação de parâmetros e reconciliação de dados (NARASIMHAN e JORDACHE, 1999). A reconciliação de dados (DR) é uma técnica baseada em modelos que reduz os erros de medição, fazendo uso de redundâncias nos dados de processo. Essa técnica tem como objetivo estimar o estado verdadeiro da planta, com base no ajuste das medidas do processo para satisfazer um conjunto de restrições, o que é conseguido minimizando-se algum tipo de desvio entre valores medidos e corrigidos. Ao contrário de outras técnicas de filtragem, a DR faz uso explícito do modelo de processo para que os valores reconciliados sejam consistentes com as relações conhecidas entre as variáveis de processo, conforme definido pelas restrições, e ainda mais precisos que os valores medidos (MARTINI *et al.*, 2013; NOUNOU e BAKSHI, 1999).

A reconciliação de dados também pode ser usada como uma ferramenta eficiente de monitoramento, produzindo estimativas para variáveis não-medidas e parâmetros de processo. Em muitos sistemas e aplicações, a estimação de parâmetros é a etapa seguinte à reconciliação de dados, em que os valores reconciliados das variáveis medidas são usados para estimar e atualizar os parâmetros do modelo (MARLIN e HRYMAK, 1997; NAYSMITH e DOUGLAS, 1995; ZHANG *et al.*, 1995b). Todavia, procedimentos de reconciliação de dados e estimação simultânea de parâmetros (DRPE) são igualmente válidos e eficientes (HLAVÁČEK, 1977; MACDONALD e HOWAT, 1988; MCBRAYER *et al.*, 1998; PICCOLO *et al.*, 1996; WEISS *et al.*, 1996). É importante ressaltar que as medidas de variáveis independentes são tipicamente utilizadas como se fossem livres de erros. Neste contexto, a DR também pode ser aplicada quando tanto as variáveis dependentes quanto as independentes contêm erros, situação em que alguns autores chamam de *error-in-variables method* (EVM) (KAZEMI *et al.*, 2013; KIM *et al.*, 1990). A técnica EVM fornece estimativas de ambos parâmetros e valores reconciliados que sejam consistentes com o modelo.

Além disso, técnicas para reconciliação de dados e estimação simultânea de parâmetros em conjunto com detecção de erros grosseiros surgiram com base em estatística robusta, que são especialmente importantes para aplicações industriais *on-line* (PRATA *et al.*, 2010). Muitos estimadores robustos foram propostos para reduzir o efeito de erros grosseiros e produzir estimativas menos tendenciosas (ARORA e BI-

EGLER, 2001; FABER *et al.*, 2007; PRATA *et al.*, 2008c; TJOA e BIEGLER, 1991; ZHANG *et al.*, 2010). Abordagens alternativas para a detecção de erros grosseiros envolvem o uso de análise de componentes principais (PCA) (NARASIMHAN e BHATT, 2015; TONG e CROWE, 1995), análise de cluster (CHEN e ROMAGNOLI, 1998), redes neuronais artificiais (VACHHANI *et al.*, 2001) e estimadores robustos (ÖZYURT e PIKE, 2004) e baseados em correntropia (CHEN *et al.*, 2013). Recentemente, foram apresentadas uma metodologia para DRPE em sistemas com múltiplas condições de operação, com um estimador baseado em correntropia (ZHANG *et al.*, 2014) e uma estrutura unificada para aplicação de PCA e DR (NARASIMHAN e BHATT, 2015).

A literatura relativa a reconciliação de dados é composta por uma quantidade expressiva de trabalhos e boas compilações das técnicas disponíveis podem ser encontradas na literatura aberta (CROWE, 1996; ÖZYURT e PIKE, 2004; TAMHANE e MAH, 1985), bem como em livros (MADRON, 1992; MAH, 1990; NARASIMHAN e JORDACHE, 1999; ROMAGNOLI e SÁNCHEZ, 2000; VEVERKA e MADRON, 1997). Começando com problemas em estado estacionário com restrições lineares para o caso em que todas as variáveis são medidas (KUEHN e DAVIDSON, 1961), os avanços em técnicas de otimização permitiram o uso de formulações mais complexas, envolvendo reconciliação de dados estacionário não linear (DU *et al.*, 1997; KIM *et al.*, 1997; PAGES *et al.*, 1994; TAKIYAMA *et al.*, 1991; TJOA e BIEGLER, 1991), dinâmico linear (DAROUACH e ZASADZINSKI, 1991; FILLON *et al.*, 1995, 1996) e dinâmico não linear (ALBUQUERQUE e BIEGLER, 1996; CHEN e ROMAGNOLI, 1998; LIEBMAN *et al.*, 1992; MCBRAYER *et al.*, 1998; PRATA *et al.*, 2006). Em cada uma dessas formulações, o modelo de processo, representado pelo sistema de equações algébricas ou diferencial, é uma restrição do modelo de programação matemática que representa o problema de estimação, sendo que a solução das equações do modelo é necessária para resolver o problema de otimização.

Aplicada na maior parte das indústrias modernas, a DR está comercialmente disponível em ferramentas de *software* como Aspen Advisor (Aspen Technology, Inc., Bedford, MA, EUA), VALI (Belsim, Awans, Bélgica), VisualMesa (Soteica, Houston, TX, EUA), SimSci DATACON (Schneider Electric Software, Lake Forest, CA, EUA, LLC), Sigmafine (OSIsoft, San Leandro, CA, EUA), entre outras (BAGAJEWICZ e ROLLINS, 2011). Em conjunto com sistemas desenvolvidos internamente nas empresas, essas ferramentas podem ser utilizadas para reconciliar medidas de vazão, temperatura e composição, de modo a satisfazer os balanços de massa e energia em cada unidade de uma planta de processo, bem como para estimar parâmetros (KYRIAKOPOULOU e KALITVENTZEFF, 1996; REIMERS *et al.*, 2008; SARABIA *et al.*, 2012). Além disso, o aumento da disponibilidade e o custo decrescente de computadores de alto desempenho permitiram a implementação de aplicações mais

robustas e complexas, em que medidas reconciliadas são usadas em aplicações *on-line*, como monitoramento, controle de processo e otimização em tempo real (BHAT e SARAF, 2004; SANCHEZ *et al.*, 1992; TRIERWEILER, 2014).

Dada a importância das técnicas de DR, a quantidade expressiva de trabalhos nesse tema e a disponibilidade da técnica em ferramentas comerciais, descrições detalhadas dos procedimentos de reconciliação de dados e seu uso em aplicações industriais reais são relativamente escassas na literatura. Esta aparente discrepância pode ser parcialmente atribuída a razões proprietárias e restrições confidenciais. Os trabalhos relacionados às implementações industriais de DR estão principalmente focados nos desafios relacionados à obtenção de uma representação matemática precisa do processo, dada a informação disponível. Eles geralmente se concentram no desenvolvimento de procedimentos e estratégias computacionais para novos sistemas, que geralmente envolve a formulação do problema de DR com base em um modelo rigoroso do processo e sua resolução, usando um algoritmo de otimização típico, adequado para problemas de programação matemática de alta dimensão, bem como esquemas de detecção de erros mais robustos. Embora tenham um valor técnico inegável, a implementação e o avanço das aplicações *on-line* de DR exigem a escolha da configuração mais adequada para a implementação.

Com efeito, a disponibilidade de um modelo preciso do processo e um conjunto adequado de variáveis medidas não é suficiente para resolver de maneira satisfatória o problema de DR. Poucos estudos apresentam análises mais detalhadas das configurações de diferentes elementos, como formulação do problema (JOHNSTON e KRAMER, 1995; ÖZYURT e PIKE, 2004), método de solução (HU e SHAO, 2006; ISLAM *et al.*, 1994; WEISS *et al.*, 1996) e abordagem (KRIST *et al.*, 1994; PIERRUCCI *et al.*, 1996), algoritmo de otimização (KELLY, 2004; POKU *et al.*, 2004; PRATA *et al.*, 2009b; SARABIA *et al.*, 2012), e estratégias de solução (FABER *et al.*, 2007). Configurações distintas, mesmo aquelas representadas por pequenas diferenças, podem conferir desempenhos diferentes, mas comumente não são discutidas. No entanto, elas devem ser analisadas durante o projeto de uma aplicação de DR, tentando responder se o problema de otimização está sendo resolvido satisfatoriamente. Embora a reconciliação de dados gere benefícios, ela pode exercer fortes efeitos negativos quando alguma hipótese é violada ou o procedimento não encontra uma solução viável, caso em que o uso dos valores reconciliados pode produzir resultados piores que usar as medições diretamente (CÂMARA *et al.*, 2016; QUELHAS *et al.*, 2013; ÉRIC POULIN *et al.*, 2010).

Neste trabalho, buscou-se fornecer uma avaliação numérica sobre o desempenho de diferentes configurações na resolução da reconciliação de dados estacionária não-linear. É mostrado que configurações diferentes exercem um grande impacto no desempenho dos procedimentos de DR e que uma configuração adequada deve ser

encontrada. A formulação do problema, relativo à presença de parâmetros estimados na função objetivo, bem como a abordagem da solução, sequencial ou simultânea, foram consideradas quando se aplicam algoritmos de programação não-linear (NLP) como método de solução. Investigou-se o impacto de métodos para estimar gradientes da função objetivo exigidos por algoritmos determinísticos de otimização, o que é geralmente ignorado em estudos e aplicações de DR. Neste trabalho, um algoritmo de ponto interior (IP) é aplicado. O desempenho de um algoritmo híbrido que combina o algoritmo metaheurístico de otimização por enxame de partículas (PSO) e IP também é avaliado e comparado. Como uma variável conhecida que afeta o desempenho dos algoritmos de otimização, analisou-se também a influência da estimativa inicial na solução do problema de DR. A avaliação baseia-se na robustez da solução, na violação de restrições na convergência e no custo computacional. Para controlar melhor os resultados, apenas dados simulados foram utilizados no estudo e a detecção de erros grosseiros não foi considerada no escopo deste trabalho. O estudo baseia-se em simulações de um modelo rigoroso e validado de um sistema real de produção de petróleo *offshore*. O modelo de processo está disponível em simulador proprietário confiável, amplamente utilizado na Petrobras (LIPORACE *et al.*, 2009; NIEDERBERGER *et al.*, 2005, 2009). É importante enfatizar que poucos trabalhos se concentram no uso de procedimentos de reconciliação de dados em sistemas de produção de petróleo (TAYLOR e DEL PILAR MORENO, 2013).

O presente capítulo está organizado da seguinte forma. A Seção 5.2 apresenta uma visão geral da reconciliação de dados em aplicações industriais. A Seção 5.3 descreve o sistema *offshore* de produção de petróleo e água, para o qual o estudo se baseia, e também detalha os experimentos de simulação realizados para avaliar diferentes configurações da resolução da DR não linear. Os resultados são apresentados e discutidos na Seção 5.4 e a Seção 5.5 apresenta as conclusões.

5.2 Visão Geral da Reconciliação de Dados em Aplicações Industriais

5.2.1 Formulação do Problema

Tal como foi apresentado na Seção 2.4.3 (p. 55), a essência da reconciliação de dados consiste em, considerando as medidas \mathbf{z}_{ms} da planta, estimar o estado do processo que satisfaz as relações conhecidas entre as variáveis de processo \mathbf{z} conforme definido pelo modelo do processo.

Dependendo da natureza das equações do modelo do processo, o problema de DR pode ser linear ou não linear, e estático ou dinâmico. Conforme apresentado na Tabela 5.1, construída após exaustiva análise de trabalhos publicados na literatura,

a maioria das aplicações encontradas está relacionada a reconciliação de dados estacionária não linear. Neste contexto, considerando problemas de DR não lineares em estado estacionário, com as equações do modelo do processo convenientemente representadas por

$$\mathbf{f}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \quad ,$$

as formulações do problema de DR normalmente seguem a Equação (2.42) com um estimador WLS baseado na suposição (implícita) de erros de medição independentes e normalmente distribuídos, resultando no seguinte programa não linear com restrições

$$\begin{aligned} \hat{\mathbf{z}} := & \arg \min_{\mathbf{z}} (\mathbf{z}_{ms} - \mathbf{z}_m)^T \mathbf{V}_z^{-1} (\mathbf{z}_{ms} - \mathbf{z}_m) \\ & \text{sujeito a } \mathbf{f}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \\ & \mathbf{g}(\mathbf{z}, \boldsymbol{\theta}) \leq \mathbf{0} \end{aligned} \quad . \quad (5.1)$$

A Equação (5.1) é facilmente estendida para um problema DRPE pela inclusão dos parâmetros $\boldsymbol{\theta}$ (ou um subconjunto) no conjunto de variáveis de decisão, conforme apresentado na Equação 2.46 (p. 58).

Tabela 5.1: Aplicações industriais reais de técnicas de reconciliação de dados ou estudos de reconciliação de dados envolvendo dados reais de planta.

Referência	Natureza	Estimador	Método de solução	Algoritmo- Pacote / Simulador	Processo industrial
HODOUIN e EVERELL (1980)	estático não linear	GL	Multiplicadores de Lagrange	Gradiente conjugado	Planta de processamento de minérios
SERTH <i>et al.</i> (1987)	estático não linear	GL	Multiplicadores de Lagrange, Lin. Suc.	BSOLVE	Metalurgia
HOLLY <i>et al.</i> (1989)	estático linear	GL	Multiplicadores de Lagrange	analítica	Extração química
SANCHEZ <i>et al.</i> (1992)	estático não linear	WLS	algoritmo NLP	SQP / PLADAT	Planta de etileno
MEYER <i>et al.</i> (1993)	estático não linear	GL	Multiplicadores de Lagrange	Newton-Raphson	Destilação de álcool em bebidas
ISLAM <i>et al.</i> (1994)	estático não linear	GL	Solução linear, Lin. Suc., algoritmo NLP	SQP	Reator de pirólise
KRIST <i>et al.</i> (1994)	<i>on-line</i> estático não linear	LS	Multiplicadores de Lagrange	Speedup/AspenPlus	Planta de benzeno
BUSSANI <i>et al.</i> (1995)	<i>on-line</i> estático não linear	WLS de erros relativos	algoritmo NLP	SQP / ORO	Planta de hidrogênio
ZHANG <i>et al.</i> (1995b)	<i>on-line</i> estático não linear	DR: Normal contaminada, PE: LS	algoritmo NLP	MINOS-GAMS / Aspen Plus	Planta de ácido sulfúrico
ZHANG <i>et al.</i> (1995a)	<i>on-line</i> estático não linear	Normal contaminada	algoritmo NLP	MINOS-GAMS / Aspen	Planta de ácido sulfúrico

Continua na próxima página

Tabela 5.1 – continuação da página anterior

Referência	Natureza	Estimador	Método de solução	Algoritmo- Pacote / Simulador	Processo industrial
WEISS <i>et al.</i> (1996)	estático não linear	GL	Solução linear, Lin. Suc., algoritmo NLP	Gauss-Marquardt, SQP	Reator de pirólise
SÁNCHEZ <i>et al.</i> (1996)	estático não linear	GL	Solução linear, Lin. Suc., algoritmo NLP	SQP	Reator de pirólise
SÁNCHEZ e ROMAGNOLI (1996)	estático (bi)linear	GL	Multiplicadores de Lagrange	Projeção de matrizes, decomposição QR	Plantas de etileno e amônia
PIERUCCI <i>et al.</i> (1996)	<i>on-line</i> estático não linear	WLS	algoritmo NLP	SQP, Levenberg Marquardt, LNSI / ORO	Forno industrial
PICCOLO <i>et al.</i> (1996)	estático não linear	WLS	algoritmo NLP	SQP/Apen Plus, MINOS5-GAMS	Separação de gás em refinaria
HEYEN <i>et al.</i> (1996)	estático não linear	GL	Multiplicadores de Lagrange	NA	Produção de energia
KYRIAKOPOULOU e KALITVENTZEFF (1996)	estático não linear	GL	algoritmo NLP	SQPIP, SOLDG-BELSIM, SQPHP / VALI-BELSIM	Rede de produção de vapor
CHIARI <i>et al.</i> (1997)	<i>on-line</i> estático não linear	WLS de erros relativos	algoritmo NLP	SQP / ORO	Plantas de hidrogênio e enxofre
CHRISTIANSEN <i>et al.</i> (1997)	estático não linear	GL	Lin. Suc., algoritmo NLP	SQP	Produção de gás de síntese
HEYEN e KALITVENTZEFF (1997)	estático não linear	GL	NLP, MINLP	SQPIP, MINLP	Redes de fornecimento e distribuição de energia
MCBRAYER <i>et al.</i> (1998)	dinâmico não linear	WLS	algoritmo NLP	SQP	Processo químico
DEMPF e LIST (1998)	<i>on-line</i> estático não linear	GL	NA	NA / VALI	Plantas químicas
PLÁCIDO e LOUREIRO (1998)	estático não linear	GL	Multiplicadores de Lagrange	Projeção de matrizes	Planta de amônia
BOUROUIS <i>et al.</i> (1998)	estático não linear	Normal contaminada	algoritmo NLP	NA	Planta de dessalinização
CHEN <i>et al.</i> (1998b)	<i>on-line</i> estático não linear	Normal contaminada, Lorenziana (GED-DR), WLS (DRPE)	algoritmo NLP	CONOPT-GAMS	Planta de ácido sulfúrico
LEE <i>et al.</i> (1998)	<i>on-line</i> estático (não) linear	GL	algoritmo NLP	SQP	Planta de utilidades
BAZIN <i>et al.</i> (1998)	estático não linear	GL	algoritmo NLP	Gradiente conjugado	Secador rotativo
SODERSTROM <i>et al.</i> (2000)	dinâmico não linear	WLS	algoritmo NLP	SQP-NOVA	Processo químico

Continua na próxima página

Tabela 5.1 – continuação da página anterior

Referência	Natureza	Estimador	Método de solução	Algoritmo- Pacote / Simulador	Processo industrial
LI <i>et al.</i> (2001)	<i>on-line</i> estático não linear	WLS	Lin. Suc.	NA	Unidade de destilação de óleo cru
ZHANG <i>et al.</i> (2001a)	estático linear	GL	Multiplicadores de Lagrange	NA	Refinaria
EKSTEEN <i>et al.</i> (2002)	estático bilinear	1. GL, 2. Resíduo de balanços	1. algoritmo NLP, 2. Multiplicadores de Lagrange	1. GRG, 2. NA	Forno industrial
SUNDE e ÖLVIND BERG (2003)	estático não linear	GL	NA	NA / TEMPO system	Turbina de produção de energia a vapor (3300 MW)
BAGAJEWICZ e CABRERA (2003)	estático não linear	GL	algoritmo NLP	GAMS / SIMSCI	Gasoduto
VIEIRA <i>et al.</i> (2003)	dinâmico não linear	GL	Multiplicadores de Lagrange	Gauss-Newton	Reator de polimerização
BHAT e SARAF (2004)	estático linear	GL	Multiplicadores de Lagrange	analítica	Unidade de destilação de óleo cru
CHATTERJEE e SARAF (2004)	estático não linear	LS	algoritmo NLP	SQP-NPSOL	Unidade de destilação de óleo cru
ÖZYURT e PIKE (2004)	estático não linear	Normal contaminação, Cauchy, Fair, Hampel, Logística, e Lorenziana	algoritmo NLP	Conopt2, Minos5	Planta de ácido sulfúrico
YI e HAN (2004)	estático MILP	GL	MILP	MILP	Redes de gases
CHEN e ANDERSEN (2005)	estático não linear	GL no resíduo de balanços	Multiplicadores de Lagrange	NA	Turbina a gás
DE AN-DRADE LIMA (2006)	estático não linear	GL	algoritmo NLP	SQP-NPSOL	Extração de ouro
AL-ARFAJ (2006)	estático linear	GL	Multiplicadores de Lagrange	Abordagem QR	Planta de metil-terc-butil-éter
FABER <i>et al.</i> (2006)	<i>on-line</i> estático linear e não linear	Estimador robusto de LI <i>et al.</i> (2000), função Fair	algoritmo NLP	SQP - IMSL	Purificação de gás de coqueria
HU e SHAO (2006)	estático não linear	GL	Multiplicadores de Lagrange, Lin. Suc.	NA	Planta de coque
PRATA <i>et al.</i> (2006)	dinâmico não linear	GL	algoritmo NLP	Gauss-Newton	Reator de polimerização
FABER <i>et al.</i> (2007)	<i>on-line</i> estático linear e não linear	Estimador robusto de LI <i>et al.</i> (2000), função Fair	algoritmo NLP	SQP - IMSL	Purificação de gás de coqueria
SCHLADT e HU (2007)	estático não linear	Normal contaminação	algoritmo NLP	SQP-SNOPT	Coluna de destilação

Continua na próxima página

Tabela 5.1 – continuação da página anterior

Referência	Natureza	Estimador	Método de solução	Algoritmo- Pacote / Simulador	Processo industrial
BELLEÇ <i>et al.</i> (2007)	<i>on-line</i> estático bilinear	WLS	filtro de Kalman	Sigmafine (OSI)	Planta de celulose
LID e SKO-GESTAD (2008a)	estático não linear	Normal contamina	algoritmo NLP	Fmincon-MATLAB	Reformador de nafta
REIMERS <i>et al.</i> (2008)	estático linear	GL	Multiplicadores de Lagrange	Gauss-Jordan / SolidSim	Planta de carvão
EMBIRUÇU <i>et al.</i> (2008)	dinâmico não linear	GL	Multiplicadores de Lagrange	Gauss-Newton	Polimerização de etileno
PRATA <i>et al.</i> (2008b)	dinâmico não linear	GL	Multiplicadores de Lagrange	Gauss-Newton	Reator de polimerização
PUIG <i>et al.</i> (2008)	estático linear	GL	NA	Macrobal	Planta de tratamento de efluentes
PRATA <i>et al.</i> (2009b)	dinâmico não linear	GL	1. Multiplicadores de Lagrange, 2. algoritmo NLP	1. Gauss-Newton, 2. PSO	Reator de polimerização
PRATA <i>et al.</i> (2010)	dinâmico não linear	GL, estimador Welsch	algoritmo NLP	PSO	Reator de polipropileno
SARABIA <i>et al.</i> (2012)	<i>on-line</i> estático não linear	GL	algoritmo NLP	SQP-NAG/EcosimPro, em refinaria IPOPT-GAMS	Redes de hidrogênio
QUELHAS <i>et al.</i> (2013)	<i>on-line</i> estático não linear	GL	algoritmo NLP	Gauss-Newton	Produção de etileno
ZHANG <i>et al.</i> (2014)	estático não linear	GL, estimador de correntropia	NA	NA	Destilação de ar
CICCIOTTI <i>et al.</i> (2014)	estático não linear	GL	algoritmo NLP	NA	Compressor centrífugo
JIANG <i>et al.</i> (2014b)	<i>on-line</i> estático não linear	GL	Lin. Suc.	NA	Planta de produção de energia (1000 MW)
EGHBAL AH-MADÍ (2015)	estático não linear	GL	algoritmo NLP	SQP / Aspen HySys	Unidade de recuperação de enxofre
GUO <i>et al.</i> (2016a)	estático não linear	GL	algoritmo NLP	NA-MATLAB	Planta de produção de energia (1000 MW)
GUO <i>et al.</i> (2016b)	estático não linear	GL	algoritmo NLP	SQP-fmincon/MATLAB	Planta de produção de energia (1000 MW)
BEHNAMI <i>et al.</i> (2016)	estático linear	WLS	Multiplicadores de Lagrange	MATLAB	Tratamento de efluentes de petroquímica
RAFIEE e BEH-ROUZSHAD (2016)	estático não linear	GL	algoritmo NLP	Fmincon-MATLAB	Planta de gás natural
CÂMARA <i>et al.</i> (2016)	<i>on-line</i> estático não linear	WLS	algoritmo NLP	NA	Destilação de petróleo

WLS – Mínimos quadrados ponderados; GL – Verossimilhança gaussiana (WLS); SQP – Programação quadrática sucessiva; NLP – Programação não linear; NA – Não informado.

Para lidar com desvios desta hipótese dos erros, como é o caso quando o conjunto de dados contém erros grosseiros e *outliers*, esquemas de detecção de erros grosseiros (GED) foram propostos, como a abordagem simultânea para a reconciliação de

dados e detecção de erros grosseiros. Esta técnica é baseada em estatística robusta e consiste em encontrar estimadores (ou seja, uma função objetivo) cuja natureza matemática, especialmente a sua função de influência, seja insensível a desvios das hipóteses ideais sobre os erros, valorizando menos a contribuição de erros grandes no problema de estimação (JOHNSTON e KRAMER, 1995; PRATA *et al.*, 2008c). O resultado esperado é um efeito reduzido dos eventuais erros grosseiros presentes, produzindo estimativas menos tendenciosas. No entanto, o uso de estimadores robustos não garante estimativas eficientes e não tendenciosas, uma vez que ainda haverá dependência com a distribuição dos erros de medição, de acordo com a Teoria de Máxima Verossimilhança (BARD, 1974; SCHWAAB e PINTO, 2007). Neste contexto, uma abordagem de reconciliação de dados robusta generalizada foi proposta (WANG e ROMAGNOLI, 2003).

Mesmo com alternativas avançadas, o estimador predominante entre aplicações industriais é o WLS, cujos pesos são determinados tanto pelo método de máxima verossimilhança quanto pelo usuário, de acordo com alguma regra (heurística), seguindo a abordagem de regressão. No primeiro, a teoria estatística para a DR exige o conhecimento da matriz de covariância dos erros \mathbf{V}_z , que depende de vários aspectos, incluindo características do sensor, amplificador, cabeamento e circuitos de digitalização. Geralmente, as variâncias σ_i^2 não são conhecidas, mas podem ser estimadas com dados históricos assumindo $\mathbf{V}_z = \text{diag} [\sigma_1^2, \sigma_2^2, \dots, \sigma_{n_{ZM}}^2]$.

Poucos casos não seguem o método de máxima verossimilhança ao definir o estimador. Na etapa de reconciliação do pacote de reconciliação e otimização *on-line* (*On-line Reconciliation and Optimization*, ORO) descrito por BUSSANI *et al.* (1995), a DR é realizada minimizando a soma ponderada do quadrado dos resíduos entre valores medidos e calculados, normalizada pelos valores medidos. Além disso, algumas aplicações formulam o problema de DRPE, considerando os parâmetros juntamente com variáveis medidas na função objetivo. Em uma aplicação para um sistema de separação de gás de uma refinaria de petróleo, PICCOLO *et al.* (1996) incluíram na função objetivo os parâmetros de eficiência de duas colunas separadoras e a carga térmica de um refeedor, que foram determinados no mesmo problema de DRPE como parâmetros estimados. Da mesma forma, CÂMARA *et al.* (2016) relataram a presença de um parâmetro de caracterização de carga na função objetivo para a DRPE de uma unidade de destilação de óleo cru.

5.2.2 Métodos de Solução

Diferentes métodos da solução foram empregados em aplicações industriais de DR. Os métodos dependem do tipo de formulação usada na modelagem do problema e muitas vezes exploram algumas características particulares, a fim de alcançar cálcu-

los mais rápidos e eficientes.

Quando o problema de programação matemática resultante é uma estimação de mínimos quadrados ponderados com restrições de lineares ou bilineares, soluções analíticas e procedimentos iterativos específicos foram usados (BHAT e SARAF, 2004; HOLLY *et al.*, 1989; SÁNCHEZ e ROMAGNOLI, 1996; ZHANG *et al.*, 2001a). No entanto, à medida que mais relações, além das equações gerais de balanço de massa, são necessárias para modelos rigorosos, bem como limites das variáveis e restrições de viabilidade, o uso de tais metodologias é limitado. Além disso, o desenvolvimento de soluções analíticas para problemas lineares e bilineares no contexto de aplicações reais também pode ser motivado por limitações computacionais e de programação, o que mudou notavelmente com a maior disponibilidade e custo reduzido de computadores de alto desempenho.

Na ausência de restrições de desigualdade $\mathbf{g}(\cdot)$, a solução de problemas não lineares de DR foi obtida pelo uso do método de multiplicadores de Lagrange (EKSTEEN *et al.*, 2002; HU e SHAO, 2006; MEYER *et al.*, 1993), em que as derivadas parciais da função de Lagrange do problema são igualadas a zero e o sistema de equações resultante é resolvido por um algoritmo de resolução simultânea ou pelo método de linearização sucessiva (CHRISTIANSEN *et al.*, 1997; HU e SHAO, 2006; ISLAM *et al.*, 1994; WEISS *et al.*, 1996). Nesse último, aproximações lineares das restrições são obtidas por uma expansão em série Taylor em torno da iteração anterior e uma série de problemas de reconciliação de dados lineares é resolvida, até que se obtenha uma solução que satisfaça as restrições não lineares. Cada solução intermediária é o ponto ótimo para as restrições lineares. A linearização sucessiva apresenta certa simplicidade e cálculos rápidos, embora não seja capaz de lidar com limites nas variáveis e possa apresentar falhas de convergência (ISLAM *et al.*, 1994; ROMAGNOLI e SÁNCHEZ, 2000).

O método mais amplamente utilizado para resolver problemas de DR não-lineares, como o da Equação (5.1), consiste na aplicação de algoritmos genéricos e bem estabelecidos de NLP. Com efeito, vários trabalhos mostraram que o uso de NLP, em vez de linearização sucessiva, melhorou significativamente os resultados da reconciliação (BARBOSA JR *et al.*, 2000; LIEBMAN e EDGAR, 1988), de modo que as técnicas de programação não-linear emergiram como as mais eficientes. Esse método permite o uso de uma função objetivo geral e não linear, como estimadores robustos, e pode lidar explicitamente com restrições de desigualdade. As restrições de desigualdade, como limites das variáveis, são geralmente necessárias no modelo de DR para garantir estimativas viáveis, como estimativas não negativas para fluxos ou composições. Além disso, em conjunto com a DR estacionária, métodos que aplicam os algoritmos NLP são usados também na solução de problemas de otimização dinâmica, como os resultantes da DR dinâmica não linear (MCBRAYER *et al.*,

1998; SODERSTROM *et al.*, 2000). Assim, o método NLP é capaz de resolver uma formulação mais completa da reconciliação de dados.

ISLAM *et al.* (1994) testaram os três métodos em um pacote de reconciliação de dados para um reator de pirólise industrial. No entanto, a discussão é limitada à influência da estimativa inicial em cada método. Observou-se que os valores de estimativa inicial longe dos valores de estado estacionário fizeram com que o método de linearização sucessiva falhasse, convergindo para valores de estado estacionário incorretos. O método NLP baseado em programação quadrática sucessiva (SQP), por sua vez, exigiu um maior tempo computacional para obter um resultado comparável, mas funcionou bem em todos os casos com limites superiores e inferiores para todas as variáveis reconciliadas. O uso de SQP na ausência de limites nas variáveis não foi relatado. WEISS *et al.* (1996) também avaliou estes três métodos de resolução de reconciliação de dados em um reator industrial de pirólise. Os autores observaram resultados semelhantes entre os métodos, o que incentivou o uso da técnica de linearização sucessiva, já que o grande tempo computacional exigido pelo método NLP não pôde ser justificado. No entanto, resultados semelhantes entre os métodos não são observados em outros trabalhos. No trabalho de CHRISTIANSEN *et al.* (1997), aplicando DR à produção industrial do gás da síntese, a linearização sucessiva foi aplicada somente quando os balanços de massa foram considerados. Para o uso de um modelo mais rigoroso, incluindo o balanço de energia, bem como em estudos para investigar a planta ao longo de um período de tempo, o método de NLP foi acoplado a um algoritmo SQP com limites nas variáveis independentes.

Ambos os métodos de multiplicadores de Lagrange e NLP foram avaliados por EKSTEEN *et al.* (2002) na DR para um forno de fundição de arco aberto. Desvios sistemáticos foram observados nos resultados, sendo maiores quando o método de multiplicadores de Lagrange foi usado. Entretanto, esta comparação não foi necessariamente justa, pois os métodos não foram avaliados para uma mesma formulação do problema; o método de multiplicadores de Lagrange resolveu a minimização do resíduo das equações de balanço de massa ponderadas pela variância, enquanto o método de NLP resolveu a minimização da soma ponderada de erros quadráticos. Em outros trabalhos, os multiplicadores de Lagrange e o método de linearizações sucessivas foram empregados para resolver a DR não linear em uma planta de coqueamento, com base nos balanços de massa por componentes e no balanço de vazões (HU e SHAO, 2006). Argumentou-se que a linearização sucessiva tem a vantagem de ser relativamente simples e permitiu cálculos mais rápidos e que, em geral, o método de multiplicadores de Lagrange tendia a deslocar com maior frequência os valores reconciliados em uma direção.

Mesmo com a maior disponibilidade de computadores de alto desempenho, lidar com o grande número de variáveis de decisão continua a ser um desafio quando

se usam métodos NLP em aplicações industriais, bem como a natureza não linear e não convexa da função objetivo, do modelo do processo e das outras restrições. Observou-se que alguns algoritmos de NLP se tornam ineficientes quando o número de graus de liberdade é alto (LUCIA e XU, 1990; LUCIA *et al.*, 1993) e a dimensão do problema de DR para modelos de grande dimensão pode impedir que técnicas usuais de otimização resolvam corretamente o problema (FABER *et al.*, 2007). Como resultado, muitos autores propuseram abordagens de decomposição do problema, como uma estratégia de solução para superar essa dificuldade ao empregar algoritmos padrão de NLP, ou propuseram algoritmos de otimização adaptados para problemas com muitos graus de liberdade.

5.2.3 Estratégias de Solução

A estratégia mais simples para resolver o problema de DR não-linear é usar algoritmos de programação não-linear para estimar os parâmetros e as variáveis reconciliadas simultaneamente. No entanto, quando a dimensão do problema prejudica significativamente o uso de algoritmos padrão de NLP, abordagens de decomposição podem ser usadas, dividindo o problema original em vários subproblemas de dimensão reduzida.

ROD e HANČIL (1980) propuseram um algoritmo iterativo que determina parâmetros e variáveis independentes separadamente. REILLY e PATINO-LEAL (1981) usaram um método no qual a reconciliação de dados é um subproblema da etapa de estimação de parâmetros. DOVÍ e PALADINO (1989) apresentaram um algoritmo de variação restrito, em que as variáveis dependentes são eliminadas pela resolução das equações do modelo, eliminando assim a necessidade de resolver o problema para ambos os parâmetros e valores reais das variáveis medidas, mas requerendo estimativas para a derivada segunda do modelo. KIM *et al.* (1990) compararam três estratégias para o problema EVM não linear: EVM de um estágio para reconciliação de dados com estimação simultânea de parâmetros, EVM de dois estágios e EVM aninhado. Os critérios de comparação adotados foram custo computacional e valores após convergência para problemas simulados da literatura. Ao avaliar essas estratégias, foram usados os algoritmos de gradiente generalizado reduzido (GRG) e SQP. Os resultados indicaram que a estratégia em dois estágios superou as outras alternativas, permanecendo como a única estratégia recomendada para solução.

Considerando aplicações industriais de DR, a estratégia predominante é a solução direta de apenas um problema de DR não linear, em que os parâmetros são estimados juntamente com variáveis reconciliadas (DRPE/EVM) quando presentes. QUELHAS *et al.* (2013), no entanto, chamou atenção para o fato de que o problema de estimação deve ser cuidadosamente concebido. Lidando com um RTO em duas

etapas e aplicado a uma planta industrial de produção de etileno, mostrou-se que as trajetórias típicas dos coeficientes de transferência de calor estimados indicam que alguns dos parâmetros do modelo não são estimáveis quando a estimação de parâmetros é executada de forma independente e simultânea com o procedimento de reconciliação de dados. Isso pode ser ocasionado pela correlação com outros parâmetros e dados de processo ou porque os coeficientes de transferência de calor não afetam a função objetivo significativamente.

Ao estabelecer que a complexidade dos problemas de otimização de elevada dimensão pode ser reduzida por meio da partição do critério a ser otimizado, HODOUIN e EVERELL (1980) desenvolveram uma estratégia na qual o problema original é particionado em subsistemas que não são funções das mesmas variáveis. A partição é realizada de tal modo que os subsistemas são mais fáceis de resolver.

Na ferramenta de otimização *on-line* descrita por ZHANG *et al.* (1995b), a reconciliação de dados com detecção simultânea de erros grosseiros e a estimação de parâmetros foram obtidas separadamente em um NLP dedicado para cada função. No entanto, nenhum argumento foi fornecido para justificar esta decisão. A ferramenta foi criada usando um programa baseado em fluxogramas (Aspen Plus), com recursos de otimização e subrotinas definidas pelo usuário para otimização de processos e estimação de parâmetros. A reconciliação de dados e a detecção de erros grosseiros foram conduzidas com o estimador da distribuição Normal Contaminada (TJOA e BIEGLER, 1991) implementada em GAMS/MINOS.

WEISS *et al.* (1996) utilizou um EVM de dois estágios não linear, separando a estimação de parâmetros em um laço externo e a reconciliação de dados em um laço interno, usando o método descrito por VALKÓ e VAJDA (1987). Além disso, a linearização sucessiva foi usada para resolver o laço interno da reconciliação de dados.

LEE *et al.* (1998) propuseram a reconciliação de dados e otimização *on-line* para uma planta industrial de utilidades com base na abordagem de decomposição hierárquica, semelhante à estratégia aplicada por HODOUIN e EVERELL (1980). A estratégia decompõe um sistema grande e complexo em vários subsistemas, admitindo-se que a manipulação de subsistemas pequenos no lugar do problema original fornece redução do custo computacional, soluções mais robustas, simplicidade na formulação do problema e fácil manutenção. Em um camada superior, variáveis relacionadas a restrições não lineares ou compartilhadas entre dois ou mais subsistemas são determinadas por algoritmos NLP. Quando os valores dessas variáveis são determinados na camada superior, tornam-se parâmetros pré-especificados dos subsistemas da camada mais baixa. Entre outros resultados, LEE *et al.* (1998) mostraram que a estratégia de decomposição alcançou um valor de função objetivo 37% menor que o obtido com SQP, gastando apenas 5,5% do tempo de computação do SQP, ao

considerar resultados médios de reconciliação de dados para 20 conjuntos de dados para o subsistema de distribuição de vapor.

FABER *et al.* (2003) propuseram uma estrutura aninhada em três estágios, em que o estágio superior é um NLP para estimação de parâmetros, o estágio intermediário consiste em vários NLPs, onde as variáveis independentes de cada conjunto de dados individuais são estimadas, e no estágio inferior as variáveis dependentes são avaliadas por meio de uma etapa de simulação. Argumentaram que são necessários apenas os gradientes das variáveis dependentes em relação aos parâmetros nessa estratégia. Ao usar a condição de otimalidade do estágio intermediário, argumentaram que as derivadas de segunda ordem não são necessárias no estágio superior, reduzindo significativamente o custo computacional. Esta estratégia foi posteriormente validada com dados de uma planta piloto de purificação de gás de coqueria (FABER *et al.*, 2004, 2006, 2007).

5.2.4 Algoritmos de Otimização para Programação Não Linear

Vários algoritmos de NLP têm sido usados para resolver o problema de otimização definido pela Equação (5.1), quer seja em aplicações industriais, quer seja em estudos de DR envolvendo dados reais de plantas. Estes incluem métodos determinísticos, baseados em gradiente, como Gauss-Newton ou Gauss-Marquardt (*e.g.*, (EMBIRUÇU *et al.*, 2008; PIERUCCI *et al.*, 1996; PRATA *et al.*, 2008b; VIEIRA *et al.*, 2003), métodos de gradiente reduzido (CHEN *et al.*, 1998b; EKSTEEN *et al.*, 2002; ÖZYURT e PIKE, 2004; PICCOLO *et al.*, 1996; ZHANG *et al.*, 1995a,b), e programação quadrática sucessiva, que é o mais utilizado entre os trabalhos industriais de DR, bem como algoritmos metaheurísticos de busca aleatória, sem informação de gradientes, como a otimização por enxame de partículas (PSO) (PRATA *et al.*, 2009b, 2010). Os algoritmos determinísticos aplicáveis a problemas de otimização não lineares em elevada dimensão estão disponíveis tanto na forma comercial quanto gratuita (BAGAJEWICZ e ROLLINS, 2011; CHEN *et al.*, 1996; LEYFFER e MAHAJAN, 2010; MORÉ e WRIGHT, 1993) e geralmente incorporam informações de derivada segunda do modelo de otimização, exploram a dispersão da matriz KKT e a estrutura do problema, e lidam eficientemente com grandes conjuntos de restrições ativas, considerando restrições dependentes e curvatura negativa.

Algoritmos de lagrangiano aumentado foram adaptados a problemas de grande dimensão, incluindo o uso de informações de segunda ordem e um método de região de confiança restrita para lidar com a curvatura negativa, como no caso de LANCELOT (CONN *et al.*, 1992). Métodos de gradiente reduzido, como MINOS (MURTAGH e SAUNDERS, 1998) e CONOPT (DRUD, 1994), aproximam informações

de segunda ordem usando atualizações densas do tipo quasi-Newton no espaço reduzido. No entanto, o esforço computacional cresce polinomialmente com o tamanho do problema e a seleção do conjunto ativo constitui um processo combinatório que pode ser custoso, especialmente em problemas degenerados (POKU *et al.*, 2004). A técnica de resolver sucessivamente uma série de subproblemas de programação quadrática (*i.e.*, SQP) provou ser adequada para resolver problemas de otimização em grande dimensão. Ela envolve a definição de programação quadrática (QP) em cada iteração, com base em uma aproximação quadrática da função de Lagrange, e uma aproximação linear das restrições do programa não-linear original. Isso resulta em uma direção de busca e um tamanho de passo para a busca em linha, que determina a próxima iteração (GILL *et al.*, 2005). O algoritmo SQP tem sido amplamente utilizado na indústria de processos (ALKAYA *et al.*, 2001), mas a identificação do conjunto de restrições de desigualdade ativas pode impor sérias dificuldades durante a solução dos QPs na presença de muitas restrições de desigualdade, uma vez que elas podem aumentar exponencialmente com o aumento do tamanho do problema. Além disso, o grande número de graus de liberdade no problema de DR pode tornar a estratégia de decomposição SQP menos adequada.

Extensões desta técnica para lidar com muitos graus de liberdade foram desenvolvidas (BETTS e FRANK, 1994; LUCIA e XU, 1990; LUCIA *et al.*, 1993; SARGENT e DING, 2001; TJOA e BIEGLER, 1992), explorando a estrutura natural do problema e a decomposição do espaço dos problemas de programação quadrática em espaços densos reduzidos. Exemplos de algoritmos de barreira no espaço completo (ou ponto interior) são SQPIP (KYRIAKOPOULOU e KALITVENTZEFF, 1996) e IPOPT (POKU *et al.*, 2004). SQPIP é um algoritmo de ponto interior de espaço completo do SQP que envolve três etapas principais: o subproblema QP, um procedimento de busca em linha, e um método de aproximação da Hessiana da função de Lagrange, em que uma vantagem importante é a garantia de viabilidade das restrições de desigualdade. IPOPT, por sua vez, é um algoritmo de ponto interior primal-dual que incorpora informações de segunda ordem e uma estratégia de busca em linha que assegura a convergência do problema de barreira. Descrições completas e uma análise abrangente dos métodos de ponto interior podem ser encontradas na literatura (FORSGREN *et al.*, 2002).

Embora esses métodos tenham sido usados em aplicações industriais e existam *softwares* comerciais de otimização, sistemas com grau de não linearidade elevado não são fáceis de resolver e muitas vezes revelam problemas de convergência. Além disso, estes métodos são locais e não oferecem qualquer garantia de que o mínimo global no problema de otimização foi encontrado, uma vez que o problema de otimização não linear é geralmente não convexo. Algoritmos determinísticos de otimização global, que podem fornecer uma garantia de solução global dentro de determinado

domínio de busca, também foram propostos para lidar com a possível existência de múltiplos mínimos locais (ESPOSITO e FLOUDAS, 1998; GAU e STADTHERR, 2002). No entanto, até onde foi possível determinar, seu uso em aplicações industriais não foi relatado.

Algoritmos de otimização devem ser eficientes e robustos, sendo capazes de evitar mínimos locais e lidar com estimativas iniciais diversas, problemas mal dimensionados e (algumas) imprecisões nos gradientes, mantendo um tempo de computação viável (BIEGLER e CUTHRELL, 1985; LEE *et al.*, 1998). Com efeito, aplicações reais de DR mostram que fatores como a normalização de variáveis e equações, estimativas iniciais, padrões da estrutura esparsa e aproximações termodinâmicas são importantes ao resolver problemas de NLP (BAILEY *et al.*, 1993; LID e SKOGESTAD, 2008a). Como indicado por KELLY (2004), uma vez que as maneiras de determinar os incrementos das variáveis de decisão para a próxima iteração são diferentes entre os algoritmos de NLP, é interessante incluir vários algoritmos independentes para resolver problemas industriais, de modo que, se um método não convergir ou não encontrar uma solução razoável, outro método pode ser usado na esperança de encontrar uma solução melhor. Considerando que cada algoritmo tem seu próprio conjunto de vantagens e desvantagens e que não há um algoritmo de otimização com desempenho melhor que qualquer outro para qualquer classe de problemas (KÖPPEN, 2004; WOLPERT e MACREADY, 1997), poucos trabalhos apresentam comparações entre algoritmos de NLP em aplicações industriais. Na verdade, poucos trabalhos abordam detalhadamente as questões numéricas que surgem ao projetar e aplicar a reconciliação de dados em sistemas reais.

Uma exceção é o trabalho de PIERUCCI *et al.* (1996), que apresentou um teste entre algoritmos, sendo um SQP, uma modificação de Levenberg-Marquardt (LM) e outro baseado em Newton (NB). Os autores relataram que SQP e LM tiveram comportamentos muito semelhantes, enquanto que NB era mais lento e menos eficiente. Outra exceção é devida a KYRIAKOPOULOU e KALITVENTZEFF (1996), com base no *software* VALI (Belsim). Os algoritmos testados foram o SQPIP, uma formulação Lagrangiana resolvida por um método *dogleg* (SOLDOG), e um SQP baseado em matrizes densas (SQPHP). Menos avaliações da matriz jacobiana e dos resíduos foram exigidos pelos algoritmos SQPIP e SQPHP em relação ao SOLDOG, enquanto SOLDOG foi mais rápido que ambos pois requeria menos fatorizações totais que SQPIP e o sistema das condições de otimalidade era mais esparso que no SQPHP. SARABIA *et al.* (2012), por sua vez, aplicou um SQP (NAG Library – The Numerical Algorithms Group Ltd, Oxford, UK) e um método de ponto interior (IPOPT-GAMS) com o objetivo de comparar as abordagens de solução sequencial e simultânea. Os autores relataram apenas que o tempo de CPU para resolver o problema de reconciliação foi de até 10 min na abordagem sequencial com SQP e 1

min na abordagem simultânea com IPOPT.

Em muitas aplicações reais, o modelo do processo é um modelo “caixa preta”, disponível por meio de um pacote de simulação que não permite a aplicação de técnicas destinadas a melhorias numéricas, como a técnica proposta por LID e SKOGESTAD (2008b) e aplicada por LID e SKOGESTAD (2008a). Aplicações industriais de DR são tipicamente baseadas em plataformas de simulação para os quais algoritmos de otimização foram interfaciados e onde os próprios modelos são construídos e calculados por meio de procedimentos numéricos, em oposição às implementações de código aberto (*open-source*) feitas em linguagens declarativas abertas. Neste contexto, a maioria destas plataformas não fornece derivadas analíticas para o algoritmo de otimização, em que gradientes são frequentemente aproximados por meio de diferenças finitas (BIEGLER, 2014). Os resultados do cálculo, portanto, dependem da técnica utilizada na estimação dos gradientes, sendo que o tamanho ideal do passo usado na diferença finita pode depender da aplicação. No entanto, esse efeito raramente é mencionado, de modo que não é possível afirmar inequivocamente que os procedimentos de reconciliação de dados estão sendo usados com sucesso em ambientes industriais reais (EMBIRUÇU *et al.*, 2008; VIEIRA *et al.*, 2003).

Nas situações em que é inconveniente ou custoso aproximar as derivadas por diferenças finitas ou quando as técnicas de diferenciação automáticas não podem ser aplicadas, algoritmos de NLP convencionais podem não ser adequados. Nesses casos, métodos que não dependem da informação dos gradientes da função objetivo ou restrições devem ser considerados (CONN *et al.*, 2009). Neste contexto, os métodos metaheurísticos de otimização, como o recozimento simulado e o algoritmo genético, podem ser usados para superar as dificuldades de obtenção das derivadas. Esses métodos também apresentam algumas vantagens adicionais, incluindo a capacidade de evitar mínimos locais, apresentando, assim, um caráter global e simplicidade de implementação. Mesmo que esses métodos estocásticos ainda não forneçam qualquer garantia de que o ponto ótimo global seja encontrado, eles fornecem formas que facilitam sua determinação (GAU e STADTHERR, 2002). Além disso, os valores calculados da função objetivo podem ser usados para análises estatísticas rigorosas das regiões de confiança das estimativas, que também podem constituir um benefício importante desses algoritmos (SCHWAAB *et al.*, 2008). O algoritmo PSO, especificamente, pode ser usado efetivamente para solução de problemas de otimização complexos devido à capacidade de lidar com funções objetivo multimodais, não-lineares e descontínuas, juntamente com sistemas de elevada dimensão, bem como ao menor esforço computacional e implementação mais simples, quando comparado a outros procedimentos metaheurísticos, tais como o recozimento simulado e o algoritmo genético.

PRATA *et al.* (2009b) apresentaram um procedimento para a DRPE dinâmica e

não linear e sua aplicação a um reator industrial de polipropileno, em que a técnica foi desenvolvida e validada em tempo real. O desempenho dos algoritmos PSO e Gauss-Newton (GN) foram comparados na resolução do DRPE dinâmico não linear *on-line*. Embora os resultados tenham mostrado que o método PSO era mais lento que o GN, o que não excluía o uso do PSO *on-line*, observou-se que GN levava a piores valores finais da função objetivo para os casos em que os algoritmos apresentaram grandes diferenças entre si. Isto foi atribuído à convergência a pontos de mínimo local. Além disso, foi mostrado que algoritmos de otimização não determinísticos podem ser usados com êxito em aplicações em tempo real, pois eles podem fornecer informações confiáveis sem qualquer frequência de falha significativa. Com base no trabalho anterior, PRATA *et al.* (2010) apresentaram um procedimento robusto para executar o problema de DRPE dinâmico não linear *on-line* com detecção simultânea de erros grosseiros. Usando estimadores robustos, o trabalho foi o primeiro a aplicar a técnica PSO para a solução simultânea do problema DRPE e detecção de erros grosseiros em um sistema industrial real.

5.2.4.1 Abordagem de Solução

Métodos que aplicam algoritmos de NLP podem ser separados em dois grupos: estratégias sequenciais e simultâneas. A diferença entre elas depende principalmente da forma como as restrições do modelo são consideradas. A abordagem sequencial segue uma estratégia iterativa de dois estágios, que resolve o sistema de equações como um laço interno para avaliar a função objetivo, enquanto um algoritmo de otimização constitui o laço externo. Como a solução e convergência do modelo é obtida em cada etapa de otimização, essa abordagem pertence à classe de métodos conhecida como “caminho viável”. A abordagem simultânea, por outro lado, inclui o problema de convergência do modelo dentro do problema de otimização, pertencendo à classe de métodos de “caminho inviável”.

Essas classes são uma grande preocupação quando se pretende resolver problemas de otimização dinâmica. Por exemplo, a abordagem para resolver problemas de otimização dinâmica usando o método de colocação em elementos finitos, que converte o sistema de equações em um sistema algébrico e o resolve simultaneamente com o problema de otimização, segue a abordagem simultânea (BIEGLER, 1984). LIEBMAN *et al.* (1992) também usaram SQP com colocação em elementos finitos e uma abordagem de horizonte móvel para resolver o problema de otimização resultante. Mesmo que a aplicação desta abordagem NLP em aplicações reais seja um pouco limitada devido à modelagem do sistema com equações algébricas ou diferenciais explícitas, ela foi aplicada a sistemas industriais (MCBRAYER *et al.*, 1998; SODERSTROM *et al.*, 2000). SAFDARNEJAD *et al.* (2015) mostraram uma comparação entre várias estratégias de inicialização para otimização de sistemas

dinâmicos, usando colocação em elementos finitos. Estas estratégias incluem a inicialização a partir de uma solução prévia obtida pela decomposição estrutural, uma abordagem simultânea e uma sequencial, que foram demonstradas para um sistema industrial de remoção de CO₂ integrado com unidades de geração de energia de carvão, gás, e eólica, resultando em um grande sistema com 768 variáveis de decisão. Comparações dos tempos computacionais entre os algoritmos IPOPT, APOPT (HEDENGREN *et al.*, 2012) e SNOPT (GILL *et al.*, 2005) também foram feitas. O algoritmo APOPT também foi usado no estudo de SAFDARNEJAD *et al.* (2016) para resolver a reconciliação dinâmica de dados e otimização de colunas de destilação em escala laboratorial, em que a colocação em elementos finitos foi usada para converter o sistema de equações para um sistema algébrico. Mais detalhes sobre abordagens de solução para sistemas dinâmicos podem ser encontradas na literatura (BIEGLER e GROSSMANN, 2004; BIEGLER *et al.*, 2002; BIEGLER, 2017).

A escolha entre as duas classes de abordagens pode depender tanto do tamanho do problema quanto da dificuldade para resolver o problema de convergência. Uma regra aproximada sugerida por PIERUCCI *et al.* (1996) indica quando um caminho viável é preferível, estando associada ao número de iterações necessárias para resolver o modelo estático do processo, incluindo também as iterações necessárias para resolver o modelo durante a fase de avaliação dos gradientes. BUSSANI *et al.* (1995) argumentaram que a escolha entre métodos de caminho viável e inviável depende tanto da proporção entre o número de variáveis de decisão e o número de variáveis iterativas, quanto da velocidade global de convergência.

No trabalho de EKSTEEN *et al.* (2002), tanto o método simultâneo (usando o método de multiplicadores de Lagrange) quanto o método sequencial (usando o algoritmo de gradiente reduzido generalizado - GRG) foram avaliados. Desvios sistemáticos foram identificados por meio de métodos estatísticos tradicionais e, em geral, pareciam maiores quando a abordagem simultânea foi utilizada. FABER *et al.* (2003) sugeriram a solução da estimação de parâmetros e do problema de reconciliação de dados para modelos não lineares de grande dimensão com uma abordagem sequencial, que foi aplicada à DR de uma planta piloto de purificação de gás de coqueria (FABER *et al.*, 2004, 2006, 2007). MARTÍNEZ-MARADIAGA *et al.* (2013) aplicaram um método similar, embora considerando apenas o problema de reconciliação de dados, para um refrigerador de absorção amônia-água de único efeito.

Em um trabalho envolvendo redes de hidrogênio em uma refinaria de petróleo, SARABIA *et al.* (2012) implementaram abordagens simultâneas e sequenciais. O primeiro foi baseado no algoritmo de ponto interior IPOPT-GAMS, em que alguns aprimoramentos foram feitos para evitar problemas de viabilidade, enquanto o último consistia de um modelo simulado no *software* EcosimPro ligado a um algo-

ritmo SQP-NAG. Como anteriormente discutido, foi relatado que o tempo de CPU necessário para resolver o problema de reconciliação foi de 1 min na abordagem simultânea com GAMS e até 10 min na abordagem sequencial com a biblioteca NAG. Infelizmente, nenhum comentário adicional foi fornecido.

Como pode ser observado em estudos anteriores, não há uma avaliação detalhada das abordagens de solução nem um consenso sobre o que é melhor para uma determinada condição.

5.3 Avaliação Numérica de Diferentes Configurações para a Resolução da Reconciliação de Dados Não Linear em Estado Estacionário

O sistema utilizado no presente estudo é uma planta de produção de petróleo, que extrai óleo e gás do reservatório para a refinaria, baseada em uma unidade real. O tratamento primário é composto por dois trens de produção e um trem de teste, como representado na Figura 5.2. Cada trem é equipado com um *manifold*, um trocador de calor e um separador trifásico. As correntes de saída de óleo em cada separador trifásico são unidas para alimentar um separador atmosférico. Em uma operação real, os trens A e B recebem a mistura resultante dos poços de produção, enquanto o trem de teste, que é responsável pela aquisição de um conjunto de medições de processo, recebe a corrente de um único poço de cada vez. No separador trifásico, a maior parte do gás é separada do óleo e da água, enquanto o separador atmosférico finaliza a purificação, separando gás residual, óleo e água na maior medida possível.

O sistema foi modelado usando um simulador de processo desenvolvido pela Petrobras e validado ao longo de anos de operação de produção (LIPORACE *et al.*, 2009; NIEDERBERGER *et al.*, 2005, 2009). Além disso, o modelo incorpora o modelo para trocadores de calor do *software* HTRI Xchanger (Heat Transfer Research, Inc., Navasota, TX, EUA). O modelo do sistema tem um total de 24 variáveis de processo, todas elas medidas, juntamente com cinco parâmetros, como mostrado na Tabela 5.2, em que os tipos 1 e 2 representam variáveis de entrada reconciliadas e parâmetros estimados, respectivamente, e o tipo 3 indica variáveis de saída medidas.

Um estado estacionário simulado que representa a média diária de um período típico de operação do sistema real forneceu o conjunto de valores verdadeiros (\mathbf{z}^*) e foi usado para gerar o conjunto de valores medidos (\mathbf{z}_{ms}). A simulação foi realizada com base em valores médios diários para melhor representar uma implementação no sistema real, em que é desejável eliminar o efeito das variações nos poços e minimizar o efeito dos inventários residuais. Neste contexto, a reconciliação de dados é realizada para um dia completo de operação, independentemente da frequência com

que os dados estão disponíveis. As amostras medidas foram obtidas por meio de um gerador de números aleatórios de distribuição normal, admitindo um desvio padrão específico para cada variável e uma média igual ao valor verdadeiro. O coeficiente de variação (CV, ou também desvio padrão relativo) das variáveis foi selecionado para simular a precisão estimada do instrumento que fornece a medida de cada variável, como apresentado na Tabela 5.2. Foram criados cinco conjuntos de valores medidos para cada variável, representando diferentes amostras do mesmo estado estacionário. A distribuição acumulada de probabilidades da variabilidade em torno dos valores verdadeiros de cada conjunto de valores medidos é apresentada na Figura 5.3.

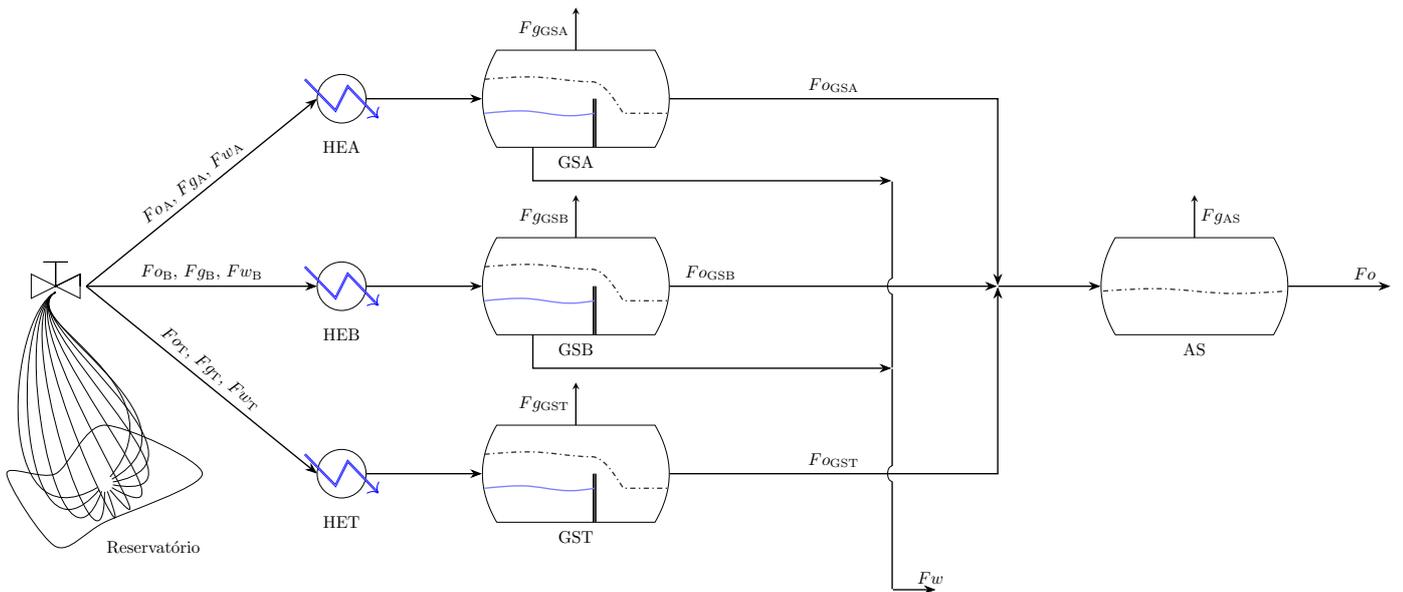


Figura 5.2: Diagrama simplificado da planta de produção de óleo analisada.

5.3.1 Configurações Testadas

Com base nos trabalhos que relatam a aplicação industrial de DR, foi estabelecida uma série de experimentos de simulação de reconciliação de dados em estado estacionário, adotando diferentes configurações. Cada configuração abrange abordagem de solução do problema de DR, função objetivo (estimador), algoritmo de otimização, técnica de aproximação de gradientes e estimativa inicial, como indicado na Tabela 5.3. O procedimento para teste das configurações consistiu em resolver o problema de reconciliação de dados para cada um dos cinco conjuntos de dados amostrados e analisar os resultados com métricas de desempenho. O processo de reconciliação de dados foi implementado em MATLAB (2010), onde o *script* gerencia as chamadas do algoritmo de otimização e do modelo de processo. O modelo estava disponível como uma biblioteca de vínculo dinâmico (DLL) e, conseqüentemente, não foi possível acessar o código fonte do modelo, inspecionar as equações

implementadas, editar ou implementar qualquer tipo de estratégia para melhorar suas propriedades numéricas, como adimensionalização das variáveis e equações. Os cálculos foram realizados em um sistema MS Windows 7 64 bit com CPU Intel Core i7 2,0 GHz e 6 GB de RAM.

Tabela 5.2: Informações sobre variáveis e parâmetros do modelo.

Variável	Tipo	Unidade	CV (%)
Corrente de entrada de óleo do HEA	1	m ³ /dia	8
Corrente de entrada de gás do HEA	1	m ³ /dia	8
Corrente de entrada de água do HEA	1	m ³ /dia	8
Corrente de entrada de óleo do HEB	1	m ³ /dia	8
Corrente de entrada de gás do HEB	1	m ³ /dia	8
Corrente de entrada de água do HEB	1	m ³ /dia	8
Corrente de entrada de óleo do HET	1	m ³ /dia	8
Corrente de entrada de gás do HET	1	m ³ /dia	8
Corrente de entrada de água do HET	1	m ³ /dia	8
Corrente de alimentação de água do HEA	1	m ³ /dia	8
Corrente de alimentação de água do HEB	1	m ³ /dia	8
Corrente de alimentação de água do HET	1	m ³ /dia	8
Coefficiente global de troca térmica do HEA	2	kcal m ⁻² h ⁻¹ °C ⁻¹	8
Coefficiente global de troca térmica do HEB	2	kcal m ⁻² h ⁻¹ °C ⁻¹	8
Coefficiente global de troca térmica do HET	2	kcal m ⁻² h ⁻¹ °C ⁻¹	8
Coefficiente de separação do SGA	2	%	8
Coefficiente de separação do SGB	2	%	8
Corrente de saída de gás do SGA	3	m ³ /dia	0,2
Corrente de saída de gás do SGB	3	m ³ /dia	0,2
Corrente de saída de gás do SGT	3	m ³ /dia	0,2
Corrente de saída de gás do AS	3	m ³ /dia	8
Corrente de água descartada	3	m ³ /dia	2
Corrente de óleo exportado	3	m ³ /dia	0,2
Temperatura da corrente de saída de água do HEA	3	°C	8
Temperatura da corrente de saída de água do HEB	3	°C	8
Temperatura da corrente de saída de água do HET	3	°C	8
Temperatura da corrente de saída de óleo do HEA	3	°C	0,4
Temperatura da corrente de saída de óleo do HEB	3	°C	0,4
Temperatura da corrente de saída de óleo do HET	3	°C	0,4

Tabela 5.3: Configurações dos experimentos de simulação.

Configuração	Abordagem de Solução	Diferença Finita	Estimativa inicial	$\theta \in J_{DRPE}$	Graus de Liberdade	Algoritmo de Otimização
1	simultânea	central	medida	sim	12	IP
2a	sequencial	central	medida	sim	12	IP
2b	sequencial	central	valor verdadeiro	sim	12	IP
2c	sequencial	central	valor verdadeiro + 10%	sim	12	IP
2d	sequencial	central	valor verdadeiro - 10%	sim	12	IP
2e	sequencial	à frente	medida	sim	12	IP
2f	sequencial	à frente	valor verdadeiro	sim	12	IP
2g	sequencial	à frente	valor verdadeiro + 10%	sim	12	IP
2h	sequencial	à frente	valor verdadeiro - 10%	sim	12	IP
3	sequencial	central	medida	não	7	IP
4	sequencial	central	medida $\pm 3\sigma$	não	7	PSO/IP

Primeiro, a reconciliação de dados foi resolvida com a rotina `fmincon` do MATLAB, usando o algoritmo IP. A rotina `fmincon` destina-se à resolução de problemas multivariáveis não lineares com restrições e está disponível no pacote MATLAB Optimization Toolbox. Em relação à abordagem de solução, o problema de programação matemática foi resolvido por ambas as abordagens sequencial e simultânea. Na primeira, a simulação de processo e a minimização da função objetivo seguiu uma abordagem de dois estágios, em que o primeiro estágio de simulação calculava as variáveis de estado, garantindo que as restrições do modelo seriam sempre satisfeitas. Ao fazer isso, apenas a função objetivo e as restrições de desigualdade foram manipuladas pelo algoritmo de otimização. Na última, o algoritmo de otimização manipulou a função objetivo e as restrições, incluindo-se as restrições de igualdade do modelo de processo. Na prática, a abordagem simultânea foi conseguida passando as restrições de igualdade do modelo de processo como argumento definido pelo parâmetro `nonlcon` da rotina `fmincon`.

Em geral, os valores medidos são usados como estimativas iniciais para as variáveis medidas do processo (ROMAGNOLI e SÁNCHEZ, 2000). Apesar disso, dada a influência da estimativa inicial no desempenho de técnicas de otimização, quatro diferentes estimativas iniciais foram usadas para testar o desempenho geral do procedimento de DR.

Além disso, a rotina `fmincon` admite uma opção para definir o tipo de diferença finita usada na aproximação numérica dos gradientes, podendo ser diferenças à frente (padrão) ou diferenças centrais. Ambos os tipos foram testados ao executar os experimentos DR.

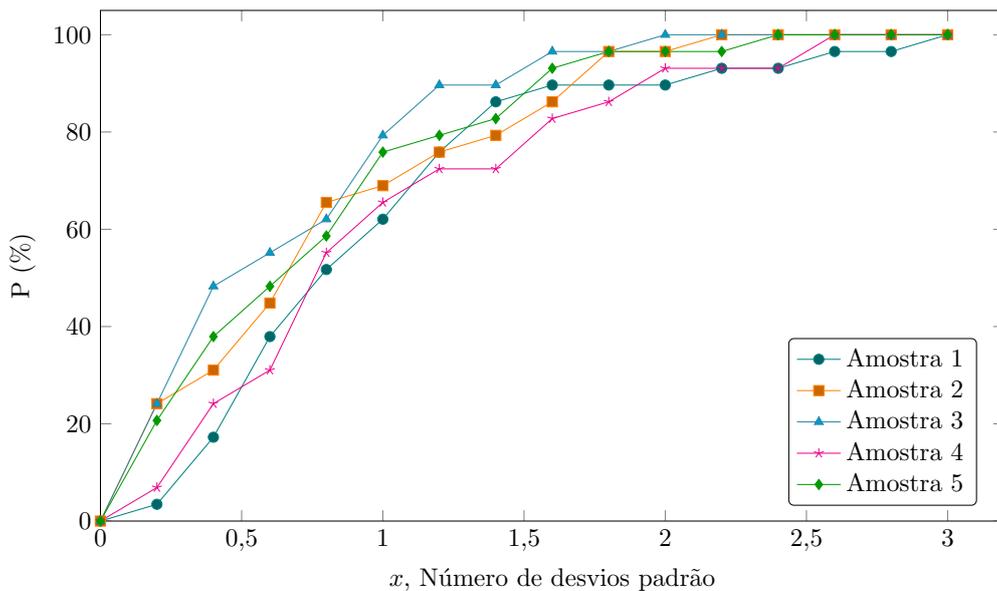


Figura 5.3: Função de distribuição cumulativa da variabilidade em torno dos valores verdadeiros para cada conjunto de amostras. $P = P \left[\frac{\text{abs}(z_{ms} - z_m^*)}{\sigma} \leq x \right]$.

Deve-se enfatizar que em todos os experimentos simulados a condição de saída da função `fmincon` foi o estado em que alterações nas variáveis de decisão e a violação máxima das restrições foram inferiores às tolerâncias definidas, a não ser que seja mencionado o contrário. Simulações da reconciliação de dados usaram os valores padrões das opções da rotina `fmincon`, exceto para aquelas listadas na Tabela 5.4. Simulações anteriores usando valores padrões para todas as opções apresentaram diferenças significativas em relação à convergência e motivaram essas alterações.

Um algoritmo híbrido de otimização foi então usado nas configurações ajustadas, combinando a técnica determinística de IP com o algoritmo metaheurístico PSO. Eles foram combinados de tal forma que o procedimento iniciado com o PSO para algumas iterações era concluído com a técnica IP, sendo a solução fornecida pelo PSO transmitida como estimativa inicial do IP. Ao fazer isso, pretendia-se melhorar a robustez da solução com base no caráter global de busca do PSO e, posteriormente, melhorar a velocidade de convergência com o IP. O algoritmo PSO usado, originalmente proposto por EBERHART e KENNEDY (1995) e adaptado por SCHWAAB *et al.* (2008) e PRATA *et al.* (2009b) para a estimação de parâmetros e reconciliação de dados, respectivamente, segue as Equações (4.26)-(4.27). Os parâmetros para o algoritmo PSO estão listados na Tabela 5.5 e foram baseados nos valores usados por PRATA *et al.* (2009b).

Tabela 5.4: Valores modificados das opções usadas na rotina de otimização `fmincon` nos experimentos de DR.

Parâmetro	Descrição	Valor
<code>TolFun</code>	Tolerância final no valor da função obj.	1×10^{-4}
<code>TolX</code>	Tolerância final das variáveis de decisão	1×10^{-9}
<code>MaxIter</code>	Número máximo de iterações	5000
<code>MaxFunEvals</code>	Número máximo de avaliações da função obj.	10.000

O problema da retificação de dados foi posto como um problema de máxima verossimilhança, em que a probabilidade do estado estimado da planta era maximizada para o conjunto de medidas. Neste contexto, o estimador apropriado, considerando média nula e erros de medição independentes e normalmente distribuídos, é o WLS. O problema da reconciliação de dados com estimação simultânea de parâmetros

Tabela 5.5: Valores dos parâmetros do algoritmo PSO.

Parâmetro	Valor
Número de gerações	60
Número de partículas	5
Parâmetro de busca local (c_1)	1
Parâmetro de busca global (c_2)	1
Peso inercial (ω)	0,7
Tolerância	$1,0 \cdot 10^{-3}$

(DRPE) foi então declarado na forma da Equação (5.1):

$$\begin{aligned}
 \{\hat{\mathbf{z}}, \hat{\boldsymbol{\theta}}\} := & \arg \min_{\mathbf{z}_m, \boldsymbol{\theta}} J_{\text{DRPE}} = (\mathbf{z}_{ms} - \mathbf{z}_m)^\top \mathbf{V}_z^{-1} (\mathbf{z}_{ms} - \mathbf{z}_m) \\
 \text{sujeito a } & \mathbf{f}(\mathbf{z}, \boldsymbol{\theta}) = \mathbf{0} \\
 & \mathbf{g}(\mathbf{z}, \boldsymbol{\theta}) \leq \mathbf{0} \\
 & \mathbf{z}_m^L \leq \mathbf{z}_m \leq \mathbf{z}_m^U \\
 & \boldsymbol{\theta}^L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}^U
 \end{aligned} \quad , \quad (5.2)$$

em que $(\mathbf{z}_m^L, \mathbf{z}_m^U)$ e $(\boldsymbol{\theta}^L, \boldsymbol{\theta}^U)$ são limites inferiores e superiores das variáveis reconciliadas e estimadas e dos parâmetros, respectivamente. Apesar desta formulação, o estimador WLS da Equação (5.2) foi modificado pela introdução dos parâmetros $\boldsymbol{\theta}$ (isto é, variáveis não medidas) na função objetivo, resultando em um estimador WLS modificado como:

$$J_{\text{DRPE}} = (\mathbf{z}_{ms} - \mathbf{z}_m)^\top \mathbf{V}_z^{-1} (\mathbf{z}_{ms} - \mathbf{z}_m) + (\boldsymbol{\theta}_{ms} - \boldsymbol{\theta})^\top \mathbf{V}_\theta^{-1} (\boldsymbol{\theta}_{ms} - \boldsymbol{\theta}) \quad , \quad (5.3)$$

em que \mathbf{V}_θ é a matriz de ponderação dos parâmetros e $\boldsymbol{\theta}_{ms}$ são valores de referência para os parâmetros estimados. Desse modo, ambos os estimadores WLS e WLS modificado foram analisados.

A qualidade dos resultados obtidos pelo procedimento DRPE usando diferentes configurações foi quantificada, analisando-se a robustez da solução, a violação das restrições na convergência e o custo computacional, *i.e.*, o tempo consumido na resolução do problema. A principal métrica de desempenho em relação à robustez da solução baseia-se na teoria estatística do método de máxima verossimilhança. Uma vez que as medidas são normalmente distribuídas em torno dos valores verdadeiros e nenhum erro grosseiro está presente, a função objetivo segue uma distribuição chi quadrado χ^2 no ponto mínimo. Com base nisso, propôs-se o uso do teste global (ou coletivo) (REILLY e PATINO-LEAL, 1981), não para detectar erros grosseiros, mas para indicar a solução satisfatória do problema de minimização. Neste teste, a magnitude da função objetivo é comparada com o limite de probabilidade da

distribuição chi quadrado para os graus de liberdade (DF) correspondentes a cada configuração e para um determinado nível de confiança ($1 - \alpha$). Assim, a hipótese H1 é verdadeira se $J_{DRPE} > \chi_{1-\alpha, DF}^2$, o que significa que a função objetivo não alcançou estatisticamente o valor mínimo (o que equivale à ocorrência de erro grosseiro). Caso contrário, a hipótese H0 é mantida, o que significa que o problema de otimização foi resolvido de forma satisfatória (o que equivale à ausência de erro grosseiro). O número de graus de liberdade foi calculado de acordo com

$$DF = n_{zm} - (n_u + n_\theta),$$

em que $\chi_{99\%, 12}^2 = 26.22$ e $\chi_{99\%, 7}^2 = 18.48$.

Mais uma vez, com base em erros de medição normalmente distribuídos, espera-se que os valores reconciliados estejam dentro do intervalo compreendido entre $\mathbf{z}_m^* \pm 3\sigma$ com um nível de confiança 99,7%. Portanto, a diferença absoluta entre os valores reconciliado e verdadeiro normalizada pelo desvio padrão da variável também foi usada como métrica de desempenho.

5.4 Resultados

5.4.1 Abordagem de Solução

De início, foi analisada a influência da abordagem de solução para resolver a reconciliação de dados e estimação simultânea de parâmetros para o sistema de produção de petróleo.

De acordo com as configurações 1 e 2a listadas na Tabela 5.3, o problema DRPE foi resolvido para cada um dos cinco conjuntos de medição. Os resultados são apresentados na Tabela 5.6 e mostram uma vantagem absoluta da abordagem sequencial sobre a abordagem simultânea para as condições apresentadas em termos do valor final da função objetivo e do custo computacional. Além disso, a abordagem simultânea apresentou violação das restrições, o que significa que algumas equações do modelo não foram satisfeitas. Esse é um resultado importante, dada a falta de consenso na literatura acerca da escolha entre as abordagens. Não obstante, é possível identificar alguma preferência pelo método simultâneo (ALBUQUERQUE e BIEGLER, 1995), principalmente no contexto dos problemas de reconciliação de dados dinâmica. Contudo, trabalhos anteriores realizados pelo grupo de pesquisa do PEQ/COPPE identificaram, para reconciliação de dados dinâmica, que a abordagem sequencial tem resultados melhores no âmbito dos problemas reais estudados (PRATA *et al.*, 2009b, 2010), corroborando os resultados obtidos no presente trabalho. Embora a violação de restrições seja realmente considerada em *softwares*

e trabalhos da literatura, nenhum dos estudos revisados de aplicação industrial de DR/DRPE discutiu a violação de restrições, especialmente no caso das restrições de igualdade do modelo do processo. Tal violação das restrições representa um conflito óbvio com os objetivos da reconciliação de dados e não deve ser aceitável.

Considerando os resultados obtidos, apenas a abordagem sequencial foi utilizada para a definição das configurações subseqüentes.

Tabela 5.6: Valores da função objetivo e custo computacional em função da abordagem de solução.

Amostra	J_{DRPE}			Tempo (min)		
	Simultânea (Config. 1)	Sequencial (Config. 2a)	Diferença Relativa (%)	Simultânea (Config. 1)	Sequencial (Config. 2a)	Diferença Relativa (%)
1	347	266	-23	54	9	-83
2	* 633	482	-24	14	9	-34
3	* 62	4	-93	63	29	-54
4	* 3352	192	-94	77	8	-89
5	990	448	-55	50	9	-81

* Nenhum ponto viável foi encontrado, conforme indicado pelo valor -2 da variável de saída `exitflag` da rotina `fmincon`.

5.4.2 Estimativa Inicial e Estimação de Gradientes

Na seqüência da abordagem de solução, foi analisada a influência da estimativa inicial e da parametrização do algoritmo, especificamente a estimação de gradientes. Mesmo sendo conhecido que tais fatores influenciam a solução de problemas de programação matemática, muito pouco é discutido quando se trata de reconciliação de dados. A Figura 5.4 mostra os valores da função objetivo (J_{DRPE}) obtidos para cada uma das cinco amostras de medição (réplicas), de acordo com as configurações 2a a 2h (conforme descrito na Tabela 5.3). Os eixos da figura foram ajustados para facilitar sua interpretação visual. Como estimativa inicial foram testados os conjuntos de valores medidos (\mathbf{z}_{ms}), valores verdadeiros (\mathbf{z}_m^*), e valores verdadeiros com ruído de $\pm 10\%$.

Em ambos os gráficos da Figura 5.4 (especialmente gráfico A), a função objetivo foi sempre maior que o valor de qui-quadrado do teste global, o que sugere a existência de um erro grosseiro nos dados. Este não é o caso, porém, quando o problema foi inicializado com valores verdadeiros. Neste contexto, propõe-se a utilização do teste global para detectar inconsistências na formulação e na solução do problema. Como os dados não contêm erros grosseiros, o fato de que três dos casos apresentam valores finais da função objetivo mais altos que o valor de qui-quadrado está relacionado a uma solução local do problema de otimização, que depende do algoritmo de otimização. Adicionalmente, dois outros fatores surgem: a estimativa inicial e a técnica de diferenças finitas aplicada para a estimativa dos gradientes.

Para diferenças finitas à frente, a Figura 5.4 mostra que diferentes estimativas iniciais resultaram em diferentes desempenhos numéricos do algoritmo de otimização, em que o uso dos valores verdadeiros como estimativa inicial forneceu o melhor resultado, como esperado. Além disso, o uso de valores verdadeiros +10% apresentou os piores resultados, refletido pelos altos valores de J_{DRPE} .

Conforme apontado por SCHLADT e HU (2007), os valores iniciais das variáveis medidas e não medidas são significativos para o desempenho do problema de

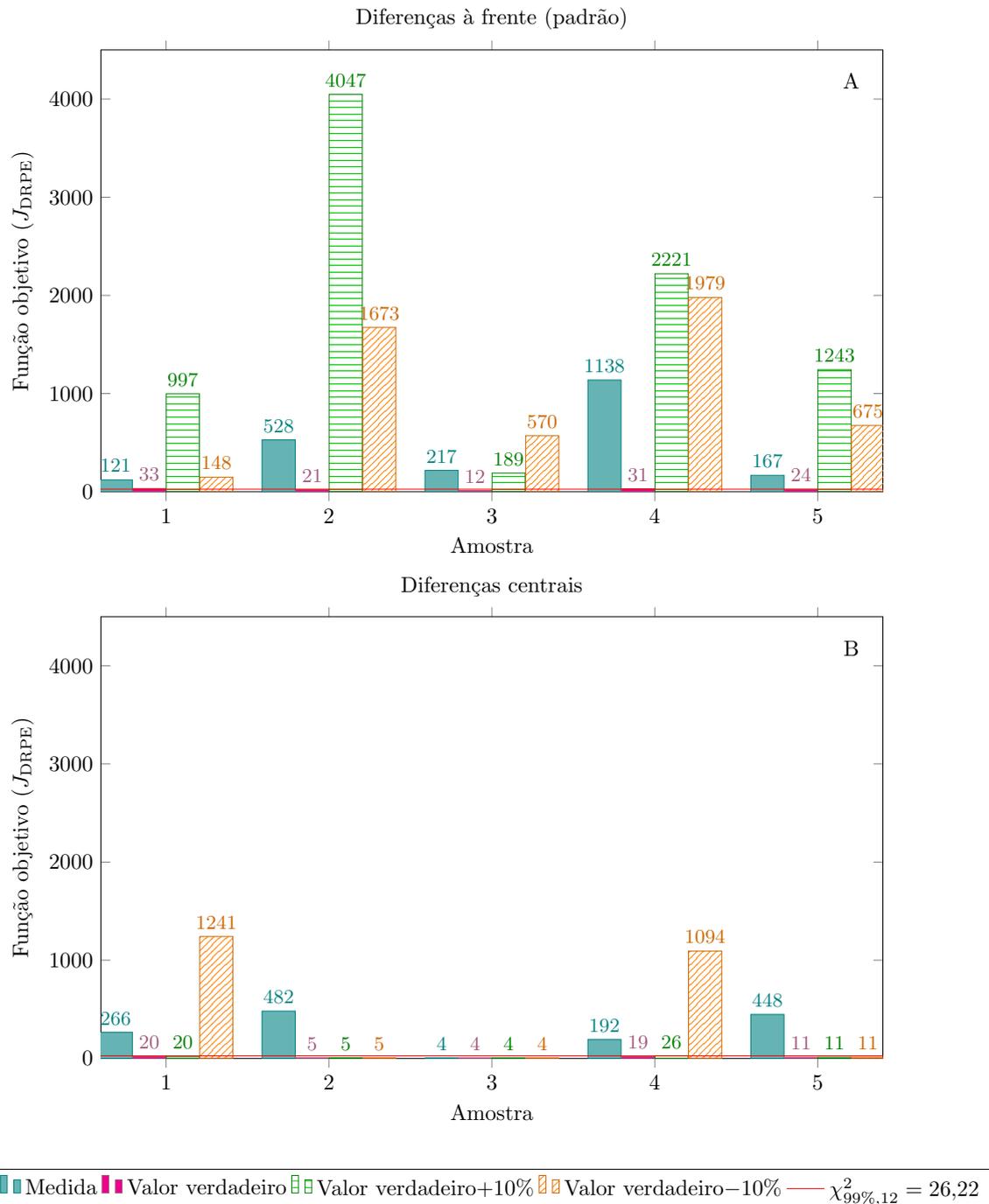


Figura 5.4: Valores finais da função objetivo em função da estimativa inicial. (A) Diferença finita à frente; (B) Diferença finita central.

otimização, em que a resolução será mais rápida e mais confiável à medida que as variáveis estiverem mais próximas de seus valores ótimos. Não obstante, dependendo do algoritmo de otimização, os autores também afirmaram que uma estimativa inicial muito próxima do valor ótimo pode não convergir para a solução, devido aos pequenos valores dos gradientes. Tais casos podem ser tratados, somando um ruído aleatório às estimativas iniciais. Um método alternativo foi proposto por KELLY (2004), em que os valores medidos foram usados como estimativas iniciais para as variáveis medidas, enquanto as estimativas iniciais das variáveis não medidas foram aproximadas iterativamente. Embora o uso de valores medidos tenha fornecido resultados razoáveis em relação às alternativas testadas, que consiste na prática mais comum entre as aplicações reais, tais resultados ainda refletem uma solução local do problema de otimização.

É interessante notar na Figura 5.4B, no entanto, que a simples mudança no tipo de diferenças finitas tenha melhorado dramaticamente os resultados para a maioria dos casos. Resultados detalhados dos valores finais da função objetivo são apresentados na Tabela 5.7. Como geralmente admite-se que a disponibilidade de boas estimativas iniciais seja suficiente para garantir um bom desempenho das técnicas de otimização (LEE *et al.*, 1998), tais resultados são surpreendentes, embora mostrem que isso nem sempre é verdade. De fato, esse tema tem recebido pouca atenção e o trabalho de PRATA *et al.* (2008b) é o único a mencionar algo sobre o tipo de diferenças finitas, sendo a diferença finita central usada.

DAVE *et al.* (2003) relataram que o SQP-NPSOL foi muito sensível ao valor da perturbação dada às variáveis para o cálculo do gradiente, já que a determinação precisa do gradiente exigia o uso de pequenas perturbações. Foi relatado que a existência de um compromisso, em que valores muito pequenos tornam o simulador insensível às mudanças e grandes valores da perturbação, por outro lado, aumentam a imprecisão dos gradientes. Determinou-se que o valor de $5,00 \times 10^{-3}$ vezes o valor das variáveis fornecia gradientes razoavelmente precisos e estáveis. No entanto, observações adicionais sobre o tipo de diferenças finitas não foram fornecidas.

Pode-se esperar que tal melhoria não fosse possível sem um aumento no custo computacional. Como pode ser visto na Tabela 5.8, não foi o caso para metade das estimativas iniciais testadas. Da mesma forma, a Tabela 5.9 mostra a melhora no custo computacional ao considerar as médias de diferentes amostras. Mesmo para aqueles casos que apresentaram maiores tempos de resolução, o custo computacional está longe de ser proibitivo.

Relembrando que os resultados são oriundos de um mesmo estado estacionário, fica evidente que a precisão numérica da aproximação dos gradientes pode não permitir um desempenho adequado do algoritmo IP na resolução do problema DRPE (em adição à estimativa inicial, à presença de erros grosseiros e erros de modela-

gem), fornecendo resultados fora de limites estatísticos aceitáveis. Nesses casos, os resultados podem ser surpreendentemente dependentes do esquema de diferenciação usado na estimação dos gradientes, especialmente quando as derivadas são calculadas numericamente, o que constitui um problema real e comum quando o modelo de processo não está acessível ao usuário, como no presente caso. Isso é percebido a partir dos resultados obtidos para ambos os valores de função objetivo e variáveis de decisão do problema de otimização. Resultados aceitáveis para todas as variáveis de decisão foram obtidos apenas para as configurações 2b, 2c e 2f.

Usando AspenPlus com um SQP, PICCOLO *et al.* (1996) relataram que os valores padrões do AspenPlus foram adequados para problemas simples, mas em casos complexos, em que o sistema é muito sensível ou altamente não-linear, alguns ajustes específicos eram necessários. Infelizmente, não foram fornecidos comentários adicionais sobre tais ajustes. A partir desta perspectiva, a precisão numérica da estimativa dos gradientes poderia ser um fator em potencial a ser investigado. No caso de uma solução DRPE insatisfatória (ou um erro grosseiro global detectado), o resultado do problema de otimização deve ser pelo menos marcado como não confiável. No estudo de SCHLADT e HU (2007), por exemplo, a abordagem aplicada foi de descartar o resultado quando detectava-se um erro grosseiro global. Outra abordagem menos severa seria repetir a resolução do problema DR(PE) até que fosse atingida uma solução satisfatória.

Os resultados apresentados não deixam dúvidas quanto à importância da estimativa inicial e da técnica utilizada para a estimação de gradientes em problemas de reconciliação de dados (e problemas de otimização em geral). Algoritmos de otimização devem ser eficientes e robustos, sendo capazes de lidar não apenas com estimativas iniciais “ruins”, mas também com variáveis de diferentes ordens de grandeza e algumas imprecisões nos gradientes (BIEGLER e CUTHRELL, 1985). Não obstante, a literatura apresenta pouca discussão sobre o tema e, até onde foi possível determinar, o trabalho de PRATA *et al.* (2008b) é o único a detalhar a estratégia de cálculo dos gradientes. Dessa maneira, estudos envolvendo a resolução de problemas de otimização devem não apenas considerar esses efeitos, mas relatar as estratégias usadas para a definição da estimativa inicial e para o cálculo dos gradientes. Ignorar esses efeitos não promove uma avaliação completa e pode, ao mesmo tempo, resultar em conclusões tendenciosas.

5.4.3 Formulação do Problema e Algoritmo de Otimização

Dada a utilização mais provável dos valores medidos como estimativa inicial em aplicações reais, testes de configurações subsequentes consideraram essa situação, a fim de melhorar a solução do problema DRPE. Nesta seção, os testes relativos a dois

Tabela 5.7: Valores finais da função objetivo após resolução do problema de DR em função da estimativa inicial e do tipo de diferenças finitas.

Configuração	2e		2a		2f		2b		2g		2c		2h		2d	
Estimativa inicial	Medida				Valor Verdadeiro				Valor Verdadeiro $\times 1.1$				Valor Verdadeiro $\times 0.9$			
Diferença Finita Amostra \	À Frente		Central		À Frente		Central		À Frente		Central		À Frente		Central	
1	121		266		33		20		997		20		148		1241	
2	528		482		21		5		4047		5		1673		5	
3	217		4		12		4		189		4		570		4	
4	1138		192		31		19		2221		26		1979		1094	
5	167		448		24		11		1243		11		675		11	
Média	434		279		24		12		1739		13		1009		471	
Diferença relativa (%)			-36				-50						-99			

Tabela 5.8: Tempo decorrido (min) para resolução do problema de DR em função da estimativa inicial e do tipo de diferenças finitas.

Configuração	2e		2a		2f		2b		2g		2c		2h		2d	
Estimativa inicial	Medida				Valor Verdadeiro				Valor Verdadeiro $\times 1.1$				Valor Verdadeiro $\times 0.9$			
Diferença Finita Amostra \	À Frente		Central		À Frente		Central		À Frente		Central		À Frente		Central	
1	33		9		6		22		13		24		29		8	
2	70		9		18		21		3		29		5		24	
3	25		29		31		20		26		31		5		23	
4	6		8		3		14		27		24		22		14	
5	60		9		2		25		12		27		33		26	
Média	39		13		12		21		16		27		19		19	
Diferença relativa (%)			-67				72						65			

Tabela 5.9: Resultados médios para valores da função objetivo e custo computacional de diferentes amostras de acordo com o tipo de diferenças finitas.

Amostra	J_{DRPE}			Tempo (min)		
	À Frente	Central	Diferença Relativa (%)	À Frente	Central	Diferença Relativa (%)
1	325	387	19	20	16	-22
2	1567	125	-92	24	21	-15
3	247	4	-98	22	26	19
4	1342	333	-75	14	15	6
5	527	120	-77	27	22	-18

estimadores diferentes são apresentados, uma vez que a formulação do problema é uma das possíveis fontes de resultados tendenciosos.

Resultados mostrados na Tabela 5.10 foram obtidos para as configurações 2a e 3, conforme descrito na Tabela 5.3. Em relação ao valor final da função objetivo,

os resultados obtidos com a modificação da formulação do problema indicam uma maior dificuldade do algoritmo de otimização em encontrar uma solução adequada. Apesar de apresentar menor custo computacional, as amostras 1, 2 e 5 apresentaram valores de função objetivo mais elevadas. No entanto, o desempenho da DRPE para a amostra 3 foi ainda melhor com a modificação do estimador. A Figura 5.3 mostra que a amostra 3 tem a menor variabilidade entre todas as amostras, o que poderia estar contribuindo para a obtenção de bons resultados.

Quando os valores para a função objetivo são analisados em conjunto com a métrica em relação à região de confiança dos valores reconciliados, os resultados ficam ainda mais interessantes. Embora tenha havido um aumento no valor final da função objetivo para a amostra 1, a Tabela 5.11 mostra que menos variáveis estimadas estiveram fora da região de confiança de 99,7%. Resultado semelhante foi observado para a amostra 2, em que um aumento de 116% no valor da função objetivo não afetou a qualidade dos valores reconciliados. As amostras 3, 4 e 5 apresentaram um comportamento regular, com exceção da diminuição de 11% no valor da função objetivo da amostra 4, responsável por uma melhoria pronunciada na qualidade dos valores reconciliados.

Com base nessa discussão, é evidente o efeito dos estimadores nos resultados da DRPE. A inclusão de variáveis não medidas entre variáveis da função objetivo induziu a ocorrência de desvios nas variáveis e parâmetros estimados. Contudo, o que não foi possível distinguir a partir dos resultados apresentados é a razão dos resultados obtidos para o estimador de máxima verossimilhança ainda apresentarem variáveis fora de seus respectivos limites de confiança. Para responder a esse questionamento, atenção foi dada ao algoritmo de otimização.

Tabela 5.10: Valores de função objetivo e custo computacional em função da formulação do problema.

Amostra	J_{DRPE}			Tempo (min)		
	$\theta \in J_{DRPE}$ (Config. 2a)	$\theta \notin J_{DRPE}$ (Config. 3)	Diferença Relativa (%)	$\theta \in J_{DRPE}$ (Config. 2a)	$\theta \notin J_{DRPE}$ (Config. 3)	Diferença Relativa (%)
1	266	485	82	9	6	-32
2	482	1042	116	9	4	-52
3	4	3	-38	29	24	-17
4	192	171	-11	8	9	8
5	448	528	18	9	5	-41

A Tabela 5.12 apresenta uma comparação entre as configurações 3 e 4, que diferem apenas no algoritmo de otimização empregado. Na configuração 4, o procedimento de minimização começou com o algoritmo PSO para explorar o espaço de busca, aproveitando seu caráter de busca global, utilizando a região de $\pm 3\sigma$ em torno dos valores medidos para a geração das partículas do PSO (conforme indicado

Tabela 5.11: Número de variáveis cujo valor reconciliado está fora da região de confiança de 99,7%.

Amostra	Config. 2a	Config. 3
1	6	5
2	5	5
3	0	0
4	6	3
5	6	8

na Tabela 5.3). Depois de algumas iterações, a solução fornecida pelo PSO foi transmitida como estimativa inicial do IP, com a expectativa de que esse encontrasse uma solução melhor e aumentasse a velocidade de convergência. É possível observar que o uso do algoritmo híbrido apresentou resultados melhores, em que os valores finais da função objetivo estão dentro da região de confiança para todas as amostras. Isso também é indicado pela região de confiança de valores reconciliados, em que todas as variáveis estão dentro da região de confiança de 99,7%, mostrando que a solução DRPE foi eficiente e robusta. A Figura 5.5 fornece resultados detalhados sobre a função de probabilidades acumulada da variabilidade em torno dos valores reais para cada conjunto de amostras, na qual se observa que a configuração 4 melhorou significativamente a precisão das variáveis reconciliadas.

No contexto de aplicações reais, os custos computacionais podem mostrar-se proibitivos em determinadas circunstâncias. Por isso, a literatura valoriza sobremaneira o uso de algoritmos determinísticos. Todavia, com base nos resultados discutidos previamente, pode-se argumentar que os mínimos locais foram evitados quando o método PSO foi usado, levando a valores reconciliados mais confiáveis. Quando o problema de otimização não é resolvido adequadamente, os resultados de DRPE podem produzir estimativas incorretas e introduzir desvios significativos na reconciliação das outras medidas e na estimação dos parâmetros. Embora essa melhoria nos resultados seja obtida a um custo computacional total mais elevado, a contribuição do PSO é constante e previsível, o que ajuda a controlar essa variável. A contribuição de IP para o custo computacional, por sua vez, é muito semelhante às dos demais experimentos de DRPE mostrados anteriormente. Esse desempenho é muito importante, pois, a despeito do fato da otimização ser conduzida por um algoritmo de busca aleatória, demonstra a viabilidade da utilização desse algoritmo em implementações reais. Por esse motivo, e considerando os resultados apresentados nas seções anteriores, recomenda-se o uso de algoritmos híbridos para a resolução de problemas de otimização envolvidos na reconciliação de dados.

5.4.4 Comentários Sobre a Métrica de Desempenho

O erro relativo entre os dados medidos e reconciliados é uma métrica quantitativa geralmente usada para avaliar como a reconciliação de dados diminui a variabilidade dos dados medidos. No entanto, esta métrica pode induzir uma má interpretação

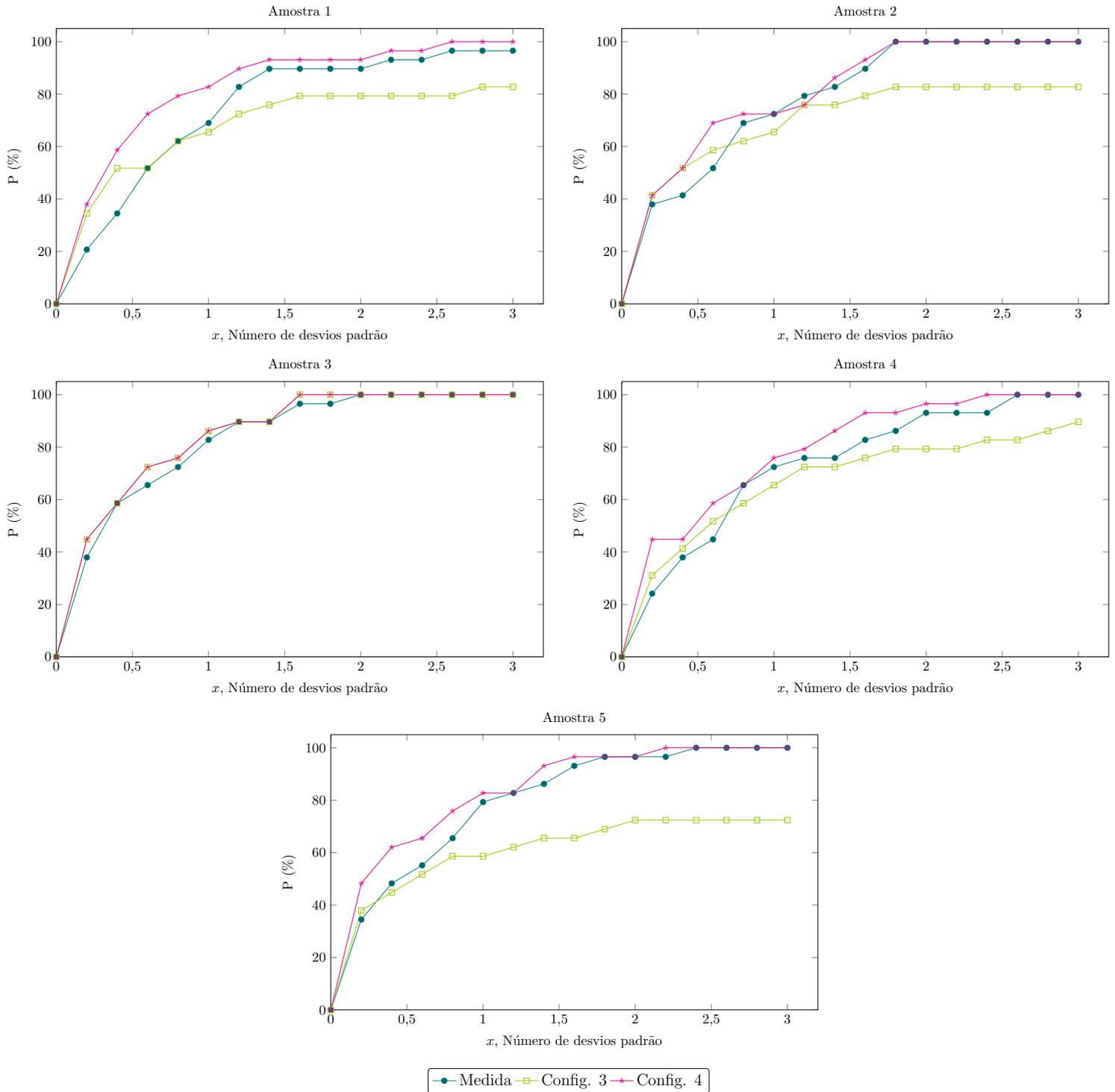


Figura 5.5: Distribuição de probabilidades acumulada da variabilidade em torno do valores verdadeiros do conjunto de amostras com os resultados de DRPE obtidos para as configurações 3 e 4. $P = P[\text{abs}(z_{ms} - z_m^*)/\sigma \leq x]$.

Tabela 5.12: Valores de função objetivo e custo computacional em função do algoritmo de otimização.

Amostra	J_{DRPE}			Tempo (min)		
	IP (Config. 3)	PSO + IP (Config. 4)	Diferença Relativa (%)	IP (Config. 3)	PSO + IP (Config. 4)	Diferença Relativa (%)
1	485	748/14	-97	6	7 + 21 = 28	354
2	1042	89/4	-99.7	4	6 + 11 = 17	309
3	3	578/3	0	24	6 + 22 = 28	16
4	171	291/17	-90	9	6 + 11 = 17	91
5	528	686/11	-98	5	6 + 13 = 19	242

dos resultados da DR. Se for admitido que os erros relativos entre os dados medidos e reconciliados abaixo de 10% indicam um problema de DR bem resolvido, os resultados obtidos para a configuração 4 não poderiam ser reconhecidos como coerentes, de acordo com a Figura 5.6A. Ao considerar os resultados da Figura 5.6B, por sua vez, observa-se que os resultados são consistentes para todas as variáveis, cuja variabilidade está dentro da região de confiança de 99,7%. Portanto, espera-se que o resíduo normalizado, como uma variável aleatória, varie no intervalo de 3σ com 99,7% de confiança.

A técnica de reconciliação de dados é altamente dependente da validade das hipóteses sobre a (co)variância dos erros de medição. Quando se trata de questões legais e jurídicas, informações sobre a variância dos erros de medição são, inclusive, obrigatórias (ANP, 2000). Em muitas situações, no entanto, pouco esforço é feito para caracterizar adequadamente os erros de medição, sendo que a suposição de erros aleatórios, normalmente distribuídos e independentes é a mais comum. Na ausência de informações suficientes sobre a variância dos erros de medição, a caracterização aproximada e hipóteses frágeis sobre os erros de medição podem levar a conclusões ambíguas ou incorretas. Isto exige que as incertezas das medições sejam avaliadas de alguma forma, o que nem sempre é fácil. No que diz respeito aos dados operacionais medidos em linha, que podem conter erros grosseiros, o desvio padrão dos dados medidos pode não ser adequado à avaliação das incertezas e algumas técnicas foram propostas para caracterização em linha de dados de processo para aplicações em tempo real (FEITAL e PINTO, 2015; FEITAL *et al.*, 2014). Uma alternativa para avaliar o desvio padrão do erro com base na precisão nominal do instrumento de medição é:

$$\sigma(\bar{x}_i) = \frac{\xi_i}{1.96\sqrt{N_s}}$$

em que ξ_i representa o erro admissível de determinado instrumento pela sua classe de precisão, com um intervalo de confiança de 95%, e N_s representa o número de sensores que medem o mesmo parâmetro. Essa alternativa foi utilizada por JIANG *et al.* (2014a).

5.5 Conclusões

A reconciliação de dados é uma ferramenta essencial em muitas indústrias e implementações reais são descritas na literatura, sendo que a maioria dos trabalhos está relacionada com a reconciliação de dados estacionária e não linear. Apesar

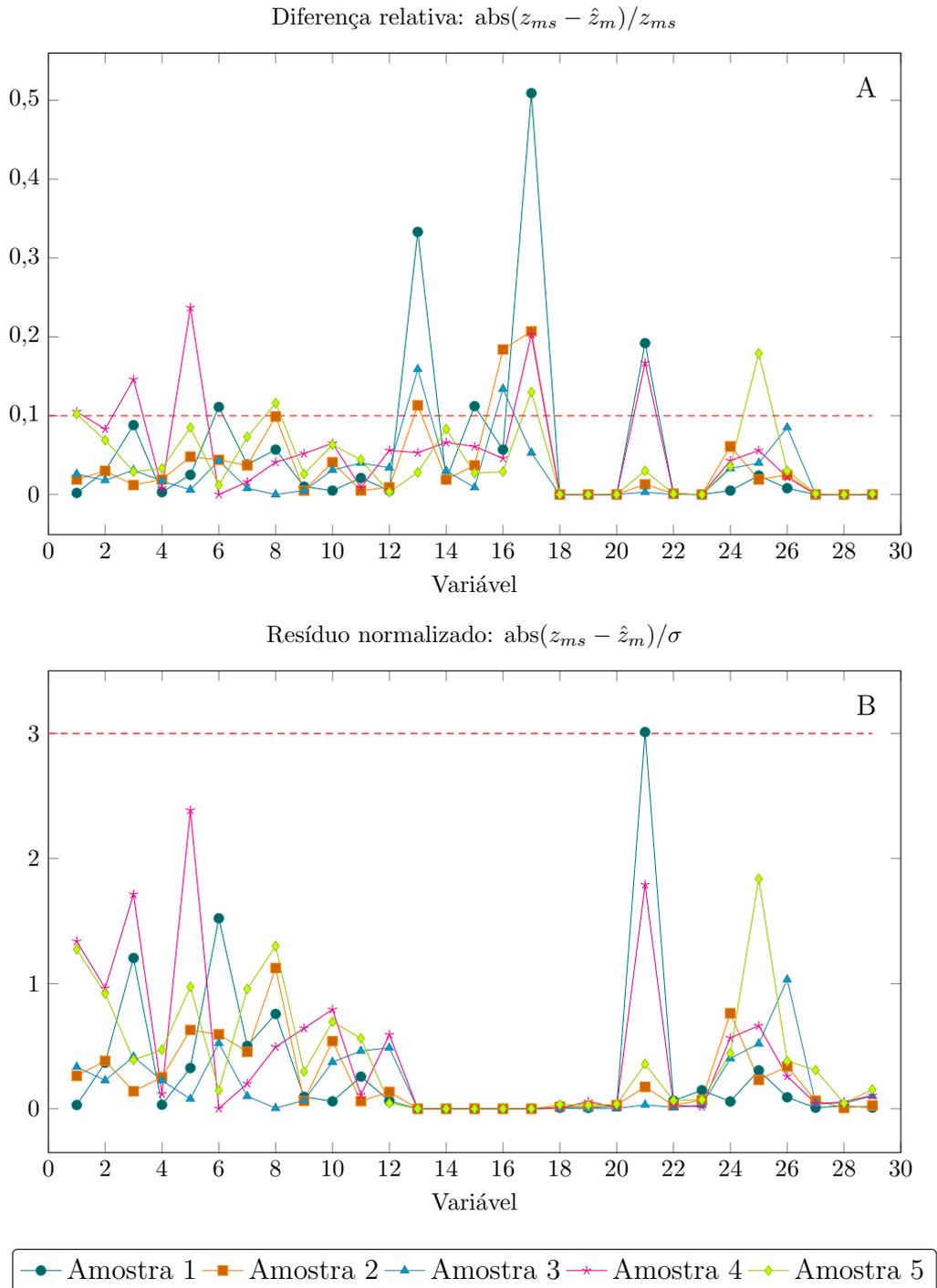


Figura 5.6: Comparação entre métricas para avaliação dos resultados reconciliados: (A) Diferença relativa entre dados medidos e reconciliados; (B) Resíduo normalizado. Resultados da reconciliação de dados para a configuração 4.

disso, poucos trabalhos avaliam aspectos numéricos importantes deste problema. Configurações distintas, mesmo aquelas representadas por diferenças sutis, podem apresentar desempenhos diferentes e devem ser analisadas durante o projeto de um sistema de DR, tentando responder se o problema de otimização está sendo resolvido satisfatoriamente. Embora a estimação de parâmetros seja usada principalmente em aplicações *off-line*, geralmente realizada durante a modelagem empírica de processos, a reconciliação de dados e a atualização do modelo constituem etapas importantes na otimização em tempo real.

Considerando as estratégias de construção de modelos de referência apresentadas, em que o Capítulo 3 explorou um cenário com excesso de informação e o Capítulo 4 um cenário com escassez de informação, o presente Capítulo mostrou que a disponibilidade de dados e de um modelo não garantem a identificação bem sucedida de um sistema. Com efeito, a resolução do problema de otimização envolvido na identificação é um elemento crítico.

Assim, este trabalho teve como objetivo fornecer uma avaliação numérica do desempenho de diferentes configurações para resolver o problema de reconciliação de dados não linear em estado estacionário. A avaliação baseou-se em um processo real de produção de petróleo *offshore*, que apresenta um comportamento altamente complexo e não linear. Foram discutidos alguns componentes que fazem parte de uma aplicação industrial de reconciliação de dados, mostrando-se como esses aspectos podem interferir na solução apropriada do problema de programação matemática resultante. Nesse caso, viu-se que, além dos dados e do modelo, a estrutura do problema de identificação é fundamental.

Foi mostrado que a precisão numérica da estimação de gradientes envolvida em algoritmos de otimização que usam informação de derivadas afeta a solução da DRPE, de forma que a solução resultante não satisfaz as suposições estatísticas adotadas. Também foi mostrado que a abordagem da solução pode afetar significativamente a qualidade da solução da DRPE, evidenciada pela obtenção de soluções inviáveis, *i.e.*, que violam as restrições. A solução do problema de DRPE usando um algoritmo de ponto interior por meio da abordagem simultânea (*i.e.*, método de relaxação Lagrangiana) forneceu resultados que violam as restrições do problema, o que ocorre devido à ativação de algumas restrições e à precisão finita dos métodos numéricos. Uma maior qualidade e resultados mais confiáveis foram obtidos quando algoritmos de otimização estocásticos foram combinados com um algoritmo determinístico em procedimentos híbridos. O algoritmo híbrido aplicado consistiu em iniciar a otimização com o algoritmo de busca aleatória PSO, aproveitando seu caráter de busca global para explorar o espaço de busca e, posteriormente, aplicar o algoritmo determinístico IP, visando à melhoria da solução e ao aumento da velocidade de convergência. Embora tal combinação de algoritmos já tenha sido

proposta (FLORES-TLACUAHUAC *et al.*, 2008; PARSOPOULOS e VRAHATIS, 2002; THANGARAJ *et al.*, 2011), a avaliação quantitativa em aplicações de reconciliação de dados é quase ausente. O procedimento híbrido combinado com a abordagem sequencial foi capaz de fornecer soluções robustas e viáveis para todas as amostras analisadas, em que ambos os valores finais da função objetivo e valores reconciliados estavam dentro da região de confiança. Em outras palavras, mostrou-se que o problema de otimização foi resolvido satisfatoriamente.

Extensões deste trabalho incluem o uso do procedimento de reconciliação com dados industriais reais em conjunto com uma estratégia de detecção de falha e erros grosseiros e a possível reformulação da função objetivo.

Capítulo 6

Conclusões

O amplo acesso a aplicações com uso de internet, sensores inteligentes, comunicação sem fio, dispositivos móveis, *etc.*, têm impulsionado o desenvolvimento de tecnologias na área de ciência de dados, notadamente *big data*. Esse movimento mundial, chamado de “Quarta Revolução Industrial”, tem atraído a atenção de empresas de diferentes setores e diversas indústrias têm investido nesse tipo de tecnologia. Muito se tem falado sobre os benefícios da “Indústria 4.0”, mas pouco se comenta sobre as questões práticas que viabilizam a implementação de um estágio tão avançado. O nível de automação é um aspecto muito importante nesse contexto e não oferece garantias de sucesso, somente diferencia os desafios a serem superados. Essa constatação foi o ponto de partida para os estudos realizados nessa pesquisa, em que buscou-se abordar alguns dos desafios relacionados com a otimização em tempo real tanto para processos com elevado nível de automação, bem como para aqueles com baixo nível de automação.

A otimização de processos e o desenvolvimento de sistemas automáticos para operação em tempo real são impactados por diversos desafios. É necessário conhecer o processo, quais variáveis estão relacionadas à sua operação e as diversas perturbações que o afetam. A modelagem matemática é uma ferramenta muito útil a diversas finalidades, notadamente para a otimização, pois é capaz de comprimir, e devolver com facilidade, uma extensa quantidade de informações sobre o comportamento da planta. No entanto, modelos matemáticos não são uma ferramenta absoluta e definitiva, mas também apresentam limitações que restringem sua aplicação. Para propósitos de otimização da operação, é desejável usar modelos simplificados que reflitam os principais fenômenos do processo e que permitam sua resolução em tempo real. Todavia, qualquer sistema de otimização baseado em modelo deve tratar diretamente o problema de incertezas no modelo e em seus parâmetros. Se por um lado, informações desconhecidas podem ser observadas por meio de medições, a elevada dimensão dos processos industriais e o tempo necessário para obtenção dessas respostas são restrições importantes que limitam o conhecimento do processo.

Em virtude da variedade, complexidade e constante mudança dos processos químicos industriais, a etapa mais árdua do desenvolvimento de um sistema para auxílio à tomada de decisão é, geralmente, a confecção de um modelo ideal para o processo. Dificuldades com processos de dimensão elevada constituem problemas antigos e estratégias envolvendo a identificação recursiva de modelos dinâmicos empíricos *on-line* (BAMBERGER e ISERMANN, 1978; BHATTACHARYA e JOSEPH, 1982; GARCIA e MORARI, 1981; MCFARLANE e BACON, 1989), bem como a incorporação de informações fenomenológicas no esquema adaptativo, já foram propostas (GOLDEN e YDSTIE, 1989), mas apresentavam problemas com a sintonia de parâmetros presentes na implementação e com imprecisões na etapa de identificação. Neste sentido, o Capítulo 3 objetivou, principalmente, a construção de sistemas de monitoramento de processos, em que se mostrou o potencial de uma estratégia via identificação de modelos dinâmicos baseados em dados, em conjunto com o adequado monitoramento do processo e da qualidade do modelo. Tecnologias de *big data* foram empregadas para viabilizar a construção rápida de modelos dinâmicos não lineares e o cálculo contínuo dos alarmes para aplicações *on-line* em processos com milhares de variáveis medidas, envolvendo um procedimento de detecção e diagnóstico de falhas com base em técnicas de controle estatístico de processos. O procedimento mostrou-se capaz de detectar o comportamento anormal e de fornecer informações relevantes para o diagnóstico da causa provável, auxiliando processos de tomada de decisão e análises de causalidade. Particularmente, a visão integrada do monitoramento de processos, para caracterização do desempenho de processos reais e identificação de anormalidades que impeçam o uso de modelos fenomenológicos em um contexto de predição e otimização, constitui a inovação dessa aplicação, que combina atividades de modelagem baseada em dados, técnicas de controle estatístico de processos e FDD, técnicas massivamente paralelas para aplicações *big data*, e estratégias de visualização dos resultados dos alarmes por diagramas de cores.

Por outro lado, conforme se mostrou no Capítulo 4, a ausência de informações também se apresenta como um desafio para a identificação de processos, o que afeta diretamente estratégias de otimização em tempo real. Focando em um processo de fermentação, mostrou-se como informações muito úteis para reduzir o consumo de energia e aumentar a produção podem ser reveladas a partir uma estratégia relativamente simples para monitoramento do processo. O estudo mostrou que mesmo os modelos fenomenológicos simples, como as cinéticas de Michaelis-Menten e Monod, podem ser usados para inferir o estado operacional da fermentação em batelada, um processo sabidamente complexo, e identificar o modelo com as medidas mais recentes. Vale destacar também a aplicação de uma metodologia sistemática para sintonia do procedimento de identificação e o importante desempenho observado com a aplicação do algoritmo metaheurístico de otimização PSO, mostrando a via-

bilidade da utilização desse algoritmo em implementações reais. Com poucas medidas de processo, o custo de instalação e manutenção do equipamento envolvido no sensor virtual é compensado pelo benefício das inferências para a estratégia de monitoramento e controle, permitindo uma maior reprodutibilidade do processo para controle em laço aberto com otimização de batelada em batelada. Como exemplos de oportunidades criadas pela aplicação da estratégia de monitoramento proposta, a estratégia de amostragem poderia ser simplificada e as atividades dos operadores repensada; simulações poderiam indicar cenários críticos e antecipar desvios do processo; os dados *on-line* que refletem atividades de enzimas reais poderiam ser usados para redução de custos com enzimas e outros aditivos; e novos conhecimentos sobre a operação do processo poderiam auxiliar a concepção de diferentes estratégias de controle, aumentando a produtividade.

Por fim, o Capítulo 5 discutiu alguns componentes que fazem parte de uma aplicação industrial de reconciliação de dados que não apresenta deficiências em relação ao conjunto de dados e ao modelo para a identificação do processo, mostrando como a resolução do problema de otimização pode ser problemática. Foi mostrado que a estimativa inicial e a precisão numérica da estimação de gradientes envolvida em algoritmos de otimização que usam informação de derivadas é uma questão importante, que afeta sobremaneira a solução do problema de reconciliação de dados. Com isso, estudos envolvendo a resolução de problemas de otimização que ignoram esses efeitos não realizam uma avaliação completa e podem apresentar conclusões tendenciosas. Foi mostrado que a abordagem sequencial apresentou melhores resultados em comparação com a abordagem simultânea. Uma maior qualidade e resultados mais confiáveis foram obtidos quando algoritmos de otimização estocásticos foram combinados com um algoritmo determinístico em um procedimento híbrido e, em conjunto com a abordagem sequencial e uma função objetivo adequada, forneceram soluções robustas que não violaram as restrições e satisfizeram as hipóteses sobre a natureza dos ruídos. Particularmente, mostrar a contribuição da abordagem sequencial para a obtenção de melhores resultados e a viabilidade da aplicação de algoritmos metaheurísticos em aplicações reais constituem resultados importantes.

Além dessas propostas, outra contribuição deste trabalho é a revisão sobre o tema da otimização em tempo real, em que é possível perceber a diversidade de temas, estratégias e metodologias envolvidas e desenvolvidas na otimização estática em tempo real. Em muitas situações, a necessidade de uso de ferramentas sistemáticas para diagnóstico, projeto e avaliação de desempenho em qualquer aspecto e/ou componente de sistemas RTO se mostrou imperativa. Além disso, ficou também evidente que, embora o ideal determinístico seja útil e embase cada uma das metodologias, sistemas RTO estão imersos em uma realidade governada por incertezas e fenômenos aleatórios. Apesar das dificuldades inerentes à realidade, há

aspectos de difícil percepção na concepção do sistema, como ocorre com o eventual conflito de interesse existente entre o objetivo de algum componente do sistema e o objetivo final do RTO, *i.e.*, a maximização do índice de desempenho do processo na otimização econômica.

Do ponto de vista acadêmico, o principal objetivo com o estudo de estratégias RTO consiste no desenvolvimento de métodos capazes de determinar o ótimo verdadeiro da planta e que, tendo encontrado esse ponto, apresentem rápida convergência. Diversos autores têm proporcionado importantes avanços para que esse objetivo seja alcançado. Todavia, as garantias fornecidas pelo rigor matemático da formulação acaba, também, representando uma fraqueza. As informações necessárias para que possam ser garantidas as propriedades teóricas de otimalidade e viabilidade da ferramenta de otimização encontram inúmeras dificuldades de serem atendidas nas aplicações industriais. Esse importante conflito entre teoria e realidade é claramente notado na otimização em tempo real. No entanto, isso ainda não impede a existência de sistemas RTO na indústria, mas indica que ainda há bastante trabalho de engenharia com potencial para tornar sistemas RTO mais confiáveis. A convicção de que ainda há espaço para o desenvolvimento de melhorias, as quais possibilitem aos sistemas RTO industriais atenderem algumas expectativas baseadas em estudos teóricos, deve motivar futuras pesquisas.

Referências Bibliográficas

- ADETOLA, V., GUAY, M., 2010, “Integration of real-time optimization and model predictive control”, *Journal of Process Control*, v. 20, n. 2, pp. 125 – 133. doi: 10.1016/j.jprocont.2009.09.001.
- AGARWAL, M., 1997, “Feasibility of On-Line Reoptimization in Batch Processes”, *Chemical Engineering Communications*, v. 158, n. 1, pp. 19–29. doi: 10.1080/00986449708936579.
- AGUIRRE, L. A., 2007, *Introdução à identificação de sistemas—Técnicas lineares e não-lineares aplicadas a sistemas reais*. Editora UFMG.
- AHLGREN, G., 1987, “Temperature Functions in Biology and Their Application to Algal Growth Constants”, *Oikos*, v. 49, n. 2, pp. 177–190. ISSN: 00301299, 16000706. Disponível em: <<http://www.jstor.org/stable/3566025>>.
- AL-ARFAJ, M. A., 2006, “Shortcut data reconciliation technique: Development and industrial application”, *AIChE Journal*, v. 52, n. 1, pp. 414–417. ISSN: 1547-5905. doi: 10.1002/aic.10627.
- ALAMO, T., FERRAMOSCA, A., GONZÁLEZ, A. H., et al., 2012, “A Gradient-Based Strategy for Integrating Real Time Optimizer (RTO) with Model Predictive Control (MPC)”. In: *Preprints of 4th IFAC Nonlinear Model Predictive Control Conference, Noordwijkerhout, NL*, v. 4, pp. 33–38. doi: 10.3182/20120823-5-NL-3013.00064.
- ALBERTON, K. P., ALBERTON, A. L., DI MAGGIO, J. A., et al., 2013, “Accelerating the parameters identifiability procedure: Set by set selection”, *Computers & Chemical Engineering*, v. 55, pp. 181 – 197. doi: 10.1016/j.compchemeng.2013.04.014.
- ALBUQUERQUE, J. S., BIEGLER, L. T., 1996, “Data reconciliation and gross-error detection for dynamic systems”, *AIChE Journal*, v. 42, n. 10, pp. 2841–2856. ISSN: 1547-5905. doi: 10.1002/aic.690421014.

- ALBUQUERQUE, J., BIEGLER, L., 1995, “Decomposition algorithms for on-line estimation with nonlinear models”, *Computers & Chemical Engineering*, v. 19, n. 10, pp. 1031 – 1039. ISSN: 0098-1354. doi: 10.1016/0098-1354(94)00107-Y.
- ALFORD, J. S., 2006, “Bioprocess control: Advances and challenges”, *Computers & Chemical Engineering*, v. 30, n. 10–12, pp. 1464 – 1475. doi: 10.1016/j.compchemeng.2006.05.039.
- ALKAYA, D., VASANTHARAJAN, S., BIEGLER, L., 2001, “Successive Quadratic Programming: Applications in the Process Industry”. In: Floudas, C. A., Pardalos, P. M. (Eds.), *Encyclopedia of Optimization*, Springer US, pp. 2574–2587. doi: 10.1007/0-306-48332-7_515.
- ALLGÖWER, F., ZHENG, A., 2000, *Nonlinear Model Predictive Control*, v. 26. Birkhäuser Basel. doi: 10.1007/978-3-0348-8407-5.
- ALMEIDA NETO, E., 2011, *Otimização dinâmica em tempo real : arquitetura de software, diagnóstico e análise de inviabilidades*. Tese de Doutorado, Universidade Federal do Rio Grande do Sul. Escola de Engenharia. Programa de Pós-Graduação em Engenharia Química. Disponível em: <<http://hdl.handle.net/10183/36008>>. Acesso em: 20 mar. 2014.
- ALMEIDA NT, E., SECCHI, A. R., 2011, “Dynamic optimization of a FCC converter unit: numerical analysis”, *Brazilian Journal of Chemical Engineering*, v. 28 (03), pp. 117 – 136. ISSN: 0104-6632. doi: 10.1590/S0104-66322011000100014.
- ALSTAD, V., SKOGESTAD, S., 2007, “Null Space Method for Selecting Optimal Measurement Combinations as Controlled Variables”, *Industrial & Engineering Chemistry Research*, v. 46, n. 3, pp. 846–853. doi: 10.1021/ie060285+.
- ALSTAD, V., SKOGESTAD, S., HORI, E. S., 2009, “Optimal measurement combinations as controlled variables”, *Journal of Process Control*, v. 19, n. 1, pp. 138 – 148. doi: 10.1016/j.jprocont.2008.01.002.
- ALTINTAŞ, M. M., KIRDAR, B., ÖNSAN, Z. İ., et al., 2002, “Cybernetic modeling of growth and ethanol production in a recombinant *Saccharomyces cerevisiae* strain secreting a bifunctional fusion protein”, *Process Biochemistry*, v. 37, n. 12, pp. 1439 – 1445. doi: 10.1016/S0032-9592(02)00031-6.

- ANDERSON, T. F., ABRAMS, D. S., GRENS, E. A., 1978, “Evaluation of parameters for nonlinear thermodynamic models”, *AIChE Journal*, v. 24, n. 1, pp. 20–29. ISSN: 1547-5905. doi: 10.1002/aic.690240103.
- ANG, C. P., TOPER, B., GAMBHIR, A., 2016, “Financial impacts of UK’s energy and climate change policies on commercial and industrial businesses”, *Energy Policy*, v. 91, pp. 273 – 286. doi: 10.1016/j.enpol.2015.12.015.
- ANP, AGÊNCIA NACIONAL DO PETRÓLEO, G. N. E. B., 2000. “PORTARIA CONJUNTA n. 01, de 19 de Junho de 2000. Aprova o Regulamento Técnico de Medição de Petróleo e Gás Natural, que estabelece as condições e requisitos mínimos para os sistemas de medição de petróleo e gás natural, com vistas a garantir resultados acurados e completos. DOU 20.6.2000, Brasília, DF.” Junho. Disponível em: <http://nxt.anp.gov.br/NXT/gateway.dll/leg%2Ffolder_portarias%2Fportarias_conj%2F2000%2Fpconj%201%20-%202000.xml>. Acesso em: 28 jun. 2017.
- ANURADHA, R., SURESH, A., VENKATESH, K., 1999, “Simultaneous saccharification and fermentation of starch to lactic acid”, *Process Biochemistry*, v. 35, n. 3–4, pp. 367 – 375. doi: 10.1016/S0032-9592(99)00080-1.
- ARIYUR, K. B., KRSTIC, M., 2003, *Real-time optimization by extremum-seeking control*. John Wiley & Sons.
- ARKUN, Y., STEPHANOPOULOS, G., 1980, “Studies in the synthesis of control structures for chemical processes: Part IV. Design of steady-state optimizing control structures for chemical process units”, *AIChE Journal*, v. 26, n. 6, pp. 975–991. doi: 10.1002/aic.690260613.
- ARKUN, Y., STEPHANOPOULOS, G., 1981, “Studies in the synthesis of control structures for chemical processes. Part V: Design of steady-state optimizing control structures for integrated chemical plants”, *AIChE Journal*, v. 27, n. 5, pp. 779–793. doi: 10.1002/aic.690270512.
- ARORA, N., BIEGLER, L. T., 2001, “Redescending estimators for data reconciliation and parameter estimation”, *Computers & Chemical Engineering*, v. 25, n. 11–12, pp. 1585 – 1599. ISSN: 0098-1354. doi: 10.1016/S0098-1354(01)00721-9.
- ASPEN PLUS USER GUIDE, 1998, *Volume 2, version 10*. Aspen Technology, Inc.

- ASSIS, A. J. D., MACIEL FILHO, R., 2000, “Soft sensors development for on-line bioreactor state estimation”, *Computers & Chemical Engineering*, v. 24, n. 2–7, pp. 1099 – 1103. doi: 10.1016/S0098-1354(00)00489-0.
- BACKX, T., BOSGRA, O., MARQUARDT, W., 1998, “Towards intentional dynamics in supply chain conscious process operations”. In: *Proceedings of Third International Conference on Foundations of Computer-Aided Process Operations (FOCAPO)*, Snowbird, Utah, USA, p. 5.
- BACKX, T., BOSGRA, O., MARQUARDT, W., 2000, “Integration of model predictive control and optimization of processes”. In: *IFAC Symposium Advanced Control of Chemical Processes, ADCHEM 2000, Pisa, Italy*, pp. 249–260.
- BACOVSKY, D., LUDWICZEK, N., OGNISSANTO, M., et al., 2013, *Status of Advanced Biofuels Demonstration Facilities in 2012. Report T39-P1b*. Relatório técnico, IEA Bioenergy Task 39. Disponível em: <<http://bit.ly/1rQ3wPq>>. Acesso em: 10 mai. 2016.
- BAGAJEWICZ, M. J., CABRERA, E., 2003, “Data Reconciliation in Gas Pipeline Systems”, *Industrial & Engineering Chemistry Research*, v. 42, n. 22, pp. 5596–5606. doi: 10.1021/ie020774j.
- BAGAJEWICZ, M. J., JIANG, Q., 1998, “Gross error modeling and detection in plant linear dynamic reconciliation”, *Computers & Chemical Engineering*, v. 22, n. 12, pp. 1789 – 1809. doi: 10.1016/S0098-1354(98)00248-8.
- BAGAJEWICZ, M. J., ROLLINS, S. D. K., 2011, “Data Reconciliation and Software Methods for Bias Detection”. In: Lipták, B. G., Eren, H. (Eds.), *Instrument Engineers’ Handbook, Volume 3: Process Software and Digital Networks*, fourth edition ed., CRC Press, cap. 22, pp. 364–381, Aug. doi: 10.1201/b11093-26.
- BAILEY, J., HRYMAK, A., TREIBER, S., et al., 1993, “Nonlinear optimization of a hydrocracker fractionation plant”, *Computers & Chemical Engineering*, v. 17, n. 2, pp. 123 – 138. doi: 10.1016/0098-1354(93)80010-K.
- BAINBRIDGE, L., 1983, “Ironies of automation”, *Automatica*, v. 19, n. 6, pp. 775 – 779. doi: 10.1016/0005-1098(83)90046-8.
- BAMBERGER, W., ISERMANN, R., 1978, “Adaptive on-line steady-state optimization of slow dynamic processes”, *Automatica*, v. 14, n. 3, pp. 223 – 230. doi: 10.1016/0005-1098(78)90087-0.

- BARBOSA JR, V., WOLF, M., MACIEL FO, R., 2000, “Development of data reconciliation for dynamic nonlinear system: application the polymerization reactor”, *Computers & Chemical Engineering*, v. 24, n. 2–7, pp. 501 – 506. ISSN: 0098-1354. doi: 10.1016/S0098-1354(00)00516-0.
- BARD, Y., 1974, *Nonlinear parameter estimation*, v. 513. Academic Press New York.
- BARTEE, J. F., MACHARIA, M. A., NOLL, P. D., et al., 2013. “Nonlinear Model Predictive Control Of A Biofuel Fermentation Process”. US 8,571,690 B2. Disponível em: <https://www.lens.org/lens/patent/US_8571690_B2>. Acesso em: 16 set. 2016.
- BARTUSIAK, R. D., NICHOLSON, D. H., 1998. “Plant parameter detection by monitoring of power spectral densities”. US Patent 5,790,413. Disponível em: <https://www.lens.org/lens/patent/US_5790413_A>. US Patent 5,790,413.
- BASAK, K., ABHILASH, K. S., GANGULY, S., et al., 2002, “On-Line Optimization of a Crude Distillation Unit with Constraints on Product Properties”, *Industrial & Engineering Chemistry Research*, v. 41, n. 6, pp. 1557–1568. doi: 10.1021/ie010059u.
- BASTIN, G., DOCHAIN, D., 1990, *On-line Estimation and Adaptive Control of Bioreactors*. Process Measurement and Control, 1. Elsevier, Amsterdam. ISBN: 0-444-88430-0.
- BAUSA, J., DÜNNEBIER, G., 2006, “Life Cycle Modelling in the chemical industries: Is there any reuse of models in automation and control?” In: Marquardt, W., Pantelides, C. (Eds.), *16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering*, v. 21, *Computer Aided Chemical Engineering*, Elsevier, pp. 3 – 8. doi: 10.1016/S1570-7946(06)80021-0.
- BAZARAA, M. S., SHERALI, H. D., SHETTY, C. M., 2006, *Nonlinear programming: theory and algorithms*. John Wiley & Sons.
- BAZIN, C., HODOUIN, D., DUCHESNE, C., et al., 1998, “Reconciliation of Mass and Energy Data Measurements: Application to a Rotary Dryer”, *Canadian Metallurgical Quarterly*, v. 37, n. 3-4, pp. 333–342. doi: 10.1179/cmq.1998.37.3-4.333.

- BECERRA, V. M., ROBERTS, P. D., GRIFFITHS, G. W., 1998, “Novel developments in process optimisation using predictive control”, *Journal of Process Control*, v. 8, n. 2, pp. 117 – 138. doi: 10.1016/S0959-1524(97)00046-2.
- BEHNAMI, A., SHAKERKHATIBI, M., DEHGHANZADEH, R., et al., 2016, “The implementation of data reconciliation for evaluating a full-scale petrochemical wastewater treatment plant”, *Environmental Science and Pollution Research*, v. 23, n. 22, pp. 22586–22595. doi: 10.1007/s11356-016-7484-5.
- BELLEÇ, S., JIANG, T., KERR, B., et al., 2007, “On-line processing and steady-state reconciliation of pulp and paper mill process data”, *Pulp & Paper Canada*, v. 108, n. 6, pp. 36–40.
- BELLMAN, R., ÅSTRÖM, K., 1970, “On structural identifiability”, *Mathematical Biosciences*, v. 7, n. 3–4, pp. 329 – 339. doi: 10.1016/0025-5564(70)90132-X.
- BELLMAN, R., 1957, *Dynamic Programming*. 1 ed. Princeton, NJ, USA, Princeton University Press.
- BELLMAN, R. E., DREYFUS, S. E., 1962, *Applied dynamic programming*. Princeton University Press, Princeton, NJ.
- BEN-TAL, A., EL GHAOU, L., NEMIROVSKI, A., 2009, *Robust optimization*. Princeton University Press. ISBN: 9780691143682. Disponível em: <<http://press.princeton.edu/titles/9099.html>>.
- BENGIO, Y., COURVILLE, A., VINCENT, P., 2013, “Representation learning: A review and new perspectives”, *IEEE transactions on pattern analysis and machine intelligence*, v. 35, n. 8, pp. 1798–1828. doi: 10.1109/TPAMI.2013.50.
- BERNARDO, F., SARAIVA, P., PISTIKOPOULOS, E., 2003, “Process design under uncertainty: Robustness criteria and value of information”. In: Asprey, S., Macchietto, S. (Eds.), *Dynamic Model Development*, v. 16, *Computer Aided Chemical Engineering*, Elsevier, pp. 175 – 208. doi: 10.1016/S1570-7946(03)80075-5.
- BERTSIMAS, D., BROWN, D. B., CARAMANIS, C., 2011, “Theory and Applications of Robust Optimization”, *SIAM Review*, v. 53, n. 3, pp. 464–501. doi: 10.1137/080734510.

- BETTS, J. T., FRANK, P. D., 1994, “A sparse nonlinear optimization algorithm”, *Journal of Optimization Theory and Applications*, v. 82, n. 3, pp. 519–541. ISSN: 1573-2878. doi: 10.1007/BF02192216.
- BEUTEL, S., HENKEL, S., 2011, “In situ sensor techniques in modern bioprocess monitoring”, *Applied Microbiology and Biotechnology*, v. 91, n. 6, pp. 1493–1505. doi: 10.1007/s00253-011-3470-5.
- BEYER, H.-G., SENDHOFF, B., 2007, “Robust optimization – A comprehensive survey”, *Computer Methods in Applied Mechanics and Engineering*, v. 196, n. 33–34, pp. 3190 – 3218. doi: 10.1016/j.cma.2007.03.003.
- BHAT, S. A., SARAF, D. N., 2004, “Steady-State Identification, Gross Error Detection, and Data Reconciliation for Industrial Process Units”, *Industrial & Engineering Chemistry Research*, v. 43, n. 15, pp. 4323–4336. doi: 10.1021/ie030563u.
- BHATTACHARYA, A., JOSEPH, B., 1982, “On-Line Optimization of Chemical Processes”. In: *American Control Conference, 1982*, pp. 334–337, June.
- BIEGLER, L. T., 1984, “Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation”, *Computers & Chemical Engineering*, v. 8, n. 3, pp. 243 – 247. ISSN: 0098-1354. doi: 10.1016/0098-1354(84)87012-X.
- BIEGLER, L. T., 2010, *Nonlinear programming: concepts, algorithms, and applications to chemical processes*. Society for Industrial and Applied Mathematics and Mathematical Optimization Society. doi: 10.1137/1.9780898719383.
- BIEGLER, L. T., GROSSMANN, I. E., 2004, “Retrospective on optimization”, *Computers & Chemical Engineering*, v. 28, n. 8, pp. 1169 – 1192. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2003.11.003.
- BIEGLER, L. T., CERVANTES, A. M., WÄCHTER, A., 2002, “Advances in simultaneous strategies for dynamic process optimization”, *Chemical Engineering Science*, v. 57, n. 4, pp. 575 – 593. doi: 10.1016/S0009-2509(01)00376-1.
- BIEGLER, L., 2014, “Recent Advances in Chemical Process Optimization”, *Chemie Ingenieur Technik*, v. 86, n. 7, pp. 943–952. doi: 10.1002/cite.201400033.
- BIEGLER, L., 2017, “Advanced Optimization Strategies For Integrated dynamic Process Operations”. In: *Proceedings of Foundations of Computer-Aided*

Process Operations (FOCAPO) and Chemical Process Control (CPC) 2017, Tucson, Arizona.

- BIEGLER, L., CUTHRELL, J., 1985, “Improved infeasible path optimization for sequential modular simulators-II: the optimization algorithm”, *Computers & Chemical Engineering*, v. 9, n. 3, pp. 257 – 267. ISSN: 0098-1354. doi: 10.1016/0098-1354(85)80004-1.
- BIEGLER, L., GROSSMANN, I., WESTERBERG, A., 1985, “A note on approximation techniques used for process optimization”, *Computers & Chemical Engineering*, v. 9, n. 2, pp. 201 – 206. ISSN: 0098-1354. doi: 10.1016/0098-1354(85)85010-9.
- BINDER, T., BLANK, L., BOCK, H. G., et al., 2001, “Introduction to model based optimization of chemical processes on moving horizons”. In: *Online optimization of large scale systems*, Springer, pp. 295–339.
- BIRGE, J. R., LOUVEAUX, F., 2011, *Introduction to stochastic programming*. Springer-Verlag New York. ISBN: 978-1-4614-0237-4. doi: 10.1007/978-1-4614-0237-4.
- BLUCK, D., KARBHARI, P. R., LIN, W.-J., 2014. “Simulated Fermentation Process”. US Patent 8,818,562. Disponível em: <https://www.lens.org/lens/patent/US_8818562_B2>. Acesso em: 16 set. 2016.
- BONVIN, D., SRINIVASAN, B., RUPPEN, D., 2001, “Dynamic Optimization in the Batch Chemical Industry”. In: *Chemical Process Control-VI*. Disponível em: <<http://infoscience.epfl.ch/record/28363>>. Acesso em: 21 mai. 2015.
- BONVIN, D., SRINIVASAN, B., 2013, “On the role of the necessary conditions of optimality in structuring dynamic real-time optimization schemes”, *Computers & Chemical Engineering*, v. 51, n. 0, pp. 172 – 180. doi: 10.1016/j.compchemeng.2012.07.012.
- BOTELHO, V. R., TRIERWEILER, L. F., TRIERWEILER, J. O., 2012, “A new approach for practical identifiability analysis applied to dynamic phenomenological models”. In: *8th IFAC International Symposium on Advanced Control of Chemical Processes, Furama Riverfront, Singapore*, v. 8, pp. 691–696. doi: 10.3182/20120710-4-SG-2026.00150.
- BOTHAST, R. J., SCHLICHER, M. A., 2005, “Biotechnological processes for conversion of corn into ethanol”, *Applied Microbiology and Biotechnology*, v. 67, n. 1, pp. 19–25. doi: 10.1007/s00253-004-1819-8.

- BOTTURA, C. P., 2013, “Memórias e reflexões de um professor brasileiro de controle de sistemas dinâmicos”, *TEMA (São Carlos)*, v. 14, n. 1, pp. 23 – 42. doi: 10.5540/tema.2013.014.01.0023.
- BOUROUIS, M., PIBOULEAU, L., FLOQUET, P., et al., 1998, “Simulation and data validation in multistage flash desalination plants”, *Desalination*, v. 115, n. 1, pp. 1 – 14. ISSN: 0011-9164. doi: 10.1016/S0011-9164(98)00021-6.
- BOX, G. E. P., 1954, “The Exploration and Exploitation of Response Surfaces: Some General Consideration and Examples”, *Biometrics*, v. 10, n. 1, pp. 16–60. ISSN: 0006341X. Disponível em: <<http://www.jstor.org/stable/3001663>>.
- BOX, G. E. P., WILSON, K. B., 1951, “On the Experimental Attainment of Optimum Conditions”, *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 13, n. 1, pp. 1–45. ISSN: 00359246. Disponível em: <<http://www.jstor.org/stable/2983966>>.
- BOX, G. E. P., YOULE, P. V., 1955, “The Exploration and Exploitation of Response Surfaces: An Example of the Link between the Fitted Surface and the Basic Mechanism of the System”, *Biometrics*, v. 11, n. 3, pp. 287–323. doi: 10.2307/3001769.
- BOX, G. E. P., 1957, “Evolutionary Operation: A Method for Increasing Industrial Productivity”, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, v. 6, n. 2, pp. 81–101. ISSN: 00359254. Disponível em: <<http://www.jstor.org/stable/2985505>>.
- BOX, G. E., DRAPER, N. R., 1969, *Evolutionary operation – A statistical method for process improvement*. John Wiley & Sons, New York.
- BRDYŚ, M., ROBERTS, P. D., 1987, “Convergence and optimality of modified two-step algorithm for integrated system optimization and parameter estimation”, *International Journal of Systems Science*, v. 18, n. 7, pp. 1305–1322. doi: 10.1080/00207728708967111.
- BRDYŚ, M., TATJEWSKI, P., 1994, “An algorithm for steady-state optimising dual control of uncertain plants”. In: *1st IFAC Workshop on New Trends in Design of Control Systems, Smolenice, Slovakia*, pp. 249–254.
- BRDYŚ, M., TATJEWSKI, P., 2005, *Iterative Algorithms for Multilayer Optimizing Control*. Imperial College Press, London UK. ISBN: 978-1-86094-514-4.

- BRDYŚ, M., CHEN, S., ROBERTS, P. D., 1986, “An extension to the modified two-step algorithm for steady-state system optimization and parameter estimation”, *International Journal of Systems Science*, v. 17, n. 8, pp. 1229–1243. doi: 10.1080/00207728608926883.
- BRDYŚ, M., ELLIS, J., ROBERTS, P., 1987, “Augmented integrated system optimisation and parameter estimation technique: derivation, optimality and convergence”, *Control Theory and Applications, IEE Proceedings D*, v. 134, n. 3, pp. 201–209. doi: 10.1049/ip-d:19870030.
- BRITT, H. I., LUECKE, R. H., 1973, “The Estimation of Parameters in Nonlinear, Implicit Models”, *Technometrics*, v. 15, n. 2, pp. 233–247. doi: 10.1080/00401706.1973.10489037.
- BROOKS, S. H., 1958, “A Discussion of Random Methods for Seeking Maxima”, *Operations Research*, v. 6, n. 2, pp. 244–251. doi: 10.1287/opre.6.2.244.
- BROSILOW, C., ZHAO, G., 1988, “A linear programming approach to constrained multivariable process control”, *Control and dynamic systems*, v. 27, pp. 141–181.
- BROWN, P., RHINEHART, R., 2000, “Automated steady-state identification in multivariable systems”, *Hydrocarbon Processing*, v. 79, n. 9, pp. 79–83.
- BĂNICĂ, F.-G., 2012, “What are Chemical Sensors?” In: *Chemical Sensors and Biosensors*, pp. 1–20, John Wiley & Sons, Ltd. ISBN: 9781118354162. doi: 10.1002/9781118354162.ch1.
- BUNIN, G., FRANCOIS, G., BONVIN, D., 2012a, “Exploiting local quasiconvexity for gradient estimation in modifier-adaptation schemes”. In: *American Control Conference (ACC), Montréal, Canada*, pp. 2806–2811, a. doi: 10.1109/ACC.2012.6314902.
- BUNIN, G., FRANÇOIS, G., BONVIN, D., 2013a, “Iterative Controller Tuning by Real-Time Optimization”. In: *Preprints of the 10th IFAC-DYCOPS Symposium*, n. 21-26, a.
- BUNIN, G., FRANÇOIS, G., BONVIN, D., 2013b, “Sufficient Conditions for Feasibility and Optimality of Real-Time Optimization Schemes - I. Theoretical Foundations”, Acesso em: 26 ago. 2015, b. Disponível em: <<http://infoscience.epfl.ch/record/188079>>.

- BUNIN, G., FRANÇOIS, G., BONVIN, D., 2013c, “Sufficient Conditions for Feasibility and Optimality of Real-Time Optimization Schemes - II. Implementation Issues”, Acesso em: 26 ago. 2015, c. Disponível em: <<http://infoscience.epfl.ch/record/188080>>.
- BUNIN, G., FRANÇOIS, G., BONVIN, D., 2013d, *The SCFO Real-Time Optimization Solver: Users’ Guide (version 0.9.4)*. Relatório técnico, d. Disponível em: <<http://ccapprox.info/soft.php>>. Acesso em: 26 ago. 2015.
- BUNIN, G. A., 2014, “On the equivalence between the modifier-adaptation and trust-region frameworks”, *Computers & Chemical Engineering*, v. 71, pp. 154–157.
- BUNIN, G. A., FRANÇOIS, G., SRINIVASAN, B., et al., 2011, “Input filter design for feasibility in constraint-adaptation schemes”. In: *18th IFAC World Congress, Milan*, pp. 5585–5590.
- BUNIN, G. A., WUILLEMIN, Z., FRANÇOIS, G., et al., 2012b, “Experimental real-time optimization of a solid oxide fuel cell stack via constraint adaptation”, *Energy*, v. 39, n. 1, pp. 54 – 62. doi: 10.1016/j.energy.2011.04.033.
- BUNIN, G. A., FRANÇOIS, G., BONVIN, D., 2013e, “From Discrete Measurements to Bounded Gradient Estimates: A Look at Some Regularizing Structures”, *Industrial & Engineering Chemistry Research*, v. 52, n. 35, pp. 12500–12513. doi: 10.1021/ie303309a.
- BUNIN, G. A., FRANÇOIS, G., BONVIN, D., 2014, “Feasible-side global convergence in experimental optimization”, *arXiv:1406.4063 [math.OC]*.
- BUSSANI, G., CHIARI, M., GROTTOLI, M., et al., 1995, “Application of data-reconciliation and optimisation procedure to hydrogen plant”, *Computers & Chemical Engineering*, v. 19, Supplement 1, pp. 299 – 304. doi: 10.1016/0098-1354(95)87053-9.
- CAMACHO, E. F., BORDONS, C., 1999, *Model Predictive Control*. Springer-Verlag, London.
- CAO, S., RHINEHART, R., 1995, “An efficient method for on-line identification of steady state”, *Journal of Process Control*, v. 5, n. 6, pp. 363 – 374. ISSN: 0959-1524. doi: 10.1016/0959-1524(95)00009-F.

- CAO, Y., 2004, “Constrained self-optimizing control via differentiation”. In: *7th International Symposium on Advanced Control of Chemical Processes (AD-CHEM), Hong Kong*, pp. 63–70.
- CAO, Y., KARIWALA, V., 2008, “Bidirectional branch and bound for controlled variable selection: Part I. Principles and minimum singular value criterion”, *Computers & Chemical Engineering*, v. 32, n. 10, pp. 2306 – 2319. doi: 10.1016/j.compchemeng.2007.11.011.
- CARDONA, C. A., SÁNCHEZ, O. J., 2007, “Fuel ethanol production: Process design trends and integration opportunities”, *Bioresource Technology*, v. 98, n. 12, pp. 2415 – 2457. doi: 10.1016/j.biortech.2007.01.002.
- CASTRO, A., ALMEIDA, F. G., AMORIM, P., et al., 2009, “A wavelet based method for steady-state detection in anesthesia”. In: *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 954–957. IEEE.
- CAUMO, L., KEMPF, A., TRIERWEILER, J., 2006, “Steady-state detection for multivariate systems based on PCA and wavelets”. In: *IFAC Symposium, ADCHEM '06, Gramado, Brazil*, pp. 693–698.
- CAVALCANTI, M. J. R., PINTO, J. C., 1997, “Modeling and optimization of suspension SAN polymerization reactors”, *Journal of Applied Polymer Science*, v. 65, n. 9, pp. 1683–1701. doi: 10.1002/(SICI)1097-4628(19970829)65:9<1683::AID-APP5>3.0.CO;2-B.
- CHACHUAT, B., MARCHETTI, A., BONVIN, D., 2008, “Process optimization via constraints adaptation”, *Journal of Process Control*, v. 18, n. 3, pp. 244–257.
- CHACHUAT, B., SRINIVASAN, B., BONVIN, D., 2009, “Adaptation strategies for real-time optimization”, *Computers & Chemical Engineering*, v. 33, n. 10, pp. 1557–1567.
- CHANTHAWONG, A., DHAKAL, S., 2016, “Stakeholders’ perceptions on challenges and opportunities for biodiesel and bioethanol policy development in Thailand”, *Energy Policy*, v. 91, pp. 189 – 206. doi: 10.1016/j.enpol.2016.01.008.
- CHATTERJEE, T., SARAF, D. N., 2004, “On-line estimation of product properties for crude distillation units”, *Journal of Process Control*, v. 14, n. 1, pp. 61 – 77. ISSN: 0959-1524. doi: 10.1016/S0959-1524(03)00036-2.

- CHAUDHURI, P., DIWEKAR, U., 1999, “Synthesis approach to the determination of optimal waste blends under uncertainty”, *AIChE Journal*, v. 45, n. 8, pp. 1671–1687. doi: 10.1002/aic.690450807.
- CHEN, C. Y., JOSEPH, B., 1987, “On-line optimization using a two-phase approach: an application study”, *Industrial & engineering chemistry research*, v. 26, n. 9, pp. 1924–1930.
- CHEN, J., ROMAGNOLI, J., 1998, “A strategy for simultaneous dynamic data reconciliation and outlier detection”, *Computers & Chemical Engineering*, v. 22, n. 4–5, pp. 559 – 562. ISSN: 0098-1354. doi: 10.1016/S0098-1354(97)00233-0.
- CHEN, J., BANDONI, A., ROMAGNOLI, J., 1998a, “Outlier detection in process plant data”, *Computers & Chemical Engineering*, v. 22, n. 4–5, pp. 641 – 646. doi: 10.1016/S0098-1354(97)00224-X.
- CHEN, J., LIU, K.-C., 2002, “On-line batch process monitoring using dynamic PCA and dynamic PLS models”, *Chemical Engineering Science*, v. 57, n. 1, pp. 63 – 75. doi: 10.1016/S0009-2509(01)00366-9.
- CHEN, J., PENG, Y., MUNOZ, J. C., 2013, “Correntropy estimator for data reconciliation”, *Chemical Engineering Science*, v. 104, pp. 1019 – 1027. ISSN: 0009-2509. doi: 10.1016/j.ces.2013.09.034.
- CHEN, P.-C., ANDERSEN, H., 2005, “The Implementation of the Data Validation Process in a Gas Turbine Performance Monitoring System”. In: *Turbo Expo: Power for Land, Sea, and Air*, pp. 609–616. American Society of Mechanical Engineers, June. doi: 10.1115/GT2005-68429.
- CHEN, X., RAO, K., YU, J., et al., 1996, “Comparison of GAMS, AMPL, and MINOS for optimization”, *Chemical Engineering Education*, v. 30, pp. 220–227.
- CHEN, X., PIKE, R. W., HERTWIG, T. A., et al., 1998b, “Optimal implementation of on-line optimization”, *Computers & Chemical Engineering*, v. 22, Supplement 1, pp. S435 – S442. ISSN: 0098-1354. doi: 10.1016/S0098-1354(98)00085-4. European Symposium on Computer Aided Process Engineering-8.
- CHENG, J.-H., ZAFIRIOU, E., 2000, “Robust Model-Based Iterative Feedback Optimization of Steady State Plant Operations”, *Industrial & Engineering Chemistry Research*, v. 39, n. 11, pp. 4215–4227. doi: 10.1021/ie000305m.

- CHIARI, M., BUSSANI, G., GROTTOLI, M., et al., 1997, “On-line data reconciliation and optimisation: Refinery applications”, *Computers & Chemical Engineering*, v. 21, Supplement, pp. S1185 – S1190. ISSN: 0098-1354. doi: 10.1016/S0098-1354(97)87663-6.
- CHRISTIANSEN, L. J., BRUNICHE-OLSEN, N., CARSTENSEN, J. H., et al., 1997, “Performance evaluation of catalytic processes”, *Computers & Chemical Engineering*, v. 21, Supplement, pp. S1179 – S1184. ISSN: 0098-1354. doi: 10.1016/S0098-1354(97)87662-4.
- CHRISTOFIDES, P. D., EL-FARRA, N. H., 2014, “Economic nonlinear model predictive control”, *Journal of Process Control*, v. 24, n. 8, pp. 1155. doi: 10.1016/j.jprocont.2014.06.012.
- CHUNG FU, T., 2011, “A review on time series data mining”, *Engineering Applications of Artificial Intelligence*, v. 24, n. 1, pp. 164 – 181. ISSN: 0952-1976. doi: 10.1016/j.engappai.2010.09.007.
- CICCIOTTI, M., XENOS, D. P., BOUASWAIG, A. E., et al., 2014, “Simultaneous Nonlinear Reconciliation and Update of Parameters for Online Use of First-Principles Models: An Industrial Case-Study on Compressors”. In: Jivří Jaromír Klemevs, P. S. V., Liew, P. Y. (Eds.), *24th European Symposium on Computer Aided Process Engineering*, v. 33, *Computer Aided Chemical Engineering*, Elsevier, pp. 457 – 462. doi: 10.1016/B978-0-444-63456-6.50077-6.
- CLAES, E. J., VAN IMPE, F. J., 2000, “Combining yield coefficients and exit-gas analysis for monitoring of the baker’s yeast fed-batch fermentation”, *Bioprocess Engineering*, v. 22, n. 3, pp. 195–200. doi: 10.1007/PL00009111.
- CLEMENTSCHITSCH, F., BAYER, K., 2006, “Improvement of bioprocess monitoring: development of novel concepts”, *Microbial Cell Factories*, v. 5, n. 1, pp. 1–11. doi: 10.1186/1475-2859-5-19.
- COBELLI, C., DISTEFANO, J. J., 1980, “Parameter and structural identifiability concepts and ambiguities: a critical review and analysis”, *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, v. 239, n. 1, pp. R7–R24.
- CONN, A. R., GOULD, G. I. M., TOINT, P. L., 1992, *LANCELOT: a Fortran package for large-scale nonlinear optimization (Release A)*, v. 17, *Springer Series in Computational Mathematics*. Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-662-12211-2.

- CONN, A. R., SCHEINBERG, K., VICENTE, L. N., 2009, *Introduction to derivative-free optimization*, v. 8, *MPS-SIAM series on optimization*. Society for Industrial and Applied Mathematics and Mathematical Optimization Society.
- COSTELLO, S., 2015, *Real-Time Optimization via Directional Modifier Adaptation, with Application to Kite Control*. Tese de Doutorado, STI, Lausanne.
- COSTELLO, S., FRANÇOIS, G., BONVIN, D., 2013a, “Real-time optimization when the plant and the model have different inputs”. In: *Preprints of the 10th IFAC-DYCOPS Symposium, Mumbai*, pp. 39–44, a.
- COSTELLO, S., FRANÇOIS, G., BONVIN, D., 2013b, “Real-Time Optimization for Kites”. In: *Proceedings of the 5th IFAC-PSYCO Workshop*, pp. 64–69, b. Disponível em: <<http://infoscience.epfl.ch/record/182808>>. Acesso em: 16 ago. 2015.
- COSTELLO, S., FRANÇOIS, G., BONVIN, D., et al., 2014, “Modifier adaptation for constrained closed-loop systems”. In: *Proceedings of the 19th IFAC World Congress*, n. EPFL-CONF-190599.
- COSTELLO, S., FRANÇOIS, G., BONVIN, D., 2015, “Directional Real-Time Optimization Applied to a Kite-Control Simulation Benchmark”. In: *European Control Conference*.
- CROWE, C. M., 1989, “Observability and redundancy of process data for steady state reconciliation”, *Chemical Engineering Science*, v. 44, n. 12, pp. 2909 – 2917. doi: 10.1016/0009-2509(89)85100-0.
- CROWE, C. M., 1996, “Data reconciliation – Progress and challenges”, *Journal of Process Control*, v. 6, n. 2–3, pp. 89 – 98. ISSN: 0959-1524. doi: 10.1016/0959-1524(96)00012-1.
- CUBILLOS, F. A., ACUÑA, G., LIMA, E., 2007, “Real-time process optimization based on grey-box neural models”, *Brazilian Journal of Chemical Engineering*, v. 24 (09), pp. 433 – 443. doi: 10.1590/S0104-66322007000300012.
- CULLICK, A. S., JOHNSON, W. D., 2007. “Methods, systems, and computer-readable media for real-time oil and gas field production optimization using a proxy simulator”. US Patent Application 2007/0192072 A1. Disponível em: <<http://www.freepatentsonline.com/y2007/0192072.html>>. Acesso em: 16 set. 2016.

- CUTLER, C. R., RAMAKER, B. L., 1980, “Dynamic matrix control - A computer control algorithm”. In: *Joint Automatic Control Conference, San Francisco, CA*, v. 17, p. 72. doi: 10.1109/JACC.1980.4232009.
- CUTLER, C., PERRY, R., 1983, “Real time optimization with multivariable control is required to maximize profits”, *Computers & chemical engineering*, v. 7, n. 5, pp. 663–667.
- CÂMARA, M. M., SOARES, R. M., FEITAL, T., et al., 2017a, “On-line identification of fermentation processes for ethanol production”, *Bioprocess and Biosystems Engineering*, v. 40, n. 7 (Jul), pp. 989–1006. ISSN: 1615-7605. doi: 10.1007/s00449-017-1762-6.
- CÂMARA, M. M., QUELHAS, A. D., PINTO, J. C., 2016, “Performance Evaluation of Real Industrial RTO Systems”, *Processes*, v. 4, n. 4, pp. 44. ISSN: 2227-9717. doi: 10.3390/pr4040044.
- CÂMARA, M. M., SOARES, R. M., FEITAL, T., et al., 2017b, “Numerical Aspects of Data Reconciliation in Industrial Applications”, *Processes*, v. 5, n. 4, pp. 56. ISSN: 2227-9717. doi: 10.3390/pr5040056.
- DAI, W., WORD, D. P., HAHN, J., 2014, “Modeling and dynamic optimization of fuel-grade ethanol fermentation using fed-batch process”, *Control Engineering Practice*, v. 22, pp. 231 – 241. doi: 10.1016/j.conengprac.2013.01.005.
- DANTZIG, G. B., THAPA, M. N., 2003, *Linear programming 2: Theory and Extensions*. Springer-Verlag New York. doi: 10.1007/b97283.
- DARBY, M. L., NIKOLAOU, M., 2012, “MPC: Current practice and challenges”, *Control Engineering Practice*, v. 20, n. 4, pp. 328 – 342. doi: 10.1016/j.conengprac.2011.12.004.
- DARBY, M. L., HARMSE, M., NIKOLAOU, M., 2009, “MPC: current practice and challenges”. In: *Proceedings of the 7th IFAC International Symposium on Advanced Control of Chemical Processes, Turkey*, v. 7, pp. 86–98. doi: 10.3182/20090712-4-TR-2008.00014.
- DARBY, M. L., NIKOLAOU, M., JONES, J., et al., 2011, “RTO: An overview and assessment of current practice”, *Journal of Process Control*, v. 21, n. 6, pp. 874 – 884. ISSN: 0959-1524. doi: 10.1016/j.jprocont.2011.03.009.
- DARLINGTON, J., PANTELIDES, C., RUSTEM, B., et al., 1999, “An algorithm for constrained nonlinear optimization under uncertainty”, *Automatica*, v. 35, n. 2, pp. 217 – 228. doi: 10.1016/S0005-1098(98)00150-2.

- DAROUACH, M., ZASADZINSKI, M., 1991, “Data reconciliation in generalized linear dynamic systems”, *AIChE Journal*, v. 37, n. 2, pp. 193–201. ISSN: 1547-5905. doi: 10.1002/aic.690370205.
- DAVE, D. J., DABHIYA, M. Z., SATYADEV, S., et al., 2003, “Online tuning of a steady state crude distillation unit model for real time applications”, *Journal of Process Control*, v. 13, n. 3, pp. 267 – 282. ISSN: 0959-1524. doi: 10.1016/S0959-1524(02)00026-4.
- DAVIS, J., EDGAR, T., PORTER, J., et al., 2012, “Smart manufacturing, manufacturing intelligence and demand-dynamic performance”, *Computers & Chemical Engineering*, v. 47, n. Supplement C, pp. 145 – 156. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2012.06.037. FOCAPO 2012.
- DE ANDRADE LIMA, L., 2006, “Nonlinear data reconciliation in gold processing plants”, *Minerals Engineering*, v. 19, n. 9, pp. 938 – 951. ISSN: 0892-6875. doi: 10.1016/j.mineng.2005.10.018.
- DE ARAÚJO, A. C., GOVATSMARK, M., SKOGESTAD, S., 2007, “Application of plantwide control to the HDA process. I – steady-state optimization and self-optimizing control”, *Control Engineering Practice*, v. 15, n. 10, pp. 1222 – 1237. doi: 10.1016/j.conengprac.2006.10.014.
- DE GOUVÊA, M., 1997, *Uso de um algoritmo SQP na otimização de processos químicos contínuos em tempo real*. Tese de Doutorado, Escola Politécnica, Universidade de São Paulo.
- DE GOUVÊA, M. T., ODLOAK, D., 1998, “One-layer real time optimization of LPG production in the FCC unit: procedure, advantages and disadvantages”, *Computers & Chemical Engineering*, v. 22, Supplement 1, pp. S191 – S198. doi: 10.1016/S0098-1354(98)00054-4.
- DE HENNIN, S. R., PERKINS, J. D., BARTON, G. W., 1994, “Structural decisions in on-line optimization”. In: *Proc. Int. Conf. on Process Sys. Eng. PSE '94*, pp. 297–302.
- DE HENNIN, S. R., 1994, *Structural decisions in on-line process optimization*. Tese de Doutorado, Imperial College London (University of London).
- DE SOUZA, G., ODLOAK, D., ZANIN, A. C., 2010, “Real time optimization (RTO) with model predictive control (MPC)”, *Computers & Chemical Engineering*, v. 34, n. 12, pp. 1999 – 2006. doi: 10.1016/j.compchemeng.2010.07.001.

- DEAN, J., GHEMAWAT, S., 2008, “MapReduce: Simplified Data Processing on Large Clusters”, *Commun. ACM*, v. 51, n. 1 (jan), pp. 107–113. ISSN: 0001-0782. doi: 10.1145/1327452.1327492.
- DEMPF, D., LIST, T., 1998, “On-line data reconciliation in chemical plants”, *Computers & Chemical Engineering*, v. 22, pp. S1023 – S1025. doi: 10.1016/S0098-1354(98)00205-1.
- DEVANTIER, R., PEDERSEN, S., OLSSON, L., 2005, “Characterization of very high gravity ethanol fermentation of corn mash. Effect of glucoamylase dosage, pre-saccharification and yeast strain”, *Applied Microbiology and Biotechnology*, v. 68, n. 5, pp. 622–629. doi: 10.1007/s00253-005-1902-9.
- DEYOU, X., 2014. “Intelligent monitoring device of fermentation process”. CN 204028679 U. Disponível em: <https://www.lens.org/lens/patent/CN_204028679_U>. Acesso em: 16 set. 2016.
- DIEHL, M. M., 2001, *Real-Time Optimization for Large Scale Nonlinear Processes*. Tese de Doutorado, Heidelberg University. Disponível em: <<http://www.ub.uni-heidelberg.de/archiv/1659>>. Acesso em: 26 ago. 2015.
- DISCENZO, F. M., 2014. “Microbial Monitoring And Prediction”. US 8,645,076 B2. Disponível em: <https://www.lens.org/lens/patent/US_8645076_B2>. Acesso em: 16 set. 2016.
- DIWEKAR, U., 2008, *Introduction to Applied Optimization*. Springer US. doi: 10.1007/978-0-387-76635-5.
- DOCHAIN, D., PERRIER, M., 1997, “Biotreatment, Downstream Processing and Modelling”. cap. Dynamical modelling, analysis, monitoring and control design for nonlinear bioprocesses, pp. 147–197, Berlin, Heidelberg, Springer Berlin Heidelberg. doi: 10.1007/BFb0103032.
- DOVÍ, V., PALADINO, O., 1989, “Fitting of experimental data to implicit models using a constrained variation algorithm”, *Computers & Chemical Engineering*, v. 13, n. 6, pp. 731 – 735. ISSN: 0098-1354. doi: 10.1016/0098-1354(89)80011-0.
- DRUD, A. S., 1994, “CONOPT – A Large-Scale GRG Code”, *ORSA Journal on Computing*, v. 6, n. 2, pp. 207–216. doi: 10.1287/ijoc.6.2.207.
- DU, Y.-G., HODOUIN, D., THIBAUT, J., 1997, “Use of a novel autoassociative neural network for nonlinear steady-state data reconciliation”,

AIChE Journal, v. 43, n. 7, pp. 1785–1796. ISSN: 1547-5905. doi: 10.1002/aic.690430714.

DUFFIELD, J. A., JOHANSSON, R., MEYER, S., 2015, *U.S. Ethanol: An Examination of Policy, Production, Use, Distribution, and Market Interactions*. Relatório técnico, Office of Energy Policy and New Uses. Office of the Chief Economist. U.S. Department of Agriculture.

DURBECK, R. C., 1965, *Principles for Simplification of Optimizing Control Models*. Tese de Doutorado, Case Institute of Technology.

EATON, A. N., BEAL, L. D., THORPE, S. D., et al., 2017, “Real time model identification using multi-fidelity models in managed pressure drilling”, *Computers & Chemical Engineering*, v. 97, n. Supplement C, pp. 76 – 84. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2016.11.008.

EBERHART, R., KENNEDY, J., 1995, “A new optimizer using particle swarm theory”. In: *Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan*, pp. 39–43. doi: 10.1109/MHS.1995.494215.

EBRAHIMPOUR, M., MATIAS, J. O. A., GRACIANO, J. E. A., et al., 2015, “Desenvolvimento de protótipo de RTO”, *Anais do XX Congresso Brasileiro de Engenharia Química - COBEQ 2014 [= Blucher Chemical Engineering Proceedings, v.1, n.2]*, v. 1, n. 2, pp. 11358 – 11365. doi: 10.5151/chemeng-cobeq2014-0262-26285-158945.

EDGAR, T. F., 2004, “Control and operations: when does controllability equal profitability?” *Computers & Chemical Engineering*, v. 29, n. 1, pp. 41 – 49. doi: 10.1016/j.compchemeng.2004.07.013. PSE 2003.

EDGAR, T. F., HIMMELBLAU, D. M., LASDON, L. S., 2001, *Optimization of chemical processes*. New York, McGraw-Hill. ISBN: 0-07-039359-1.

EDLER, J., NIKIFORUK, P. N., TINKER, E. B., 1970, “A comparison of the performance of techniques for direct, on-line optimization”, *The Canadian Journal of Chemical Engineering*, v. 48, n. 4, pp. 432–440. doi: 10.1002/cjce.5450480415.

EGHBAL AHMADI, M. H., 2015, “Data reconciliation and gross error detection: application in chemical processes”, *Cumhuriyet Science Journal*, v. 36, n. 3, pp. 1905–1913.

- EKSTEEN, J., FRANK, S., REUTER, M., 2002, “Dynamic structures in variance based data reconciliation adjustments for a chromite smelting furnace”, *Minerals Engineering*, v. 15, n. 11, Supplement 1, pp. 931 – 943. ISSN: 0892-6875. doi: 10.1016/S0892-6875(02)00131-0.
- ELGSÆTER, S. M., SLUPPHAUG, O., JOHANSEN, T. A., 2010, “A structured approach to optimizing offshore oil and gas production with uncertain models”, *Computers & Chemical Engineering*, v. 34, n. 2, pp. 163 – 176. doi: 10.1016/j.compchemeng.2009.07.011.
- ELLIS, J., KAMBHAMPATI, C., SHENG, G., et al., 1988, “Approaches to the optimizing control problem”, *International Journal of Systems Science*, v. 19, n. 10, pp. 1969–1985. doi: 10.1080/00207728808964092.
- ELLIS, M., DURAND, H., CHRISTOFIDES, P. D., 2014, “A tutorial review of economic model predictive control methods”, *Journal of Process Control*, v. 24, n. 8, pp. 1156 – 1178. doi: 10.1016/j.jprocont.2014.03.010.
- EMBIRUÇU, M., PRATA, D. M., LIMA, E. L., et al., 2008, “Continuous Soluble Ziegler-Natta Ethylene Polymerizations in Reactor Trains, 2 – Estimation of Kinetic Parameters from Industrial Data”, *Macromolecular Reaction Engineering*, v. 2, n. 2, pp. 142–160. ISSN: 1862-8338. doi: 10.1002/mren.200700046.
- ENGELL, S., 2007, “Feedback control for optimal process operation”, *Journal of Process Control*, v. 17, n. 3, pp. 203 – 219. doi: 10.1016/j.jprocont.2006.10.011. Special Issue ADCHEM 2006 Symposium.
- ESPOSITO, W. R., FLOUDAS, C. A., 1998, “Global Optimization in Parameter Estimation of Nonlinear Algebraic Models via the Error-in-Variables Approach”, *Industrial & Engineering Chemistry Research*, v. 37, n. 5, pp. 1841–1858. doi: 10.1021/ie970852g.
- FABER, R., LI, P., WOZNY, G., 2003, “Sequential Parameter Estimation for Large-Scale Systems with Multiple Data Sets. 1. Computational Framework”, *Industrial & Engineering Chemistry Research*, v. 42, n. 23, pp. 5850–5860. doi: 10.1021/ie030296s.
- FABER, R., LI, P., WOZNY, G., 2004, “Sequential Parameter Estimation for Large-Scale Systems with Multiple Data Sets. 2. Application to an Industrial Coke-Oven-Gas Purification Process”, *Industrial & Engineering Chemistry Research*, v. 43, n. 15, pp. 4350–4362. doi: 10.1021/ie0305622.

- FABER, R., LI, B., LI, P., et al., 2006, “Data reconciliation for real-time optimization of an industrial coke-oven-gas purification process”, *Simulation Modelling Practice and Theory*, v. 14, n. 8, pp. 1121 – 1134. ISSN: 1569-190X. doi: 10.1016/j.simpat.2006.09.005. {SIMS} 2004.
- FABER, R., ARELLANO-GARCIA, H., LI, P., et al., 2007, “An optimization framework for parameter estimation of large-scale systems”, *Chemical Engineering and Processing: Process Intensification*, v. 46, n. 11, pp. 1085 – 1095. ISSN: 0255-2701. doi: 10.1016/j.cep.2007.02.019. Special Issue on Process Optimization and Control in Chemical Engineering and Processing.
- FAULWASSER, T., BONVIN, D., 2014, “On the use of second-order modifiers for real-time optimization”. In: *Proceedings of the 19th IFAC World Congress*.
- FEITAL, T., PINTO, J. C., 2015, “Use of variance spectra for in-line validation of process measurements in continuous processes”, *The Canadian Journal of Chemical Engineering*, v. 93, n. 8, pp. 1426–1437. doi: 10.1002/cjce.22219.
- FEITAL, T., KRUGER, U., DUTRA, J., et al., 2013, “Modeling and performance monitoring of multivariate multimodal processes”, *AIChE Journal*, v. 59, n. 5, pp. 1557–1569. ISSN: 1547-5905. doi: 10.1002/aic.13953.
- FEITAL, T., PRATA, D. M., PINTO, J. C., 2014, “Comparison of methods for estimation of the covariance matrix of measurement errors”, *The Canadian Journal of Chemical Engineering*, v. 92, n. 12, pp. 2228–2245. doi: 10.1002/cjce.22063.
- FILLON, M., MEYER, M., PINGAUD, H., et al., 1995, “Data reconciliation based on elemental balances applied to batch experiments”, *Computers & Chemical Engineering*, v. 19, pp. 293 – 298. ISSN: 0098-1354. doi: 10.1016/0098-1354(95)87052-0.
- FILLON, M., MEYER, M., PINGAUD, H., et al., 1996, “An Efficient Formulation for Batch Reactor Data Reconciliation”, *Industrial & Engineering Chemistry Research*, v. 35, n. 7, pp. 2288–2298. doi: 10.1021/ie950152g.
- FINDEISEN, W., BRDYS, M., MALINOWSKI, K., et al., 1978, “On-line hierarchical control for steady-state systems”, *Automatic Control, IEEE Transactions on*, v. 23, n. 2 (Apr), pp. 189–209. doi: 10.1109/TAC.1978.1101702.
- FINN, B., HARVEY, L. M., MCNEIL, B., 2006, “Near-infrared spectroscopic monitoring of biomass, glucose, ethanol and protein content in a high cell

- density baker's yeast fed-batch bioprocess", *Yeast*, v. 23, n. 7, pp. 507–517. doi: 10.1002/yea.1371.
- FLEHMIG, F., MARQUARDT, W., 2008, "Inference of multi-variable trends in unmeasured process quantities", *Journal of Process Control*, v. 18, n. 5, pp. 491–503.
- FLEHMIG, F., WATZDORF, R., MARQUARDT, W., 1998, "Identification of trends in process measurements using the wavelet transform", *Computers & chemical engineering*, v. 22, pp. S491–S496.
- FLEHMIG, F. U., MARQUARDT, W., 2006, "Detection of multivariable trends in measured process quantities", *Journal of Process Control*, v. 16, n. 9, pp. 947–957.
- FLETCHER, R., 2000, *Practical Methods of Optimization, 2nd ed.* John Wiley & Sons Ltd. ISBN: 0-471-49463-1.
- Flickinger, M. C. (Ed.), 2013, *Upstream industrial biotechnology*, v. 2. Hoboken, New Jersey, John Wiley & Sons, Inc.
- FLORES-TLACUAHUAC, A., MORENO, S. T., BIEGLER, L. T., 2008, "Global Optimization of Highly Nonlinear Dynamic Systems", *Industrial & Engineering Chemistry Research*, v. 47, n. 8, pp. 2643–2655. doi: 10.1021/ie070379z.
- FOORD, A. G., 1974, *On-line optimisation of a petrochemical complex.* Tese de Doutorado, University of Cambridge.
- FORBES, F. J., 1994, *Model Structure and Adjustable Parameter Selection for Operations Optimization.* Tese de Doutorado, McMaster University. Disponível em: <<http://hdl.handle.net/11375/7726>>. Acesso em: 20 mar. 2014.
- FORBES, J. F., MARLIN, T. E., 1994, "Model accuracy for economic optimizing controllers: the bias update case", *Industrial & engineering chemistry research*, v. 33, n. 8, pp. 1919–1929.
- FORBES, J. F., MARLIN, T. E., 1996, "Design cost: a systematic approach to technology selection for model-based real-time optimization systems", *Computers & chemical engineering*, v. 20, n. 6, pp. 717–734.
- FORBES, J., MARLIN, T., MACGREGOR, J., 1994, "Model adequacy requirements for optimizing plant operations", *Computers & chemical engineering*, v. 18, n. 6, pp. 497–510.

- FORSGREN, A., GILL, P. E., WRIGHT, M. H., 2002, “Interior Methods for Nonlinear Optimization”, *SIAM Review*, v. 44, n. 4, pp. 525–597. doi: 10.1137/S0036144502414942.
- FORTUNA, L., GRAZIANI, S., RIZZO, A., et al., 2007, *Soft Sensors for Monitoring and Control of Industrial Processes*. Advances in Industrial Control. Springer-Verlag London. doi: 10.1007/978-1-84628-480-9.
- FRALEIGH, L., GUAY, M., FORBES, J., 2003, “Sensor selection for model-based real-time optimization: relating design of experiments and design cost”, *Journal of Process Control*, v. 13, n. 7, pp. 667 – 678. doi: 10.1016/S0959-1524(02)00122-1.
- FRANÇOIS, G., 2014, *Optimisation en Temps Réel: Optimiser les Performances des Procédés Chimiques malgré l’Incertitude et les Erreurs de Modélisation*. Relatário técnico, Nancy. France. Thèse d’Habilitation à Diriger les Recherches de l’Université de Lorraine, Génie des Procédés et Produits.
- FRANÇOIS, G., BONVIN, D., 2013a, “Chapter One - Measurement-Based Real-Time Optimization of Chemical Processes”. In: Pushpavanam, S. (Ed.), *Control and Optimisation of Process Systems*, v. 43, *Advances in Chemical Engineering*, Academic Press, pp. 1 – 50, a. doi: 10.1016/B978-0-12-396524-0.00001-5.
- FRANÇOIS, G., BONVIN, D., 2013b, “Use of Convex Model Approximations for Real-Time Optimization via Modifier Adaptation”, *Industrial & Engineering Chemistry Research*, v. 52, n. 33, pp. 11614–11625. doi: 10.1021/ie3032372.
- FRANÇOIS, G., BONVIN, D., 2014, “Use of Transient Measurements for the Optimization of Steady-State Performance via Modifier Adaptation”, *Industrial & Engineering Chemistry Research*, v. 53, n. 13, pp. 5148–5159. doi: 10.1021/ie401392s.
- FRANÇOIS, G., SRINIVASAN, B., BONVIN, D., 2005, “Use of measurements for enforcing the necessary conditions of optimality in the presence of constraints and uncertainty”, *Journal of Process Control*, v. 15, n. 6, pp. 701–712.
- FRANÇOIS, G., SRINIVASAN, B., BONVIN, D., 2012, “Comparison of Six Implicit Real-Time Optimization Schemes”, *Journal Européen des Systèmes Automatisés*, v. 46, n. 2-3, pp. 291–305. doi: 10.3166/JESA.46.291-305.

- FRANÇOIS, G., COSTELLO, S., BONVIN, D., 2015, “Application of Real-Time Optimization Methods to Energy Systems in the Presence of Uncertainties and Disturbances.” *TMC Academic Journal*, v. 9, n. 2, pp. 19–40.
- FULLER, W. A., 1987, *Measurement error models*. John Wiley & Sons, Inc.
- GABREL, V., MURAT, C., THIELE, A., 2014, “Recent advances in robust optimization: An overview”, *European Journal of Operational Research*, v. 235, n. 3, pp. 471 – 483. doi: 10.1016/j.ejor.2013.09.036.
- GANI, R., GERNAEY, K. V., SIN, G., 2000, “Process Systems Engineering, 1. Introduction”. In: *Ullmann’s Encyclopedia of Industrial Chemistry*, cap. 1, Wiley-VCH Verlag GmbH & Co. KGaA. doi: 10.1002/14356007.b06_317.pub2.
- GAO, W., ENGELL, S., 2005a, “Comparison of iterative set-point optimisation strategies under structural plant-model mismatch”. In: *Proceedings of the 16th IFAC World Congress, Prague, Czech Republic*, v. 16, pp. 401–406, a. doi: 10.3182/20050703-6-CZ-1902.00402.
- GAO, W., ENGELL, S., 2005b, “Iterative set-point optimization of batch chromatography”, *Computers & chemical engineering*, v. 29, n. 6, pp. 1401–1409.
- GAO, W., WENZEL, S., ENGELL, S., 2016, “A reliable modifier-adaptation strategy for real-time optimization”, *Computers & Chemical Engineering*, v. 91, pp. 318 – 328. doi: 10.1016/j.compchemeng.2016.03.019. PSE-2015/ESCAPE-25, 31 May - 4 June 2015, Copenhagen, Denmark.
- GARCIA, C. E., MORARI, M., 1984, “Optimal operation of integrated processing systems: Part II: Closed-loop on-line optimizing control”, *AIChE Journal*, v. 30, n. 2, pp. 226–234. doi: 10.1002/aic.690300209.
- GARCIA, C. E., MORARI, M., 1981, “Optimal operation of integrated processing systems. Part I: Open-loop on-line optimizing control”, *AIChE Journal*, v. 27, n. 6, pp. 960–968. doi: 10.1002/aic.690270612.
- GARCIA, C. E., MORARI, M., 1982, “Internal model control. 1. A unifying review and some new results”, *Industrial & Engineering Chemistry Process Design and Development*, v. 21, n. 2, pp. 308–323. doi: 10.1021/i200017a016.
- GARCIA, C. E., MORARI, M., 1985, “Internal model control. 2. Design procedure for multivariable systems”, *Industrial & Engineering Chemistry Process Design and Development*, v. 24, n. 2, pp. 472–484. doi: 10.1021/i200029a043.

- GAU, C.-Y., STADTHERR, M. A., 2002, “Deterministic global optimization for error-in-variables parameter estimation”, *AIChE Journal*, v. 48, n. 6, pp. 1192–1197. ISSN: 1547-5905. doi: 10.1002/aic.690480607.
- GELADI, P., KOWALSKI, B. R., 1986, “Partial least-squares regression: a tutorial”, *Analytica Chimica Acta*, v. 185, n. Supplement C, pp. 1 – 17. ISSN: 0003-2670. doi: 10.1016/0003-2670(86)80028-9.
- GELETU, A., LI, P., 2014, “Recent Developments in Computational Approaches to Optimization under Uncertainty and Application in Process Systems Engineering”, *ChemBioEng Reviews*, v. 1, n. 4, pp. 170–190. doi: 10.1002/cben.201400013.
- GEORGAKIS, C., 2009, “A model-free methodology for the optimization of batch processes: Design of dynamic experiments”. In: *7th IFAC International Symposium on Advanced Control of Chemical Processes, Advanced Control of Chemical Processes, Turkey*, v. 7, pp. 536–541. doi: 10.3182/20090712-4-TR-2008.00086.
- GILL, P. E., MURRAY, W., SAUNDERS, M. A., 2005, “SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization”, *SIAM Review*, v. 47, n. 1, pp. 99–131. doi: 10.1137/S0036144504446096.
- GNANADESIKAN, R., 1997, *Methods for statistical data analysis of multivariate observations*. 2 ed. New York, John Wiley & Sons.
- GOLDEN, M. P., YDSTIE, B. E., 1989, “Adaptive extremum control using approximate process models”, *AIChE Journal*, v. 35, n. 7, pp. 1157–1169. doi: 10.1002/aic.690350711.
- GOLSHAN, M., PISHVAIE, M. R., BOOZARJOMEHRY, R. B., 2008, “Stochastic and global real time optimization of Tennessee Eastman challenge problem”, *Engineering Applications of Artificial Intelligence*, v. 21, n. 2, pp. 215 – 228. doi: 10.1016/j.engappai.2007.04.004.
- GOMES, M. V. D. C., 2007, *Otimização sequencial por aproximações – Uma aplicação em tempo real para o refino de petróleo*. Tese de Doutorado, Programa de Engenharia Química, COPPE, UFRJ.
- GOVATSMARK, M. S., SKOGESTAD, S., 2005, “Selection of Controlled Variables and Robust Setpoints”, *Industrial & Engineering Chemistry Research*, v. 44, n. 7, pp. 2207–2217. doi: 10.1021/ie049750y.

- GRACIANO, J., MENDOZA, D., LE ROUX, G., 2014, “Performance comparison of parameter estimation techniques for unidentifiable models”, *Computers & Chemical Engineering*, v. 64, pp. 24 – 40. doi: 10.1016/j.compchemeng.2014.01.009.
- GRACIANO, J. E. A., JÄSCHKE, J., LE ROUX, G. A., et al., 2015, “Integrating self-optimizing control and real-time optimization using zone control MPC”, *Journal of Process Control*, v. 34, pp. 35 – 48. doi: 10.1016/j.jprocont.2015.07.003.
- GRAEFF, T. R., HARMON, S., 2002, “Collecting and using personal data: consumers’ awareness and concerns”, *Journal of Consumer Marketing*, v. 19, n. 4, pp. 302–318. doi: 10.1108/07363760210433627.
- GRAY, S. R., PERETTI, S. W., LAMB, H. H., 2013, “Real-time monitoring of high-gravity corn mash fermentation using in situ raman spectroscopy”, *Biotechnology and Bioengineering*, v. 110, n. 6, pp. 1654–1662. doi: 10.1002/bit.24849.
- GRIVA, I., NASH, S. G., SOFER, A., 2009, *Linear and Nonlinear Optimization*, 2nd ed. SIAM, Philadelphia, PA. ISBN: 978-0-898716-61-0.
- GROS, S., SRINIVASAN, B., BONVIN, D., 2009, “Optimizing control based on output feedback”, *Computers & Chemical Engineering*, v. 33, n. 1, pp. 191 – 198. doi: 10.1016/j.compchemeng.2008.07.017.
- GROSSMANN, I. E., SARGENT, R. W. H., 1978, “Optimum design of chemical plants with uncertain parameters”, *AIChE Journal*, v. 24, n. 6, pp. 1021–1028. doi: 10.1002/aic.690240612.
- GUAY, M., ZHANG, T., 2003, “Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties”, *Automatica*, v. 39, n. 7, pp. 1283–1293.
- GUO, M., SONG, W., BUHAIN, J., 2015, “Bioenergy and biofuels: History, status, and perspective”, *Renewable and Sustainable Energy Reviews*, v. 42, pp. 712 – 725. doi: 10.1016/j.rser.2014.10.013.
- GUO, S., LIU, P., LI, Z., 2016a, “Estimation of exhaust steam enthalpy and steam wetness fraction for steam turbines based on data reconciliation with characteristic constraints”, *Computers & Chemical Engineering*, v. 93, pp. 25 – 35. doi: 10.1016/j.compchemeng.2016.05.019.

- GUO, S., LIU, P., LI, Z., 2016b, “Inequality constrained nonlinear data reconciliation of a steam turbine power plant for enhanced parameter estimation”, *Energy*, v. 103, pp. 215 – 230. ISSN: 0360-5442. doi: 10.1016/j.energy.2016.02.158.
- HALVORSEN, I. J., SKOGESTAD, S., MORUD, J. C., et al., 2003, “Optimal Selection of Controlled Variables”, *Industrial & Engineering Chemistry Research*, v. 42, n. 14, pp. 3273–3284. doi: 10.1021/ie020833t.
- HANMANDLU, M., PURKAYASTHA, P., PAL, J., 1985, “On the use of nonlinear programming in real time control in process industries”, *Annual Review in Automatic Programming*, v. 13, pp. 71–78.
- HANSSEN, K. G., FOSS, B., 2015, “Production Optimization under Uncertainty - Applied to Petroleum Production”, *IFAC-PapersOnLine*, v. 48, n. 8, pp. 217 – 222. doi: 10.1016/j.ifacol.2015.08.184.
- HARMS, P., KOSTOV, Y., RAO, G., 2002, “Bioprocess monitoring”, *Current Opinion in Biotechnology*, v. 13, n. 2, pp. 124 – 127. doi: 10.1016/S0958-1669(02)00295-1.
- HEDENGREN, J., MOJICA, J., COLE, W., et al., 2012, “APOPT: MINLP solver for differential and algebraic systems with benchmark testing”. In: *Proceedings of the INFORMS National Meeting, Phoenix, AZ, USA*, v. 1417.
- HELBIG, A., ABEL, O., MARQUARDT, W., 2000, “Structural Concepts for Optimization Based Control of Transient Processes”. In: Allgöwer, F., Zheng, A. (Eds.), *Nonlinear Model Predictive Control*, v. 26, *Progress in Systems and Control Theory*, Birkhäuser Basel, pp. 295–311. doi: 10.1007/978-3-0348-8407-5_16.
- HELTON, J., 1997, “Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty”, *Journal of Statistical Computation and Simulation*, v. 57, n. 1-4, pp. 3–76. doi: 10.1080/00949659708811803.
- HEYEN, G., MARÉCHAL, E., KALITVENTZEFF, B., 1996, “Sensitivity calculations and variance analysis in plant measurement reconciliation”, *Computers & Chemical Engineering*, v. 20, pp. S539 – S544. ISSN: 0098-1354. doi: 10.1016/0098-1354(96)00099-3.
- HEYEN, G., KALITVENTZEFF, B., 1997, “Methodology for optimization of operation to reduce site-scale energy use in production plants”, *Applied Thermal Engineering*, v. 17, n. 8, pp. 1005 – 1014. ISSN: 1359-4311. doi: 10.1016/S1359-4311(97)00017-3.

- HIMMELBLAU, D., 1970, *Process Analysis by Statistical Methods*. John Wiley & Sons, New York, NY.
- HLAVÁČEK, V., 1977, “Analysis of a complex plant-steady state and transient behavior”, *Computers & Chemical Engineering*, v. 1, n. 1, pp. 75 – 100. doi: 10.1016/0098-1354(77)80011-2.
- HODGE, V., AUSTIN, J., 2004, “A Survey of Outlier Detection Methodologies”, *Artificial Intelligence Review*, v. 22, n. 2, pp. 85–126. doi: 10.1007/s10462-004-4304-y.
- HODOUIN, D., EVERELL, M., 1980, “A hierarchical procedure for adjustment and material balancing of mineral processes data”, *International Journal of Mineral Processing*, v. 7, n. 2, pp. 91 – 116. ISSN: 0301-7516. doi: 10.1016/0301-7516(80)90002-2.
- HOFFERBER, J. A., WEBB, P. U., 1978. “Fluid catalytic cracking with automatic temperature control”. US Patent 4,092,722. Disponível em: <https://www.lens.org/lens/patent/US_4092722_A>. Acesso em: 16 set. 2016.
- HOFFMANN, F., SCHMIDT, M., RINAS, U., 2000, “Simple technique for simultaneous on-line estimation of biomass and acetate from base consumption and conductivity measurements in high-cell density cultures of *Escherichia coli*”, *Biotechnology and Bioengineering*, v. 70, n. 3, pp. 358–361. doi: 10.1002/1097-0290(20001105)70:3<358::AID-BIT14>3.0.CO;2-T.
- Holder, A. (Ed.), 2006–14, *Mathematical Programming Glossary*. <http://glossary.computing.society.informs.org>, INFORMS Computing Society. Originally authored by Harvey J. Greenberg, 1999–2006.
- HOLLY, W., COOK, R., CROWE, C. M., 1989, “Reconciliation of mass flow rate measurements in a chemical extraction plant”, *The Canadian Journal of Chemical Engineering*, v. 67, n. 4, pp. 595–601. ISSN: 1939-019X. doi: 10.1002/cjce.5450670411.
- HOLMES, D., 2003, “Model-free optimization in cement plants”. In: *Cement Industry Technical Conference, 2003. Conference Record. IEEE-IAS/PCA 2003*, pp. 159–173, May. doi: 10.1109/CITCON.2003.1204717.
- HOTELLING, H., 1941, “Experimental Determination of the Maximum of a Function”, *Ann. Math. Statist.*, v. 12, n. 1 (03), pp. 20–45. doi: 10.1214/aoms/1177731784.

- HU, M., SHAO, H., 2006, “Theory Analysis of Nonlinear Data Reconciliation and Application to a Coking Plant”, *Industrial & Engineering Chemistry Research*, v. 45, n. 26, pp. 8973–8984. doi: 10.1021/ie060077h.
- HU, W., UMAR, L. M., XIAO, G., et al., 2012, “Local self-optimizing control of constrained processes”, *Journal of Process Control*, v. 22, n. 2, pp. 488 – 493. doi: 10.1016/j.jprocont.2011.11.003.
- HUANG, L. P., JIN, B., LANT, P., et al., 2005, “Simultaneous saccharification and fermentation of potato starch wastewater to lactic acid by *Rhizopus oryzae* and *Rhizopus arrhizus*”, *Biochemical Engineering Journal*, v. 23, n. 3, pp. 265 – 276. doi: 10.1016/j.bej.2005.01.009.
- HUBER, P. J., 2014, “Robust Statistics”. In: Lovric, M. (Ed.), *International Encyclopedia of Statistical Science*, Springer Berlin Heidelberg, pp. 1248–1251. doi: 10.1007/978-3-642-04898-2_594.
- HUBER, P. J., RONCHETTI, E., 2009, *Robust statistics, second edition*. John Wiley & Sons, Inc., Hoboken, New Jersey. ISBN: 978-0-470-12990-6.
- HUNTER, W. G., KITTRELL, J. R., 1966, “Evolutionary Operation: A Review”, *Technometrics*, v. 8, n. 3, pp. 389–397. doi: 10.1080/00401706.1966.10490372.
- IERAPETRITOU, M., ACEVEDO, J., PISTIKOPOULOS, E., 1996, “An optimization approach for process engineering problems under uncertainty”, *Computers & Chemical Engineering*, v. 20, n. 6–7, pp. 703 – 709. doi: 10.1016/0098-1354(95)00203-0. Fifth International Symposium on Process Systems Engineering.
- IPCC, 2014, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]*. IPCC, Geneva, Switzerland. ISBN: 978-92-9169-143-2.
- ISERMANN, R., 1997, “Supervision, fault-detection and fault-diagnosis methods – An introduction”, *Control Engineering Practice*, v. 5, n. 5, pp. 639 – 652. ISSN: 0967-0661. doi: 10.1016/S0967-0661(97)00046-4.
- ISERMANN, R., BALLÉ, P., 1997, “Trends in the application of model-based fault detection and diagnosis of technical processes”, *Control Engineering Practice*, v. 5, n. 5, pp. 709 – 719. doi: 10.1016/S0967-0661(97)00053-1.

- ISLAM, K., WEISS, G., ROMAGNOLI, J., 1994, “Non-linear data reconciliation for an industrial pyrolysis reactor”, *Computers & Chemical Engineering*, v. 18, Supplement 1, pp. S217 – S221. doi: 10.1016/0098-1354(94)80036-7. European Symposium on Computer Aided Process Engineering—325th European Symposium of the Working Party on Computer Aided Process Engineering and 494th Event of the European Federation of Chemical Engineering (EFChE).
- JANG, M.-F., CHOU, Y.-S., 2013, “Modeling and optimization of bioethanol production via a simultaneous saccharification and fermentation process using starch”, *Journal of Chemical Technology & Biotechnology*, v. 88, n. 6, pp. 1164–1174. doi: 10.1002/jctb.3961.
- JANG, S.-S., JOSEPH, B., MUKAI, H., 1987, “On-line optimization of constrained multivariable chemical processes”, *AIChE journal*, v. 33, n. 1, pp. 26–35.
- JÄSCHKE, J., SKOGESTAD, S., 2011, “NCO tracking and self-optimizing control in the context of real-time optimization”, *Journal of Process Control*, v. 21, n. 10, pp. 1407 – 1416. doi: 10.1016/j.jprocont.2011.07.001.
- JÄSCHKE, J., SKOGESTAD, S., 2012, “Optimal controlled variables for polynomial systems”, *Journal of Process Control*, v. 22, n. 1, pp. 167 – 179. doi: 10.1016/j.jprocont.2011.09.005.
- JÄSCHKE, J., SKOGESTAD, S., 2013, “Using Process Data for Finding Self-Optimizing Controlled Variables”. In: *Preprints of the 10th IFAC International Symposium on Dynamics and Control of Process Systems, Mumbai, India*, pp. 451–456.
- JESUS, N. J. C. D., 2011, *Otimização em Tempo Real em um Processo Industrial de Produção de Etileno*. Tese de Doutorado, UFRJ/ COPPE/ Programa de Engenharia Química.
- JIANG, T., CHEN, B., HE, X., et al., 2003, “Application of steady-state detection method based on wavelet transform”, *Computers & chemical engineering*, v. 27, n. 4, pp. 569–578.
- JIANG, X., LIU, P., LI, Z., 2014a, “Data reconciliation and gross error detection for operational data in power plants”, *Energy*, v. 75, pp. 14 – 23. ISSN: 0360-5442. doi: 10.1016/j.energy.2014.03.024.
- JIANG, X., LIU, P., LI, Z., 2014b, “Data reconciliation for steam turbine on-line performance monitoring”, *Applied Thermal Engineering*, v. 70, n. 1, pp. 122 – 130. doi: 10.1016/j.applthermaleng.2014.05.007.

- JOHNSON, A., 1987, “The control of fed-batch fermentation processes – A survey”, *Automatica*, v. 23, n. 6, pp. 691 – 705. doi: 10.1016/0005-1098(87)90026-4.
- JOHNSTON, L. P. M., KRAMER, M. A., 1995, “Maximum likelihood data rectification: Steady-state systems”, *AIChE Journal*, v. 41, n. 11, pp. 2415–2426. ISSN: 1547-5905. doi: 10.1002/aic.690411108.
- JUN, Z., ZHI-ZHONG, M., RUN-DA, J., 2015, “Real-time optimization based on {SCFO} for gold cyanidation leaching process”, *Chemical Engineering Science*, v. 134, pp. 467 – 476. doi: 10.1016/j.ces.2015.05.020.
- KABBAJ, N., POLIT, M., DAHHOU, B., et al., 2001, “Adaptive observers based fault detection and isolation for an alcoholic fermentation process”. In: *Emerging Technologies and Factory Automation, 2001. Proceedings. 8th IEEE International Conference on*, v. 2, pp. 669–673. doi: 10.1109/ETFA.2001.997754.
- KADAM, J., MARQUARDT, W., SCHLEGEL, M., et al., 2003, “Towards integrated dynamic real-time optimization and control of industrial processes”, *Proceedings foundations of computer-aided process operations (FOCAPO2003)*, pp. 593–596.
- KADAM, J., MARQUARDT, W., 2007, “Integration of Economical Optimization and Control for Intentionally Transient Process Operation”. In: Findeisen, R., Allgöwer, F., Biegler, L. T. (Eds.), *Assessment and Future Directions of Nonlinear Model Predictive Control*, v. 358, *Lecture Notes in Control and Information Sciences*, Springer Berlin Heidelberg, pp. 419–434. doi: 10.1007/978-3-540-72699-9_34.
- KADAM, J., SCHLEGEL, M., SRINIVASAN, B., et al., 2005, “Dynamic real-time optimization: from off-line numerical solution to measurement-based implementation”. In: *16th IFAC World Congress, Prague*.
- KADLEC, P., GABRYS, B., STRANDT, S., 2009, “Data-driven Soft Sensors in the process industry”, *Computers & Chemical Engineering*, v. 33, n. 4, pp. 795 – 814. doi: 10.1016/j.compchemeng.2008.12.012.
- KALL, P., WALLACE, S. W., 1994, *Stochastic Programming*. John Wiley & Sons Ltd, Chichester. ISBN: 9780471951087.
- KARIWALA, V., CAO, Y., 2010, “Bidirectional Branch and Bound for Controlled Variable Selection Part III: Local Average Loss Minimization”, *Industrial Informatics, IEEE Transactions on*, v. 6, n. 1, pp. 54–61. doi: 10.1109/TII.2009.2037494.

- KARIWALA, V., 2007, “Optimal Measurement Combination for Local Self-Optimizing Control”, *Industrial & Engineering Chemistry Research*, v. 46, n. 11, pp. 3629–3634. doi: 10.1021/ie0610187.
- KARIWALA, V., CAO, Y., 2009, “Bidirectional branch and bound for controlled variable selection. Part II: Exact local method for self-optimizing control”, *Computers & Chemical Engineering*, v. 33, n. 8, pp. 1402 – 1412. doi: 10.1016/j.compchemeng.2009.01.014.
- KARIWALA, V., CAO, Y., JANARDHANAN†, S., 2008, “Local Self-Optimizing Control with Average Loss Minimization”, *Industrial & Engineering Chemistry Research*, v. 47, n. 4, pp. 1150–1158. doi: 10.1021/ie070897+.
- KASSMANN, D., BADGWELL, T., HAWKINS, R., 2002. “Robust Steady-state Target Calculation For Model Predictive Control”. US Patent 6,381,505 B1. Disponível em: <https://www.lens.org/lens/patent/US_6381505_B1>. Acesso em: 16 set. 2016.
- KASSMANN, D. E., BADGWELL, T. A., HAWKINS, R. B., 2000, “Robust steady-state target calculation for model predictive control”, *AIChE Journal*, v. 46, n. 5, pp. 1007–1024. doi: 10.1002/aic.690460513.
- KAZEMI, N., DUEVER, T. A., PENLIDIS, A., 2013, “A powerful estimation scheme with the error-in-variables-model for nonlinear cases: Reactivity ratio estimation examples”, *Computers & Chemical Engineering*, v. 48, pp. 200 – 208. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2012.08.015.
- KELLY, J. D., 2004, “Techniques for solving industrial nonlinear data reconciliation problems”, *Computers & Chemical Engineering*, v. 28, n. 12, pp. 2837 – 2843. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2004.06.009.
- KENNEDY, J., EBERHART, R., 1995, “Particle swarm optimization”. In: *Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia*, pp. 1942–1948, November.
- KIM, I.-W., LIEBMAN, M., EDGAR, T., 1991, “A sequential error-in-variables method for nonlinear dynamic systems”, *Computers & Chemical Engineering*, v. 15, n. 9, pp. 663 – 670. doi: 10.1016/0098-1354(91)87028-8.
- KIM, I.-W., LIEBMAN, M. J., EDGAR, T. F., 1990, “Robust error-in-variables estimation using nonlinear programming techniques”, *AIChE Journal*, v. 36, n. 7, pp. 985–993. ISSN: 1547-5905. doi: 10.1002/aic.690360703.

- KIM, I.-W., KANG, M. S., PARK, S., et al., 1997, “Robust data reconciliation and gross error detection: The modified MIMT using NLP”, *Computers & Chemical Engineering*, v. 21, n. 7, pp. 775 – 782. ISSN: 0098-1354. doi: 10.1016/S0098-1354(96)00304-3.
- KIM, M., YOON, S. H., DOMANSKI, P. A., et al., 2008, “Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner”, *International Journal of Refrigeration*, v. 31, n. 5, pp. 790–799.
- KIPARISSIDES, C., VERROS, G., PERTSINIDIS, A., 1994, “On-line optimization of a high-pressure low-density polyethylene tubular reactor”, *Chemical engineering science*, v. 49, n. 24, pp. 5011–5024.
- KLATT, K.-U., MARQUARDT, W., 2009, “Perspectives for process systems engineering – Personal views from academia and industry”, *Computers & Chemical Engineering*, v. 33, n. 3, pp. 536–550.
- KLAUS, S., 2016, *The Fourth Industrial Revolution*. World Economic Forum, Geneva, Switzerland.
- KLEINSCHRODT III, F., JONES, J., 1996, “Industrial vision for process optimization”, *Computers & Chemical Engineering*, v. 20, Supplement 1, n. 0, pp. S473 – S483. doi: 10.1016/0098-1354(96)00089-0. European Symposium on Computer Aided Process Engineering-6.
- KLIR, G., FOLGER, T., 1988, *Fuzzy sets, Uncertainty and Information*. Prentice Hall, Englewood Cliffs.
- KOMPALA, D. S., RAMKRISHNA, D., TSAO, G. T., 1984, “Cybernetic modeling of microbial growth on multiple substrates”, *Biotechnology and Bioengineering*, v. 26, n. 11, pp. 1272–1281. doi: 10.1002/bit.260261103.
- KONINCKX, J., 1988, *On-Line Optimization of Chemical Plants Using Steady State Models*. Tese de Doutorado, University of Maryland. Disponível em: <<http://hdl.handle.net/1903/4793>>.
- KOOKOS, I. K., 2003, “Optimal operation of batch processes under uncertainty: A Monte Carlo simulation-deterministic optimization approach”, *Industrial & engineering chemistry research*, v. 42, n. 26, pp. 6815–6822.
- KOOKOS, I. K., 2005, “Real-time regulatory control structure selection based on economics”, *Industrial & engineering chemistry research*, v. 44, n. 11, pp. 3993–4000.

- KÖPPEN, M., 2004, “No-Free-Lunch theorems and the diversity of algorithms”. In: *Evolutionary Computation, CEC2004. Congress on*, v. 1, pp. 235–241 Vol.1. doi: 10.1109/CEC.2004.1330862.
- KÖPPEN-SELIGER, B., FRANK, P., WOLFF, A., 1995, “Residual evaluation for fault detection and isolation with RCE neural networks”. In: *American Control Conference, Proceedings of the*, v. 5, pp. 3264–3268, Seattle, WA, USA, 21-23 June. IEEE. doi: 10.1109/ACC.1995.532206.
- KOUVARITAKIS, B., CANNON, M., 2001, *Nonlinear Predictive Control Theory and Practice*. IEE Control Engineering Series 61. The Institution of Electrical Engineers, London, UK.
- KRASŁAWSKI, A., 1989, “Review of applications of various types of uncertainty in chemical engineering”, *Chemical Engineering and Processing: Process Intensification*, v. 26, n. 3, pp. 185 – 191. doi: 10.1016/0255-2701(89)80016-9.
- KREYSZIG, E., 1989, *Introductory Functional Analysis with Applications*. John Wiley & Sons., Inc. ISBN: 0-471-03729-x.
- KRISHNAN, S., BARTON, G., PERKINS, J., 1992, “Robust parameter estimation in on-line optimization—part I. methodology and simulated case study”, *Computers & Chemical Engineering*, v. 16, n. 6, pp. 545 – 562. doi: 10.1016/0098-1354(92)80066-I.
- KRISHNAN, S., BARTON, G., PERKINS, J., 1993, “Robust parameter estimation in on-line optimization—part 2. Application to an industrial process”, *Computers & Chemical Engineering*, v. 17, n. 7, pp. 663 – 669. doi: 10.1016/0098-1354(93)80054-Q.
- KRIST, J., LAPÈRE, M., WASSINK, S. G., et al., 1994, “Generic System for On-Line Optimization & the Implementation in a Benzene plant”, *Computers & Chemical Engineering*, v. 18, pp. S517 – S524. ISSN: 0098-1354. doi: 10.1016/0098-1354(94)80084-7.
- KRIST, J. H., LAPERE, M. R., WASSINK, S. G., et al., 1996. “System For Real Time Optimization”. US Patent 5,486,995. Disponível em: <https://www.lens.org/lens/patent/US_5486995_A>. Acesso em: 16 set. 2016.
- KROUMOV, A. D., MÓDENES, A. N., TAIT, M. C. D. A., 2006, “Development of new unstructured model for simultaneous saccharification and fermentation of starch to ethanol by recombinant strain”, *Biochemical Engineering Journal*, v. 28, n. 3, pp. 243 – 255. doi: 10.1016/j.bej.2005.11.008.

- KU, W., STORER, R. H., GEORGAKIS, C., 1995, “Disturbance detection and isolation by dynamic principal component analysis”, *Chemometrics and Intelligent Laboratory Systems*, v. 30, n. 1, pp. 179 – 196. ISSN: 0169-7439. doi: 10.1016/0169-7439(95)00076-3. InCINC '94 Selected papers from the First International Chemometrics Internet Conference.
- KUEHN, D., DAVIDSON, H., 1961, “Computer control II. Mathematics of control”, *Chemical Engineering Progress*, v. 57, n. 6, pp. 44–47.
- KVÅLSETH, T. O., 1985, “Cautionary Note about R 2”, *The American Statistician*, v. 39, n. 4, pp. 279–285. doi: 10.1080/00031305.1985.10479448.
- KYRIAKOPOULOU, D., KALITVENTZEFF, B., 1996, “Validation of measurement data using an interior point SQP”, *Computers & Chemical Engineering*, v. 20, pp. S563 – S568. ISSN: 0098-1354. doi: 10.1016/0098-1354(96)00103-2.
- LÄMMEL, R., 2008, “Google’s MapReduce programming model – Revisited”, *Science of Computer Programming*, v. 70, n. 1, pp. 1 – 30. ISSN: 0167-6423. doi: 10.1016/j.scico.2007.07.001.
- LARENTIS, A. L., BENTES, JR., A. M. P., RESENDE, N. S. D., et al., 2003, “Analysis of experimental errors in catalytic tests for production of synthesis gas”, *Applied Catalysis A: General*, v. 242, n. 2, pp. 365 – 379. doi: 10.1016/S0926-860X(02)00525-2.
- LARSSON, T., GOVATSMARK, M. S., SKOGESTAD, S., et al., 2003, “Control Structure Selection for Reactor, Separator, and Recycle Processes”, *Industrial & Engineering Chemistry Research*, v. 42, n. 6, pp. 1225–1234. doi: 10.1021/ie0200860.
- LARSSON, T., SKOGESTAD, S., 2000, “Plantwide control - A review and a new design procedure”, *Modeling, Identification and Control*, v. 21, n. 4, pp. 209–240. doi: 10.4173/mic.2000.4.2.
- LARSSON, T., HESTETUN, K., HOVLAND, E., et al., 2001, “Self-Optimizing Control of a Large-Scale Plant: The Tennessee Eastman Process”, *Industrial & Engineering Chemistry Research*, v. 40, n. 22, pp. 4889–4901. doi: 10.1021/ie000586y.
- LASDON, L. S., 1970, *Optimization theory for large systems*. Macmillan, New York.

- LAUKS, U., VASBINDER, R., VALKENBURG, P., et al., 1992, “On-line optimization of an ethylene plant”, *Computers & Chemical Engineering*, v. 16, Supplement 1, n. 0, pp. S213 – S220. doi: 10.1016/S0098-1354(09)80025-2.
- ŁAWRYŃZUK, M., MARUSAK, P., TATJEWSKI, P., 2008, “Cooperation of model predictive control with steady-state economic optimisation”, *Control Cybern.*, v. 37, pp. 133–158.
- LEE, C. G., KIM, C. H., RHEE, S. K., 1992, “A kinetic model and simulation of starch saccharification and simultaneous ethanol fermentation by amyloglucosidase and *Zymomonas mobilis*”, *Bioprocess Engineering*, v. 7, n. 8, pp. 335–341. doi: 10.1007/BF00369488.
- LEE, J. H., 2014, “From robust model predictive control to stochastic optimal control and approximate dynamic programming: A perspective gained from a personal journey”, *Computers & Chemical Engineering*, v. 70, pp. 114 – 121. doi: 10.1016/j.compchemeng.2013.10.014. Manfred Morari Special Issue.
- LEE, J. H., COOLEY, B. L., 2000, “Min–max predictive control techniques for a linear state-space system with a bounded set of input matrices”, *Automatica*, v. 36, n. 3, pp. 463 – 473. doi: 10.1016/S0005-1098(99)00178-8.
- LEE, M. H., LEE, S. J., HAN, C., et al., 1998, “Hierarchical on-line data reconciliation and optimization for an industrial utility plant”, *Computers & Chemical Engineering*, v. 22, Supplement 1, pp. S247 – S254. ISSN: 0098-1354. doi: 10.1016/S0098-1354(98)00061-1.
- LENNOX, B., MONTAGUE, G., HIDEN, H., et al., 2001, “Process monitoring of an industrial fed-batch fermentation”, *Biotechnology and Bioengineering*, v. 74, n. 2, pp. 125–135. doi: 10.1002/bit.1102.
- LEONTARITIS, I. J., BILLINGS, S. A., 1985, “Input-output parametric models for non-linear systems Part I: deterministic non-linear systems”, *International Journal of Control*, v. 41, n. 2, pp. 303–328. doi: 10.1080/0020718508961129.
- LEYFFER, S., MAHAJAN, A., 2010, “Software For Nonlinearly Constrained Optimization”. In: Cochran, J. J., Cox, L. A., Keskinocak, P., et al. (Eds.), *Wiley Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, Inc. ISBN: 9780470400531. doi: 10.1002/9780470400531.eorms0570.

- LI, B., CHEN, B., WANG, J., et al., 2001, “Steady-state online data reconciliation in a crude oil distillation unit”, *Hydrocarbon Processing*, v. 80, n. 3, pp. 61–64. ISSN: 00188190.
- LI, B., CHEN, B., DUAN, D., et al., 2000, “Correction coefficient method for gross error detection based on temporal redundancy”, *Journal of Tsinghua University (Science And Technology)*, v. 40, n. 10, pp. 65–69.
- LI, P., ARELLANO-GARCIA, H., WOZNY, G., 2008, “Chance constrained programming approach to process optimization under uncertainty”, *Computers & Chemical Engineering*, v. 32, n. 1-2, pp. 25–45. doi: 10.1016/j.compchemeng.2007.05.009.
- LID, T., SKOGESTAD, S., 2008a, “Data reconciliation and optimal operation of a catalytic naphtha reformer”, *Journal of Process Control*, v. 18, n. 3–4, pp. 320 – 331. ISSN: 0959-1524. doi: 10.1016/j.jprocont.2007.09.002. Festschrift honouring Professor Dale Seborg.
- LID, T., SKOGESTAD, S., 2008b, “Scaled steady state models for effective on-line applications”, *Computers & Chemical Engineering*, v. 32, n. 4–5, pp. 990 – 999. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2007.04.003. Festschrift devoted to Rex Reklaitis on his 65th Birthday.
- LID, T., STRAND, S., 1997, “Real-time optimization of a cat cracker unit”, *Computers & Chemical Engineering*, v. 21, Supplement, n. 0, pp. S887 – S892. doi: 10.1016/S0098-1354(97)87614-4.
- LID, T., STRAND, S., SKOGESTAD, S., 2002, “On-line optimization of a crude unit heat exchanger network”. In: *AICHE SYMPOSIUM SERIES*, pp. 403–407. New York; American Institute of Chemical Engineers; 1998.
- LIEBMAN, M., EDGAR, T., 1988, “Data reconciliation for nonlinear processes”. In: *AIChE Annual Meeting, Washington, DC*.
- LIEBMAN, M., EDGAR, T., LASDON, L., 1992, “Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques”, *Computers & Chemical Engineering*, v. 16, n. 10, pp. 963 – 986. ISSN: 0098-1354. doi: 10.1016/0098-1354(92)80030-D.
- LIN, B., RECKE, B., KNUDSEN, J. K., et al., 2007, “A systematic approach for soft sensor development”, *Computers & Chemical Engineering*, v. 31, n. 5–6, pp. 419 – 425. doi: 10.1016/j.compchemeng.2006.05.030. Selected Papers from the 15th European Symposium on Computer Aided Process Engineering, Barcelona, Spain, May 29-June 1, 2005.

- LIPORACE, F. S., GOMES, M. V., KATATA, A. C., et al., 2009, “PETROBRAS Experience Implementing Real Time Optimization”. In: de Brito Alves, R. M., do Nascimento, C. A. O., Biscaia, E. C. (Eds.), *10th International Symposium on Process Systems Engineering: Part A*, v. 27, *Computer Aided Chemical Engineering*, Elsevier, pp. 1245 – 1250. doi: 10.1016/S1570-7946(09)70598-X.
- LIU, D., XU, L., XIONG, W., et al., 2014, “Fermentation Process Modeling with Levenberg-Marquardt Algorithm and Runge-Kutta Method on Ethanol Production by *Saccharomyces cerevisiae*”, *Mathematical Problems in Engineering*, v. 2014, pp. 10. doi: 10.1155/2014/289492.
- LIU, J., 2013, “Moving horizon state estimation for nonlinear systems with bounded uncertainties”, *Chemical Engineering Science*, v. 93, pp. 376 – 386. doi: 10.1016/j.ces.2013.02.030.
- LOCHER, G., SONNLEITNER, B., FIECHTER, A., 1992, “High-Tech Analyses in Modern Biotechnology On-line measurement in biotechnology: Techniques”, *Journal of Biotechnology*, v. 25, n. 1, pp. 23 – 53. doi: 10.1016/0168-1656(92)90108-L.
- LOEBLEIN, C., PERKINS, J. D., 1999, “Structural design for on-line process optimization: I. Dynamic economics of MPC”, *AIChE Journal*, v. 45, n. 5, pp. 1018–1029. doi: 10.1002/aic.690450511.
- LOEBLEIN, C., PERKINS, J., 1996, “Economic analysis of different structures of on-line process optimization systems”, *Computers & Chemical Engineering*, v. 20, Supplement 1, n. 0, pp. S551 – S556. doi: 10.1016/0098-1354(96)00101-9. European Symposium on Computer Aided Process Engineering-6.
- LOEBLEIN, C., PERKINS, J., 1998, “Economic analysis of different structures of on-line process optimization systems”, *Computers & Chemical Engineering*, v. 22, n. 9, pp. 1257 – 1269. doi: 10.1016/S0098-1354(98)00010-6.
- LOEBLEIN, C., PERKINS, J., SRINIVASAN, B., et al., 1999, “Economic performance analysis in the design of on-line batch optimization systems”, *Journal of Process Control*, v. 9, n. 1, pp. 61 – 78. doi: 10.1016/S0959-1524(98)00034-1.
- LOURENÇO, N. D., LOPES, J. A., ALMEIDA, C. F., et al., 2012, “Bioreactor monitoring with spectroscopy and chemometrics: a review”, *Analytical*

and Bioanalytical Chemistry, v. 404, n. 4, pp. 1211–1237. doi: 10.1007/s00216-012-6073-9.

- LU, J. Z., 2013. “Apparatus and method for real-time sequential quadratic programming in industrial process control systems”. US 2013/0282146 A1. Disponível em: <https://www.lens.org/lens/patent/US_2013_0282146_A1>. Acesso em: 16 set. 2016.
- LUCIA, A., XU, J., 1990, “Chemical process optimization using Newton-like methods”, *Computers & Chemical Engineering*, v. 14, n. 2, pp. 119 – 138. ISSN: 0098-1354. doi: 10.1016/0098-1354(90)87072-W.
- LUCIA, A., XU, J., D’COUTO, G. C., 1993, “Sparse quadratic programming in chemical process optimization”, *Annals of Operations Research*, v. 42, n. 1, pp. 55–83. ISSN: 1572-9338. doi: 10.1007/BF02023172.
- LUYBEN, M. L., TYREUS, B. D., LUYBEN, W. L., 1997, “Plantwide control design procedure”, *AIChE Journal*, v. 43, n. 12, pp. 3161–3174. doi: 10.1002/aic.690431205.
- MACDONALD, R. J., HOWAT, C. S., 1988, “Data reconciliation and parameter estimation in plant performance analysis”, *AIChE Journal*, v. 34, n. 1, pp. 1–8. ISSN: 1547-5905. doi: 10.1002/aic.690340102.
- MACGREGOR, J., BRUWER, M., MILETIC, I., et al., 2015, “Latent Variable Models and Big Data in the Process Industries”, *IFAC-PapersOnLine*, v. 48, n. 8, pp. 520 – 524. ISSN: 2405-8963. doi: 10.1016/j.ifacol.2015.09.020. 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015.
- MADRON, F., 1992, *Process plant performance: Measurement and data processing for optimization and retrofits*. Ellis Horwood.
- MAGALHÃES, O. I. D. B., 2010, *Desenvolvimento de um Sistema de Otimização Dinâmica em Tempo Real*. Tese de Doutorado, UFRJ/ COPPE/ Programa de Engenharia Química.
- MAH, R. S., STANLEY, G. M., DOWNING, D. M., 1976, “Reconciliation and Rectification of Process Flow and Inventory Data”, *Industrial & Engineering Chemistry Process Design and Development*, v. 15, n. 1, pp. 175–183. doi: 10.1021/i260057a030.

- MAH, R. S., 1990, *Chemical Process Structures and Information Flows*. Butterworth-Heinemann. ISBN: 978-0-409-90175-7. doi: 10.1016/B978-0-409-90175-7.50006-2.
- MAHULI, S. K., RHINEHART, R. R., RIGGS, J. B., 1992, “Experimental demonstration of non-linear model-based in-line control of pH”, *Journal of Process Control*, v. 2, n. 3, pp. 145 – 153. doi: 10.1016/0959-1524(92)85004-G.
- MAIA, J. G. S. S., DEMUNER, R. B., SECCHI, A. R., et al., 2018, “Process Modeling and Simulation of an Industrial-Scale Plant for Green Ethylene Production”, *Industrial & Engineering Chemistry Research*. doi: 10.1021/acs.iecr.8b00776.
- MAIORELLA, B., BLANCH, H. W., WILKE, C. R., 1983, “Byproduct inhibition effects on ethanolic fermentation by *Saccharomyces cerevisiae*”, *Biotechnology and Bioengineering*, v. 25, n. 1, pp. 103–121. doi: 10.1002/bit.260250109.
- MANDUR, J., BUDMAN, H., 2014, “Robust optimization of chemical processes using Bayesian description of parametric uncertainty”, *Journal of Process Control*, v. 24, n. 2, pp. 422 – 430. doi: 10.1016/j.jprocont.2013.10.004. ADCHEM 2012 Special Issue.
- MANSOUR, M., ELLIS, J., 2003, “Comparison of methods for estimating real process derivatives in on-line optimization”, *Applied Mathematical Modelling*, v. 27, n. 4, pp. 275 – 291. doi: 10.1016/S0307-904X(02)00124-5.
- MANSOUR, M., ELLIS, J., 2008, “Methodology of on-line optimisation applied to a chemical reactor”, *Applied Mathematical Modelling*, v. 32, n. 2, pp. 170–184. doi: 10.1016/j.apm.2006.11.014.
- MANUM, H., SKOGESTAD, S., 2012, “Self-optimizing control with active set changes”, *Journal of Process Control*, v. 22, n. 5, pp. 873 – 883. doi: 10.1016/j.jprocont.2012.02.015.
- MANYIKA, J., CHUI, M., BROWN, B., et al., 2011, “Big Data: The Next Frontier for Innovation, Competition, and Productivity”, *McKinsey Global Institute Report*, (June). Disponível em: <http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation>. Acesso em: 2 jun. 2015.

- MARCHETTI, A., CHACHUAT, B., BONVIN, D., 2010, “A dual modifier-adaptation approach for real-time optimization”, *Journal of Process Control*, v. 20, n. 9, pp. 1027–1037.
- MARCHETTI, A., 2013, “A new dual modifier-adaptation approach for iterative process optimization with inaccurate models”, *Computers & Chemical Engineering*, v. 59, pp. 89–100.
- MARCHETTI, A., CHACHUAT, B., BONVIN, D., 2009, “Modifier-adaptation methodology for real-time optimization”, *Industrial & engineering chemistry research*, v. 48, n. 13, pp. 6022–6033.
- MARCHETTI, A. G., 2009, *Modifier-adaptation methodology for real-time optimization*. Tese de Doutorado, STI, Lausanne.
- MARLIN, T. E., HRYMAK, A. N., 1997, “Real-time operations optimization of continuous processes”. In: *AIChE Symposium Series*, v. 93, pp. 156–164. New York, NY: American Institute of Chemical Engineers, 1971-c2002.
- MARONNA, R. A., MARTIN, R. D., YOHAI, V. J., 2006, *Robust Statistics: Theory and Methods*. John Wiley & Sons, Ltd. ISBN: 0-470-01092-4.
- MARTIN, R., 1999, *Large Scale Linear and Integer Optimization: A Unified Approach*. Springer US. doi: 10.1007/978-1-4615-4975-8.
- MARTINI, A., SORCE, A., TRAVERSO, A., et al., 2013, “Data Reconciliation for power systems monitoring: Application to a microturbine-based test rig”, *Applied Energy*, v. 111, pp. 1152 – 1161. ISSN: 0306-2619. doi: 10.1016/j.apenergy.2012.12.045.
- MARTÍNEZ-MARADIAGA, D., BRUNO, J. C., CORONAS, A., 2013, “Steady-state data reconciliation for absorption refrigeration systems”, *Applied Thermal Engineering*, v. 51, n. 1–2, pp. 1170 – 1180. ISSN: 1359-4311. doi: 10.1016/j.applthermaleng.2012.10.027.
- MATLAB, 2010, *version 7.10.0 (R2010a)*. Natick, Massachusetts, The MathWorks Inc.
- MAYNE, D. Q., 2014, “Model predictive control: Recent developments and future promise”, *Automatica*, v. 50, n. 12, pp. 2967 – 2986. doi: 10.1016/j.automatica.2014.10.128.
- MAYNE, D., RAWLINGS, J., RAO, C., et al., 2000, “Constrained model predictive control: Stability and optimality”, *Automatica*, v. 36, n. 6, pp. 789 – 814. doi: 10.1016/S0005-1098(99)00214-9.

- MCBRAYER, K. F., EDGAR, T. F., 1995, “Bias detection and estimation in dynamic data reconciliation”, *Journal of Process Control*, v. 5, n. 4, pp. 285 – 289. doi: 10.1016/0959-1524(95)00020-Q.
- MCBRAYER, K. F., SODERSTROM, T. A., EDGAR, T. F., et al., 1998, “The application of nonlinear dynamic data reconciliation to plant data”, *Computers & Chemical Engineering*, v. 22, n. 12, pp. 1907 – 1911. ISSN: 0098-1354. doi: 10.1016/S0098-1354(98)00224-5.
- MCFARLANE, R. C., BACON, D. W., 1989, “Empirical strategies for open-loop on-line optimization”, *The Canadian Journal of Chemical Engineering*, v. 67, n. 4, pp. 665–677. doi: 10.1002/cjce.5450670421.
- MEJÍA, R. I., DUARTE, M. B., TRIERWEILER, J. O., 2010, “Avaliação do desempenho de métodos de identificação de estado estacionário”. In: *XVIII Congresso Brasileiro de Automática-CBA2010*, pp. 4634–4641.
- MELEIRO, L., MACIEL FILHO, R., 2000, “A self-tuning adaptive control applied to an industrial large scale ethanol production”, *Computers & Chemical Engineering*, v. 24, n. 2–7, pp. 925 – 930. doi: 10.1016/S0098-1354(00)80005-8.
- MENDOZA, D. F., GRACIANO, J. E. A., DOS SANTOS LIPORACE, F., et al., 2016, “Assessing the reliability of different real-time optimization methodologies”, *The Canadian Journal of Chemical Engineering*, v. 94, n. 3, pp. 485–497. ISSN: 1939-019X. doi: 10.1002/cjce.22402.
- MERCANGÖZ, M., DOYLE, III, F. J., 2008, “Real-time optimization of the pulp mill benchmark problem”, *Computers & Chemical Engineering*, v. 32, n. 4–5, pp. 789 – 804. doi: 10.1016/j.compchemeng.2007.03.004. Festschrift devoted to Rex Reklaitis on his 65th Birthday.
- MESAROVIC, M. D., MACKO, D., TAKAHARA, Y., 1970, *Theory of Hierarchical, Multilevel, Systems*. Academic Press, New York.
- MESFIN, G., SHUHAIMI, M., 2010, “A chance constrained approach for a gas processing plant with uncertain feed conditions”, *Computers & Chemical Engineering*, v. 34, n. 8, pp. 1256 – 1267. doi: 10.1016/j.compchemeng.2010.03.009.
- MEYER, M., KOEHRET, B., ENJALBERT, M., 1993, “Data reconciliation on multicomponent network process”, *Computers & Chemical Engineering*, v. 17, n. 8, pp. 807 – 817. doi: 10.1016/0098-1354(93)80065-U.

- MILETIC, I., MARLIN, T., 1998, “On-line statistical results analysis in real-time operations optimization”, *Industrial & engineering chemistry research*, v. 37, n. 9, pp. 3670–3684.
- MILETIC, I., MARLIN, T., 1996, “Results analysis for real-time optimization (RTO): Deciding when to change the plant operation”, *Computers & Chemical Engineering*, v. 20, Supplement 2, n. 0, pp. S1077 – S1082. doi: 10.1016/0098-1354(96)00187-1. European Symposium on Computer Aided Process Engineering-6.
- MINGFANG, K., BINGZHEN, C., BO, L., 2000, “An Integral approach to dynamic data rectification”, *Computers & Chemical Engineering*, v. 24, n. 2–7, pp. 749 – 753. doi: 10.1016/S0098-1354(00)00378-1.
- MORAD, K., YOUNG, B. R., SVRCEK, W. Y., 2005, “Rectification of plant measurements using a statistical framework”, *Computers & Chemical Engineering*, v. 29, n. 5, pp. 919 – 940. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2004.07.019.
- MORARI, M., ZAFIRIOU, E., 1989, *Robust Process Control*. Prentice Hall, Englewood Cliffs.
- MORARI, M., LEE, J. H., 1999, “Model predictive control: past, present and future”, *Computers & Chemical Engineering*, v. 23, n. 4–5, pp. 667 – 682. doi: 10.1016/S0098-1354(98)00301-9.
- MORARI, M., ARKUN, Y., STEPHANOPOULOS, G., 1980, “Studies in the synthesis of control structures for chemical processes: Part I: Formulation of the problem. Process decomposition and the classification of the control tasks. Analysis of the optimizing control structures”, *AIChE Journal*, v. 26, n. 2, pp. 220–232. doi: 10.1002/aic.690260205.
- MORENO, R. D. P., 2010, *Steady state detection, data reconciliation, and gross error detection: Development for industrial processes*. Tese de Mestrado, University of New Brunswick. Disponível em: <http://www.ece.unb.ca/jtaylor/Theses/Pilar_thesis.pdf>. Acesso em: 20 mar. 2014.
- MORSHEDI, A., CUTLER, C., SKROVANEK, T., 1985, “Optimal Solution of Dynamic Matrix Control with Linear Programming Techniques (LDMC)”. In: *American Control Conference*, pp. 199–208.
- MORÉ, J. J., WRIGHT, S. J., 1993, “12. Software Classification”. In: *Optimization Software Guide*, cap. 12, pp. 67–149. doi: 10.1137/1.9781611970951.pt2.

- MOSIER, N. S., ILELEJI, K. E., 2014, “Bioenergy: Biomass to Biofuels”. 1 ed., cap. How fuel ethanol is made from corn, pp. 379–384, Academic Press. ISBN: 9780124079090.
- MUKAI, H., SINGH, J., SPARE, J., et al., 1981, “A Reevaluation of the Normal Operating State Control of the Power System Using Computer Control and System Theory Part II: Dispatch Targeting”, *Power Apparatus and Systems, IEEE Transactions on*, v. PAS-100, n. 1 (Jan), pp. 309–317. doi: 10.1109/TPAS.1981.316844.
- MUKHERJEE, V., STEENSELS, J., LIEVENS, B., et al., 2014, “Phenotypic evaluation of natural and industrial *Saccharomyces* yeasts for different traits desirable in industrial bioethanol production”, *Applied Microbiology and Biotechnology*, v. 98, n. 22, pp. 9483–9498. doi: 10.1007/s00253-014-6090-z.
- MÜLLER, D., ESCHE, E., LÓPEZ, C., D. C., et al., 2014, “An algorithm for the identification and estimation of relevant parameters for optimization under uncertainty”, *Computers & Chemical Engineering*, v. 71, pp. 94 – 103. doi: 10.1016/j.compchemeng.2014.07.007.
- MULVEY, J. M., VANDERBEI, R. J., ZENIOS, S. A., 1995, “Robust Optimization of Large-Scale Systems”, *Operations Research*, v. 43, n. 2, pp. 264–281. doi: 10.1287/opre.43.2.264.
- MURTAGH, B., SAUNDERS, M., 1998. “MINOS 5.5 user’s guide”. Jul. Disponível em: <<http://www.sbsi-sol-optimize.com/manuals/Minos%20Manual.pdf>>. Acesso em: 6 mai. 2017.
- MURTHY, G. S., RAUSCH, K. D., JOHNSTON, D. B., et al., 2011, “RESEARCH: Industrial evaluation of a dynamic controller for simultaneous saccharification and fermentation process”, *Industrial Biotechnology*, v. 7, pp. 298–307. doi: 10.1089/ind.2011.7.298.
- MURTHY, G. S., JOHNSTON, D. B., RAUSCH, K. D., et al., 2012a, “A simultaneous saccharification and fermentation model for dynamic growth environments”, *Bioprocess and Biosystems Engineering*, v. 35, n. 4, pp. 519–534. doi: 10.1007/s00449-011-0625-9.
- MURTHY, G. S., JOHNSTON, D. B., RAUSCH, K. D., et al., 2012b, “Design and Evaluation of an Optimal Controller for Simultaneous Saccharification and Fermentation Process”, *Applied Biochemistry and Biotechnology*, v. 166, n. 1, pp. 87–111. doi: 10.1007/s12010-011-9406-9.

- MURTHY, G. S., SINGH, V., 2011. “Dynamic Fermentation Controller”. US 7,862,992 B2. Disponível em: <https://www.lens.org/lens/patent/US_7862992_B2>. Acesso em: 16 set. 2016.
- MUSKE, K., 1997, “Steady-state target optimization in linear model predictive control”. In: *American Control Conference, 1997. Proceedings of the 1997*, v. 6, pp. 3597–3601. doi: 10.1109/ACC.1997.609493.
- MYERS, R. H., MONTGOMERY, D. C., ANDERSON-COOK, C. M., 2009, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments. – 3rd ed.* Wiley series in probability and statistics. John Wiley & Sons, Inc., Hoboken, New Jersey.
- NARASIMHAN, S., MAH, R., 1989, “Treatment of general steady state process models in gross error identification”, *Computers & Chemical Engineering*, v. 13, n. 7, pp. 851 – 853. doi: 10.1016/0098-1354(89)85056-2.
- NARASIMHAN, S., KAO, C. S., MAH, R., 1987, “Detecting changes of steady states using the mathematical theory of evidence”, *AIChE journal*, v. 33, n. 11, pp. 1930–1932.
- NARASIMHAN, S., BHATT, N., 2015, “Deconstructing principal component analysis using a data reconciliation perspective”, *Computers & Chemical Engineering*, v. 77, pp. 74 – 84. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2015.03.016.
- NARASIMHAN, S., JORDACHE, C., 1999, *Data reconciliation and gross error detection: An intelligent use of process data.* Gulf Professional Publishing. ISBN: 978-0-88415-255-2.
- NARRAWAY, L., PERKINS, J.D. AND BARTON, G., 1991, “Interaction between process design and process control: economic analysis of process dynamics”, *Journal of Process Control*, v. 1, n. 5, pp. 243 – 250. doi: 10.1016/0959-1524(91)85015-B.
- NATH, R., ALZEIN, Z., 2000, “On-line dynamic optimization of olefins plants”, *Computers & Chemical Engineering*, v. 24, n. 2–7, pp. 533 – 538. doi: 10.1016/S0098-1354(00)00525-1.
- NATIONAL RESEARCH COUNCIL, 2013, *Frontiers in Massive Data Analysis.* Washington, DC, The National Academies Press. ISBN: 978-0-309-28778-4. doi: 10.17226/18374.

- NAVIA, D., MARTÍ, R., SARABIA, D., et al., 2012, “Handling infeasibilities in dual modifier-adaptation methodology for real-time optimization”. In: *8th IFAC Symposium on Advanced Control of Chemical Processes, Singapore*, pp. 537–542.
- NAVIA, D., GUTIÉRREZ, G., DE PRADA, C., 2013, “Nested Modifier-Adaptation for RTO in the Otto Williams Reactor”. In: *Proc. IFAC Symp. DYCOPS, Mumbai, India*, pp. 123–128.
- NAYSMITH, M. R., DOUGLAS, P. L., 1995, “Review of Real Time Optimization in the Chemical Process Industries”, *Developments in Chemical Engineering and Mineral Processing*, v. 3, n. 2, pp. 67–87. ISSN: 1932-2143. doi: 10.1002/apj.5500030202.
- NELDER, J. A., MEAD, R., 1965, “A Simplex Method for Function Minimization”, *The Computer Journal*, v. 7, n. 4, pp. 308–313. doi: 10.1093/comjnl/7.4.308.
- NEMHAUSER, G., RINOY KAN, A., TODD, M., 1989, “Optimization”. v. 1, *Handbooks in Operations Research and Management Science*, Elsevier. Disponível em: <<http://www.sciencedirect.com/science/handbooks/09270507/1>>.
- NIEDERBERGER, J., ZECH, I. A., DA SILVA, J. A., et al., 2005, “PETROX – PETROBRAS’ PROCESS SIMULATOR”. In: *Proceedings of the EN-PROMER 2005 (2nd Mercosur Congress on Chemical Engineering, 4th Mercosur Congress on Process Systems Engineering)*, Costa Verde - RJ, Brasil, pp. 1 – 10.
- NIEDERBERGER, J., GAMA, M. S., DOS SANTOS, L. C., et al., 2009, “PETROX – PETROBRAS Technology in Process Simulation”. In: de Brito Alves, R. M., do Nascimento, C. A. O., Biscaia, E. C. (Eds.), *10th International Symposium on Process Systems Engineering: Part A*, v. 27, *Computer Aided Chemical Engineering*, Elsevier, pp. 675 – 680. doi: 10.1016/S1570-7946(09)70333-5.
- NOCEDAL, J., WRIGHT, S., 2006, *Numerical Optimization*. Springer-Verlag New York. doi: 10.1007/978-0-387-40065-5.
- NORKIN, M., COLOME, S. D., 2014. “Real-time Monitor For Wine Fermentation”. US Patent 8,794,049. Disponível em: <https://www.lens.org/lens/patent/US_8794049_B1>. Acesso em: 16 set. 2016.

- NOUNOU, M. N., BAKSHI, B. R., 1999, “On-line multiscale filtering of random and gross errors without process models”, *AIChE Journal*, v. 45, n. 5, pp. 1041–1058. ISSN: 1547-5905. doi: 10.1002/aic.690450513.
- OCHOA, S., YOO, A., REPKE, J.-U., et al., 2007, “Modeling and Parameter Identification of the Simultaneous Saccharification-Fermentation Process for Ethanol Production”, *Biotechnology Progress*, v. 23, n. 6, pp. 1454–1462. ISSN: 1520-6033. doi: 10.1021/bp0702119.
- ÖDMAN, P., JOHANSEN, C. L., OLSSON, L., et al., 2009, “On-line estimation of biomass, glucose and ethanol in *Saccharomyces cerevisiae* cultivations using in-situ multi-wavelength fluorescence and software sensors”, *Journal of Biotechnology*, v. 144, n. 2, pp. 102 – 112. doi: 10.1016/j.jbiotec.2009.08.018.
- OLSSON, L., SCHULZE, U., NIELSEN, J., 1998, “On-line bioprocess monitoring – an academic discipline or an industrial tool?” *TrAC Trends in Analytical Chemistry*, v. 17, n. 2, pp. 88 – 95. doi: 10.1016/S0165-9936(97)00125-8.
- ÖNÖZ, B., BAYAZIT, M., 2003, “The power of statistical tests for trend detection”, *Turkish Journal of Engineering and Environmental Sciences*, v. 27, n. 4, pp. 247–251.
- OSTERGAARD, S., OLSSON, L., NIELSEN, J., 2000, “Metabolic Engineering of *Saccharomyces cerevisiae*”, *Microbiology and Molecular Biology Reviews*, v. 64, n. 1, pp. 34–50. doi: 10.1128/MMBR.64.1.34-50.2000.
- OSTROVSKY, G., ZIYATDINOV, N., LAPTEVA, T., 2010, “One-stage optimization problem with chance constraints”, *Chemical Engineering Science*, v. 65, n. 7, pp. 2373 – 2381. doi: 10.1016/j.ces.2009.09.072.
- ÖZYURT, D. B., PIKE, R. W., 2004, “Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes”, *Computers & Chemical Engineering*, v. 28, n. 3, pp. 381 – 402. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2003.07.001.
- PAGES, A., PINGAUD, H., MEYER, M., et al., 1994, “A strategy for simultaneous data reconciliation and parameter estimation on process flowsheets”, *Computers & Chemical Engineering*, v. 18, pp. S223 – S227. ISSN: 0098-1354. doi: 10.1016/0098-1354(94)80037-5.
- PAIANO, A., LAGIOIA, G., 2016, “Energy potential from residual biomass towards meeting the EU renewable energy and climate targets. The Italian case”, *Energy Policy*, v. 91, pp. 161 – 173. doi: 10.1016/j.enpol.2015.12.039.

- PAIS, T. M., FOULQUIÉ-MORENO, M. R., HUBMANN, G., et al., 2013, “Comparative Polygenic Analysis of Maximal Ethanol Accumulation Capacity and Tolerance to High Ethanol Levels of Cell Proliferation in Yeast”, *PLoS Genet*, v. 9, n. 6 (06), pp. 1–18. doi: 10.1371/journal.pgen.1003548.
- PARSOPOULOS, K., VRAHATIS, M., 2002, “Recent approaches to global optimization problems through Particle Swarm Optimization”, *Natural Computing*, v. 1, n. 2, pp. 235–306. ISSN: 1572-9796. doi: 10.1023/A:1016568309421.
- PASCAL, F., CORRIOU, J., DAGOT, C., et al., 1993, “Dynamic simulation of an industrial fed-batch alcoholic workshop”, *Computers & Chemical Engineering*, v. 17, pp. S215 – S220. ISSN: 0098-1354. doi: 10.1016/0098-1354(93)80232-C. European Symposium on Computer Aided Process Engineering—2.
- PASCAL, F., DAGOT, C., PINGAUD, H., et al., 1995, “Modeling of an industrial alcohol fermentation and simulation of the plant by a process simulator”, *Biotechnology and Bioengineering*, v. 46, n. 3, pp. 202–217. doi: 10.1002/bit.260460304.
- PATI, Y. C., REZAIIFAR, R., KRISHNAPRASAD, P. S., 1993, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition”. In: *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, pp. 40–44. IEEE, 1-3 Nov. doi: 10.1109/ACSSC.1993.342465.
- PEREIRA, M. G., CAMACHO, C. F., FREITAS, M. A. V., et al., 2012, “The renewable energy market in Brazil: Current status and potential”, *Renewable and Sustainable Energy Reviews*, v. 16, n. 6, pp. 3786 – 3802. doi: 10.1016/j.rser.2012.03.024.
- PERKINS, J. D., 1998, “Plant-wide Optimization: Opportunities and Challenges”. In: *AIChE Symposium Series*, v. 94, pp. 15–26. New York, NY: American Institute of Chemical Engineers, 1971-c2002.
- PFAFF, G., FORBES, J. F., MCLELLAN, P. J., 2006, “Generating information for real-time optimization”, *Asia-Pacific Journal of Chemical Engineering*, v. 1, n. 1-2, pp. 32–43. doi: 10.1002/apj.5.
- PIBOULEAU, L., FLOQUET, P., DOMENECH, S., et al., 1999, “A survey of optimization tools through ESCAPE symposia”, *Computers & Chemical Engineering*

neering, v. 23, Supplement, pp. S495 – S498. doi: 10.1016/S0098-1354(99)80122-7.

- PICCOLO, M., DOUGLAS, P. L., LEE, P. L., 1996, “Data Reconciliation Using AspenPlus”, *Developments in Chemical Engineering and Mineral Processing*, v. 4, n. 3-4, pp. 157–182. ISSN: 1932-2143. doi: 10.1002/apj.5500040303.
- PIERUCCI, S., BRANDANI, P., RANZI, E., et al., 1996, “An industrial application of an on-line data reconciliation and optimization problem”, *Computers & Chemical Engineering*, v. 20, Supplement 2, pp. S1539 – S1544. ISSN: 0098-1354. doi: 10.1016/0098-1354(96)00262-1. European Symposium on Computer Aided Process Engineering-6.
- PINTO, J. C., 1998, “On the costs of parameter uncertainties. Effects of parameter uncertainties during optimization and design of experiments”, *Chemical engineering science*, v. 53, n. 11, pp. 2029–2040.
- PINTO, J. C., 2001, “On the costs of parameter uncertainties. Part 2: Impact of EVOP procedures on the optimization and design of experiments”, *The Canadian Journal of Chemical Engineering*, v. 79, n. 3, pp. 412–421. ISSN: 1939-019X. doi: 10.1002/cjce.5450790314.
- PISTIKOPOULOS, E., 1995, “Uncertainty in process design and operations”, *Computers & Chemical Engineering*, v. 19, pp. 553–563.
- PISTIKOPOULOS, E., IERAPETRITOU, M., 1995, “Novel approach for optimal process design under uncertainty”, *Computers & Chemical Engineering*, v. 19, n. 10, pp. 1089 – 1110. doi: 10.1016/0098-1354(94)00093-4.
- PLÁCIDO, J., LOUREIRO, L., 1998, “Industrial application of data reconciliation”, *Computers & Chemical Engineering*, v. 22, pp. S1035 – S1038. doi: 10.1016/S0098-1354(98)00208-7.
- POKU, M. Y. B., BIEGLER, L. T., KELLY, J. D., 2004, “Nonlinear Optimization with Many Degrees of Freedom in Process Engineering”, *Industrial & Engineering Chemistry Research*, v. 43, n. 21, pp. 6803–6812. doi: 10.1021/ie0341000.
- POLI, R., KENNEDY, J., BLACKWELL, T., 2007, “Particle swarm optimization”, *Swarm Intelligence*, v. 1, n. 1, pp. 33–57. ISSN: 1935-3812. doi: 10.1007/s11721-007-0002-0.

- PONS, M.-N., 1993, “Biosensors for fermentation control”, *Current Opinion in Biotechnology*, v. 4, n. 2, pp. 183 – 187. doi: 10.1016/0958-1669(93)90122-D.
- PONTON, J., 1995, “Process systems engineering: Halfway through the first century”, *Chemical engineering science*, v. 50, n. 24, pp. 4045–4059.
- PONTRYAGIN, L., BOLTYANSKII, V., GAMKRELIDZE, R., et al., 1962, *The Mathematical Theory of Optimal Processes*. John Wiley & Sons, Inc.
- PORFÍRIO, C., ODLOAK, D., 2011, “Optimizing model predictive control of an industrial distillation column”, *Control Engineering Practice*, v. 19, n. 10, pp. 1137 – 1146. doi: 10.1016/j.conengprac.2011.06.003.
- PRATA, D. M., PINTO, J. C., LIMA, E. L., 2008a, “Análise Comparativa de Estimadores Robustos em Reconciliação de Dados Dinâmica e Não Linear”. In: *XVII Congresso Brasileiro de Engenharia Química - COBEQ, Recife - PE*, a.
- PRATA, D. M., 2009, *Reconciliação robusta de dados para monitoramento em tempo real*. Tese de Doutorado, UFRJ/ COPPE/ Programa de Engenharia Química.
- PRATA, D. M., LIMA, E. L., PINTO, J. C., 2006, “Simultaneous Data Reconciliation and Parameter Estimation in Bulk Polypropylene Polymerizations in Real Time”, *Macromolecular Symposia*, v. 243, n. 1, pp. 91–103. ISSN: 1521-3900. doi: 10.1002/masy.200651110.
- PRATA, D. M., LIMA, E. L., PINTO, J. C., 2008b, “In-Line Monitoring of Bulk Polypropylene Reactors Based on Data Reconciliation Procedures”, *Macromolecular Symposia*, v. 271, n. 1, pp. 26–37. ISSN: 1521-3900. doi: 10.1002/masy.200851104.
- PRATA, D. M., PINTO, J. C., LIMA, E. L., 2008c, “Comparative analysis of robust estimators on nonlinear dynamic data reconciliation”. In: Braunschweig, B., Joulia, X. (Eds.), *18th European Symposium on Computer Aided Process Engineering*, v. 25, *Computer Aided Chemical Engineering*, pp. 501 – 506. Elsevier, c. doi: 10.1016/S1570-7946(08)80088-0.
- PRATA, D. M., LIMA, E. L., PINTO, J. C., 2009a, “Nonlinear Dynamic Data Reconciliation in Real Time in Actual Processes”. In: Rita Maria de Brito Alves, C. A. O. d. N., Biscaia, E. C. (Eds.), *10th International Symposium on Process Systems Engineering: Part A*, v. 27, Elsevier, pp. 47 – 54, a. doi: 10.1016/S1570-7946(09)70228-7.

- PRATA, D. M., SCHWAAB, M., LIMA, E. L., et al., 2009b, “Nonlinear dynamic data reconciliation and parameter estimation through particle swarm optimization: Application for an industrial polypropylene reactor”, *Chemical Engineering Science*, v. 64, n. 18, pp. 3953 – 3967. ISSN: 0009-2509. doi: 10.1016/j.ces.2009.05.028.
- PRATA, D. M., SCHWAAB, M., LIMA, E. L., et al., 2010, “Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor”, *Chemical Engineering Science*, v. 65, n. 17, pp. 4943 – 4954. ISSN: 0009-2509. doi: 10.1016/j.ces.2010.05.017.
- PRÉKOPA, A., 1995, *Stochastic Programming*. Springer Netherlands. doi: 10.1007/978-94-017-3087-7.
- PUIG, S., VAN LOOSDRECHT, M., COLPRIM, J., et al., 2008, “Data evaluation of full-scale wastewater treatment plants by mass balance”, *Water Research*, v. 42, n. 18, pp. 4645 – 4655. doi: 10.1016/j.watres.2008.08.009.
- QIN, S. J., 2012, “Survey on data-driven industrial process monitoring and diagnosis”, *Annual Reviews in Control*, v. 36, n. 2, pp. 220 – 234. ISSN: 1367-5788. doi: 10.1016/j.arcontrol.2012.09.004.
- QIN, S. J., 2014, “Process data analytics in the era of big data”, *AIChE Journal*, v. 60, n. 9, pp. 3092–3100. ISSN: 1547-5905. doi: 10.1002/aic.14523.
- QIN, S. J., BADGWELL, T. A., 2003, “A survey of industrial model predictive control technology”, *Control Engineering Practice*, v. 11, n. 7, pp. 733 – 764. doi: 10.1016/S0967-0661(02)00186-7.
- QUELHAS, A. D., 2013, *Análise teórica e prática de Implementações de sistemas de otimização em tempo real (RTO)*. Tese de Doutorado, Programa de Engenharia Química, COPPE, UFRJ.
- QUELHAS, A. D., DE JESUS, N. J. C., PINTO, J. C., 2013, “Common vulnerabilities of RTO implementations in real chemical processes”, *The Canadian Journal of Chemical Engineering*, v. 91, n. 4, pp. 652–668. ISSN: 1939-019X. doi: 10.1002/cjce.21738.
- RAFIEE, A., BEHROUZSHAD, F., 2016, “Data reconciliation with application to a natural gas processing plant”, *Journal of Natural Gas Science and Engineering*, v. 31, pp. 538 – 545. doi: 10.1016/j.jngse.2016.03.071.

- RAMKRISHNA, D., KOMPALA, D. S., TSAO, G. T., 1987, “Are Microbes Optimal Strategists?” *Biotechnology Progress*, v. 3, n. 3, pp. 121–126. doi: 10.1002/btpr.5420030302.
- RAO, C. V., RAWLINGS, J. B., 1999, “Steady states and constraints in model predictive control”, *AIChE Journal*, v. 45, n. 6, pp. 1266–1278. ISSN: 1547-5905. doi: 10.1002/aic.690450612.
- RAWLINGS, J. B., AMRIT, R., 2009, “Optimizing Process Economic Performance Using Model Predictive Control”. In: Magni, L., Raimondo, D. M., Allgöwer, F. (Eds.), *Nonlinear Model Predictive Control*, v. 384, *Lecture Notes in Control and Information Sciences*, Springer Berlin Heidelberg, pp. 119–138. doi: 10.1007/978-3-642-01094-1_10.
- RAWLINGS, J. B., JI, L., 2012, “Optimization-based state estimation: Current status and some new results”, *Journal of Process Control*, v. 22, n. 8, pp. 1439 – 1444. doi: 10.1016/j.jprocont.2012.03.001. Ken Muske Special Issue.
- RAWLINGS, J. B., MAYNE, D. Q., 2009, *Model Predictive Control: Theory and Design*. Nob Hill Publishing, LLC. Disponível em: <<http://jbrwww.che.wisc.edu/home/jbraw/mpc/>>.
- REILLY, P. M., PATINO-LEAL, H., 1981, “A Bayesian Study of the Error-in-Variables Model”, *Technometrics*, v. 23, n. 3, pp. 221–231. doi: 10.1080/00401706.1981.10486290.
- REIMERS, C., WERTHER, J., GRUHN, G., 2008, “Flowsheet simulation of solids processes: Data reconciliation and adjustment of model parameters”, *Chemical Engineering and Processing: Process Intensification*, v. 47, n. 1, pp. 138 – 158. ISSN: 0255-2701. doi: 10.1016/j.cep.2007.07.015.
- REIS, M. S., GINS, G., 2017, “Industrial Process Monitoring in the Big Data/Industry 4.0 Era: from Detection, to Diagnosis, to Prognosis”, *Processes*, v. 5, n. 3, pp. 35. ISSN: 2227-9717. doi: 10.3390/pr5030035.
- RFA, 2016, *Ethanol Industry Outlook*. Relatório técnico, Renewable Fuel Association. Disponível em: <http://www.ethanolrfa.org/wp-content/uploads/2016/02/RFA_2016_full_final.pdf>. Acesso em: 6 mai. 2016.
- RICH, R. L., MYSZKA, D. G., 2003, “A survey of the year 2002 commercial optical biosensor literature”, *Journal of Molecular Recognition*, v. 16, n. 6, pp. 351–382. doi: 10.1002/jmr.649.

- RICHALET, J., RAULT, A., TESTUD, J., et al., 1978, “Model predictive heuristic control: Applications to industrial processes”, *Automatica*, v. 14, n. 5, pp. 413 – 428. doi: 10.1016/0005-1098(78)90001-8.
- ROBERTS, J. D., 1980, “Linear model reduction and solution of the algebraic Riccati equation by use of the sign function”, *International Journal of Control*, v. 32, n. 4, pp. 677–687. doi: 10.1080/00207178008922881.
- ROBERTS, P., 1978, “Algorithms for integrated system optimisation and parameter estimation”, *Electronics Letters*, v. 14, n. 6, pp. 196–197.
- ROBERTS, P., 1979, “An algorithm for steady-state system optimization and parameter estimation”, *International Journal of Systems Science*, v. 10, n. 7, pp. 719–734.
- ROBERTS, P., 1995, “Coping with model-reality differences in industrial process optimisation – a review of integrated system optimisation and parameter estimation (ISOPE)”, *Computers in Industry*, v. 26, n. 3, pp. 281–290.
- ROBERTS, P., WILLIAMS, T., 1981, “On an algorithm for combined system optimisation and parameter estimation”, *Automatica*, v. 17, n. 1, pp. 199–209.
- ROBERTSON, D. G., LEE, J. H., RAWLINGS, J. B., 1996, “A moving horizon-based approach for least-squares estimation”, *AIChE Journal*, v. 42, n. 8, pp. 2209–2224. doi: 10.1002/aic.690420811.
- ROD, V., HANČIL, V., 1980, “Iterative estimation of model parameters when measurements of all variables are subject to error”, *Computers & Chemical Engineering*, v. 4, n. 2, pp. 33 – 38. ISSN: 0098-1354. doi: 10.1016/0098-1354(80)80011-1.
- RODGER, E., 2010, *Dual modifier adaptation methodology for the on-line optimization of uncertain processes*. Tese de Mestrado, McMaster University. Disponível em: <<http://hdl.handle.net/11375/9191>>. Acesso em: 20 mar. 2014.
- RODGER, E. A., CHACHUAT, B., 2011, “Design methodology of modifier adaptation for on-line optimization of uncertain processes”. In: *Proceedings of 18th IFAC World Congress, Milano, Italy*, pp. 4113–4118.
- RODRÍGUEZ-BLANCO, T., SARABIA, D., PITARCH, J., et al., 2017, “Modifier Adaptation methodology based on transient and static measurements for

- RTO to cope with structural uncertainty”, *Computers & Chemical Engineering*, v. 106, n. Supplement C, pp. 480 – 500. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2017.07.001. ESCAPE-26.
- ROMAGNOLI, J. A., SÁNCHEZ, M. C., 2000, *Data processing and reconciliation for chemical process operations*, v. 2. Academic Press. ISBN: 978-0-12-594460-1.
- ROTAVA, O., ZANIN, A. C., 2005, “Multivariable control and real-time optimization – an industrial practical view”, *Hydrocarbon Processing*, v. 84, n. 6, pp. 61–71.
- RUCKAERT, M., MARTENS, X., DESARNAUTS, J., 1978, “Ethylene plant optimization by geometric programming”, *Computers & Chemical Engineering*, v. 2, n. 2–3, pp. 93 – 97. doi: 10.1016/0098-1354(78)80013-1.
- RYAN, T. P., 2011, “Evolutionary Operation”. In: *Statistical Methods for Quality Improvement, 3rd Edition*, cap. 15, John Wiley & Sons, Inc., Hoboken, New Jersey.
- SAFDARNEJAD, S. M., HEDENGREN, J. D., LEWIS, N. R., et al., 2015, “Initialization strategies for optimization of dynamic systems”, *Computers & Chemical Engineering*, v. 78, n. Supplement C, pp. 39 – 50. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2015.04.016.
- SAFDARNEJAD, S. M., GALLACHER, J. R., HEDENGREN, J. D., 2016, “Dynamic parameter estimation and optimization for batch distillation”, *Computers & Chemical Engineering*, v. 86, n. Supplement C, pp. 18 – 32. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2015.12.001.
- SAHINIDIS, N. V., 2004, “Optimization under uncertainty: state-of-the-art and opportunities”, *Computers & Chemical Engineering*, v. 28, n. 6, pp. 971–983.
- SANCHEZ, M., BANDONI, A., ROMAGNOLI, J., 1992, “Pladat: A package for process variable classification and plant data reconciliation”, *Computers & Chemical Engineering*, v. 16, pp. S499 – S506. doi: 10.1016/S0098-1354(09)80059-8.
- SARABIA, D., DE PRADA, C., GÓMEZ, E., et al., 2012, “Data reconciliation and optimal management of hydrogen networks in a petrol refinery”, *Control Engineering Practice*, v. 20, n. 4, pp. 343 – 354. ISSN: 0967-0661. doi: 10.1016/j.conengprac.2011.06.009. Special Section: {IFAC} Symposium on Advanced Control of Chemical Processes - {ADCHEM} 2009.

- SARGENT, R. W. H., DING, M., 2001, “A New SQP Algorithm for Large-Scale Nonlinear Programming”, *SIAM Journal on Optimization*, v. 11, n. 3, pp. 716–747. doi: 10.1137/S1052623496297012.
- SAYED, A. H., 2008, “Performance of SignError LMS”. In: *Adaptive Filters*, cap. 18, pp. 257–261, Wiley-Blackwell. ISBN: 9780470374122. doi: 10.1002/9780470374122.ch27.
- SCHLADT, M., HU, B., 2007, “Soft sensors based on nonlinear steady-state data reconciliation in the process industry”, *Chemical Engineering and Processing: Process Intensification*, v. 46, n. 11, pp. 1107–1115.
- SCHÖLKOPF, B., SMOLA, A., MÜLLER, K.-R., 1998, “Nonlinear Component Analysis as a Kernel Eigenvalue Problem”, *Neural Computation*, v. 10, n. 5, pp. 1299–1319. doi: 10.1162/089976698300017467.
- SCHUBERT, U., KRUGER, U., ARELLANO-GARCIA, H., et al., 2011, “Unified model-based fault diagnosis for three industrial application studies”, *Control Engineering Practice*, v. 19, n. 5, pp. 479 – 490. ISSN: 0967-0661. doi: 10.1016/j.conengprac.2011.01.009.
- SCHÜGERL, K., 2001, “Progress in monitoring, modeling and control of bioprocesses during the last 20 years”, *Journal of Biotechnology*, v. 85, n. 2, pp. 149 – 173. doi: 10.1016/S0168-1656(00)00361-8. Twenty years of the European Federation of Biotechnology.
- SCHULTZ, E., FARENZENA, M., 2015, “Seria Real-Time Optimization o Limite Superior Para Self-Optimizing Control?” *Anais do XX Congresso Brasileiro de Engenharia Química - COBEQ 2014 [= Blucher Chemical Engineering Proceedings, v.1, n.2]*, v. 1, n. 2, pp. 12155 – 12162. doi: 10.5151/chemeng-cobeq2014-1165-20699-176861.
- SCHWAAB, M., PINTO, J. C., 2007, *Análise de Dados Experimentais: I. Fundamentos de Estatística e Estimação de Parâmetros*, v. 1, Escola Piloto em Engenharia Química. Rio de Janeiro: E-papers. ISBN: 978-85-7650-136-7.
- SCHWAAB, M., PINTO, J. C., 2008, “Optimum reparameterization of power function models”, *Chemical Engineering Science*, v. 63, n. 18, pp. 4631 – 4635. doi: 10.1016/j.ces.2008.07.005.
- SCHWAAB, M., BISCAIA JR., E. C., MONTEIRO, J. L., et al., 2008, “Nonlinear parameter estimation through particle swarm optimization”, *Chemical Engineering Science*, v. 63, n. 6, pp. 1542 – 1552. doi: 10.1016/j.ces.2007.11.024.

- SEBORG, D. E., MELLICHAMP, D. A., EDGAR, T. F., et al., 2010, *Process dynamics and control*. John Wiley & Sons.
- SEQUEIRA, S. E., 2003, *Real Time Evolution (RTE) for on-line optimisation of continuous and semi-continuous chemical processes*. Tese de Doutorado, Universitat Politècnica de Catalunya.
- SEQUEIRA, S. E., GRAELLS, M., PUIGJANER, L., 2002, “Real-Time Evolution for On-line Optimization of Continuous Processes”, *Industrial & Engineering Chemistry Research*, v. 41, n. 7, pp. 1815–1825. doi: 10.1021/ie010464l.
- SERRALUNGA, F. J., MUSSATI, M. C., AGUIRRE, P. A., 2013, “Model Adaptation for Real-Time Optimization in Energy Systems”, *Industrial & Engineering Chemistry Research*, v. 52, n. 47, pp. 16795–16810. doi: 10.1021/ie303621j.
- SERRALUNGA, F. J., AGUIRRE, P. A., MUSSATI, M. C., 2014, “Including Disjunctions in Real-Time Optimization”, *Industrial & Engineering Chemistry Research*, v. 53, n. 44, pp. 17200–17213. doi: 10.1021/ie5004619.
- SERTH, R., VALERO, C., HEENAN, W., 1987, “Detection Of Gross Errors In Non-linearly Constrained Data: A Case Study”, *Chemical Engineering Communications*, v. 51, n. 1-6, pp. 89–104. doi: 10.1080/00986448708911836.
- SHANRUI, D., 2016. “A intelligent control system of anaerobic fermentation tank for producing fermented feed”. CN 205103615 U. Disponível em: <<https://goo.gl/QAHuu3>>. Acesso em: 16 set. 2016.
- SHAPIRO, V., MARKEVICH, I., KHOTS, D., 2015. “Method of large scale process optimization and optimal planning based on real time dynamic simulation”. US Patent 9,008,807 B2. Disponível em: <https://www.lens.org/lens/patent/US_9008807_B2>. Acesso em: 16 set. 2016.
- SHEN, J., AGBLEVOR, F. A., 2010, “The Operable Modeling of Simultaneous Saccharification and Fermentation of Ethanol Production from Cellulose”, *Applied Biochemistry and Biotechnology*, v. 160, n. 3, pp. 665–681. doi: 10.1007/s12010-009-8650-8.
- SHOBRYS, D. E., WHITE, D. C., 2002, “Planning, scheduling and control systems: why cannot they work together”, *Computers & Chemical Engineering*, v. 26, n. 2, pp. 149 – 160. doi: 10.1016/S0098-1354(01)00737-2.

- SHULER, M., 1985, “Invited Review On The Use Of Chemically Structured Models For Bioreactors”, *Chemical Engineering Communications*, v. 36, n. 1-6, pp. 161–189. doi: 10.1080/00986448508911252.
- SIROLA, J., EDGAR, T., 2012, “Process energy systems: Control, economic, and sustainability objectives”, *Computers & Chemical Engineering*, v. 47, pp. 134 – 144. doi: 10.1016/j.compchemeng.2012.06.019.
- SINGH, V., 2012, “Effect of corn quality on bioethanol production”, *Biocatalysis and Agricultural Biotechnology*, v. 1, n. 4, pp. 353 – 355. doi: 10.1016/j.bcab.2012.06.001.
- SIVAKESAVA, S., IRUDAYARAJ, J., DEMIRCI, A., 2001, “Monitoring a bioprocess for ethanol production using FT-MIR and FT-Raman spectroscopy”, *Journal of Industrial Microbiology and Biotechnology*, v. 26, n. 4, pp. 185–190. doi: 10.1038/sj.jim.7000124.
- SKOGESTAD, S., 2000a, “Plantwide control: the search for the self-optimizing control structure”, *Journal of Process Control*, v. 10, n. 5, pp. 487 – 507. doi: 10.1016/S0959-1524(00)00023-8.
- SKOGESTAD, S., 2000b, “Self-optimizing control: the missing link between steady-state optimization and control”, *Computers & Chemical Engineering*, v. 24, n. 2, pp. 569–575.
- SKOGESTAD, S., 2004, “Control structure design for complete chemical plants”, *Computers & Chemical Engineering*, v. 28, n. 1, pp. 219–234.
- SODERSTROM, T. A., EDGAR, T. F., RUSSO, L. P., et al., 2000, “Industrial Application of a Large-Scale Dynamic Data Reconciliation Strategy”, *Industrial & Engineering Chemistry Research*, v. 39, n. 6, pp. 1683–1693. doi: 10.1021/ie990798z.
- SPENDLEY, W., HEXT, G. R., HIMSWORTH, F. R., 1962, “Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation”, *Technometrics*, v. 4, n. 4, pp. 441–461. doi: 10.1080/00401706.1962.10490033.
- SRINIVASAN, B., BONVIN, D., 2007, “Real-Time Optimization of Batch Processes by Tracking the Necessary Conditions of Optimality”, *Industrial & Engineering Chemistry Research*, v. 46, n. 2, pp. 492–504. doi: 10.1021/ie0600487.

- SRINIVASAN, B., PRIMUS, C., BONVIN, D., et al., 2001, “Run-to-run optimization via control of generalized constraints”, *Control Engineering Practice*, v. 9, n. 8, pp. 911–919.
- SRINIVASAN, B., BONVIN, D., VISSER, E., et al., 2003, “Dynamic optimization of batch processes: II. Role of measurements in handling uncertainty”, *Computers & Chemical Engineering*, v. 27, n. 1, pp. 27–44. doi: 10.1016/S0098-1354(02)00117-5.
- SRINIVASAN, B., BIEGLER, L. T., BONVIN, D., 2008, “Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function”, *Computers & Chemical Engineering*, v. 32, n. 3, pp. 572–579.
- STANLEY, G. M., MAH, R. S. H., 1977, “Estimation of flows and temperatures in process networks”, *AIChE Journal*, v. 23, n. 5, pp. 642–650. doi: 10.1002/aic.690230504.
- STEPHANOPOULOS, G., NG, C., 2000, “Perspectives on the synthesis of plant-wide control structures”, *Journal of Process Control*, v. 10, n. 2–3, pp. 97 – 111. doi: 10.1016/S0959-1524(99)00023-2.
- STEPHANOPOULOS, G., REKLAITIS, G. V., 2011, “Process systems engineering: From Solvay to modern bio- and nanotechnology. A history of development, successes and prospects for the future”, *Chemical Engineering Science*, v. 66, n. 19, pp. 4272 – 4306. doi: 10.1016/j.ces.2011.05.049. Multiscale Simulation.
- STRAND, S., VELAND, L. H., 1997, “Model-predictive control and real-time optimization of a cat cracker unit”, *Modeling, Identification and Control*, v. 18, n. 2, pp. 155–175. doi: 10.4173/mic.1997.2.5.
- STRAUB, D. A., GROSSMANN, I., 1993, “Design optimization of stochastic flexibility”, *Computers & Chemical Engineering*, v. 17, n. 4, pp. 339 – 354. doi: 10.1016/0098-1354(93)80025-I.
- SUNDE, S., ÖIVIND BERG, 2003, “Data reconciliation and fault detection by means of plant-wide mass and energy balances”, *Progress in Nuclear Energy*, v. 43, n. 1, pp. 97 – 104. ISSN: 0149-1970. doi: 10.1016/S0149-1970(03)00015-5.
- SÁNCHEZ, M., SENTONI, G., SCHBIB, S., et al., 1996, “Gross measurements error detection/identification for an industrial ethylene reactor”, *Computers*

Chemical Engineering, v. 20, Supplement 2, pp. S1559 – S1564. ISSN: 0098-1354. doi: 10.1016/0098-1354(96)00265-7. European Symposium on Computer Aided Process Engineering-6.

SÁNCHEZ, M., ROMAGNOLI, J., 1996, “Use of orthogonal transformations in data classification-reconciliation”, *Computers & Chemical Engineering*, v. 20, n. 5, pp. 483 – 493. ISSN: 0098-1354. doi: 10.1016/0098-1354(95)00212-X.

TADEJ, W., TATJEWSKI, P., 2001, “Analysis of an Isope-Type Dual Algorithm for Optimizing Control and Nonlinear Optimization”, *International Journal of Applied Mathematics and Computer Science*, v. 11, n. 2, pp. 429–457.

TAKAMATSU, T., 1983, “The nature and role of process systems engineering”, *Computers & chemical engineering*, v. 7, n. 4, pp. 203–218.

TAKIYAMA, H., NAKA, Y., O’SHIMA, E., et al., 1991, “Sensor-Based Data Reconciliation Method and Application to the Pilot Plant”, *Journal of Chemical Engineering of Japan*, v. 24, n. 3, pp. 339–346. doi: 10.1252/jcej.24.339.

TAMHANE, A. C., MAH, R. S. H., 1985, “Data Reconciliation and Gross Error Detection in Chemical Process Networks”, *Technometrics*, v. 27, n. 4, pp. 409–422. doi: 10.1080/00401706.1985.10488080.

TATJEWSKI, P., 2002, “Iterative Optimizing Set-Point Control - The Basic Principle Redesigned”. In: *Proceedings of the 15th IFAC World Congress, Barcelona, Spain*, v. 15, pp. 21–26, July. doi: 10.3182/20020721-6-ES-1901.00994.

TAYLOR, J. H., DEL PILAR MORENO, R., 2013, “Nonlinear dynamic data reconciliation: In-depth case study”. In: *2013 IEEE International Conference on Control Applications (CCA)*, pp. 746–753, Aug. doi: 10.1109/CCA.2013.6662839.

TEIXEIRA, A. F., 2013, *Otimização da Produção de Poços de Petróleo com Gas Lift Contínuo*. Tese de Mestrado, Programa de Engenharia Química, COPPE, UFRJ.

TERRY, P. A., HIMMELBLAU, D. M., 1993, “Data rectification and gross error detection in a steady state process via artificial neural networks”, *Industrial & Engineering Chemistry Research*, v. 32, n. 12, pp. 3020–3028. doi: 10.1021/ie00024a010.

- THANGARAJ, R., PANT, M., ABRAHAM, A., et al., 2011, “Particle swarm optimization: Hybridization perspectives and experimental illustrations”, *Applied Mathematics and Computation*, v. 217, n. 12, pp. 5208 – 5226. doi: 10.1016/j.amc.2010.12.053.
- TJOA, I.-B., BIEGLER, L., 1992, “Reduced successive quadratic programming strategy for errors-in-variables estimation”, *Computers & Chemical Engineering*, v. 16, n. 6, pp. 523 – 533. ISSN: 0098-1354. doi: 10.1016/0098-1354(92)80064-G.
- TJOA, I., BIEGLER, L., 1991, “Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems”, *Computers & Chemical Engineering*, v. 15, n. 10, pp. 679 – 690. doi: 10.1016/0098-1354(91)85014-L.
- TONG, H., CROWE, C. M., 1995, “Detection of gross errors in data reconciliation by principal component analysis”, *AIChE Journal*, v. 41, n. 7, pp. 1712–1722. ISSN: 1547-5905. doi: 10.1002/aic.690410711.
- TØRVI, H., HERTZBERG, T., 1997, “Estimation of uncertainty in dynamic simulation results”, *Computers & Chemical Engineering*, v. 21, Supplement, n. 0, pp. S181 – S185. doi: 10.1016/S0098-1354(97)87499-6.
- TOSUKHOWONG, T., LEE, J. M., LEE, J. H., et al., 2004, “An introduction to a dynamic plant-wide optimization strategy for an integrated plant”, *Computers & Chemical Engineering*, v. 29, n. 1, pp. 199 – 208. doi: 10.1016/j.compchemeng.2004.07.028.
- TRESMONDI, A., DOMINGUES, A., MACIEL FILHO, R., 2001, “Online optimization integrated with online analyzers and multivariable predictive controller in industrial airlift reactors”. In: Gani, R., Jørgensen, S. B. (Eds.), *34th European Symposium of the Working Party on Computer Aided Process Engineering*, v. 9, Elsevier, pp. 805 – 810. doi: 10.1016/S1570-7946(01)80128-0.
- TRIERWEILER, J. O., 2014, “Real-Time Optimization of Industrial Processes”. In: Baillieul, J., Samad, T. (Eds.), *Encyclopedia of Systems and Control*, Springer London, pp. 1–11. doi: 10.1007/978-1-4471-5102-9_243-1.
- TUCKER, C. E., 2014, “Social Networks, Personalized Advertising, and Privacy Controls”, *Journal of Marketing Research*, v. 51, n. 5, pp. 546–562. doi: 10.1509/jmr.10.0355.

- ULAS, S., DIWEKAR, U. M., 2004, “Thermodynamic uncertainties in batch processing and optimal control”, *Computers & Chemical Engineering*, v. 28, n. 11, pp. 2245 – 2258. doi: 10.1016/j.compchemeng.2004.04.001.
- VACHHANI, P., RENGASWAMY, R., VENKATASUBRAMANIAN, V., 2001, “A framework for integrating diagnostic knowledge with nonlinear optimization for data reconciliation and parameter estimation in dynamic systems”, *Chemical Engineering Science*, v. 56, n. 6, pp. 2133 – 2148. doi: 10.1016/S0009-2509(00)00488-7.
- VALKÓ, P., VAJDA, S., 1987, “An extended Marquardt-type procedure for fitting error-in-variables models”, *Computers & Chemical Engineering*, v. 11, n. 1, pp. 37 – 43. ISSN: 0098-1354. doi: 10.1016/0098-1354(87)80004-2.
- VAN IMPE, J. F., 1996, “Power and limitations of model based bioprocess optimization”, *Mathematics and Computers in Simulation*, v. 42, n. 2–3, pp. 159 – 169. doi: 10.1016/0378-4754(95)00128-X. Mathematical Modelling and Simulation in Agriculture and Bio-Industries Proceedings of the 1st IMACS-IFAC Symposium Msu2SABI.
- VAN WIJK, R., POPE, M., 1992, “Advanced process control and on-line optimisation in shell refineries”, *Computers & Chemical Engineering*, v. 16, Supplement 1, pp. S69 – S80. doi: 10.1016/S0098-1354(09)80010-0.
- VAN ZYL, J. M., VAN RENSBURG, E., VAN ZYL, W. H., et al., 2011, “A Kinetic Model for Simultaneous Saccharification and Fermentation of Avicel With *Saccharomyces cerevisiae*”, *Biotechnology and Bioengineering*, v. 108, n. 4, pp. 924–933. doi: 10.1002/bit.23000.
- VANCE, A., 2009, “Hadoop, A Free Software Program, Finds Uses Beyond Search”, *The New York Times*, (March). Disponível em: <<http://www.nytimes.com/2009/03/17/technology/business-computing/17cloud.html>>. Acesso em: 19 ago. 2016.
- VANHATALO, E., KULAHCI, M., BERGQUIST, B., 2017, “On the structure of dynamic principal component analysis used in statistical process monitoring”, *Chemometrics and Intelligent Laboratory Systems*, v. 167, pp. 1 – 11. ISSN: 0169-7439. doi: 10.1016/j.chemolab.2017.05.016.
- VARNER, J., RAMKRISHNA, D., 1999, “The non-linear analysis of cybernetic models. Guidelines for model formulation”, *Journal of Biotechnology*, v. 71, n. 1–3, pp. 67 – 103. doi: 10.1016/S0168-1656(99)00016-4. Modelling in Biochemical Engineering.

- VAZ, A. I. F., VICENTE, L. N., 2007, “A particle swarm pattern search method for bound constrained global optimization”, *Journal of Global Optimization*, v. 39, n. 2, pp. 197–219. ISSN: 0925-5001. doi: 10.1007/s10898-007-9133-5.
- VENKATASUBRAMANIAN, V., RENGASWAMY, R., YIN, K., et al., 2003, “A review of process fault detection and diagnosis: Part I: Quantitative model-based methods”, *Computers & Chemical Engineering*, v. 27, n. 3, pp. 293 – 311. doi: 10.1016/S0098-1354(02)00160-6.
- VEVERKA, V. V., MADRON, F., 1997, *Material and energy balancing in the process industries: From microscopic balances to large plants*, v. 7, *Computer Aided Chemical Engineering*. Elsevier. Disponível em: <[//www.sciencedirect.com/science/bookseries/15707946/7](http://www.sciencedirect.com/science/bookseries/15707946/7)>.
- VIEIRA, R., EMBIRUÇU, M., SAYER, C., et al., 2003, “Control strategies for complex chemical processes. Applications in polymerization processes”, *Computers & Chemical Engineering*, v. 27, n. 8–9, pp. 1307 – 1327. ISSN: 0098-1354. doi: 10.1016/S0098-1354(03)00055-3. 2nd Pan American Workshop in Process Systems Engineering.
- VOJINOVIĆ, V., CABRAL, J., FONSECA, L., 2006, “Real-time bioprocess monitoring: Part I: In situ sensors”, *Sensors and Actuators B: Chemical*, v. 114, n. 2, pp. 1083 – 1091. doi: 10.1016/j.snb.2005.07.059.
- VOLPE, P. L. O., 1997, “Estudo da fermentação alcoólica de soluções diluídas de diferentes açúcares utilizando microcalorimetria de fluxo”, *Química Nova*, v. 20 (10), pp. 528 – 534. doi: 10.1590/S0100-40421997000500013.
- WANG, D., ROMAGNOLI, J. A., 2003, “A Framework for Robust Data Reconciliation Based on a Generalized Objective Function”, *Industrial & Engineering Chemistry Research*, v. 42, n. 13, pp. 3075–3084. doi: 10.1021/ie0206655.
- WANG, Y., CHU, J., ZHUANG, Y., et al., 2009, “Industrial bioprocess control and optimization in the context of systems biotechnology”, *Biotechnology Advances*, v. 27, n. 6, pp. 989 – 995. doi: 10.1016/j.biotechadv.2009.05.022. Biotechnology for the Sustainability of Human Society Invited Papers from {IBS} 2008.
- WANG, Z., ZHUGE, J., FANG, H., et al., 2001, “Glycerol production by microbial fermentation: A review”, *Biotechnology Advances*, v. 19, n. 3, pp. 201 – 223. doi: 10.1016/S0734-9750(01)00060-X.

- WEISS, G., ROMAGNOLI, J., ISLAM, K., 1996, “Data reconciliation—An industrial case study”, *Computers & Chemical Engineering*, v. 20, n. 12, pp. 1441 – 1449. doi: 10.1016/0098-1354(95)00241-3.
- WHITE, D., 1997, “Online optimization: what, where and estimating ROI: Process optimization: Process control and instrumentation”, *Hydrocarbon Processing*, v. 76, n. 6, pp. 43–51.
- WHITE, D., 1998, “Online Optimization: What have we learned?” *Hydrocarbon processing*, v. 77, n. 6, pp. 55–59.
- WHITE, D. C., 2005, “Advanced automation technology reduces refinery energy costs”, *Oil and Gas Journal*, v. 103, n. 37, pp. 45–53.
- WILLIAMS, T. J., OTTO, R. E., 1960, “A generalized chemical processing model for the investigation of computer control”, *American Institute of Electrical Engineers, Part I: Communication and Electronics, Transactions of the*, v. 79, n. 5, pp. 458–473. doi: 10.1109/TCE.1960.6367296.
- WOLPERT, D., MACREADY, W., 1997, “No free lunch theorems for optimization”, *Evolutionary Computation, IEEE Transactions on*, v. 1, n. 1, pp. 67–82. doi: 10.1109/4235.585893.
- WÜRTH, L., RAWLINGS, J., MARQUARDT, W., 2009, “Economic dynamic real-time optimization and nonlinear model predictive control”. In: *Proceedings of the 7th IFAC International Symposium on Advanced Control of Chemical Processes, Istanbul, Turkey*, pp. 219–224. doi: 10.3182/20090712-4-TR-2008.00033.
- YANG, X.-S., 2010, *Engineering Optimization: An Introduction with Metaheuristic Applications*. John Wiley & Sons, Inc. doi: 10.1002/9780470640425.
- YE, L., CAO, Y., LI, Y., et al., 2013, “Approximating Necessary Conditions of Optimality as Controlled Variables”, *Industrial & Engineering Chemistry Research*, v. 52, n. 2, pp. 798–808. doi: 10.1021/ie300654d.
- YE, L., CAO, Y., MA, X., et al., 2014, “A Novel Hierarchical Control Structure with Controlled Variable Adaptation”, *Industrial & Engineering Chemistry Research*, v. 53, n. 38, pp. 14695–14711. doi: 10.1021/ie500562t.
- YELCHURU, R., SKOGESTAD, S., 2012, “Convex formulations for optimal selection of controlled variables and measurements using Mixed Integer Quadratic Programming”, *Journal of Process Control*, v. 22, n. 6, pp. 995 – 1007. doi: 10.1016/j.jprocont.2012.04.013.

- YI, H.-S., HAN, C., 2004, “Industrial Application of MILP-Based Simultaneous Compensation to a Large-Scale Byproduct Gases Network in an Iron-and Steel-Making Plant”, *Industrial & Engineering Chemistry Research*, v. 43, n. 1, pp. 119–126. doi: 10.1021/ie030301s.
- YING, C.-M., JOSEPH, B., 1999, “Performance and stability analysis of LP-MPC and QP-MPC cascade control systems”, *AIChE Journal*, v. 45, n. 7, pp. 1521–1534. doi: 10.1002/aic.690450714.
- YIP, W. S., MARLIN, T. E., 2002, “Multiple data sets for model updating in real-time operations optimization”, *Computers & Chemical Engineering*, v. 26, n. 10, pp. 1345 – 1362. doi: 10.1016/S0098-1354(02)00016-9.
- YIP, W. S., MARLIN, T. E., 2003, “Designing plant experiments for real-time optimization systems”, *Control engineering practice*, v. 11, n. 8, pp. 837–845. doi: 10.1016/S0967-0661(02)00213-7.
- YIP, W., MARLIN, T. E., 2004, “The effect of model fidelity on real-time optimization performance”, *Computers & chemical engineering*, v. 28, n. 1, pp. 267–280.
- YIQING, L., XIGANG, Y., YONGJIAN, L., 2007, “An improved PSO algorithm for solving non-convex NLP/MINLP problems with equality constraints”, *Computers & Chemical Engineering*, v. 31, n. 3, pp. 153–162. doi: 10.1016/j.compchemeng.2006.05.016.
- YOULE, P., DUNCANSON, L., 1970, “On-line control of olefine plant”, *Chemical and Process Engineering*, v. 51, n. 5, pp. 49.
- YOUNG, R., 2006, “Petroleum refining process control and real-time optimization”, *Control Systems, IEEE*, v. 26, n. 6 (Dec), pp. 73–83. doi: 10.1109/MCS.2006.252833.
- YOUSFI, C., TOURNIER, R., 1991, “Steady State Optimization Inside Model Predictive Control”. In: *American Control Conference*, pp. 1866–1870.
- YU, J., 2012, “A Bayesian inference based two-stage support vector regression framework for soft sensor development in batch bioprocesses”, *Computers & Chemical Engineering*, v. 41, pp. 134 – 144. doi: 10.1016/j.compchemeng.2012.03.004.
- ZAFIRIOU, E., CHENG, J.-H., 2004, “Measurement Noise Tolerance and Results Analysis for Iterative Feedback Steady-State Optimization”, *Industrial &*

- Engineering Chemistry Research*, v. 43, n. 14, pp. 3577–3589. doi: 10.1021/ie034072l.
- ZANIN, A., DE GOUVÊA, M. T., ODLOAK, D., 2000, “Industrial implementation of a real-time optimization strategy for maximizing production of LPG in a FCC unit”, *Computers & Chemical Engineering*, v. 24, n. 2–7, pp. 525 – 531. doi: 10.1016/S0098-1354(00)00524-X.
- ZANIN, A., DE GOUVÊA, M. T., ODLOAK, D., 2002, “Integrating real-time optimization into the model predictive controller of the FCC system”, *Control Engineering Practice*, v. 10, n. 8, pp. 819 – 831. doi: 10.1016/S0967-0661(02)00033-3.
- ZHANG, H., ROBERTS, P. D., 1991, “Integrated system optimization and parameter estimation using a general form of steady-state model”, *International Journal of Systems Science*, v. 22, n. 10, pp. 1679–1693. doi: 10.1080/00207729108910744.
- ZHANG, H., ROBERTS, P., 1990, “On-line steady-state optimisation of nonlinear constrained processes with slow dynamics”, *Transactions of the Institute of Measurement and Control*, v. 12, n. 5, pp. 251–261. doi: 10.1177/014233129001200504.
- ZHANG, P., RONG, G., WANG, Y., 2001a, “A new method of redundancy analysis in data reconciliation and its application”, *Computers & Chemical Engineering*, v. 25, n. 7–8, pp. 941 – 949. doi: 10.1016/S0098-1354(00)00659-1.
- ZHANG, Y., FORBES, J. F., 2000, “Extended design cost: a performance criterion for real-time optimization systems”, *Computers & Chemical Engineering*, v. 24, n. 8, pp. 1829–1841. doi: 10.1016/S0098-1354(00)00561-5.
- ZHANG, Y., FORBES, J. F., 2006, “Performance Analysis of Perturbation-Based Methods for Real-Time Optimization”, *The Canadian Journal of Chemical Engineering*, v. 84, n. 2, pp. 209–218. doi: 10.1002/cjce.5450840208.
- ZHANG, Y., NADLER, D., FORBES, J. F., 2001b, “Results analysis for trust constrained real-time optimization”, *Journal of Process Control*, v. 11, n. 3, pp. 329–341. doi: 10.1016/S0959-1524(99)00072-4.
- ZHANG, Y., MONDER, D., FORBES, J. F., 2002, “Real-time optimization under parametric uncertainty: a probability constrained approach”, *Journal of Process Control*, v. 12, n. 3, pp. 373–389. doi: 10.1016/S0959-1524(01)00047-6.

- ZHANG, Y., JIANG, J., 2008, “Bibliographical review on reconfigurable fault-tolerant control systems”, *Annual Reviews in Control*, v. 32, n. 2, pp. 229 – 252. ISSN: 1367-5788. doi: 10.1016/j.arcontrol.2008.03.008.
- ZHANG, Z., PIKE, R. W., HERTWIG, T. A., 1995a, “An approach to on-line optimization of chemical plants”, *Computers & Chemical Engineering*, v. 19, Supplement 1, pp. 305 – 310. ISSN: 0098-1354. doi: 10.1016/0098-1354(95)87054-7. European Symposium on Computer Aided Process Engineering–5.
- ZHANG, Z., PIKE, R. W., HERTWIG, T. A., 1995b, “Source reduction from chemical plants using on-line optimization”, *Waste Management*, v. 15, n. 3, pp. 183 – 191. ISSN: 0956-053X. doi: 10.1016/0956-053X(95)00016-S.
- ZHANG, Z., SHAO, Z., CHEN, X., et al., 2010, “Quasi-weighted least squares estimator for data reconciliation”, *Computers & Chemical Engineering*, v. 34, n. 2, pp. 154 – 162. doi: 10.1016/j.compchemeng.2009.09.007.
- ZHANG, Z., CHUANG, Y.-Y., CHEN, J., 2014, “Methodology of data reconciliation and parameter estimation for process systems with multi-operating conditions”, *Chemometrics and Intelligent Laboratory Systems*, v. 137, pp. 110 – 119. ISSN: 0169-7439. doi: 10.1016/j.chemolab.2014.06.007.
- ÉRIC POULIN, HODOUIN, D., LACHANCE, L., 2010, “Impact of plant dynamics on the performance of steady-state data reconciliation”, *Computers & Chemical Engineering*, v. 34, n. 3, pp. 354 – 360. ISSN: 0098-1354. doi: 10.1016/j.compchemeng.2009.11.018.

Apêndice A

Questões Preliminares sobre a Formulação do Problema

A.1 Programação Matemática

Se equações algébricas e diferenciais são a forma matemática de modelar fenômenos e processos da engenharia química, a programação matemática é a forma de modelar problemas de otimização. Embora a resolução de problemas reais de otimização empreguem programas de computador desenvolvidos para esse propósito, programação em *programação matemática* significa otimização. O uso original do termo *programação* refere-se ao conceito mais geral de otimização, no sentido de planejamento ótimo, sendo herança histórica da formalização proposta por George Dantzig nos anos 1940 ao formular problemas de planejamento e distribuição de recursos (BIEGLER, 2010; DANTZIG e THAPA, 2003; MARTIN, 1999).

Em otimização, a formulação do problema consiste em identificar e expressar matematicamente os elementos correspondentes às seguintes categorias, essenciais a todo problema de otimização: i) função objetivo; ii) restrições de igualdade; e iii) restrições de desigualdade (EDGAR *et al.*, 2001). A forma geral do modelo de programação matemática usado para representar o problema de otimização é expresso como (HOLDER, 2006–14)

$$\min\{J(\mathbf{x}) : \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n, h(\mathbf{x}) = 0, g(\mathbf{x}) \leq 0\} \quad ,$$

ou na forma equivalente

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{Minimizar}} & J(\mathbf{x}) \\ \text{sujeito a} & h(\mathbf{x}) = 0 \\ & g(\mathbf{x}) \leq 0 \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{array} \quad , \quad (\text{A.1})$$

em que $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ é um vetor pertencente ao conjunto \mathbb{X} contido no espaço dos números reais de dimensão n , \mathbb{R}^n , que corresponde ao domínio das funções $J(\cdot)$, $\mathbf{h}(\cdot)$ e $\mathbf{g}(\cdot)$, que são mapeamentos no espaço dos números reais \mathbb{R} , *i.e.*, $J : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}$, e $g : \mathbb{R}^n \rightarrow \mathbb{R}$, respectivamente. A função $J(\cdot)$ é a função objetivo (também chamada de função custo), enquanto as funções $h(\cdot)$ e $g(\cdot)$ são funções de restrição de igualdade e desigualdade, respectivamente. As funções de restrição também podem ser vetores de funções, nesse caso representados como $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, n_h$ e $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, n_g$. A região que contém a coleção de valores válidos, *i.e.*, que satisfaz simultaneamente as restrições $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n$, $h(\mathbf{x}) = 0$ e $g(\mathbf{x}) \leq 0$ é chamada de região viável (HOLDER, 2006–14).

O modelo geral apresentado na Equação (A.1) descreve matematicamente que se deseja determinar o vetor de variáveis de decisão \mathbf{x} que correspondem ao extremo da função objetivo $J(\cdot)$, garantindo que os limites estabelecidos pelas restrições de igualdade $h(\cdot)$ e desigualdade $g(\cdot)$ sejam respeitados (DIWEKAR, 2008). Além disso, presume-se que o conjunto \mathbb{X} e as funções $h(\cdot)$ e $g(\cdot)$ sejam dados pela modelagem do processo e restrições operacionais. Nesse sentido, especial atenção deve ser dada para que não exista confusão entre o problema de otimização (o problema real que pretende-se resolver), sua representação como um modelo de programação matemática (modelo da Eq. (A.1)) e a técnica/ algoritmo de resolução empregado da determinação da solução (ou valor ótimo) do problema (técnica para resolver a Eq. (A.1)).

De acordo com as propriedades do conjunto \mathbb{X} e das funções $J(\cdot)$, $h(\cdot)$ e $g(\cdot)$ definidas pelo problema, o modelo da Equação (A.1) pode ser (KALL e WALLACE, 1994):

- (a) *linear*, se o conjunto \mathbb{X} é um poliedro convexo e as funções $J(\cdot)$, $h(\cdot)$ e $g(\cdot)$ são lineares;
- (b) *não linear*, se ao menos uma das funções $J(\cdot)$, $h(\cdot)$ e $g(\cdot)$ é não linear ou \mathbb{X} não é um poliedro convexo.

Entre os programas não lineares, tem-se adicionalmente dois tipos:

- (b1) *convexo*, se $\mathbb{X} \cap \{\mathbf{x} : h(\mathbf{x}) = 0, g(\mathbf{x}) \leq 0\}$ é um conjunto convexo e $J(\cdot)$ é uma função convexa (em particular se as funções $h(\cdot)$ e $g(\cdot)$ são convexas e \mathbb{X} é um conjunto convexo);
- (b2) *não convexo*, se $\mathbb{X} \cap \{\mathbf{x} : h(\mathbf{x}) = 0, g(\mathbf{x}) \leq 0\}$ ou a função objetivo $J(\cdot)$ são não convexas.

A grande divisão em programação matemática não ocorre entre problemas lineares e não lineares, mas entre problemas convexas e não convexas. Sendo um modelo de

programação do tipo convexo, é possível calcular sua solução global, pois qualquer solução local (*i.e.*, relativa) será a solução global do problema (ALLGÖWER e ZHENG, 2000). Vale ressaltar que um problema de programação linear, caso (a), é sempre convexo.

Considerando os tipos apresentados juntamente com propriedades das variáveis de decisão, os modelos de programação matemática podem receber as seguintes denominações (DIWEKAR, 2008):

- Programação Linear (*Linear Programming* - LP): função objetivo e restrições são lineares; variáveis de decisão envolvidas são contínuas.
- Programação Não Linear (*Nonlinear Programming* - NLP): função objetivo e/ou restrições são não lineares; variáveis de decisão envolvidas são contínuas.
- Programação Inteira (*Integer Programming* - IP): variáveis de decisão envolvidas são inteiras.
- Programação Linear Inteira Mista (*Mixed integer linear programming* - MILP): dentre as variáveis de decisão envolvidas, um subconjunto é do tipo inteiro, enquanto seu complemento é contínuo.
- Programação Não Linear Inteira Mista (*Mixed integer nonlinear programming* - MINLP): um problema de programação não linear envolvendo variáveis inteiras e contínuas.
- Otimização discreta: problemas envolvendo variáveis de decisão discretas (inteiras), o que inclui problemas IP, MILP e MINLP.
- Otimização multiobjetivo: problemas envolvendo mais de um objetivo, que pode ser matematicamente descrito como $J : \mathbb{R}^n \rightarrow \mathbb{R}^q$ com $q \geq 2$.

Uma outra classe de problemas que merece destaque é a otimização dinâmica, *i.e.*, problemas de otimização associados a sistemas cujo comportamento seja governado por equações algébrico-diferenciais ou de diferenças (problemas de otimização algébrico-diferenciais, *differential algebraic optimization problem* - DAOP) (RAWLINGS e MAYNE, 2009). De maneira geral, problemas de otimização algébrico

diferenciais são escritos como:

$$\begin{aligned}
\text{Minimizar}_{\mathbf{u}(t), \boldsymbol{\theta}} \quad & J[\mathbf{u}(t), \boldsymbol{\theta}] := \phi[t_f, \mathbf{x}(t_f), \boldsymbol{\theta}] + \\
& \int_{t_0}^{t_f} L[t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}] dt \\
\text{sujeito a} \quad & \mathbf{F}[t, \dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}] = \mathbf{0}, \quad \forall t \in I, \\
& \mathbf{F}_0[t_0, \dot{\mathbf{x}}(t_0), \mathbf{x}(t_0), \mathbf{u}(t_0), \boldsymbol{\theta}] = \mathbf{0} \\
& \mathbf{S}[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\theta}] \leq \mathbf{0} \\
& \mathbf{T}[\mathbf{x}(t_f), \boldsymbol{\theta}] \leq \mathbf{0}
\end{aligned} \tag{A.2}$$

em que $t \in I := [t_0, t_f] \subset [0, \infty)$, com t_0 e t_f sendo os tempos inicial e final, respectivamente; $\mathbf{x} : I \rightarrow \mathbb{R}^{n_x}$ é o perfil temporal do vetor de n_x variáveis de estado com condições iniciais \mathbf{x}_0 conhecidas, em que a notação $\dot{\mathbf{x}}$ denota a derivada do estado em relação ao tempo; $\mathbf{u} : I \rightarrow \mathbb{R}^{n_u}$ é o perfil temporal do vetor de n_u variáveis de entrada (ou variáveis de controle); $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ é o vetor de n_θ variáveis invariantes no tempo, como parâmetros do modelo; $\mathbf{F}(\cdot)$ é o conjunto de equações algébrico diferenciais que descrevem o processo escritas de forma implícita, $\mathbf{F} : I \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}^{n_x}$, e $\mathbf{F}_0(\cdot)$ é o conjunto de condições iniciais; $\mathbf{S}(\cdot)$ é o conjunto de n_S restrições de caminho, incluindo limites nas variáveis de estado, de entrada e parâmetros; $\mathbf{T}(\cdot)$ é o conjunto de n_T restrições terminais; e $J(\cdot)$ é a função objetivo, um funcional que incorpora a função $\phi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$, denominada termo de *Mayer*, e a função $L : I \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}$ é denominada integrando de *Lagrange*. Se t_f é variável de decisão e/ou \mathbf{x}_0 é desconhecido, então t_f e \mathbf{x}_0 são incluídos no conjunto $\boldsymbol{\theta}$. Assume-se que as funções $\phi(\cdot)$, $L(\cdot)$, $\mathbf{F}(\cdot)$, $\mathbf{S}(\cdot)$ e $\mathbf{T}(\cdot)$ são contínuas e duas vezes diferenciáveis em relação aos seus argumentos. O problema de otimização (A.2) é denominado problema de *Bolza*, enquanto o objetivo de encontrar o extremo da função apenas no ponto final t_f é denominado problema de *Mayer*. Na situação em que apenas as variáveis de entrada são as variáveis de decisão do problema, o DAOP corresponde ao problema de controle ótimo (*optimal control problem* - OCP) (BELLMAN e DREYFUS, 1962; BINDER *et al.*, 2001; DIWEKAR, 2008; FRANÇOIS e BONVIN, 2013a). Técnicas gerais de solução dos problemas de otimização dinâmica envolvem o uso do cálculo de variações, princípio do mínimo (ou máximo) de Pontryagin (PONTRYAGIN *et al.*, 1962) e programação dinâmica. A programação dinâmica (*dynamic programming* - DP) é mais adequada para processos que possam ser decompostos em múltiplos estágios e baseia-se no princípio de otimalidade de Bellman (BELLMAN, 1957; BELLMAN e DREYFUS, 1962), que afirma que se a decisão de um estágio for uma solução ótima, então qualquer decisão remanescente será ótima em relação ao estado resultante dessa decisão; em outras palavras, o princípio de Bellman afirma que o valor máximo (ou mínimo)

da função depende do tempo e estado iniciais (DIWEKAR, 2008).

A.2 Otimização de Processos Baseada em Modelo

A otimização de processos visa a determinação de um conjunto de valores das variáveis de entrada, que garantam o atendimento ao conjunto de restrições existentes de forma a minimizar o valor de uma função dita função objetivo ou função custo. Nesse contexto, as variáveis de entrada são definidas no sentido dado pela teoria de sistemas e não são necessariamente as entradas físicas do processo, como vazões de alimentação, mas toda grandeza manipulável e independente que influencia o estado do sistema (FRANÇOIS, 2014).

A otimização de processos é usualmente realizada por meio de técnicas baseadas em modelos matemáticos, cujo objeto de interesse, ou realidade, é um processo químico denominado de “planta”, enquanto “modelo” é um conjunto de equações algébricas, diferenciais ou algébrico-diferenciais que o descreve (BONVIN e SRINIVASAN, 2013). A otimização propriamente dita também é uma abstração da realidade, *i.e.*, um modelo, mas representada por um problema de programação matemática, cuja função objetivo é um índice de desempenho, *e.g.*, um índice econômico, enquanto o modelo do processo ou equipamento e demais exigências de operação e qualidade correspondem às restrições de igualdade e desigualdade (EDGAR *et al.*, 2001).

Dessa forma, a otimização de um processo estático ou em regime permanente é feita por meio de um modelo estacionário e conduz a um problema de otimização estática, geralmente formulado como um problema de programação não linear (NLP) (mas também programação linear - LP, quadrática - QP ou inteira mista - MIP) que possui um número finito de variáveis de decisão. Considere, por exemplo, um reator contínuo perfeitamente agitado e o interesse em se determinar a melhor temperatura de operação do reator, ou as taxas de alimentação de um ou mais dos reagentes. Se, por outro lado, houver interesse em otimizar um processo dinâmico, em que considera-se modelos dinâmicos na descrição do processo, a formulação do problema passa a ser por meio da otimização dinâmica, cujo objetivo é determinar não mais um único valor, mas uma trajetória ótima para cada variável de decisão considerada (como perfis de alimentação) e que possuem dimensão infinita (FRANÇOIS e BONVIN, 2013a).

A.2.1 Otimização Estática

O objetivo mais comum perseguido pela otimização de processos contínuos operando em estado estacionário consiste em determinar as condições de operação (geralmente *set-points* em um sistema de controle em multicamadas) que proporcionem o melhor

desempenho do processo em termos de algum índice, como o mínimo custo operacional ou a máxima taxa de produção, respeitando as restrições de processo existentes. Uma formulação geral para esse problema de otimização estática e adequada a muitas aplicações práticas consiste no seguinte problema de programação matemática (FRANÇOIS e BONVIN, 2013a), conforme apresentado na Equação (2.1):

$$\begin{aligned} \mathbf{u}_p^* := \arg \min_{\mathbf{u}} \quad & J_p := \varphi_p(\mathbf{u}, \mathbf{y}_p) \\ \text{sujeito a} \quad & \mathbf{g}_p(\mathbf{u}, \mathbf{y}_p) \leq \mathbf{0} \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned} \quad , \quad (\text{A.3})$$

em que $\varphi_p : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}$ é a função escalar a ser minimizada; \mathbf{u}_p^* é o argumento que minimiza $J_p(\cdot)$, *i.e.*, a solução do problema; $\mathbf{u} \in \mathbb{R}^{n_u}$ é o vetor de n_u variáveis de decisão (ou de entrada); $\mathbf{y}_p \in \mathbb{R}^{n_y}$ é o vetor de n_y variáveis de saída; $g_{p,i} : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}$, $i = 1, \dots, n_g$ é o conjunto de funções de restrição de desigualdade; e \mathbf{u}^L e \mathbf{u}^U são os limites das variáveis de decisão.

Visto que a relação entre as variáveis de entrada e saída nunca pode ser perfeitamente conhecida, utiliza-se um modelo matemático que, embora sendo uma aproximação, seja capaz de descrever com alguma acurácia a relação entre essas variáveis no estado estacionário, representado por

$$\mathbf{f}(\mathbf{u}, \mathbf{x}, \boldsymbol{\theta}) = \mathbf{0} \quad (\text{A.4})$$

$$\mathbf{y} = \mathbf{h}(\mathbf{u}, \mathbf{x}) \quad , \quad (\text{A.5})$$

em que $\mathbf{f} \in \mathbb{R}^{n_x}$ é o conjunto de equações do modelo do processo, que inclui balanços de massa e energia, relações termodinâmicas, *etc.*; $\mathbf{x} \in \mathbb{R}^{n_x}$ é o vetor de n_x variáveis de estado do modelo; $\mathbf{y} \in \mathbb{R}^{n_y}$ é o vetor de n_y variáveis de saída previstas pelo modelo, as quais são determinadas por relações envolvendo as variáveis de estado e de entrada, representadas pelo conjunto de funções $\mathbf{h}(\cdot)$; $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ é o vetor de n_θ parâmetros do modelo. Desse modo, o subscrito p simboliza quantidades relacionadas à realidade, *i.e.*, à planta, indicando que funções e variáveis correspondem a relações e quantidades experimentais, que não são perfeitamente conhecidas e podem ser aproximadas apenas por meio de um modelo; o modelo será representado com os mesmos símbolos da planta, mas sem o subscrito p (BUNIN, 2014).

Assumindo que as variáveis de saída possam ser representadas explicitamente apenas em termos de \mathbf{u} e $\boldsymbol{\theta}$ na forma $\mathbf{y} = \mathbf{F}(\mathbf{u}, \boldsymbol{\theta})$, o problema de otimização para a planta pode ser formulado como um problema NLP com restrições (MARCHETTI,

2009)

$$\begin{aligned} \min_{\mathbf{u}} \quad & J := \varphi [\mathbf{u}, \mathbf{F}(\mathbf{u}, \boldsymbol{\theta})] \\ \text{sujeito a} \quad & \mathbf{g} [\mathbf{u}, \mathbf{F}(\mathbf{u}, \boldsymbol{\theta})] \leq \mathbf{0} \\ & \mathbf{u}^L \leq \mathbf{u} \leq \mathbf{u}^U \end{aligned} \quad , \quad (\text{A.6})$$

ou de maneira equivalente,

$$\begin{aligned} \min_{\mathbf{u}} \quad & J := \phi(\mathbf{u}, \boldsymbol{\theta}) \\ \text{sujeito a} \quad & \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) \leq \mathbf{0} \end{aligned} \quad , \quad (\text{A.7})$$

com

$$\phi(\mathbf{u}, \boldsymbol{\theta}) := \varphi [\mathbf{u}, \mathbf{y}(\mathbf{u}, \boldsymbol{\theta})] \quad \mathbf{G}(\mathbf{u}, \boldsymbol{\theta}) := \begin{bmatrix} \mathbf{g} [\mathbf{u}, \mathbf{y}(\mathbf{u}, \boldsymbol{\theta})] \\ \mathbf{u}^L - \mathbf{u} \\ \mathbf{u} - \mathbf{u}^U \end{bmatrix}$$

Assume-se que as funções $\varphi(\cdot)$ e $\mathbf{g}(\cdot)$ sejam conhecidas e que a dependência com os parâmetros $\boldsymbol{\theta}$ dos valores da função objetivo e das restrições preditos com uso do modelo, *i.e.*, $\varphi(\mathbf{u}, \mathbf{y})$ e $\mathbf{g}(\mathbf{u}, \mathbf{y})$, seja dada de maneira implícita por meio das variáveis de saída do modelo \mathbf{y} .

Admitindo que a função objetivo e as restrições sejam contínuas e diferenciáveis com relação a \mathbf{u} e o domínio viável do problema seja não vazio e limitado, está garantida a existência de soluções \mathbf{u}_p^* e \mathbf{u}^* para os problemas (2.1) e (A.7)⁷, respectivamente, que podem adicionalmente apresentar múltiplas soluções locais de acordo com a não convexidade do problema (CHACHUAT *et al.*, 2009). Na situação em que a solução \mathbf{u}^* seja um ponto regular, *i.e.*, que nesse ponto seja satisfeita uma condição de regularidade (ou qualificação de restrição)⁸, também serão satisfeitas nesse ponto as condições necessárias de primeira ordem de Karush-Kuhn-Tucker (KKT) (também chamadas de condições necessárias de otimalidade, do termo em inglês *necessary conditions of optimality* - NCO) com um conjunto único de multiplicadores de Lagrange $\boldsymbol{\nu} \in \mathbb{R}^{n_g}$ (BAZARAA *et al.*, 2006). Para o problema (A.7),

⁷ Pelo Teorema de Bolzano-Weierstrass (BAZARAA *et al.*, 2006).

⁸ Essa condição pode ser a independência linear entre os gradientes das restrições ativas de desigualdade, *i.e.*, $\text{rank} [\nabla_{\mathbf{u}} \mathbf{G}^a(\mathbf{u}^*, \boldsymbol{\theta})] = n_{g,a}$, conhecida como qualificação de independência linear das restrições (do termo em inglês *linear-independence constraint qualification* - LICQ) (FRANÇOIS e BONVIN, 2013a).

as condições de KKT são

$$\begin{aligned}
 G_k(\mathbf{u}^*, \boldsymbol{\theta}) &\leq 0 \\
 \nu_k &\geq 0 \\
 \frac{\partial \mathcal{L}}{\partial \mathbf{u}}(\mathbf{u}^*, \boldsymbol{\theta}, \boldsymbol{\nu}) &= \frac{\partial \phi}{\partial \mathbf{u}}(\mathbf{u}^*, \boldsymbol{\theta}) + \sum_{k=1}^{n_g} \nu_k \frac{\partial G_k}{\partial \mathbf{u}}(\mathbf{u}^*, \boldsymbol{\theta}) = \mathbf{0} \\
 \nu_k G_k(\mathbf{u}^*, \boldsymbol{\theta}) &= 0, \quad k = 1, \dots, n_g \quad ,
 \end{aligned}$$

ou rescritas na forma vetorial

$$\mathbf{G}(\mathbf{u}^*, \boldsymbol{\theta}) \leq \mathbf{0} \quad (\text{A.8})$$

$$\boldsymbol{\nu} \geq \mathbf{0} \quad (\text{A.9})$$

$$\nabla_{\mathbf{u}} \mathcal{L} = \nabla_{\mathbf{u}} \phi(\mathbf{u}^*, \boldsymbol{\theta}) + \boldsymbol{\nu}^T \nabla_{\mathbf{u}} \mathbf{G}(\mathbf{u}^*, \boldsymbol{\theta}) = \mathbf{0} \quad (\text{A.10})$$

$$\boldsymbol{\nu}^T \mathbf{G}(\mathbf{u}^*, \boldsymbol{\theta}) = \mathbf{0} \quad (\text{A.11})$$

com a função de Lagrange definida como

$$\mathcal{L}(\mathbf{u}, \boldsymbol{\theta}, \boldsymbol{\nu}) := \phi + \boldsymbol{\nu}^T \mathbf{G} \quad (\text{A.12})$$

em que \mathbf{u}^* representa uma solução; $\boldsymbol{\nu} \in \mathbb{R}^{n_g}$ é o vetor de multiplicadores de Lagrange associado às restrições; $\nabla_{\mathbf{u}} \phi(\mathbf{u}^*, \boldsymbol{\theta})$ é o vetor linha de dimensão n_u representando o gradiente da função objetivo com relação a \mathbf{u} avaliado no ponto \mathbf{u}^* ; e $\nabla_{\mathbf{u}} \mathbf{G}(\mathbf{u}^*, \boldsymbol{\theta})$ é a matriz Jacobiana das restrições avaliada no ponto \mathbf{u}^* com dimensão $n_g \times n_u$. É comum que as condições de KKT sejam designadas como condição de viabilidade primal (A.8), condições de viabilidade dual (A.9) e (A.10), e condição de complementaridade (A.11). Assim, qualquer ponto \mathbf{u}^* para o qual exista um vetor de multiplicadores $\boldsymbol{\nu}^*$ que satisfaçam as condições (A.8) a (A.11) é chamado de *ponto KKT*; nesse contexto, a condição de regularidade se faz importante, pois implica que os multiplicadores de Lagrange $\boldsymbol{\nu}^*$ são unicamente determinados no ponto KKT \mathbf{u}^* .

A satisfação das NCOs para um dado problema exige a determinação do conjunto de restrições ativas, \mathbf{G}^a , que pode ser feita com base no sinal dos multiplicadores de Lagrange. Se o conjunto \mathbf{G}^a for conhecido, as condições de KKT podem ser simplificadas pela eliminação dos multiplicadores a duas expressões (CHACHUAT

et al., 2009):

$$\mathbf{G}^a(\mathbf{u}^*, \boldsymbol{\theta}) = \mathbf{0} \quad (\text{A.13})$$

$$\nabla_r \phi := \frac{\partial \phi}{\partial \mathbf{u}} \left[\mathbf{I} - \left(\frac{\partial \mathbf{G}^a}{\partial \mathbf{u}} \right)^+ \left(\frac{\partial \mathbf{G}^a}{\partial \mathbf{u}} \right) \right] = 0 \quad (\text{A.14})$$

em que $(\cdot)^+$ indica a pseudo-inversa e $\nabla_r \phi \in \mathbb{R}^{1 \times n_u}$ surge ao eliminar as restrições ativas.

Uma vez que os problemas de otimização estática sejam representados como modelos de programação não linear, sua resolução se dá por meio de técnicas e algoritmos desenvolvidos para esse tipo de modelo. Técnicas populares para a resolução de programas não lineares com restrições são o método de função de penalização, método do ponto interior e programação quadrática sequencial (do termo em inglês *sequential quadratic programming* - SQP). Brevemente, o método de penalização incorpora as restrições do problema na função objetivo por meio de termos de penalização, transformando o problema em um modelo sem restrições. O método do ponto interior também incorpora as restrições na função objetivo (FORSGREN *et al.*, 2002), mas com a vantagem de atuar mais como uma barreira que uma penalização, gerando valores viáveis entre as iterações, o que no método de penalização só é garantida após convergência. Por outro lado o método SQP resolve uma sequência de subproblemas rescritos em termos da função de Lagrange, com aproximações quadráticas para a função objetivo e lineares para as restrições, utilizando internamente técnicas como o método de Newton para encontrar as condições de KKT. Maiores detalhes sobre essas e outras técnicas de otimização podem ser encontrados na literatura (BIEGLER, 2010; BIEGLER e GROSSMANN, 2004; EDGAR *et al.*, 2001; FLETCHER, 2000; GRIVA *et al.*, 2009; NOCEDAL e WRIGHT, 2006).

A.3 Presença de Incertezas

Incerteza é um conceito vago que incorpora tudo aquilo que é desconhecido e/ou que não é conhecido com certeza, sendo, assim, definido por tautologia. A incerteza é característica importante e inerente à modelagem e otimização de processos; em processos químicos, ela é causada principalmente por erros de modelagem (designado na literatura como *plant-model mismatch*, refletindo-se no uso de um modelo imperfeito), erros paramétricos, variações aleatórias nos dados medidos e perturbações que afetam o processo (FRANÇOIS, 2014).

A tentativa de representação da realidade por meio de modelos matemáticos é útil em diversas aplicações tecnológicas, incluindo aplicações a processos químicos. Todavia, modelos precisos e/ou rigorosos podem não estar disponíveis, seja por falta

de conhecimento do processo, seja por incapacidade de medir e observar variáveis ou por falta de dispositivos de medida adequados (CHAUDHURI e DIWEKAR, 1999; ULAS e DIWEKAR, 2004). Com isso, a consideração de hipóteses e demais simplificações feitas podem reduzir a capacidade de representação do modelo e/ou limitar sua faixa de validade, o que constitui fonte de incerteza. Informações sobre a incerteza de modelagem são comumente obtidas por meio de dados experimentais e envolvem a determinação de elementos como constantes cinéticas, propriedades físicas e coeficientes de transferência. Por outro lado, a aquisição de respostas do sistema também constitui fonte de incerteza, pois os dados sempre estão corrompidos por algum tipo de ruído e/ou possuem variações provenientes de erros de medição, que podem ou não estar associados a fatores inerentes ao processo como variações de temperatura e vazão e flutuações na qualidade das correntes. Adicionalmente, outras incertezas são provenientes de influências externas, como oscilações na temperatura ambiente, variações na disponibilidade de matéria-prima, demandas de mercado, flutuações de preços, *etc.* (PISTIKOPOULOS, 1995). Dessa forma, identificar e caracterizar essas incertezas é etapa essencial, visto que considerar apenas as imprecisões das medidas já é suficiente para concluir que todos os resultados do modelo estarão sujeitos a incertezas (PINTO, 1998).

Diversas formas de classificar e modelar incertezas foram propostas na literatura (BERNARDO *et al.*, 2003; IERAPETRITOU *et al.*, 1996; KRASŁAWSKI, 1989; PISTIKOPOULOS, 1995; STRAUB e GROSSMANN, 1993; ULAS e DIWEKAR, 2004). Embora tenha limitações para abordar todos os aspectos das incertezas, um esquema conveniente de classificação frequentemente usado no ramo da engenharia mecânica diferencia as incertezas entre objetivas e subjetivas (BEYER e SENDHOFF, 2007). Incertezas objetivas são aquelas que possuem natureza estocástica intrínseca e irredutível, sendo por isso chamadas de incertezas aleatórias ou estocásticas (HELTON, 1997). Essa categoria representa incertezas de natureza física e que não podem ser removidas, *e.g.*, ruído em sinais elétricos, variações na umidade e temperatura ambientes e parâmetros físicos em materiais. Portanto, a formulação do problema deve estar adequada a essa condição. Devido a sua natureza probabilística, distribuições de probabilidade representam a maneira mais adequada para a modelagem matemática dessas incertezas. Opondo-se ao caráter objetivo das incertezas aleatórias, o tipo subjetivo de incerteza reflete a falta de conhecimento sobre dado elemento do problema de interesse, *i.e.*, reflete a falta de informações que, em princípio, poderia ser reduzida com algum esforço. Nessa categoria podem ser incluídas as incertezas de modelagem, condições de validade do modelo, bem como os erros introduzidos por métodos numéricos de solução, *e.g.*, erros de aproximação e problemas de convergência.

Na prática, a representação das incertezas no modelo matemático se dá por meio

de uma variável, ou parâmetro, chamado de incerto, sendo possível identificar três tipos de modelagem para as incertezas: determinística, probabilística e possibilística (BEYER e SENDHOFF, 2007). Esses três diferentes tipos são representados matematicamente por meio de conjuntos convencionais, distribuições de probabilidade e conjuntos *fuzzy*, respectivamente (KLIR e FOLGER, 1988). A abordagem determinística define domínios nos quais os parâmetros incertos podem variar. Já variáveis do tipo probabilística são definidas por meio de medidas de probabilidade que descrevem a chance com a qual determinado evento ocorre, enquanto o tipo possibilístico é descrito por meio de medidas *fuzzy* associadas a possibilidade de ocorrência de um evento. Assim, uma forma de descrever uma incerteza “determinística” consiste em definir um conjunto finito de valores discretos (cenários) que seguem determinada probabilidade de ocorrência, *e.g.*, $\Theta = \{\theta : \theta^L \leq \theta_i \leq \theta^U, i = 1, \dots, N\}$, em que θ^L e θ^U são limites especificados (GROSSMANN e SARGENT, 1978). Na abordagem estocástica, a modelagem consiste em descrever o parâmetro incerto como uma variável aleatória contínua que segue uma PDF especificada, *e.g.*, $\xi \sim p_\xi$, em que a variável aleatória ξ está distribuída com densidade de probabilidade p_ξ (PISTIKOPOULOS e IERAPETRITOU, 1995). Em ambos os casos, a escolha da PDF depende da informação disponível, em que a distribuição uniforme representa o mínimo grau de conhecimento sobre a incerteza (BERNARDO *et al.*, 2003).

Considerar todos os tipos de incertezas que podem surgir deixaria a formulação matemática intratável, o que exige admitir simplificações adequadas de forma que o modelo resultante reflita o equilíbrio entre a representação do problema e a capacidade de resolução (BAZARAA *et al.*, 2006). Assim, é imperativo determinar quais incertezas apresentam impacto significativo na formulação do modelo, *i.e.*, no desempenho do processo e suas restrições. Para isso, técnicas de análise de sensibilidade (probabilísticas) podem ser usadas e adicionalmente combinadas com estratégias de classificação e ordenamento, que permitem reduzir a dimensão das incertezas (MÜLLER *et al.*, 2014). Incertezas com pouca ou nenhuma contribuição nas respostas do sistema podem ser desprezadas ou substituídas por valores (nominais) constantes (GELETU e LI, 2014).

A.3.1 Incertezas de Medição: Ruídos Aleatórios

A descrição da realidade por meio de modelos matemáticos, por mais rigorosos que sejam, esbarra em informações desconhecidas, as quais podem ser determinadas por meio de experimentos e dependem fundamentalmente da qualidade da informação contida nos dados coletados diretamente da planta ou resultantes de análises de laboratório. Esses dados quantificam variáveis do processo como vazões, concentrações, temperaturas, níveis, pressões, e propriedades finais de interesse dos produtos.

Todavia, a observação da realidade também apresenta suas limitações, *e.g.*, técnicas de medição possuem precisão finita e limitada, condições experimentais nunca são perfeitamente controladas e perturbações desconhecidas, que sequer poderiam ser previstas, ocorrem com frequência. Diante disso, o processo de modelagem pode ser melhorado se for capaz de descrever, de maneira adequada, essas incertezas de medição.

Admitindo que todas as variáveis de saída \mathbf{y} , representadas na Equação (A.5), sejam medidas, tendo seus valores medidos designados por \mathbf{y}_{ms} e valores verdadeiros designados por \mathbf{y}_p , o erro total nas medidas, $\boldsymbol{\varepsilon}_y$, é definido como a diferença entre valores medidos e verdadeiros, *i.e.*,

$$\boldsymbol{\varepsilon}_y := \mathbf{y}_{ms} - \mathbf{y}_p \quad , \quad (\text{A.15})$$

em que $\boldsymbol{\varepsilon}_y \in \mathbb{R}^{n_y}$. Os erros de medição podem ser causados por diferentes fontes e estão sempre presentes em qualquer medida, sendo predominantemente aleatórios (NARASIMHAN e JORDACHE, 1999).

A tentativa de relacionar um modelo determinístico aos dados coletados no instante de tempo k permite representar as Equações (A.4) e (A.5) como

$$\mathbf{f}_k(\mathbf{u}_k, \mathbf{x}_k, \boldsymbol{\theta}) = \mathbf{0} \quad (\text{A.16})$$

$$\mathbf{x}_k = \mathbf{g}(\mathbf{u}_k, \boldsymbol{\theta}) \quad (\text{A.17})$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{u}_k, \mathbf{x}_k) \quad . \quad (\text{A.18})$$

Dado que o número total de equações existentes na Equação (A.18) excede o número de parâmetros desconhecidos $\boldsymbol{\theta}$, apenas sob circunstâncias ideais existirão valores de $\boldsymbol{\theta}$ que resultem em $\mathbf{y}_{ms,k} = \mathbf{y}_k$. Uma forma conveniente de modelar os erros de medição consiste em representá-los matematicamente como variáveis aleatórias e considerar que seu efeito sobre as variáveis medidas é aditivo, *i.e.*,

$$\mathbf{y}_{ms} = \mathbf{y} + \boldsymbol{\varepsilon}_y \quad , \quad (\text{A.19})$$

em que \mathbf{y} são as variáveis de saída do modelo. Se $\boldsymbol{\varepsilon}_y = \mathbf{0}$, então as medidas possuem elevada precisão, de modo que o modelo do processo de medição se torna $\mathbf{y}_{ms} = \mathbf{y}$; caso contrário, $\boldsymbol{\varepsilon}_y \neq \mathbf{0}$ representa medidas corrompidas com ruído aleatório (ROMAGNOLI e SÁNCHEZ, 2000). Em um processo perfeito de medição, as medidas possuem elevada precisão e acurácia, em que $\mathbf{y}_{ms} = \mathbf{y}_p$. Para o caso em que o ruído de medida segue distribuição Normal com média nula e covariância \mathbf{V} , representado

por $\boldsymbol{\varepsilon}_y \sim N(\mathbf{0}, \mathbf{V})$, tem-se (SCHWAAB e PINTO, 2007):

$$p(\boldsymbol{\varepsilon}_y) = \frac{1}{(2\pi)^{n/2} (\det \mathbf{V})^{1/2}} \exp \left[-\frac{1}{2} \boldsymbol{\varepsilon}_y^T \mathbf{V}^{-1} \boldsymbol{\varepsilon}_y \right] \quad , \quad (\text{A.20})$$

em que $\det \mathbf{V}$ denota o determinante da matriz \mathbf{V} .

Incertezas de medição afetam os valores das variáveis medidas na planta e todo esforço para identificar tais incertezas pode ser usado para melhorar a confiabilidade de procedimentos numéricos que façam uso dos dados medidos, notadamente os procedimentos de estimação de parâmetros e otimização. Nesse sentido, a teoria de probabilidade fornece ferramentas úteis para essas atividades (BARD, 1974; FULLER, 1987; ZHANG *et al.*, 2001b). As incertezas de variáveis medidas são caracterizadas com base em dados de processo por meio de informações estatísticas como média, variância, covariância e distribuições de probabilidade, sendo a representação probabilística bastante adequada para essas variáveis. Incertezas correlacionadas tem especial impacto no desempenho do processo e seu desconhecimento pode afetar negativamente as análises, uma vez que pequenas variações em uma variável propagam-se às outras variáveis com as quais esta esteja correlacionada. Adicionalmente, a correlação também pode estar presente em elementos do modelo do processo, notadamente nos parâmetros. Nesse sentido, é mais vantajoso lidar com dados (e/ou parâmetros em um modelo) não correlacionados e algum esforço no sentido de descorrelacionar as incertezas pode ser interessante e ou necessária (GELETU e LI, 2014; LARENTIS *et al.*, 2003; SCHWAAB e PINTO, 2008). Deve ser notado também que a natureza das incertezas depende do sistema de interesse e podem apresentar variação com o tempo, como é possível ocorrer no estudo de sistemas dinâmicos (TØRVI e HERTZBERG, 1997), requerendo que o modelo considere esse comportamento. Além disso, a caracterização serve não apenas à modelagem da incerteza, mas também a diversas outras técnicas matemáticas envolvidas no monitoramento, controle e otimização de processos, bem como a ações legais e regulatórias da operação industrial. Dada sua importância prática e caráter obrigatório em alguns casos, além das técnicas convencionais também foram propostas técnicas de caracterização de dados de processo em tempo real (ou caracterização *on-line*) (FEITAL e PINTO, 2015; FEITAL *et al.*, 2014). Em muitas situações, no entanto, pouco esforço é empregado para caracterizar os dados, sendo comum admitir que as medidas (ou mais precisamente os ruídos de medida) são variáveis aleatórias e independentes, seguindo distribuições gaussianas (ou distribuição Normal de probabilidade). Na ausência de informações suficientes, caracterizações aproximadas e hipóteses acerca de distribuições de probabilidade podem levar a resultados ambíguos ou incorretos.

A.3.2 Incertezas de Modelagem: Erros Estruturais do Modelo e Erros Paramétricos

A maioria dos problemas de otimização baseada em modelos admite que as incertezas de modelagem, embora desconhecidas, possam ser incorporadas no vetor de parâmetros incertos θ (FRANÇOIS e BONVIN, 2013a). Todavia, na presença de erros de modelagem, a solução \mathbf{u}^* de um procedimento de otimização em geral não corresponde ao ponto ótimo da planta \mathbf{u}_p^* (CHACHUAT *et al.*, 2009).

Modelos matemáticos correspondem a um conjunto de equações que tentam representar o processo real. Sua construção está baseada em leis de conservação (massa, momento e energia) e relações constitutivas que expressam condições de equilíbrio, fenômenos cinéticos e de transporte, mas sempre envolve simplificações, como desprezar determinados fenômenos físicos e/ou químicos, ou remover alguns fenômenos dinâmicos para não aumentar demasiadamente a complexidade. Adicionalmente, o processo de modelagem não garante que a estrutura das relações entre as variáveis seja unicamente determinada considerando apenas que dada estrutura matemática represente um conjunto de observações experimentais. Em outras palavras, é possível encontrar, para um dado sistema, mais de um modelo capaz de explicar o mesmo conjunto de dados, e as decisões ótimas fornecidas por esses modelos não serão necessariamente as mesmas (TAKAMATSU, 1983). Dessa forma, modelos são essencialmente limitados e sua estrutura invariavelmente difere daquela advinda de um modelo verdadeiro idealizado, capaz de representar perfeitamente a realidade, de maneira que tal erro estrutural é fonte de incerteza que afeta a qualidade das predições do modelo.

Além disso, o modelo resultante depende naturalmente de parâmetros, cujos valores não são conhecidos com precisão, mas são identificados usando as medidas do processo e, conseqüentemente, podem ser apenas conhecidos dentro de certo intervalo de confiança de acordo com uma dada probabilidade. Erros paramétricos podem estar presentes em todos os modelos, sejam estruturalmente corretos ou não, e são tipicamente causados pelas dificuldades inerentes aos problemas de identificação, como número de medidas independentes disponíveis, quantidade de informação no sinal, entre outros (FRANÇOIS, 2014; FRANÇOIS e BONVIN, 2013a).

Nesse contexto, é conveniente distinguir formalmente as imprecisões do modelo originadas por erros estruturais daquelas resultantes de erros paramétricos. Assim, MARCHETTI (2009) define que um modelo $\mathbf{y}(\mathbf{u}, \theta)$ é considerado estruturalmente correto se existirem valores para os parâmetros θ para os quais o modelo fornece uma representação precisa da planta com relação ao mapeamento existente entre

variáveis de entrada e saída, *i.e.*, existem valores dos parâmetros $\boldsymbol{\theta}$ tal que $\forall \mathbf{u} \in \mathbb{U}$,

$$\begin{aligned} \mathbf{y}(\mathbf{u}, \boldsymbol{\theta}) &= \mathbf{y}_p(\mathbf{u}) \\ \frac{\partial \mathbf{y}(\mathbf{u}, \boldsymbol{\theta})}{\partial \mathbf{u}} &= \frac{\partial \mathbf{y}_p(\mathbf{u})}{\partial \mathbf{u}} \\ \frac{\partial^2 \mathbf{y}(\mathbf{u}, \boldsymbol{\theta})}{\partial \mathbf{u}^2} &= \frac{\partial^2 \mathbf{y}_p(\mathbf{u})}{\partial \mathbf{u}^2} . \end{aligned}$$

Contrariamente, o modelo é considerado estruturalmente incorreto se tais valores para os parâmetros não existem.

Na prática, os erros estruturais de modelagem correspondem a diferenças funcionais, *e.g.*, entre as expressões de J_p e J , e entre \mathbf{G}_p e \mathbf{G} (Equações 2.1 e A.7). Se determinado modelo é estruturalmente correto e representado por $\mathbf{G}_p(\mathbf{u}_p, \boldsymbol{\theta}_p)$, erros paramétricos representam a situação em que $\boldsymbol{\theta} \neq \boldsymbol{\theta}_p$ no modelo conhecido $\mathbf{G}(\mathbf{u}, \boldsymbol{\theta})$ (FRANÇOIS, 2014; FRANÇOIS e BONVIN, 2013a).

A.4 Otimização sob Incertezas

Os problemas apresentados anteriormente são todos determinísticos, em que a solução encontrada é nominal. Todavia, toda situação envolvendo processos de tomada de decisão possui incertezas associadas em diversos elementos, notadamente nos dados e no modelo. Dessa forma, muitas são as motivações para considerar incertezas no problema de otimização, entre elas: i) erros de medição estão sempre presentes em dados reais; ii) existe a constante incapacidade de implementar na prática uma solução nominal com toda sua acurácia; e iii) em todo o tempo, é possível que incertezas, mesmo pequenas, descaracterizem totalmente a solução nominal do ponto de vista prático (BEN-TAL *et al.*, 2009). Por isso, além dos tipos mencionados anteriormente, outra classe de modelos surge quando incertezas são consideradas na formulação do modelo, caso em que o problema é classificado como otimização sob incertezas (do termo em inglês *optimization under uncertainty*)⁹. Diversas são as formas de lidar com incertezas e as principais abordagens da otimização sob incertezas incluem a programação estocástica (*stochastic programming* - SP) e a otimização robusta (*robust optimization* - RO) (BEN-TAL *et al.*, 2009; MULVEY *et al.*, 1995; SAHINIDIS, 2004).

A.4.1 Programação Estocástica

Os modelos de programação estocástica podem ser vistos como extensões dos modelos de programação matemática apresentados anteriormente, em que os parâmetros

⁹A despeito das motivações apresentadas, os modelos não deixam de ser determinísticos, mas consideram as incertezas explicitamente (BIEGLER e GROSSMANN, 2004).

incertos são representados por funções de probabilidade, sendo representados na função objetivo e/ou restrições por variáveis aleatórias (NEMHAUSER *et al.*, 1989).

A presença de variáveis aleatórias no problema (A.1) resulta tipicamente na situação em que uma decisão \mathbf{x} deve ser feita diante de um ambiente incerto descrito por um parâmetro ξ , que pode ser representado como (KALL e WALLACE, 1994; NEMHAUSER *et al.*, 1989):

$$\begin{aligned} \text{“Minimizar”} \quad & J(\mathbf{x}, \xi) \\ \text{sujeito a} \quad & h(\mathbf{x}, \xi) = 0 \\ & g(\mathbf{x}, \xi) \leq 0 \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned} \quad , \quad (\text{A.21})$$

em que ξ é um vetor de variáveis aleatórias variando no conjunto $\Xi \subset \mathbb{R}^k$. Implicitamente, reconhece-se uma variável aleatória como uma função (ou um mapeamento) $\xi : \Xi \rightarrow \mathbb{R}^k$ existente em um *espaço de probabilidades* representado por (Ξ, \mathcal{F}, P) , em que *resultados* $\xi(\omega)$ são obtidos a partir de um espaço Ξ por meio de determinada coleção de *eventos* \mathcal{F} avaliados por uma medida de probabilidade P (ou distribuição de probabilidade), uma métrica que associa a cada elemento de \mathcal{F} a probabilidade com a qual ele ocorre. Os elementos que compõem a coleção de eventos \mathcal{F} são subconjuntos $F \subset \Xi$. Admite-se que a coleção de eventos \mathcal{F} e a distribuição de probabilidades P em \mathcal{F} são dadas, de forma que para todo evento $A \in \mathcal{F}$ a probabilidade $P(A)$ é conhecida. Por consequência, as funções $J(\cdot)$, $h(\cdot)$ e $g(\cdot)$ também são variáveis aleatórias, sendo $J(x, \cdot) : \Xi \rightarrow \mathbb{R}$, $h(x, \cdot) : \Xi \rightarrow \mathbb{R}$ e $g(x, \cdot) : \Xi \rightarrow \mathbb{R}$, e que sua distribuição de probabilidade P é independente de x . No entanto, a forma apresentada do problema (A.21) não está bem definida e outras definições devem estar presentes para que se obtenha um equivalente determinístico do problema (KALL e WALLACE, 1994).

Modelos de programação estocástica são naturalmente classificados como “espere e veja” (*wait-and-see*, ou modelos adaptativos) e “aqui e agora” (*here-and-now*, ou modelos antecipativos). Na abordagem “espere e veja”, aguarda-se a realização das variáveis aleatórias e, de posse dos valores observados, resolve-se o problema de otimização determinístico (BIRGE e LOUVEAUX, 2011; NEMHAUSER *et al.*, 1989). A representação geral desse problema é

$$\begin{aligned} \text{Otimizar} \quad & J(\mathbf{x}, \tilde{\xi}) \\ \text{sujeito a} \quad & h(\mathbf{x}, \tilde{\xi}) = 0 \\ & g(\mathbf{x}, \tilde{\xi}) \leq 0 \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned} \quad , \quad (\text{A.22})$$

em que $\tilde{\xi}$ é uma possível realização do vetor aleatório ξ , *i.e.*, um elemento de Ξ . A

resolução do problema (A.22) é repetida para cada amostragem do vetor ξ , obtendo-se ao final uma representação probabilística do resultado (DIWEKAR, 2008).

Por outro lado, a abordagem “aqui e agora” apresenta uma conexão maior com processos de decisão considerando que o valor do vetor aleatório ξ será conhecido depois da seleção das variáveis de decisão \mathbf{x} , situação em que a otimização envolve alguma métrica probabilística, usualmente a média da variável (também chamado de valor esperado, momento de primeira ordem ou esperança matemática). Um problema genérico modelado pela abordagem “aqui e agora” pode ser representado como (KALL e WALLACE, 1994)

$$\begin{aligned} \underset{\mathbf{x}}{\text{Otimizar}} \quad & J(\mathbf{x}, \xi) := E_{\xi} [j(\mathbf{x}, \xi)] \\ \text{sujeito a} \quad & E_{\xi} [h(\mathbf{x}, \xi)] = 0 \\ & E_{\xi} [g(\mathbf{x}, \xi)] \leq 0 \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned} \quad , \quad (\text{A.23})$$

em que, dado o espaço de probabilidade (Ξ, \mathcal{F}, P) e a função $j : \mathbb{R}^n \times \Xi \rightarrow \mathbb{R}$ (introduzidos pelo vetor aleatório $\xi \in \mathbb{R}$), o valor esperado E_{ξ} , ou simplesmente E , é a integral com relação à medida de probabilidade P , representada por

$$E_{\xi} [j(\mathbf{x}, \xi)] = \int_{\Xi} j(\mathbf{x}, \xi) \, dP = \int_{\Xi} P(\xi) j(\mathbf{x}, \xi) \, d\xi \quad .$$

Nesse caso, a estratégia para lidar com a função objetivo aleatória consistiu em converter $j(\mathbf{x}, \xi)$ em uma função objetivo determinística aplicando o operador de média à função, *i.e.*, $E[j(\mathbf{x}, \xi)]$. Se $j(\mathbf{x}, \xi)$ é linear em \mathbf{x} , $j(\mathbf{x}, \xi) = \xi^{\top} \mathbf{x}$, então $E[j(\mathbf{x}, \xi)] = (E[\xi])^{\top} \mathbf{x}$ também é linear em \mathbf{x} . No entanto, existem outras formas de lidar com incertezas na função objetivo e outros critérios de decisão, diferentes do valor médio, podem ser usados (BERNARDO *et al.*, 2003; MULVEY *et al.*, 1995; PRÉKOPA, 1995).

Há ainda situações em que ambas as abordagens estão presentes na modelagem, incorporando tanto antecipação quanto adaptação, caso denominado programação estocástica com recurso (KALL e WALLACE, 1994; SAHINIDIS, 2004). Problemas com recurso podem envolver múltiplos estágios e dizem respeito a decisões que são tomadas antes da realização das incertezas (“aqui e agora”) e ações de recurso que podem ser tomadas quando a incerteza torna-se conhecida (“espere e veja”) para compensar possíveis violações das restrições. Um problema com recurso de dois estágios é representado por

$$\begin{aligned} \underset{\mathbf{x}}{\text{Minimizar}} \quad & J(\mathbf{x}, \xi) := f(\mathbf{x}) + E_{\xi} [Q(\mathbf{x}, \xi)] \\ \text{sujeito a} \quad & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned} \quad , \quad (\text{A.24})$$

com

$$\begin{aligned} Q(\mathbf{x}, \xi) = \min_y \quad & q(\mathbf{x}, \xi, y) \\ \text{sujeito a} \quad & g(\mathbf{x}, \xi) \leq y \\ & y \in \mathbb{Y} \subset \mathbb{R} \end{aligned} \quad , \quad (\text{A.25})$$

em que a função $Q(\cdot)$ é chamada de função de recurso e $q : \mathbb{R}^n \times \Xi \times \mathbb{R} \rightarrow \mathbb{R}$. Para cada restrição existe uma variável recurso (ou segundo estágio) y , que é determinada após a realização de ξ para compensar a violação da respectiva restrição. A função $q(\cdot)$, por sua vez, corresponde ao custo ou penalidade imposta pela violação da restrição. O problema (A.24) com a variável \mathbf{x} constitui o primeiro estágio, que é resolvido antes da realização do vetor aleatório ξ , enquanto o problema (A.25) com a variável y constitui o segundo estágio. Em resumo, o esquema consiste em: i) determinação de \mathbf{x} ; ii) observação de ξ ; e iii) determinação de y . Portanto, o modelo exige a resolução de dois problemas de otimização, em que a formulação do problema de recurso admite que \mathbf{x} e ξ são fixos (PRÉKOPA, 1995).

Voltando a atenção para as restrições e como as incertezas são consideradas nesses elementos, considere a versão modificada do problema (A.23) em que a restrição de igualdade foi reescrita como duas restrições de desigualdade, totalizando três restrições de desigualdade:

$$\begin{aligned} \text{Otimizar}_{\mathbf{x}} \quad & J(\mathbf{x}, \xi) := E_{\xi} [j(\mathbf{x}, \xi)] \\ \text{sujeito a} \quad & E_{\xi} [g_i(\mathbf{x}, \xi)] \leq 0, i = 1, 2, 3 \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned} \quad , \quad (\text{A.26})$$

As restrições podem também ser escritas na forma

$$E [-\mathbf{g}(\mathbf{x}, \xi) \mid \mathbf{g}(\mathbf{x}, \xi) > \mathbf{0}] \leq \mathbf{d} \quad . \quad (\text{A.27})$$

Ambas as expressões em (A.26) e (A.27) são restrições baseadas no valor esperado. No problema (A.26) interpreta-se que cada uma das i restrições $g_i(\cdot)$ são satisfeitas na média dos casos, o que reflete a expectativa de que essas restrições sejam violadas em menos de 50% dos casos, aproximadamente. Outra expressão possível é a restrição apresentada em (A.27), que exige que a magnitude da média das violações de cada uma das i restrições $g_i(\mathbf{x}, \xi) \leq 0$ seja inferior a um limite superior \mathbf{d} , em que a média é tomada apenas para os casos que violem a restrição. Ao limitar a magnitude média dessas violações, a expressão (A.27) ignora a frequência com que as restrições são violadas. Dessa forma, ambas as expressões em (A.26) e (A.27) devem estar presentes caso se deseje controlar tanto a magnitude quanto a frequência das violações (ZHANG *et al.*, 2002).

Outro equivalente determinístico do problema (A.21) é denominado de problema

com restrição de probabilidade ou programação probabilística (referido em inglês como *probabilistically constrained program* ou *chance constrained program*). Nessa abordagem, impõe-se que as restrições sejam satisfeitas com algum nível de probabilidade. Partindo do problema (A.26), é possível em alguns passos rescrever as restrições em termos de probabilidades (KALL e WALLACE, 1994), resultando em

$$\begin{aligned} \underset{\mathbf{x}}{\text{Otimizar}} \quad & J(\mathbf{x}, \xi) := E_{\xi} [j(\mathbf{x}, \xi)] \\ \text{sujeito a} \quad & P [\xi \mid g_i(\mathbf{x}, \xi) \leq 0, i = 1, 2, 3] \geq \alpha \quad , \quad (\text{A.28}) \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned}$$

em que $\alpha \in [0; 1]$. Nesse caso do problema (A.28), diz-se que as restrições são de probabilidade conjunta (do termo em inglês *joint probabilistic constraints* - JPC), em que todas as i restrições $g_i(\cdot)$ são simultaneamente satisfeitas a um nível de confiança maior ou igual que α . Por outro lado, se forem definidos $\alpha_i \in [0; 1], i = 1, 2, 3$, as restrições passam a ser de probabilidade individual (do termo em inglês *single, separate* ou *individual probabilistic constraints* - IPC), expressas por

$$\begin{aligned} \underset{\mathbf{x}}{\text{Otimizar}} \quad & J(\mathbf{x}, \xi) := E_{\xi} [j(\mathbf{x}, \xi)] \\ \text{sujeito a} \quad & P [\xi \mid g_i(\mathbf{x}, \xi) \leq 0] \geq \alpha_i, i = 1, 2, 3 \quad , \quad (\text{A.29}) \\ & \mathbf{x} \in \mathbb{X} \subset \mathbb{R}^n \end{aligned}$$

Em ambos os casos, essa formulação para as restrições permite algum controle sobre ambos frequência e magnitude de variação para dada distribuição de probabilidades. Problemas de programação estocástica com JPC possuem resolução difícil e/ou custosa e, por isso, mais atenção na literatura tem sido dada aos problemas com IPC (ZHANG *et al.*, 2002).

A.4.2 Otimização Robusta

Outra possível abordagem para lidar com incertezas em otimização é conhecida como *otimização robusta* (RO). Os conceitos de *robustez* e *otimização robusta* foram desenvolvidos de maneira independente em diferentes ramos da ciência¹⁰, principalmente na área conhecida como pesquisa de operações (do termo em inglês *operations research*) e em engenharia de projetos (BEYER e SENDHOFF, 2007), mas também em outras áreas como estatística e controle de processos¹¹. De maneira geral, utiliza-se

¹⁰ Maiores detalhes sobre a abordagem da otimização robusta, suas formulações, aspectos teóricos e computacionais, aplicações, *etc.*, podem ser encontrados nos trabalhos de BERTSIMAS *et al.* (2011); BEYER e SENDHOFF (2007); GABREL *et al.* (2014); GELETU e LI (2014).

¹¹ O termo *robustez* também é utilizado em outras áreas, mas nem sempre carrega exatamente o mesmo significado. Em estatística, *robustez* é usado para qualificar determinado método ou procedimento como protegido contra os efeitos negativos causados por desvios pequenos e finitos do modelo idealizado e/ou hipóteses presentes, particularmente sobre os desvios que estão próximos

o termo *otimização robusta* para qualificar a formulação que forneça soluções cujos resultados sejam insensíveis às incertezas. Dessa forma, é possível que uma solução seja considerada *robusta* se a sua aplicação garantir a viabilidade das restrições em qualquer realização das incertezas, ou ainda, se sua aplicação garantir determinado valor da função objetivo (ou outra métrica de otimalidade) dentro de uma tolerância. Outro paradigma comumente utilizado na otimização robusta é conhecido como abordagem do pior caso (do termo em inglês *worst-case approach*), em que busca-se determinar a melhor solução para a realização mais desfavorável das incertezas, *i.e.*, para o pior cenário conferido pelas incertezas, a despeito do custo que isso possa resultar¹². Portanto, é possível afirmar que a definição do que se considera robusto faz parte da modelagem do problema de otimização robusta (GABREL *et al.*, 2014).

De acordo com BEYER e SENDHOFF (2007), a noção de robustez em otimização ganhou mais atenção após o trabalho de MULVEY *et al.* (1995), que cunhou o termo *otimização robusta* para descrever a formulação que integra a descrição das incertezas como um conjunto de cenários e métodos de programação multiobjetivo com programação estocástica, além de considerar introdução de momentos de ordem superior na função objetivo. Como método de programação multiobjetivo, a

ou abaixo dos limites de detecção. Dessa forma, na situação em que um procedimento (*e.g.*, estimação de parâmetros) sofre pouca influência de desvios nas hipóteses assumidas (*e.g.*, presença de *outliers* no conjuntos de dados), este procedimento é dito robusto (PRATA, 2009; PRATA *et al.*, 2010). Robustez é conferida a procedimentos estatísticos por meio de modificações nos estimadores (*i.e.*, na função objetivo), limitando a influência de erros grosseiros (ou *outliers*) e protegendo o procedimento contra outros tipos de contaminação (HUBER, 2014; HUBER e RONCHETTI, 2009; MARONNA *et al.*, 2006). Admitindo que robustez signifique insensibilidade a pequenos desvios das suposições, é possível encontrar correspondência entre a *robustez* da otimização robusta com aquela da estatística robusta. Todavia, uma conexão precisa ainda não foi estabelecida entre estatística robusta e otimização robusta (BEN-TAL *et al.*, 2009).

Também é possível encontrar o termo *robustez* na área de estimação de estados, técnica que integra as estratégias de controle avançado e que apresenta correspondência com a reconciliação de dados. A estimação de estados se preocupa em considerar os efeitos da imperfeição do modelo. Uma proposta de estimador de estados robusto consiste no estimador de horizonte móvel (do termo em inglês *moving horizon estimator* - MHE, um estimador baseado em procedimento de otimização) que inclui na formulação uma restrição que limita o desvio de predição a uma região definida pelo desvio de medição, sendo robusto com relação a aproximações ruins dos desvios de predição na presença de erros limitados (LIU, 2013; RAWLINGS e JI, 2012). Nesse caso, a robustez está relacionada à estabilidade do sistema dinâmico. Diferentemente, a reconciliação de dados robusta admite que o modelo é perfeito e aplica estatística robusta para considerar desvios sobre as hipóteses admitidas para os erros de medida (NARASIMHAN e JORDACHE, 1999; PRATA, 2009; PRATA *et al.*, 2010; ROMAGNOLI e SÁNCHEZ, 2000).

Outra área que faz uso do termo *robustez* é a área de controle de processos. Exemplos de formas com que as incertezas podem surgir nos problemas de controle incluem perturbações desconhecidas afetando o sistema, conhecimento imperfeito do estado do sistema e/ou imperfeições no modelo do sistema. No caso de projeto de controladores, a principal preocupação é com relação à estabilidade, em que um controle é dito robusto se sua estabilidade permanecer garantida, juntamente com suas especificações de desempenho, mesmo na presença de incertezas (MORARI e LEE, 1999; RAWLINGS e MAYNE, 2009).

¹² Visto que muitas aplicações são modeladas pela abordagem do pior caso, que segue a formulação *min-max*, é comum que o termo *otimização robusta* seja definido como otimização do pior caso (HOLDER, 2006–14) ou otimização min-max (GELETU e LI, 2014).

abordagem incorpora explicitamente no modelo o conflito entre soluções que permanecem próximas do ótimo em todos os cenários (solução robusta com relação a otimalidade) e soluções que permanecem viáveis em todos os cenários (solução robusta com relação a viabilidade), com o propósito de determinar aquelas que são menos sensíveis à realização das incertezas dentro do conjunto de cenários definido.

Dessa forma, a abordagem robusta de otimização sob incertezas é derivada da programação estocástica e pode ser considerada sua complementar. Isso porque embora a programação estocástica seja bem intencionada ao considerar as incertezas do problema, sua formulação admite que essas incertezas sejam aleatórias e que sua distribuição de probabilidades seja conhecida *a priori*. Todavia, cada uma dessas suposições contribui para restringir as aplicações da abordagem, principalmente naquelas envolvendo dados medidos, em que, na maioria dos casos, pouco ou nenhum esforço é feito para caracterizar os dados e se admite que a distribuição é gaussiana ou que seus valores verdadeiros pertençam a determinada região de confiança em torno dos valores nominais. Em muitos casos, identificar com precisão distribuições de probabilidade (sobretudo as multidimensionais) requer uma quantidade de observações que raramente estará disponível, de forma que a adoção de hipóteses simplificadoras em modelos de programação estocástica, necessária para sua resolução, acaba gerando novas incertezas cuja influência é difícil de avaliar. Além disso, diversas situações exigem que a satisfação das restrições seja garantida na tomada de decisão a despeito da presença das incertezas, como, *e.g.*, ao considerar limites operacionais de segurança em problemas de interesse da engenharia química. Especialmente nesses casos, o nível de tolerância à violação das restrições é nulo: o risco associado à adoção de hipóteses sobre as distribuições de probabilidade torna-se relevante e a garantia probabilística de satisfação das restrições deixa de ser adequada ao problema. A otimização robusta, por sua vez, reconhece que a existência de hipóteses na modelagem das incertezas nos problemas de programação estocástica pode incorporar novas incertezas ao problema. Assim, enquanto a abordagem da programação estocástica modela as incertezas do problema como variáveis aleatórias de distribuição de probabilidades conhecida, a otimização robusta possibilita formulações que não pressupõem a existência dessa informação e considera que as incertezas pertencem a algum intervalo ou conjunto Ω fechado, limitado e suficientemente grande (comumente denominado de *conjunto de incertezas*) (BEN-TAL *et al.*, 2009; GELETU e LI, 2014).

Diversas são as maneiras de formular problemas usando otimização robusta. É comum que medidas de variabilidade, como a variância, sejam usadas para considerar robustez (SAHINIDIS, 2004). Nesse sentido, uma forma usual de estabelecer um critério multiobjetivo entre otimalidade e viabilidade consiste em definir como função objetivo uma combinação linear entre o valor esperado do índice de desempenho e

a sua variabilidade, com o intuito de ponderar o máximo desempenho e a robustez (MANDUR e BUDMAN, 2014), de acordo com

$$\begin{aligned} \min_{\mathbf{u}} \quad & J := w_1 E[f(\mathbf{x}, \mathbf{u})] + w_2 Var[f(\mathbf{x}, \mathbf{u})] \\ \text{sujeito a} \quad & \mathbf{g}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}) \leq \mathbf{0} \end{aligned}, \quad (\text{A.30})$$

em que $\boldsymbol{\theta}$ é o vetor de parâmetros, sendo essa a variável incerta; \mathbf{u} é o vetor de variáveis de decisão (ou variáveis manipuladas); \mathbf{x} é o vetor de variáveis de estado; $f(\cdot)$ é o índice de desempenho; \mathbf{g} é o vetor de restrições de igualdade e desigualdade, o qual inclui o modelo do processo e limites operacionais; $E(\cdot)$ e $Var(\cdot)$ são o valor esperado e a variância da métrica $f(\cdot)$, respectivamente; e w_1 e w_2 são constantes que ponderam cada um dos termos.

No âmbito dos problemas em PSE, também é comum encontrar aplicações da abordagem do pior caso, geralmente usada nas situações em que se deseja determinar condições ótimas de operação para desempenho robusto do sistema e satisfação estrita das restrições sob todas as possíveis ocorrências das incertezas em uma dada região. Em outras palavras, o problema consiste em minimizar o máximo impacto das incertezas no critério de desempenho, *i.e.*, minimizar o pior caso (GELETU e LI, 2014; LEE e COOLEY, 2000). Para essa abordagem, a modelagem segue a formulação conhecida como *min-max*, resultando em

$$\begin{aligned} \min_{\mathbf{u} \in \mathbb{U}} \max_{\boldsymbol{\theta} \in \Theta} f(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}) \\ \text{sujeito a} \quad \mathbf{g}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}) \leq \mathbf{0} \end{aligned} \quad (\text{A.31})$$

Modelos de otimização robusta do tipo min-max resultantes da abordagem do pior caso, semelhantes ao apresentado na Equação (A.31), são notadamente empregados na área de controle (LEE e COOLEY, 2000; MORARI e ZAFIRIOU, 1989; MORARI e LEE, 1999), em especial na formulação do MPC robusto, que é considerado um marco no estudo de controle robusto (LEE, 2014).