AGREMIS GUINHO BARBOSA

Desenvolvimento de um Software para Detecção de Erros Grosseiros e Reconciliação de Dados Estática e Dinâmica de Processos Químicos e Petroquímicos

Campinas, 2008

UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA QUÍMICA

AGREMIS GUINHO BARBOSA

Desenvolvimento de um Software para Detecção de Erros Grosseiros e Reconciliação de Dados Estática e Dinâmica de Processos Químicos e Petroquímicos

Orientador: Prof. Dr. Rubens Maciel Filho

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Química da Faculdade de Engenharia Química da Universidade Estadual de Campinas para obtenção do título de Doutor em Engenharia Química

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO AGREMIS GUINHO BARBOSA E ORIENTADA PELO PROF. DR. RUBENS MACIEL FILHO

Assinatura do Orientador

Campinas, 2008

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

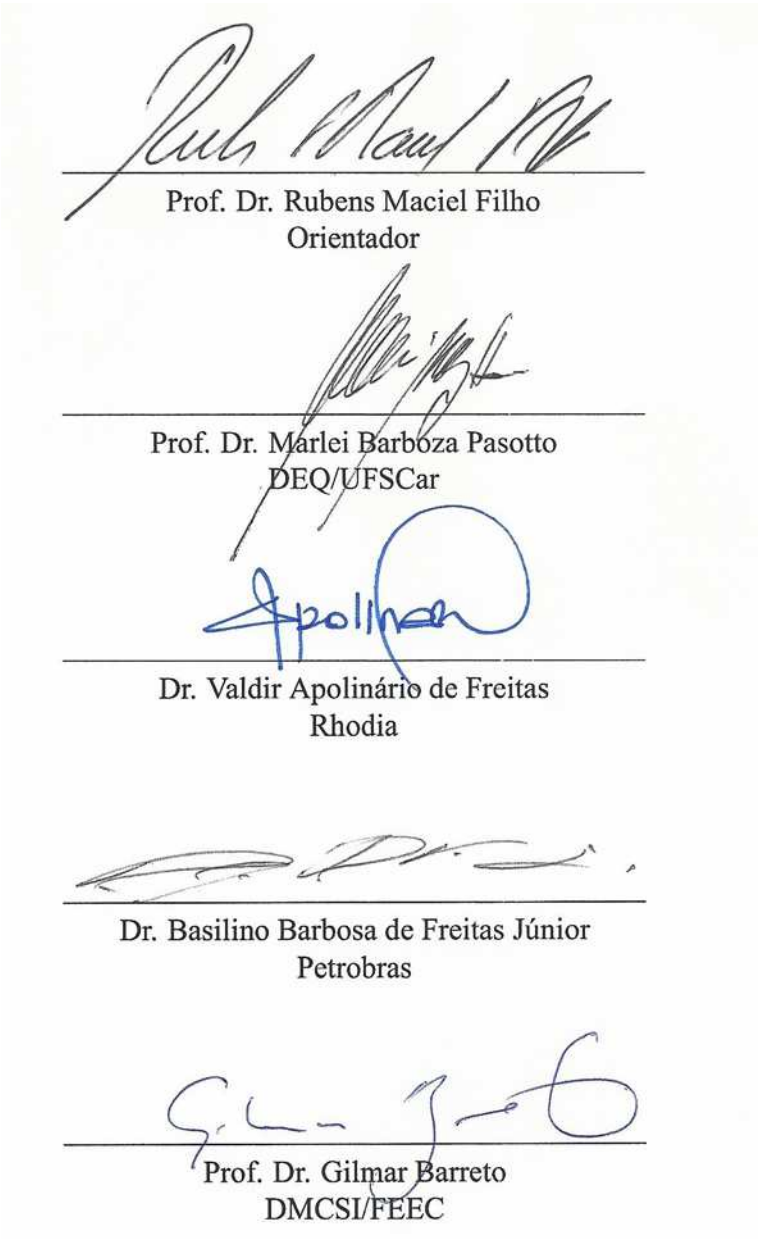
Barbosa, Agremis Guinho
B234d Desenvolvimento de um software para detecção de erros grosseiros e reconciliação de dados estática e dinâmica de processos químicos e petroquímicos / Agremis Guinho Barbosa. --Campinas, SP: [s.n.], 2008.

Orientador: Rubens Maciel Filho.
Tese de Doutorado - Universidade Estadual de Campinas, Faculdade de Engenharia Química.

1. Teoria do controle. 2. Controle de processos químicos. 3. Estimativa de parâmetro. 4. Localização de falhas (Engenharia). 5. Redundância (Engenharia). I. Maciel Filho, Rubens, 1958-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Química. III. Título.

Título em Inglês: Development of software for static and dynamic gross error detection and data reconciliation of chemical and petrochemical processes
Palavras-chave em Inglês: Control theory, Chemical process control, Parameter estimation, System fault location (Engineering), Redundancy (Engineering)
Área de concentração: Desenvolvimento de processos Químicos
Titulação: Doutor em Engenharia Química
Banca examinadora: Marlei Barboza Pasotto, Valdir Apolinário de Freitas, Basilino Barbosa Freitas Junior, Gilmar Barreto
Data da defesa: 18-12-2008
Programa de Pós Graduação: Engenharia Química

Tese de Doutorado defendida por Agremis Guinho Barbosa e aprovada em 18 de dezembro de 2008 pela banca examinadora constituída pelos doutores:



Prof. Dr. Rubens Maciel Filho
Orientador

Prof. Dr. Marlei Barbosa Pasotto
DEQ/UFSCar

Dr. Valdir Apolinário de Freitas
Rhodia

Dr. Basilino Barbosa de Freitas Júnior
Petrobras

Prof. Dr. Gilmar Barreto
DMCSI/FEEC

Agradecimentos

A vida vai sendo tecida com o fio que também se chama vida e são tantos os que compuseram a trama. Há tanta gente para agradecer! Cada conselho, cada bronca, cada mágoa, sorrisos, desesperos, alegrias enormes e tristezas sem ter fim. Todos em volta assistindo ao trabalho de urdidura num grande suspense para saber o que reserva o próximo capítulo. Se vai ser comédia hoje e drama amanhã. Ou será o inverso?

Assim, agradeço primeiro aos que primeiro lançaram fios às engrenagens: minha mãe, Maria de Lourdes Guinho Barbosa e meu pai já falecido, Durval de Lima Barbosa – eu acho que ele já sabia desse doutorado antes mesmo de eu fazer o curso técnico em química na antiga ETFAL, a Escola Técnica Federal de Alagoas.

Sigo agradecendo aos vários amigos que fiz até aqui. Alguns até da infância, com quem falo até hoje. Entre eles, amigos que fiz durante a graduação e são verdadeiros irmãos: Pleycienne Trajano Ribeiro e Sebastião Araújo Coutinho.

No laboratório onde estou desde o início do doutorado e onde esse trabalho vem sendo feito, o LOPCA, fiz também muitos amigos aos quais agradeço enormemente. Os que primeiro me recepcionaram: Edvaldo Rodrigo de Morais e Jefferson Pinto (a comunidade Shadu), e também Igor Victorino, Eduardo "Urso" Toledo, Mylene e Rodrigo Rezende, Caliane, Luiz Meleiro e Nagel Alves Costa, o sujeito mais inteligente que já conheci em minha vida e foi nosso professor na graduação em engenharia química na UFPB, em Campina Grande. Omito desta lista tantos outros nomes, também muito importantes, por questões de espaço e memória.

Agradeço também aos professores Maria Regina Wolf Maciel e Edson Tomaz da UNICAMP e Severino Rodrigues de Farias Neto e Michel François Fossy da UFPB.

Agradeço à FAPESP que vem fomentando esse trabalho desde o mestrado e através dela, agradeço também aos cidadãos do estado de São Paulo.

Agradeço muitíssimo ao caro Prof. Rubens Maciel Filho, meu orientador. Esse que me apresentou ao tema do qual nunca havia ouvido falar antes e me cedeu essa oportunidade de me aper-

feiçoar em um ambiente de liberdade. Obrigado mesmo!

Finalmente, agradeço à minha filha, Luiza da Anunciação Guinho, o fio de vida que lancei nas engrenagens da própria vida. Já vejo tanta coisa se anunciando nessa trama. É uma alegria assistir à sua confecção.

Devo ter sido injusto na lista acima. Como eu disse no começo, a vida é um tecido e são muitos os fios que nos compõem. É fácil esquecer de algum que tenha sido importante, pois todos foram e são importantes. A estes fios, o meu reconhecimento e agradecimento.

Resumo

O principal objetivo deste trabalho foi o desenvolvimento de um *software* para reconciliação de dados, detecção e identificação de erros grosseiros, estimativa de parâmetros e monitoramento da qualidade da informação em unidades industriais em estado estacionário e dinâmico. O desenvolvimento desse *software* focalizou atender aos critérios de modularidade, extensibilidade e facilidade de uso.

A reconciliação de dados é um procedimento de tratamento de medidas em plantas de processos necessário devido ao fato da inexorável presença de erros aleatórios de pequena magnitude associados aos valores obtidos dos equipamentos de medição. Além dos erros aleatórios, por vezes os dados estão associados a erros de maior magnitude e que constituem uma tendência, ou viés. Erros desta natureza podem ser qualificados e quantificados por técnicas de detecção de erros grosseiros.

É importante para aplicação de subrotinas de otimização que os dados sejam confiáveis e livres de erros tanto quanto possível. A tarefa da remoção destes erros através de modelos previamente conhecidos (reconciliação de dados) não é trivial, já sendo estudada no campo da engenharia química nos últimos 40 anos e apresenta uma crescente quantidade de trabalhos publicados.

Contudo, uma parte destes trabalhos é voltada para aplicação da reconciliação sobre equipamentos isolados, como tanques, reatores e colunas de destilação, ou pequenos conjuntos destes equipamentos e não são muitos os trabalhos que utilizam dados reais de operação. Isto pode ser atribuído à dimensão do trabalho computacional associado ao grande número de variáveis. O que se propõe neste trabalho é tomar partido da crescente capacidade computacional e das modernas ferramentas de desenvolvimento, provendo uma aplicação na qual seja facilitada a tarefa de descrever sistemas de maior dimensão, para estimar dados de qualidade superior, em tempo hábil, para sistemas de controle e otimização.

É importante frisar que a reconciliação de dados e a detecção de erros grosseiros são fundamentais para a confiabilidade de resultados de subrotinas de otimização e controle supervisório e também pode ser utilizada para a reconstrução de estados do processo.

Abstract

The main goal of this work was the development of a software for data reconciliation, gross errors detection and identification, data reconciliation, parameter estimation, and information quality monitoring in industrial units under steady state and dynamic operation. The development of this software was focused on meeting the criteria of modularity, extensibility, and user friendliness.

Data reconciliation is a procedure for measurement data treatment in process plants, which is necessary due the fact of the inexorable presence of random, small magnitude errors associated to the values obtained from measurement devices. In addition to the random errors, sometimes data are associated to major magnitude errors that lead to a trend or bias. Errors of this nature can be qualified and quantified through gross errors detection techniques.

It is important for optimization routines that data are reliable and error free as much as possible. The task of removal of these errors using previously known models (data reconciliation) is not trivial, and has been studied for the last 40 years in the field of chemical engineering, showing an increasing amount of published works.

However, part of these works is devoted to applying data reconciliation over a single equipment, such as tanks, reactors, distillation columns, or small sets of these equipments. Furthermore, not much of these published work relies on real operation data. This can be regarded to the dimension of computational work associated to the great number of variables. This work proposes to take advantage of increasing computational capacity and modern development tools to provide an application in which the task of higher dimension systems description is accomplished with ease in order to produce data estimates of superior quality, in a suitable time frame, to control and optimization systems.

It is worthwhile mentioning that data reconciliation and gross error detection are fundamental for reliability of the results from supervisory control and optimization routines, and can be used also to process state reconstruction.

Sumário

Lista de Figuras

Lista de Tabelas

Nomenclatura

Letras Latinas

$\mathbf{A}$	Matriz de ocorrência
$\mathbf{A_1}$	Matriz de ocorrência das variáveis medidas
$\mathbf{A_2}$	Matriz de ocorrência das variáveis não medidas
g	Quantidade de valores em $\mathbf{x}$
H_0	Hipótese nula
H_1	Hipótese alternativa
$\mathbf{I}$	Matriz identidade
$\mathbf{J}$	Jacobiana de ϕ
l	Quantidade de medidas em $\mathbf{y}$
L	Lagrangiano
m	Número de equações de restrição adicionais
N	Função objetivo do problemas de minimização
$\mathbf{P}$	Matriz de projeção de Crowe
$\mathbf{Q}$	Matriz ortogonal resultado da fatoração QR
$\mathbf{R}$	Matriz ortogonal resultado da fatoração QR
r	Grau de redundância
$\mathbf{r}$	Vetor dos resíduos
$\mathbf{u}$	Vetor das variáveis não medidas
$\mathbf{w}$	Erro aditivo ao modelo de restrições adicionais
$\mathbf{x}$	Vetor das variáveis do modelo (valores verdadeiros)

$\hat{\mathbf{x}}$ Vetor das estimativas das variáveis do modelo
$\mathbf{y}$ Vetor das medidas com ruído obtidas da instrumentação
$\mathbf{z}_a$ Matriz das estatísticas do teste da medida
Z_a Critério de rejeição no teste da medida
$\mathbf{z}_r$ Matriz das estatísticas do teste nodal
Z_r Critério de rejeição no teste nodal

Letras Gregas

α Nível de significância
β Nível de significância modificado
γ Estatística do teste global
δ Dimensão do viés
ε Erro aleatório aditivo às medidas
λ Autovalores de uma matriz
σ Desvio padrão
ϕ Modelo funcional de medição
φ Modelo de restrições adicionais
χ^2 Distribuição de probabilidade *chi-quadrado*
Ψ Matriz de covariância

Siglas

CSTR *Continuous Stirred Tank Reactor* – Reator Tanque de Mistura Contínua
DCS *Distributed Control System* – Sistema de Controle Distribuído
DLL *Dynamic Link Library* – Biblioteca de Vínculo Dinâmico
ERP *Enterprise Resource Planning* – Planejamento de Recursos Empresariais
FCC *Fluid Catalytic Cracking* – Craqueamento Catalítico Fluido
GAMS *General Algebraic Modeling System* — Sistema Geral de Modelagem Algébrica
GLP Gás Liqüefeito de Petróleo

GLR	*Generalized Likelihood Ratio* – Razão de Verossimilhança Generalizada
GUI	*Graphical User Interface* – Interface Gráfica com o Usuário
MT	*Measurement Test* – Teste da Medida
NT	*Nodal Test* – Teste Nodal
RAD	*Rapid Application Development* – Desenvolvimento Rápido de Aplicativos
SGBDR	Sistema Gerenciador de Banco de Dados Relacional
SQP	*Successive Quadratic Programming* – Programação Quadrática Sucessiva

1 Introdução

Este capítulo apresenta os objetivos da tese para em seguida fazer uma contextualização das metodologias de reconciliação de dados e detecção de erros grosseiros, sendo apresentados também exemplos que justificam sua implementação em sistemas industriais de informação. Finalmente, é feita uma apresentação da tese, descrevendo capítulo por capítulo os assuntos que são abordados neste trabalho.

1.1 Objetivos

O trabalho descrito nesta tese teve como objetivo principal o desenvolvimento de um *software* baseado no trabalho iniciado em Barbosa (2003), ampliando o escopo e as funcionalidades do aplicativo ***Reconciliare***, devotado à reconciliação de dados e detecção/identificação de erros grosseiros. O desenvolvimento deste aplicativo tentou contemplar alguns requisitos importantes como a **modularidade**, a **extensibilidade** e a **facilidade de uso**.

A modularidade diz respeito à maneira como o aplicativo é desenvolvido. Um aplicativo modular tem independência entre suas partes em maior ou menor grau. Isto representa um grande custo no momento do desenvolvimento, mas que é recompensado quando são necessários reparos ou ampliações. No desenvolvimento do aplicativo ***Reconciliare*** é usado o paradigma de programação orientada a objetos que propicia o desenvolvimento modular. Esse tipo de desenvolvimento permite também que partes inteiras do programa sejam trocadas sem afetar as demais.

A extensibilidade é a capacidade de, apoiada na modularidade, permitir que sejam acrescentadas novas características ao programa com o menor impacto possível. Foi desenvolvido no aplicativo ***Reconciliare*** um conjunto de interfaces que permite a criação de novos módulos, inclusive por terceiros que não precisam ter acesso ao código fonte do programa principal, para estender a aplicação, conferindo novas funcionalidades ou modificando as existentes.

A facilidade de uso foi considerada fundamental no desenvolvimento do aplicativo. Ela diz respeito não apenas à clareza ou rapidez com que as funções do aplicativo são usadas, mas principalmente ao ordenamento que é imposto ao usuário de modo a diminuir tanto quanto possível a quantidade de eventuais enganos na sua utilização. Comumente, no desenvolvimento de aplicações computacionais na área da reconciliação de dados e detecção/identificação de erros grosseiros no domínio acadêmico, a entrada de dados e a interação com o usuário é relegada a um segundo plano. O objetivo aqui foi integrar tanto quanto possível características de *softwares* comerciais e tirar proveito disso para as finalidades de pesquisa as quais esse trabalho se propôs.

Alguns objetivos específicos desta tese foram:

- Desenvolvimento de subrotinas genéricas necessárias para a reconciliação de dados e detecção de erros grosseiros;
- Desenvolvimento de ferramentas de monitoramento de dados industriais para fornecer relatórios sobre o estado da qualidade da informação vinda do sistema de aquisição de dados;
- Criação de um ambiente visual (GUI - *Graphical User Interface*) com interface "amigável" onde o usuário possa fazer desde a descrição do processo (com as respectivas ligações às fontes de dados) até o uso das subrotinas desenvolvidas com a subseqüente disponibilização das variáveis estimadas, passando pelo monitoramento das medições e dos dados ajustados;
- Desenvolvimento de módulos para a conexão com fontes de dados, tanto para receber como para disponibilizar dados.

Este trabalho objetivou também coletar e ordenar uma grande quantidade de resultados disponíveis na literatura, aprofundando no detalhamento teórico de modo a facilitar a reprodução de alguns dos resultados.

1.2 Contextualização

Nas últimas décadas, refinarias e outras indústrias de processos reconheceram a necessidade da melhoria da informação e do acesso mais rápido a ela. Para atingir esta meta, um grande investimento foi feito em DCS's (*Distributed Control System* – Sistema de Controle Distribuído), Bancos de Dados e infraestrutura de coleta de dados. O resultado de todo este investimento é que uma

quantidade cada vez maior de dados tem sido disponibilizada com o acesso cada vez mais facilitado. Em qualquer planta química moderna, como refinarias ou petroquímicas, existem milhares de variáveis, como por exemplo taxas de fluxo, temperaturas, pressões, medidas de nível e composições que são constantemente medidas e automaticamente gravadas para propósito de controle de processo, otimização *on line* e avaliação econômica. Computadores e sistemas de aquisição de dados facilitam a armazenagem e processamento de um grande volume de dados, amostrado às vezes com freqüência na ordem de minutos ou mesmo segundos. O uso de computadores não somente permite que sejam obtidos dados com uma freqüência maior mas resulta também na eliminação dos erros presentes na armazenagem manual. Por si só, este fato aumentou muito a exatidão e a validade dos dados de processo. Além disso, a grande quantidade de dados pode ser usada no aumento da exatidão e consistência através de sua verificação sistemática e tratamento.

Contudo, a grande quantidade de dados gera inconsistências, as quais resultam inevitavelmente em perda de confiança no sistema de medição. O mesmo conjunto de dados é, então, manipulado por diferentes grupos, de diferentes maneiras, o que faz com que as mesmas medidas tenham uma variedade de ajustes, todos diferentes entre si. Além disso, as medidas de um processo são inevitavelmente corrompidas por erros durante a própria medição, seu processamento e transmissão. O erro total em uma medida, que é a diferença entre o valor medido e o valor verdadeiro da variável, pode ser representado como a soma da contribuição de dois tipos de erros – erros aleatórios e erros grosseiros.

O termo erro aleatório implica que nem a magnitude nem o nível do erro pode ser predito com certeza ou, em outras palavras, se uma medida é repetida com o mesmo instrumento sob condições idênticas de processo, um valor diferente pode ser obtido dependendo do valor do erro aleatório. A única maneira de caracterizar estes erros é através de uma distribuição de probabilidades. Por outro lado, os erros grosseiros são causados por eventos não aleatórios, tais como mal funcionamento devido a instalação imprópria dos equipamentos de medição, descalibração, desgaste ou corrosão dos sensores e depósitos sólidos. A natureza não aleatória destes erros implica que num dado tempo, eles têm uma certa magnitude e sinal, os quais podem ser desconhecidos. Assim, se a medida é repetida com o mesmo instrumento sob condições idênticas, a contribuição de um erro sistemático ao valor medido tende a ser a mesma.

Erros nos dados medidos podem levar a uma significativa deterioração no desempenho da planta. Pequenos erros aleatórios e erros grosseiros podem levar à deterioração do desempenho dos sistemas de controle enquanto que erros grosseiros maiores podem anular ganhos alcançados

através da otimização do processo. Em alguns casos, dados errôneos podem conduzir o processo a um regime de operação não econômico, ou o pior, inseguro. Sendo assim, é importante reduzir, senão eliminar completamente, os efeitos tanto dos erros aleatórios quanto dos erros grosseiros.

Assim, quantidade de dados não é sinônimo de qualidade da informação. Devido aos erros inerentes à qualquer medida e a erros de origens diversas (falha humana, vazamentos, isolamento térmico ineficaz) uma grande massa de dados acaba se transformando em pouca informação ou em informação de pouca qualidade. O investimento agora é direcionado à qualidade da informação.

Desta forma, se faz necessária uma metodologia padrão a ser aplicada sobre os dados, para que estes se convertam em informação útil e de qualidade. Médias, filtros simples e ajustes baseados na experiência são alguns dos possíveis métodos de condicionamento. A pesquisa e desenvolvimento na área de condicionamento de dados levou ao projeto de filtros analógicos e digitais que podem ser usados para atenuar o efeitos do ruído de alta freqüência entre as medidas. Erros grosseiros de grande tamanho podem ser inicialmente detectados usando-se de várias verificações de validação dos dados que incluem procedimentos para analisar se o dado medido e taxa na qual ele está variando estão dentro de limites operacionais. Nos dias de hoje, estão disponíveis sensores inteligentes que podem diagnosticar a ocorrência de problemas de *hardware* e se o dado medido é aceitável.

Técnicas mais sofisticadas incluem os testes de controle estatístico de processo (CEP) que podem ser usadas para detectar erros significativos (*outliers*) entre os dados. Estas técnicas são normalmente aplicadas à cada variável separadamente de modo que mesmo aumentando a exatidão das medidas, elas não fazem uso do modelo do processo e portanto não asseguram consistência dos dados com respeito às interrelações entre as diferentes variáveis. No entanto, estas técnicas devem ser usadas como um primeiro passo para reduzir os efeitos dos erros aleatórios nos dados e eliminar erros grosseiros.

É possível reduzir mais ainda os efeitos dos erros aleatórios e também eliminar os erros grosseiros sistemáticos nos dados explorando as relações conhecidas entre as diferentes variáveis de um processo. As técnicas de reconciliação de dados e detecção de erros grosseiros que começaram a ser desenvolvidas na engenharia química na década de 1960 exploram estas relações.

A principal diferença entre a reconciliação de dados e as técnicas de filtragem é que a reconciliação faz uso explicito do modelo de processo e obtém estimativas das variáveis de processo por ajuste às medidas de modo que as estimativas satisfaçam às restrições. A reconciliação de dados consiste em um tratamento matemático objetivando a melhoria da qualidade dos dados que repre-

sentam processos experimentais ou industriais (PLÁCIDO, 1995). É uma ferramenta relativamente nova e ainda pouco abordada e conhecida, tanto no meio acadêmico quanto industrial. As técnicas de reconciliação de dados e detecção/identificação de erros grosseiros são eminentemente estatísticas mas fazem uso dos princípios de conservação de massa e energia como modelos de restrição e os resultados do seu emprego podem ser bastante expressivos, principalmente na otimização de processos, mas também no diagnóstico de falhas em uma planta.

As estimativas reconciliadas são supostamente mais exatas que as medidas e o mais importante, são também consistentes com as relações conhecidas entre as variáveis de processo definidas nas restrições. Para que a reconciliação de dados seja efetiva, não deve haver a presença de erros grosseiros nem entre as medidas nem no modelo do processo. A detecção de erros grosseiros é uma técnica intimamente associada à reconciliação de dados e foi desenvolvida para identificar e eliminar erros grosseiros.

A reconciliação de dados e a detecção de erros grosseiros obtêm a redução de erros pela exploração da redundância nas medidas. Tipicamente em qualquer processo as variáveis se relacionam umas às outras através de restrições físicas tais como leis de conservação de massa e energia. Dado um conjunto de restrições do sistema, um número mínimo de medidas livres de erros grosseiros é necessário para calcular todas as outras variáveis e parâmetros do sistema. Se houver mais medidas que este mínimo, então há a redundância nas medidas e esta pode ser explorada. Este tipo de redundância é normalmente chamada de redundância espacial e o sistema de equação é chamado de sobredeterminado.

A reconciliação de dados não pode ser realizada sem a redundância espacial. Se não houver informação extra, o sistema é chamado de determinado e nenhuma correção de medidas errôneas é possível e, além disso, se há menos variáveis que o necessário para determinar o sistema, este é dito indeterminado e o valor de algumas variáveis só pode ser estimado através de outros meios ou através da adição de algumas medidas.

Um segundo tipo de redundância nas medidas é a redundância temporal e surge do fato de que as medidas de um processo são feitas continuamente no tempo a uma taxa elevada de amostragem, produzindo mais dados que o necessário para determinar um processo em estado estacionário. Se o processo é considerado em estado estacionário, a redundância temporal é apresentada por exemplo através de uma simples média dos valores com a subseqüente aplicação da reconciliação de dados sobre a esta média.

Se o processo é dinâmico, contudo, a evolução do seu estado é descrita por equações diferenciais correspondendo aos balanços de massa e energia que capturam inerentemente tanto a redundância espacial quanto temporal das variáveis medidas. Para tais processos, a reconciliação de dados dinâmica e técnicas de detecção de erros grosseiros foram desenvolvidas para obter estimativas com maior exatidão e consistentes com o modelo do processo.

As estratégias de reconciliação de dados e detecção de erros grosseiros se combinam em uma seqüência de passos que estão esquematizados de forma simplificada na Figura 1.1.

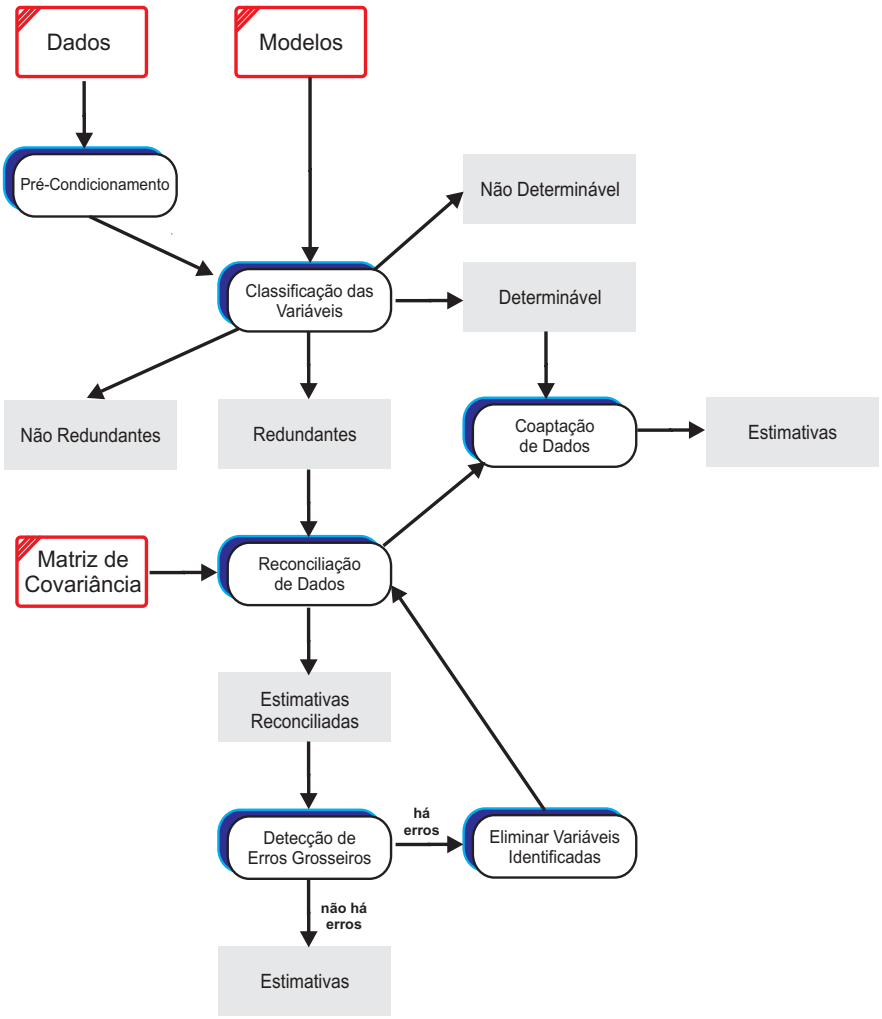


Figura 1.1: Algoritmo de reconciliação de dados/detecção de erros grosseiros

Nesta figura, os insumos básicos para a seqüência de tratamento de dados são os dados brutos, os modelos dos sistemas considerados e a descrição probabilística dos erros nos dados medidos, codificada na forma da matriz de covariância. Os dados brutos podem passar por algum pré-tratamento, como médias simples ou filtros, ou alimentarem diretamente um procedimento de

classificação de variáveis que vai separar as variáveis em quatro tipos: **não determináveis**, que não sofrem nenhum tratamento posterior, **não redundantes**, cujos valores medidos são a melhor estimativa possível para a variável, **redundantes**, que vão ser tratadas pela reconciliação de dados e as **determináveis**, que com auxílio das estimativas reconciliadas podem ser estimadas. O esquema mostra também que os dados reconciliados devem passar por uma detecção de erros grosseiros para determinar a presença de algum erro deste tipo. Se houver, uma possibilidade é definir essa variável portadora de erro como não medida e tentar obter uma estimativa para ela repetindo o procedimento de reconciliação sem a sua presença.

Ao fim dos procedimentos esquematizados na figura, os dados estimados se tornam disponíveis para outros sistemas. Esses dados podem substituir as leituras originais ou serem usados para algum tipo de métrica para análises posteriores.

O processamento de sinais e as técnicas de reconciliação de dados para redução de erros podem ser aplicadas a processos industriais como parte de uma estratégia integrada, chamada de condicionamento ou retificação de dados. A Figura 1.2 ilustra as várias operações e a posição ocupada pela reconciliação de dados no condicionamento de dados para aplicações *on line*.

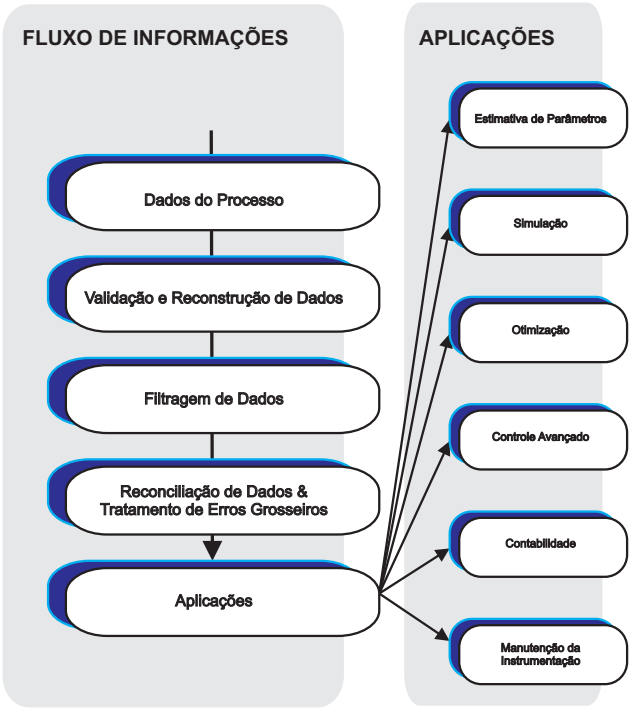


Figura 1.2: Sistema de coleta e condicionamento de dados *on-line* – adaptado de Narasimhan e Jordache (2000)

Distribuidores de *softwares* comerciais apregoam que a introdução da tecnologia de reconcili-

ação de dados resultou em milhões de dólares de economia. Perdas de petróleo vêm de cálculos inexatos de transações, de evaporação de tanques e derramamentos. Desbalanços entre as medições de entrada e saída em refinarias podem ficar na ordem de 0,5 % a 1,0 % (MILES; JELFFS, 1988) ou até maiores, dependendo da qualidade da instrumentação. Em uma refinaria típica, processando milhares de barris de óleo cru por dia, isto representa milhões de dólares por ano em petróleo que não é contabilizado. Qualquer projeto objetivando a identificação destas perdas pode chegar a este tipo de economia. Contudo, perdas não são a única questão de interesse. Uma instrumentação enviesada pode levar não apenas a uma subestimação dos produtos produzidos, mas também afetar o monitoramento, reduzindo assim a eficiência de operação da planta. Ao mesmo tempo, a identificação apropriada de uma instrumentação enviesada permite uma melhor manutenção e previne acidentes.

Segundo Narasimhan e Jordache (2000), o desenvolvimento de um programa para reconciliação de dados e detecção de erros grosseiros para um sistema e sua implementação prática é uma tarefa complexa. A justificativa para o uso destas técnicas pode vir de várias aplicações importantes para melhora de desempenho, como ilustrado na Figura 1.2, que requerem dados com maior exatidão para alcançarem os benefícios esperados. Alguns exemplos de aplicação são listados a seguir.

i. Uma aplicação direta da reconciliação de dados está na avaliação de produtos do processo ou na inferência do consumo de utilidades em diferentes unidades do processo. Valores reconciliados provêem estimativas com maior exatidão quando comparadas com o uso de medições não tratadas. Por exemplo, a reconciliação de balanço material na escala completa de uma refinaria ajuda numa melhor estimativa da produção global da mesma. De uma maneira similar, uma auditoria dos balanços energéticos usando fluxos e temperaturas reconciliados ajuda numa melhor identificação de processos e equipamentos energeticamente ineficientes.

ii. Aplicações como simulação e otimização de um equipamento de processo existente necessitam do modelo do equipamento com detalhes suficientes para sua adequada representação. Estes modelos comumente contêm parâmetros, os quais têm que ser estimados a partir de dados da planta. Isto é conhecido como *sintonia do modelo* ou *identificação do modelo*, para o qual dados com maior exatidão são essenciais. O uso de medidas errôneas na sintonia do modelo pode dar lugar a parâmetros incorretos que, eventualmente, anulem os benefícios alcançáveis através da otimização. Existem duas possibilidades de utilização da reconciliação

de dados em tais aplicações e estas são ilustradas a seguir, tomando-se como exemplo uma coluna de destilação:

- Considerando-se o problema de otimização do desempenho de uma coluna de destilação existente, pode-se obter dados operacionais, medidas de fluxo, temperatura e composição de todas as correntes de saída e entrada. Um caminho possível é reconciliar estas medidas usando somente balanços globais de massa e energia em torno da coluna. Estes dados reconciliados podem ser usados em conjunto com um modelo prato-a-prato detalhado da coluna, no sentido de se estimar parâmetros tais como eficiência de prato. O modelo sintonizado pode então ser usado para otimizar o desempenho da coluna.
- Em uma outra abordagem, pode ser feita uma reconciliação de dados simultânea com uma estimativa de parâmetros usando-se o modelo detalhado prato-a-prato da coluna. Neste caso, se as medidas de temperatura do prato e/ou composições estão disponíveis, estas também podem ser usadas e reconciliadas como parte do problema. Obviamente, a segunda abordagem leva a um significante aumento de esforço e de tempo computacional. Esta última abordagem é também referida como modelagem rigorosa on line e vem sendo incorporada em vários simuladores em estado estacionário comerciais.

iii. A reconciliação de dados pode ser muito útil no planejamento de manutenção de equipamento de processo. Dados reconciliados podem ser usados para estimar com grande exatidão o desempenho de parâmetros-chave dos equipamentos. Por exemplo, o coeficiente de troca térmica dos trocadores de calor ou o nível de atividade do catalisador nos reatores pode ser estimado e usado para determinar se o trocador deve ser limpo ou se o catalisador deve ser reposto/regenerado, respectivamente.

iv. Várias estratégias de controle avançado tais como controle baseado em modelo ou controle inferencial requerem estimativas com maior exatidão das variáveis controladas. As técnicas de reconciliação de dados dinâmica podem ser usadas para derivar estas estimativas para um melhor controle de processos.

v. A detecção de erros grosseiros não somente aumenta a exatidão das estimativas dos procedimentos de reconciliação, como também é útil na identificação de problemas de instrumentação, os quais requerem manutenção especial e correção. Uma incipiente detecção de erros grosseiros pode reduzir os custos de manutenção e promover a operação da planta de uma

maneira mais suave. Estes métodos também podem ser estendidos para detectar equipamentos defeituosos.

1.3 Organização da Tese

No Capítulo 2 são introduzidos os conceitos fundamentais e a terminologia básica para logo em seguida ser apresentada uma revisão bibliográfica com alguns trabalhos relevantes separados entre os dois temas principais: a reconciliação de dados e a detecção de erros grosseiros.

O Capítulo 3 descreve a fundamentação matemática do problema de reconciliação de dados, apresenta alguns conceitos básicos que serão usados por todo o texto desta tese, como a decomposição do problema de estimativa, e aponta abordagens para a solução de problemas de reconciliação de dados em estado estacionário para sistemas lineares.

O Capítulo 4 aborda o problema da reconciliação de dados estacionária em sistemas com a peculiaridade da não linearidade ser fruto do produto de duas variáveis. A bilinearidade é explorada na formulação do problema e nas suas estratégias de solução. São também descritas suas vantagens e desvantagens.

O Capítulo 5 descreve técnicas de reconciliação de dados não linear, iniciando com a formulação do problema e seguindo com os métodos que lidam com problemas sujeitos a restrições de igualdade e de desigualdade. Também é tratado o problema de classificação de variáveis e feita uma comparação de estratégias de otimização não linear para a solução do problema.

O Capítulo 6 apresenta conceitos e abordagens disponíveis para realizar a reconciliação de dados em sistemas transientes. É tratada a estimativa ótima de estado usando o filtro de Kalman e é feita uma analogia entre o filtro de Kalman e a reconciliação de dados.

O Capítulo 7 é dedicado à detecção de erros grosseiros, descrevendo os principais testes estatísticos usados na detecção deste tipo de erro e apresentando também técnicas para sua identificação.

O Capítulo 8 descreve a implementação de elementos da teoria vista nos capítulos anteriores e apresenta os detalhes do desenvolvimento do aplicativo ***Reconciliare***. As conclusões deste trabalho são discutidas e são dadas sugestões para trabalhos futuros.

2 *Conceitos Fundamentais e Trabalhos Relevantes em Reconciliação de Dados e Detecção de Erros Grosseiros*

Neste capítulo são apresentados os termos mais freqüentes encontrados na literatura sobre os temas abordados nessa tese, são dadas algumas justificativas da importância deste trabalho e são apresentados alguns conceitos fundamentais para a sua compreensão. Na seqüência é feita uma revisão bibliográfica sobre reconciliação de dados e detecção de erros grosseiros, onde são apresentados alguns trabalhos relevantes nas áreas da reconciliação de dados e detecção de erros grosseiros.

2.1 Informação de processo e gerenciamento de operações

A instrumentação é necessária nas plantas de processos químicos para obter dados essenciais na realização de várias atividades. Dentre as mais importantes estão o controle, a avaliação da qualidade de produtos, a contabilidade da produção e a detecção de falhas relacionadas à segurança. Além disso, esses dados permitem a obtenção de alguns parâmetros que não podem ser medidos diretamente, como a incrustação de trocadores de calor ou eficiências de coluna.

Nas últimas décadas, a indústria de processos químicos incorporou novas tecnologias na forma de vários programas computacionais que ajudam a coletar, filtrar, organizar e usar a informação de plantas para várias atividades técnicas e de gerenciamento. Através destes pacotes de *software*, a qualidade dos produtos e o controle de custos melhorou consideravelmente. A eficiência das operações, que no passado dependia da experiência dos operadores, pode agora depender de controladores supervisórios interativos baseados em computadores operando sobre toda a planta. Além disso, a contabilidade de produção, o planejamento de operações e o agendamento de manutenção se beneficiam de dados mais confiáveis e exatos. Em vista de tais avanços no processamento de

dados, até mesmo procedimentos gerenciais estão sofrendo revisão (BAGAJEWICZ, 2001).

A maioria das plantas de processos é projetada para funcionar em condições de estado estacionário. Na prática, estas condições não são estritamente alcançadas porque as plantas estão sujeitas a variações imprevisíveis. Contudo, a hipótese de estado estacionário é ainda usado com sucesso para a análise dos dados coletados, com a óbvia exceção do controle. Esta análise, seguida das tomadas de decisão, cobre várias atividades de operação das plantas, quais sejam:

Monitoramento da operação: Esta é a atividade de curto termo na qual os dados são usados como parte das malhas de controle ou para alterar os ajustes dos controladores.

Detecção de falhas: Esta atividade inclui a detecção de falhas de instrumentação e de equipamento, além da avaliação e quantificação de vazamentos.

Análise de desempenho: Esta atividade é tipicamente realizada numa base diária e cobre o que se chama comumente de contabilidade de produção ou movimentação de produtos.

Modelagem de processos: A atividade de simulação se provou uma ferramenta eficiente para o engenheiro de processos. Ela oferece meios de analisar condições de operação e de projeto alternativas, avaliar melhorias e detectar problemas operacionais.

Planejamento de operações: Atividades comuns como a programação de atividades, o monitoramento de incrustações, a limpeza de trocadores de calor e ativação de catalisadores dependem de dados de processo e sua modelagem adequada.

Planejamento de produção: Esta atividade é realizada em uma escala de tempo maior. Vários pacotes de *software* foram desenvolvidos para tratar deste problema, especialmente na indústria do petróleo.

Planejamento de manutenção: Perdas de produção aumentam com uma manutenção pobre, mas os custos crescem com o aumento da freqüência de manutenção. Um equilíbrio é alcançado quando a mínima perda de produção é obtida com o mínimo de manutenção. Se o plano de manutenção pode ser modificado pelo uso de técnicas que permitam a identificação de possíveis problemas o mais cedo possível, o custo global é então reduzido.

Estimativa de parâmetros: A estimativa de parâmetros é de suma importância prática. Tipicamente estes parâmetros são eficiências de colunas, coeficientes de transferência de calor,

eficiências de vaporização *flash*, etc. Todos são impossíveis de se medir diretamente. Em vários casos, a precisão e a localização de uma instrumentação já existente são insuficientes para obter estimativas de boa qualidade. Um exemplo comum é a predição da incrustação em fracionadoras de óleo cru, que é uma informação vital para a otimização de ciclos de limpeza.

Otimização *on line*: Os dados da planta são usados para sintonizar os parâmetros de modelos usados em simulação, que por sua vez é usada como base para otimização da operação da planta. Como resultado, um novo estado estacionário que maximiza o lucro é alcançado. Esta nova tecnologia foi introduzida na indústria nos anos 80 e resultou em milhões de dólares de economia em custos operacionais (de 2% a 20%) para diferentes cenários de aplicação: planta de fracionamento de crus (MULLICK, 1993 apud BAGAJEWICZ, 2001), destilação (SMITH, 1996 apud BAGAJEWICZ, 2001), plantas de etileno (LAUKS *et al.*, 1992 apud BAGAJEWICZ, 2001), plantas de olefinas, craqueamento catalítico, hidrocraqueamento, unidades FCC e outros casos (ZHANG *et al.*, 1995 apud BAGAJEWICZ, 2001) e (BRYDGES *et al.*, 1998 apud BAGAJEWICZ, 2001). A Figura 2.1 mostra como os ciclos de otimização em tempo real operam nas plantas (FORBES; MARLIN, 1996 apud BAGAJEWICZ, 2001).

Planejamento de Recursos Empresariais: O terno **ERP** (*Enterprise Resource Planning*) foi cunhado para a atividade de planejamento baseada principalmente nos procedimentos de otimização, que venha a abarcar várias das atividades de planejamento que foram citadas anteriormente (operações, produção e manutenção) em conjunto com o planejamento financeiro e gerencial (BADELL *et al.*; BUNCH; GROSDIDIER, 1998b, 1998, apud BAGAJEWICZ, 2001). Assim, há uma integração vertical entre negócios e operação de planta (BADELL; PUIGJANER, 1998a apud BAGAJEWICZ, 2001). Uma integração de todas estas atividades, da operação e monitoramento ao planejamento de longo termo é também conhecida como gerenciamento e controle em escala completa da planta (*Plant-wide Management and Control*) (SWANSON; STEWART; PELHAM; PHARRIS, 1994, 1996 apud BAGAJEWICZ, 2001). Algumas companhias começaram a integrar o gerenciamento e o planejamento de recursos com as atividades supracitadas (BENSON; NATORI; TJOA; HARKINS, 1995, 1998, 1999 apud BAGAJEWICZ, 2001).

Fica evidente que a qualidade dos dados é um aspecto chave em todas as atividades listadas. Algumas delas, como a otimização *on line*, não podem ser realizadas sem dados exatos e consistentes. Isto desafia os projetistas a determinar a melhor rede de sensores, em termo do número de instrumentos, localização e qualidade, que é necessária para que todo o ciclo funcione apropriadamente.

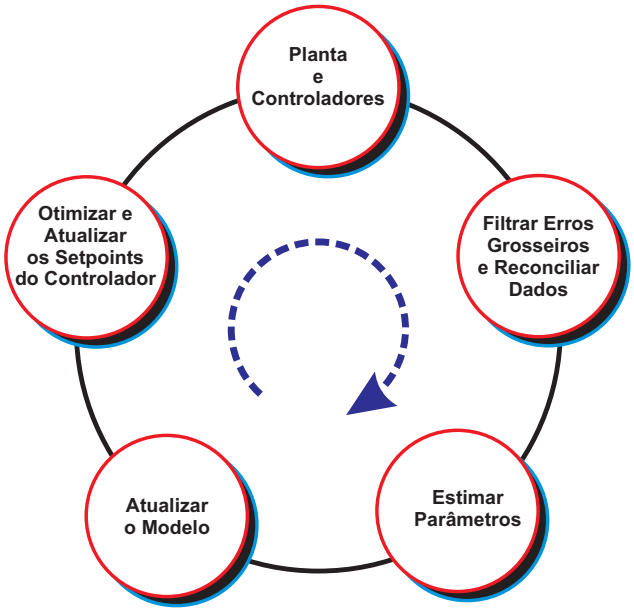


Figura 2.1: Otimização em tempo real – adaptado de Bagajewicz (2001)

A escolha da instrumentação deixou, portanto, de ser um problema localizado e independente na planta, para se tornar global e multiobjetivo (BAGAJEWICZ, 2001).

2.2 Monitoramento baseado em modelos

Segundo Bagajewicz (2001), o **monitoramento baseado em modelos** consiste do uso de uma combinação de modelos (descrições formais do comportamento do processo) e medições *on line* para atingir os seguintes objetivos:

i. Produzir estimativas de variáveis medidas e não medidas[1];

ii. Identificar instrumentação defeituosa;

iii. Identificar condições de operação falhas ou inseguras e suas origens;

iv. Identificar eventos que tenham impacto na eficiência e na qualidade dos produtos.

[1]O termo "estimativa" não se refere somente ao valor, resultado de um procedimento matemático, que "substituiria" a própria medida. Ao invés disso, a medida é também uma estimativa, em sentido amplo, do **valor verdadeiro** de uma dada variável. Os instrumentos de medição "estimam" o valor verdadeiro dessa variável que, em última instância, a natureza não dá nunca ao conhecimento.

Kramer e Mah (1993), em uma abrangente revisão sobre o assunto, discutiram vários cenários nos quais boas estimativas dos dados e a detecção de erros grosseiros são técnicas que ajudam a alcançar as metas **i** e **ii** e constituem um caso particular do conceito de retificação de dados. Ainda que a reconciliação de dados dependa de restrições analíticas e mais ainda da estimativa por mínimos quadrados, a retificação de dados pode obter estas estimativas através de técnicas como o filtro de Kalman, reconhecimento de padrões, redes neurais, análise de componentes principais e mínimos quadrados parciais. Do mesmo modo, a detecção de falhas depende de técnicas baseadas na estatística, mas várias outras técnicas podem ser usadas. As metas **iii** e **iv** são conseqüência direta do uso de um modelo.

O campo do projeto e atualização de redes de sensores tradicionalmente repousa em conceitos baseados em modelos. Quase todos os trabalhos voltados para metas de monitoramento incluem:

- A capacidade da rede de sensores de prover estimativas para as variáveis de interesse;
- A capacidade de garantir uma certa exatidão através da reconciliação de dados;
- Confiabilidade;
- A capacidade de identificação de erros grosseiros através de técnicas estatísticas baseadas em modelos;
- A capacidade de identificação de falhas no processo.

2.3 Erros aleatórios e erros grosseiros

Quando se descreve um processo, têm-se dois problemas principais: fazer com que as flutuações inerentes à dinâmica e os erros de diferentes naturezas não afastem o conjunto dos dados das restrições (balanços de massa, balanços de energia, somatório de frações molares) e, ainda, estimar dados que não são diretamente medidos. Segundo Crowe (1996), qualquer conjunto de variáveis de um processo que não obedece às leis de conservação (balanços de massa, de energia e somatórios de frações mássicas e molares) é considerado portador de erros.

O erros podem advir de uma leitura errada, ou seja, de um problema no equipamento de medição ou pode ser fruto de flutuações naturais nos dados do processo, oriundas de diversas causas, que faz com que os dados, mesmo sendo corretamente medidos, levem à fuga dos referidos balanços.

Esta análise primária é útil, pois a partir dela pode-se utilizar os procedimentos de reconciliação de dados para identificar problemas nos equipamentos de medição auxiliando a manutenção preventiva.

No tocante aos aspectos estocásticos dos erros, sabe-se que os **erros aleatórios** (também chamados de randômicos) são comuns a todo processo real de obtenção de dados, constituindo pequenos desvios em relação a um valor central. A concepção de regime permanente, rigorosamente falando, é extremamente útil na compreensão e modelagem matemática de diversos fenômenos e equipamentos na indústria de processos, mas não corresponde à realidade por mais que um átimo de tempo, haja vista as perturbações naturais ou provocadas de diversas origens e tipos, como as ações de controle, impurezas, vibrações nas bombas, sem falar nas que fogem totalmente de qualquer possibilidade de controle ou mesmo de serem evitadas, como os raios cósmicos, variações na atividade solar e alternância de marés (ação gravitacional da Lua). Assim, uma avaliação das medidas disponíveis de um processo com base na precisão nominal dos medidores é incompleta quando não contabiliza também a "vibração" do processo. (PLÁCIDO, 1995).

Ao contrário dos erros aleatórios, normalmente distribuídos, com média zero, uma variância conhecida e de pequena magnitude, os **erros grosseiros**[2] (*gross errors*) são associados a eventos não aleatórios como por exemplo:

- descalibração ou calibração deficiente dos instrumentos de medição que provocam desvios;
- vazamento insuspeito nas unidades do processo;
- mal-funcionamento ou falha nos instrumentos;
- regime transiente ou flutuações naturais do processo;
- flutuações nas condições ambientais;
- erros de amostragem e de análise;
- leituras feitas por diferentes observadores;

[2]O termo *gross* em inglês pode ser traduzido como "bruto" (no sentido de completo ou global), "grande" ou ainda "flagrante" ou "extremo". A tradução da expressão *gross errors* como "erros grosseiros" em português parece não captar o seu sentido real. Mesmo sendo amplamente usada a expressão "erros grosseiros" na literatura em lingua portuguesa, esta seria melhor traduzida como "erros flagrantes" ou "erros de grande monta ou magnitude", haja vista que "grosseiro" não tem nenhuma das acepções de *gross* mostradas acima.

- dados levantados em tempos não-simultâneos, pertencentes a regimes permanentes distintos, etc. (PLÁCIDO, 1995).

2.3.1 Detecção de falhas de processo

As falhas se propagam através de todo o processo, alterando as leituras da instrumentação (pressões, temperaturas, taxas de fluxo, etc.). Assim, os sensores devem ser capazes de determinar quando o sistema sai da operação normal. A qualidade dos dados necessária para esta tarefa é, portanto, diferente. Enquanto que na operação normal o foco é na exatidão da estimativa das variáveis chave, em situações anormais o foco é desviado para a observabilidade e o diagnóstico apropriado de falhas. Assim, um dos problemas é prevenir o falso diagnóstico devido a uma instrumentação comprometida.

2.4 Qualidade de dados

A **qualidade de dados** é um termo amplo que é freqüentemente usado para se referir a várias propriedades que um conjunto de dados deve possuir. Estas propriedades podem ser condensadas nos três atributos amplos:

- Exatidão;
- Precisão ou reprodutibilidade;
- Confiabilidade.

A **exatidão**[3] é a habilidade de um instrumento em medir o valor correto ou "verdadeiro". Por sua vez, a **precisão** pode ser definida como a habilidade de um sensor de reproduzir um valor dentro de um certo intervalo. Assim, um instrumento pode ser preciso, mas não exato. Isto ocorre quando repetidas medições da mesma variável recaem dentro de um pequeno intervalo que não contém o valor verdadeiro. Por outro lado, um instrumento exato pode apresentar uma reprodutibilidade pobre, mas o valor da média de várias medições pode ser próximo do valor verdadeiro. A Figura 2.2 ilustra estes conceitos.

[3] O termo exatidão na literatura em língua inglesa nos contextos envolvidos neste trabalho é referido invariavelmente como ***accuracy***. Aqui optou-se por usar o termo "exatidão" ao invés de "acurácia".

Figura 2.2: Exatidão e precisão – adaptado de Bagajewicz (2001))

Ainda que estes conceitos sejam bem simples, eles são por vezes confundidos. É comum observar engenheiros e pessoal de operação se referirem a sistemas precisos como "exatos" e vice-versa. E mais, segundo Bagajewicz (2001) alguns livros e manuais usam o conceito de precisão como fazendo parte de um conceito maior e mais geral de exatidão, que incluiria tanto a variância das medições quanto o desvio da média em relação ao valor verdadeiro.

A **confiabilidade dos dados** é definida como a probabilidade de os dados estarem realmente presentes durante um dado período de tempo. Por sua vez, a **disponibilidade dos dados** é a probabilidade dos dados estarem presentes em um dado instante de tempo. Desta forma, a confiabilidade é um requisito mais rigoroso (BAGAJEWICZ, 2001).

2.4.1 Redundância analítica

Somando-se às medições diretas, existem formas indiretas de determinar variáveis de processo. Portanto, passa-se a classificar aquisição de dados como:

- Aquisição de dados baseada somente em leituras de instrumentação;
- Aquisição de dados aprimorada por *software*.

Quando a aquisição de dados é puramente suportada por leituras de instrumentação, cada valor em particular de uma variável do sistema é diretamente associado à sua fonte, o instrumento que a mede. Não fossem as restrições de custo, uma instrumentação precisa e à prova de falhas seria suficiente na obtenção de estimativas confiáveis e precisas. Contudo, em quase todos os casos, a probabilidade de falha de um sensor não é desprezível e a redundância é usada como um meio para garantir a disponibilidade dos dados. Com a redundância vêm as discrepâncias entre as medições

advindas de uma instrumentação diversa relacionada à mesma variável. Não haveria problema se os ruídos associados ao sinal fossem pequenos o suficiente, mas como geralmente este não é o caso, as leituras têm que ser reconciliadas. Assim, a redundância é classificada como:

- Redundância de *hardware*;
- Redundância analítica.

A redundância de *hardware* é caracterizada por dois ou mais sensores usados para medir a mesma variável. Este é o caso para dois termopares medindo a temperatura dentro de um vaso ou uma tubulação no qual esta temperatura seja supostamente homogênea (a exemplo de uma caldeira). Um outro exemplo é o uso de dois medidores de vazão para medir o mesmo fluxo.

A redundância analítica é caracterizada por um conjunto de medições de variáveis diferentes que a priori satisfaçam um modelo matemático. Um exemplo disto é uma unidade com várias corrente de entrada e saída. Se a taxa de fluxo de todas as correntes (redé medida com um único instrumento por fluxo, não há redundância de *hardware*, mas devido à soma dos fluxos de entrada ser obrigatoriamente igual à soma dos fluxos de saída, a redundância é agora analítica. Em outras palavras, duas estimativas estão disponíveis para cada taxa de fluxo. Uma vem da medida direta e a outra é obtida pelo uso de equações de balanço material e isto constitui um conflito que precisa ser resolvido.

2.4.2 Reconciliação de dados

Segundo Bagajewicz (2001), devido às leituras dos instrumentos serem inexatas e não obedecerem sozinhas às leis básicas de conservação, existe a necessidade de determinar as melhores estimativas a partir de um conjunto conflitante de leituras. Contudo, em várias instalações, os operadores ainda assumem que as leituras são suficientemente exatas para os propósitos de monitoramento e controle. Por causa dos conflitos criados pelos desbalanços, freqüentemente o balanço é forçado através de várias técnicas heurísticas. Por exemplo, a eliminação ou correção manual de medições que na experiência do operador são menos confiáveis é uma abordagem das mais comuns. A **reconciliação de dados** lida com a tarefa de determinar as melhores estimativas estatísticas para todas as variáveis. As técnicas de reconciliação de dados foram introduzidas em meados da década de 80, em especial na indústria do petróleo. Estas técnicas não somente podem melhorar a exatidão dos dados da planta, como também são capazes de detectar e filtrar viéses de instrumentação e de

identificar vazamentos. Vários livros cobrem diversos aspectos da reconciliação de dados (MAH, 1990; MADRON, 1992; VEVERKA; MADRON, 1997; ROMAGNOLI; SÁNCHEZ, 2000; NARASIMHAN; JORDACHE, 2000; BAGAJEWICZ, 2001). Além desses, centenas de artigos são devotados ao tema.

Os dados têm que ser exatos e consistentes. A exatidão é obtida através de uma seleção apropriada da instrumentação e de sua calibração. A consistência é testada em um primeiro nível pelo operador da planta e posteriormente pelo engenheiro de processo. Subjacente à cada uma destas avaliações está um modelo baseado em conhecimento empírico da planta ou nas leis fundamentais da natureza (conservação de massa e energia). No caso da conservação total de massa, a forma matemática deste modelo é um simples conjunto de equações lineares. Quando são realizados balanços de componente, o conjunto de equações é bilinear. O balanço completo de energia envolvendo temperatura, composições, taxas de fluxo e pressões como variáveis é também não linear, em especial porque inclui a avaliação de propriedades termodinâmicas. Assim, a reconciliação de dados é um meio sistemático de realizar esta avaliação levando em consideração a precisão de cada uma das medições e fazendo uso de ferramentas estatísticas.

A grande maioria dos procedimentos de reconciliação de dados é baseada na suposição de que as medidas estão sujeitas somente a erros aleatórios. Esta suposição não é, em geral, confirmada pela realidade e as medidas dos processos estão também sujeitas a *biases* ou viéses[4] para fora da distribuição normal.

Existem ainda enfoques alternativos nos quais não seria necessário remover as variáveis portadoras de erros grosseiros para o procedimento de reconciliação, como o apresentado por Romagnoli (1983). Neste caso, o próprio modelo é modificado de modo a considerar a presença de erros sistemáticos, acrescentando-se as magnitudes dos desvios em relação às médias, desconhecidas, como incógnitas a serem calculadas. Ambos os enfoques seriam equivalentes, já que levam a sistemas de equações que têm o mesmo grau de liberdade.

O ajuste das medidas para compensar os erros aleatórios envolve a resolução de um problema de minimização sujeito a restrições, comumente do tipo de mínimos quadrados sujeito a restrições. As equações de balanço são incluídas nas restrições e estas podem ser lineares mas são geralmente não lineares. A função objetivo é geralmente quadrática com relação ao ajuste das medidas e tem a

[4] O viés é a diferença entre o valor esperado pelo estimador e o valor verdadeiro da variável sendo estimada. Um estimador não enviesado coincide sua estimativa com o valor verdadeiro da variável. O termo *bias* é amplamente usado na literatura em língua inglesa sobre reconciliação e retificação de dados e significa "tendência". Neste trabalho, no lugar de se usar *bias*, optou-se pelo termo em português "viés".

matriz da covariância[5] dos erros das medidas como fator ponderante. Assim, essa matriz é essencial na obtenção de conhecimento confiável do processo.

A estimativa de parâmetros é também uma atividade importante no projeto, avaliação e controle de processos. Devido aos dados não satisfazerem às restrições do processo, o **método dos erros nas variáveis** provê tanto estimativas de parâmetros quanto estimativas de dados reconciliados que são consistentes com respeito ao modelo. Estes problemas representam uma classe especial de problemas de otimização porque a estrutura dos mínimos quadrados pode ser explorada no desenvolvimento de métodos de otimização. Uma revisão deste assunto pode ser encontrada no trabalho de Biegler *et al.* (1986).

Além de estimativas para as variáveis medidas, a reconciliação de dados é capaz também de prover estimativas para variáveis não medidas ou descartadas. Quando estas variáveis não medidas são parâmetros do processo, a técnica recebe o nome de **estimativa de parâmetros** ou ainda **coaptação de dados**. Contudo, este procedimento propaga erros e deve ser realizado com cuidado.

Finalmente, algumas abordagens estão emergindo de dentro do problema de reconciliação de dados, tais como abordagens **bayesianas** e técnicas robustas de estimativa, bem como estratégias que usam **análise de componentes principais** (PCA – *Principal component analysis*), como encontrado em Consul (2002). Foram propostas também várias formas de validar dados sem a necessidade de usar a reconciliação de dados. Em princípio, o termo *soft sensor* foi cunhado para o uso de dados existentes de processo ou de laboratório para inferir o valor de uma certa medida (MARTIN; HONG *et al.*, 1997, 1999 apud BAGAJEWICZ, 2001) e (GONZAGA *et al.*, in press; MELEIRO; MACIEL FILHO, 2000). Estes *soft sensors* são geralmente baseados em redes neurais e decomposição em ondaletas (*wavelets*). Todas estas técnicas oferecem alternativas viáveis aos métodos tradicionais e provêem novos terrenos para posterior aperfeiçoamento.

2.4.3 Precisão e redundância analítica

A redundância analítica aumenta a precisão das estimativas. Sabe-se que a precisão das estimativas produzidas pela reconciliação de dados é maior que a precisão de medições redundantes.

[5] A variância é uma medida da dispersão estatística de uma variável, sendo a média do quadrado da distância de seus possíveis valores em relação ao valor esperado (a média). Enquanto que a média denota a localização de uma distribuição, a variância denota o seu grau de espalhamento. Já a covariância é uma medida do quão correlacionadas estão duas variáveis. A matriz da covariância (ou variância-covariância) tem o vetor variância na sua diagonal principal. Neste trabalho essa matriz é sempre denotada pelo nome de **matriz de covariância**.

Ou seja, o desvio padrão associado com as estimativas é menor que o desvio padrão de cada medida individual. Na contabilidade de produção e planejamento, esta redução de nível de incerteza é importante. Na indústria do petróleo, um pequeno percentual ganho sobre a incerteza pode levar a diferentes decisões financeiras.

2.4.4 Confiabilidade e redundância analítica

Se os dados têm que ser confiáveis, a instrumentação também tem que ser. Contudo, devido às equações de balanço e outros modelos mais complexos poderem ser usados na estimativa de variáveis, a redundância provê um meio de aumentar a confiabilidade. Portanto, através do uso da reconciliação de dados, pode-se garantir que quando um dado instrumento medindo uma variável diretamente falhe, uma estimativa desta variável esteja disponível através do modelo. Assim, pode-se permitir o uso de uma instrumentação menos confiável e obter reduções de custo, contanto que o esquema como um todo esteja a postos para realizar a estimativa analiticamente.

2.4.5 Detecção de erros grosseiros na instrumentação

Na ausência de erros grosseiros, a reconciliação de dados reduz-se a um simples problema de otimização no qual os ajustes sobre as medições são minimizados sob à condição de que estes valores ajustados satisfaçam à modelagem da planta, em geral balanços de massa e energia. É precisamente a presença de **erros grosseiros** e/ou valores espúrios que torna a tarefa da reconciliação difícil porque estes erros precisam primeiro ser identificados e eliminados.

O defeito de instrumentação é um termo genérico que está relacionado a situações que vão da descalibração à falha total. Na ausência de redundância, a descalibração ou viés não podem ser detectados a menos que o desvio seja tão grande que ser torne óbvio. A redundância, e em especial a redundância analítica, é o único meio de contrastar dados e determinar possíveis falhas deste tipo por meio da filtragem dos dados para detectar, estimar ou ainda eliminar estes viéses. Além disso, os processos freqüentemente podem apresentar vazamentos que podem ser também detectados pela redundância analítica. Esta abordagem foi chamada de **detecção de erros grosseiros** (*gross error detection*) por um grupo de pesquisadores (ROMAGNOLI; STEPHANOPOULOS, 1980a; MAH; TAMHANE, 1982; CROWE *et al.*, 1983; CROWE, 1994; BAGAJEWICZ; JIANG, 1997, 1998) e ultimamente de **validação de dados ou de sensores** (*data or sensor validation*) ou ainda **reconstrução de sinal** (*signal reconstruction*) por um outro grupo (DUNIA *et al.*, 1996; QIN *et al.*, 1997; TAY, 1996;

THAM; PARR, 1994, e outros).

A identificação de erros grosseiros é feita através do uso de diferentes técnicas. As mais populares têm suas bases no teste estatístico de hipóteses. Nestes testes, a hipótese nula é a da ausência de erros grosseiros enquanto que a hipótese alternativa é de que haja pelo menos um erro grosseiro dentro do conjunto de medições consideradas. O conjunto sob análise pode ser o sistema completo (testes globais), uma medida em particular (testes da medida) e/ou uma unidade em particular (testes nodais).

Pode-se resumir algumas diferentes contribuições ao problema da detecção de erros grosseiros na breve revisão a seguir: Reilly e Carpani (1963) foram os primeiros a propor o teste global e o teste nodal. O teste da medida foi proposto por Mah e Tamhane (1987) e Crowe *et al.* (1983). Almasy e Sztano (1975) propuseram o teste da medida da máxima potência que foi objeto do vários artigos subseqüentes. Madron (1985) propôs um teste alternativo com quase a mesma potência do teste da medida de máxima potência. O teste nodal foi introduzido por Reilly e Carpani (1963) e discutido posteriormente por Mah *et al.* (1976). Novas técnicas baseadas na razão de verossimilhança generalizada foram propostas por Narasimhan e Mah (1987) e Narasimhan e Mah (1988) que se provaram posteriormente equivalentes ao teste da medida. Johnston e Kramer (1995) propuseram uma abordagem bayesiana para o problema conjunto da reconciliação de dados e detecção de erros grosseiros. Posteriormente, a análise de componentes principais foi aplicada ao teste de erros grosseiros por Tong e Crowe (1995). Desde então, uma grande quantidade de técnicas de validação de sensores baseadas em **PCA** foram apresentadas (DUNIA *et al.*, 1996; QIN *et al.*, 1997; TAY, 1996, e outros). Finalmente, Rollins e Davis (1992) introduziram uma técnica de estimativa não enviesada que usa os testes de Bonferroni.

Em um esforço de não se deter somente em técnicas baseadas na estatística, as técnicas baseadas em redes neurais foram também propostas como ferramentas para a detecção de erros Grosseiros (GUPTA; NARASIMHAN, 1993; HIMMELBLAU, 1994; KARJALA; HIMMELBLAU, 1994, 1996; REDDY; MAVROVOUNIOTIS, 1998). Cada uma das metodologias citadas apresenta seus próprios problemas, principalmente quando vários erros grosseiros se apresentam simultaneamente pois em geral a possibilidade de falha na detecção de erros grosseiros aumenta com a quantidade desses erros presentes.

2.4.6 Estimativa de erros grosseiros

Uma vez que um único erro grosseiro ou um conjunto de erros sejam detectados, as duas questões a seguir precisam ser respondidas:

- As medições correspondentes devem ser eliminadas ou devem ser estimadas independentemente;
- Como se relaciona um erro grosseiro com um viés, especialmente quando médias de várias medições são reconciliadas? Como diferençar variações do processo de erros grosseiros? Os dados históricos devem ser usados?

Para responder a estas questões, foram desenvolvidos vários métodos. Três tipos de estratégias ajudam a identificar erros grosseiros múltiplos:

- Técnicas de eliminação serial (RIPPS, 1965; SERTH; HEENAN, 1986; ROSENBERG *et al.*, 1987) que identificam um erro grosseiro por vez pelo uso de alguma estatística de teste e eliminam a medida correspondente até que nenhum erro grosseiro seja detectado. Diferentes *softwares* comerciais usam esta técnica baseada no teste da medida e recentemente por meio de testes de componentes principais;
- Técnicas de compensação serial (NARASIMHAN; MAH, 1987), as quais identificam o erro grosseiro e sua dimensão, compensando a medida correspondente e continuando até que nenhum erro seja encontrado;
- Técnicas de compensação simultânea ou coletiva (KELLER *et al.*, 1994; KIM *et al.*, 1997; SÁNCHEZ; ROMAGNOLI, 1994, 1996; SÁNCHEZ *et al.*, 1999), que propõem a estimativa de todos os erros grosseiros simultaneamente. Além desses, Jiang e Bagajewicz (1999) propuseram uma identificação serial com uma estratégia de compensação coletiva (SICC – *Serial Identification with Collective Compensation*) para sistemas dinâmicos e em estado estacionário. Finalmente, a técnica da estimativa não enviesada (UBET – *Unbiased Estimation Technique*), proposta por Rollins e Davis (1992), faz primeiro a identificação e depois a estimativa simultânea.

Vários pesquisadores avaliaram o desempenho destas abordagens. A eliminação serial é simples mas tem o inconveniente da perda de redundância e não é aplicável a erros grosseiros que não

estejam diretamente associados com medições, a exemplo de vazamentos (MAH, 1990). A compensação serial é aplicável a todos os tipos de erros grosseiros. Contudo, seus resultados dependem completamente da exatidão na estimativa da dimensão dos erros grosseiros (ROLLINS; DAVIS, 1992). A compensação coletiva é considerada como mais correta, mas é computacionalmente intensiva e pouco prática (KELLER *et al.*, 1994), não obstante, os resultados dos métodos de compensação simultânea coletiva parecem mais exatos. Por exemplo, o método da estimativa simultânea desenvolvido por Sánchez e Romagnoli (1996) e posteriormente modificado em Sánchez *et al.* (1999) é muito exato. Contudo, ainda não é adequado para sistemas muito grandes pois torna-se combinatorialmente custoso.

2.5 Erros nas medições

Esta seção apresenta aspectos básicos dos erros presentes nas medições. São apresentadas primeiro algumas definições relacionadas com as propriedades da instrumentação e em seguida são cobertos os elementos que influenciam a qualidade dos dados como precisão, erros sistemáticos, histerese, banda morta, sensitividade e velocidade de resposta.

2.5.1 Faixas de medição, alcance e largura de faixa

O intervalo dentro do qual uma certa variável é medida ou transmitida é chamado de faixa de medição (ou ***range***) e é expresso pelos valores inferior e superior das leituras que a medida pode exibir. O alcance (ou ***span***) é a diferença entre este valor superior e o inferior e a largura de faixa ou rangeabilidade (***rangeability***) é definida como a razão do maior valor pelo menor valor que um instrumento pode medir com a mesma exatidão.

2.5.2 Variáveis de influência

Várias variáveis influenciam no desempenho de um instrumento. A temperatura ambiente e umidade, por exemplo, afetam as leituras e introduzem viéses e/ou alteram a precisão. Estas variáveis são ditas de influência (MILLER, 1996) e os seus efeitos não são linearmente aditivos.

2.5.3 Legibilidade

A legibilidade é definida como o menor incremento de escala no qual uma leitura pode ser determinada como um percentual da escala completa (PERRY; GREEN, 1984). Esta propriedade dos sensores é mais relacionada à leitura visual, mas também aparece em dispositivos analógicos e digitais de leitura.

2.6 Qualidade da medida

As medições são sempre sujeitas a erros, não importa o quanto sejam melhoradas as condições e os instrumentos. A precisão e a exatidão por vezes se confundem. A seguir são discutidos elementos que contribuem para a definição.

2.6.1 Precisão

A precisão de um instrumento é definida como a concordância de leitura entre um dado número de medições consecutivas de uma variável que mantenha o seu valor fixo. Isso não tem a ver com o valor "verdadeiro" da variável sendo medida. A Figura 2.3 ilustra o conceito com um conjunto de medições de temperatura da água fervente. As leituras da esquerda e da direita (a e b), correspondem a termômetros diferentes, sendo que o da direita tem o mesmo valor médio, mas apresenta desvios maiores da média.

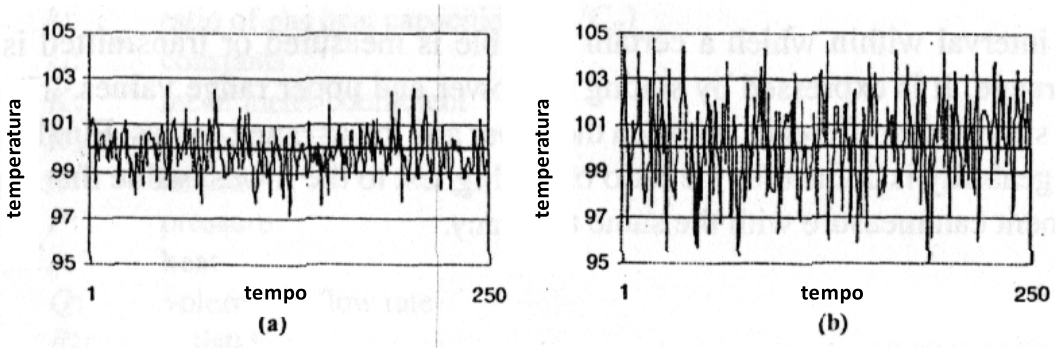


Figura 2.3: Precisão – adaptado a partir de Bagajewicz (2001)

A teoria estatística é usada para definir a precisão. Se a distribuição dos erros é pressupostamente uma gaussiana, então a variância desta distribuição é estimada pelo desvio padrão de uma amostra.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} \tag{2.1}$$

onde x_i são os valores das amostras, $\bar{x}$ é a média e s é o desvio padrão. Uma estimativa estatisticamente não enviesada da variância de uma distribuição normal σ^2 é a variância modificada $\hat{s}^2$, onde

$$\hat{s} = s\sqrt{\frac{n}{n-1}} \tag{2.2}$$

Como são tomadas n medições, a teoria estatística da estimativa diz que a média de uma população μ tem uma probabilidade de $p\%$ de se encontrar dentro de um intervalo $\bar{x} \pm z_p \hat{s}/\sqrt{n}$. O parâmetro z_p é chamado de coeficiente de confiança e é obtido pela determinação do valor de x tal que a área sob a curva no intervalo $[x, -x]$ seja igual a p. Por sua vez, p é chamado de nível de confiança, ou limite de confiança. Por exemplo, para $p = 0,95$ (95%), o coeficiente de confiança é $z_p = 1,96$, como ilustrado na Figura 2.4.

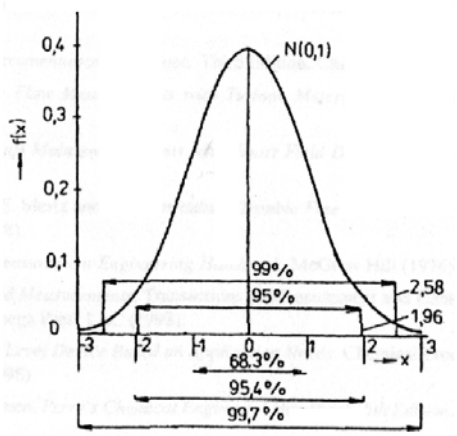


Figura 2.4: Distribuição normal – adaptado a partir de Bagajewicz (2001)

Devido ao número de medições ser finito na prática, o uso da distribuição de **t-student** é mais apropriado. O coeficiente de confiança desta distribuição (t_p) é definido como $p\%$ para $(n-1)$ graus de liberdade. Quando $n \geq 30$, as distribuições de t-student e a normal são praticamente as mesmas. A precisão (σ_P) é então definida como metade do intervalo de confiança de uma dada medida, ou seja:

$$\sigma_P = t_P \hat{s} \tag{2.3}$$

A precisão é mais formalmente chamada de repetibilidade ou de reprodutibilidade pela **ISA** (*International Society for Measurement and Control*) e outras fontes da literatura. A repetibilidade é definida pelos padrões da ISA como a proximidade de concordância entre um número consecutivo de medições de saída para uma mesma entrada, sob mesmas condições de operação (temperatura ambiente, pressão ambiente, tensão elétrica, etc.), abordadas pelo mesmo sentido. A reprodutibilidade é definida do mesmo modo, mas abordada em ambos os sentidos. Assim, a repetibilidade não inclui histerese, banda morta e efeitos de tendência, mas a reprodutibilidade sim. A precisão é desta forma um termo muito geral e não tem uma definição padronizada, sendo portanto o seu uso ambíguo pois pode se referir tanto à repetibilidade quanto à reprodutibilidade. Mesmo assim, o termo é livremente usado em toda a literatura disponível sobre os temas da reconciliação de dados, detecção de erros grosseiros e estimativa de parâmetros. Neste trabalho foi também empregado com o mesmo sentido usado na literatura, não sendo feita, salvo explicitado em contrário, as distinções acima.

2.6.2 Origens das flutuações

As flutuações das medições têm fontes variadas. Por exemplo, medições de pressão são afetadas por pequenas flutuações originadas nas bombas ou vibrações de compressores, além de outros fatores. Temperaturas no processo são afetadas por flutuações na temperatura ambiente e etc. O fluxo turbulento é por definição repleto de flutuações e mesmo fluxos laminares estão sujeitos a variações porque o fluxo é impulsionando por diferenças de pressão e depende da densidade. Assim, existem dois tipos de flutuações:

i. Inerentes ao processo e portanto à variável medida;

ii. Resultado de uma perturbação externa no processo de medição.

2.6.3 A hipótese da distribuição normal

Os erros aleatórios por hipótese seguem uma distribuição normal. Este pressuposto é baseado na teoria de que os erros são o produto de incontáveis fontes e, conseqüentemente, aplica-se o **teorema do limite central** da teoria estatística. Este teorema diz que a soma de uma grande quantidade de perturbações, cada uma tendo sua própria distribuição, tende a resultar numa perturbação com uma distribuição normal. Valores medidos são relacionados às variáveis de estado através de

uma série de transformações de sinais. Estas transformações envolvem, entre outras coisas, o uso de dispositivos de medição, transdutores[6], amplificadores eletrônicos e elementos finais de leitura. Desta forma, os sinais são distorcidos quando estas transformações não lineares são realizadas. Um exemplo é a extração da raiz quadrada realizada quando sinais de diferença de pressão são usados para medir taxa de fluxo. Mesmo que a distribuição original seja gaussiana, a distribuição resultante é distorcida. Bagajewicz (1996) provou que este efeito pode ter um impacto de dimensões mensuráveis sobre a reconciliação de dados. Além disso, o ruído é geralmente dado como sendo normalmente distribuído. Contudo, isto por vezes não é verdadeiro. Por exemplo, quando o ruído tem uma banda de freqüência curta, uma distribuição de Rayleigh seria mais apropriada (BROWN; GLAZIER, 1964). Finalmente, é bem conhecido o fato que sinais oscilantes apresentam distribuições de probabilidade diferentes da normal (HIMMELBLAU, 1970). As conseqüências destas distribuições de probabilidade não normais não serão consideradas neste trabalho, mas é digno de nota que apesar de todos estes contra exemplos, o pressuposto subjacente de que todos os erros têm uma distribuição gaussiana é comum a quase todas as abordagens de reconciliação de dados e detecção de erros grosseiros em processos químicos encontradas na literatura e vem mostrando excelentes resultados.

2.6.4 Erros sistemáticos

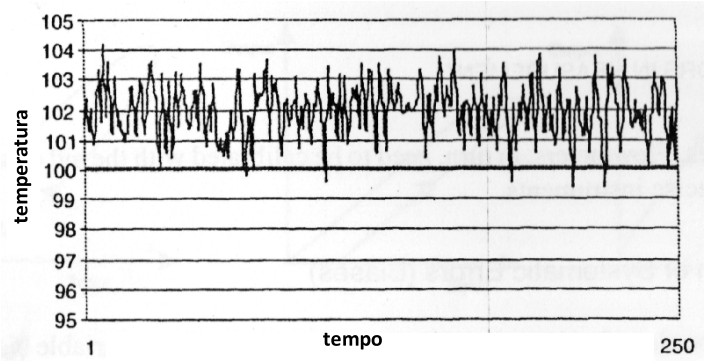


Figura 2.5: Erro sistemático – adaptado a partir de Bagajewicz (2001)

Considerando-se que uma variável mantenha o seu valor constante, a exemplo da temperatura de ebulição de um fluido puro, e que um dado número de medições consecutivas desta variável

[6]Um transdutor, é um dispositivo que transforma um tipo de energia em outro, utilizando para isso um elemento sensor que recebe os dados e os transforma. Por exemplo, o sensor pode traduzir informação não elétrica (velocidade, posição, temperatura, pH) em informação elétrica (corrente, tensão, resistência). Um exemplo de transdutor é o elaborado a partir de cristais naturais denominados piezoelétricos. Estes transdutam energia elétrica em energia mecânica na relação de 1:1 (um sinal elétrico para um sinal mecânico).

sejam feitas e sua média calculada. O **erro sistemático** do instrumento é definido como a discrepância entre o valor médio destas medições e o valor verdadeiro da variável, e é chamado também de **viés**. A Figura 2.5 ilustra o conceito com um conjunto de medições da temperatura de ebulição da água. As medições nesta figura têm um erro sistemático de cerca de $+2\,^{\circ}C$.

Quando o valor verdadeiro é conhecido, a dimensão do viés (δ) pode ser estimada pela subtração da média de todos os valores pelo valor verdadeiro, $\hat{x}$:

$$\delta = \bar{x} - \hat{x} \tag{2.4}$$

Assim, quando as medições são maiores que o valor verdadeiro, o viés é positivo e a leitura é dita "alta". Um viés negativo corresponde a uma leitura "baixa". Quando os valores verdadeiros não são conhecidos, outros instrumentos são necessários para determinar uma boa estimativa para eles. Este processo é chamado de calibração. Por exemplo, medidores de temperatura podem ser calibrados usados sistemas nos quais a temperatura seja bem conhecida, como o ponto de fusão ou ebulição de substâncias puras. Medidores de vazão, por sua vez, precisam ser calibrados com a ajuda de outros instrumento mais precisos.

2.6.5 Classificação dos erros sistemáticos

Os viéses podem ser classificados em duas categorias principais: constantes e variáveis. Fontes de viéses constantes são:

- Uso de alguma hipótese incorreta no procedimento de calibração. A presunção de comportamento de gás ideal na calibração de medidores de fluxo de gases é um exemplo.
- Não realização de correções no procedimento de calibração;
- Erros desconhecidos nos padrões de referência;
- Instalação incorreta do instrumento. Por exemplo a instalação de um medidor de fluxo próximo a um cotovelo.
- Deslocamento do zero.

Fontes de viéses variáveis:

- Derivas na tensão de alimentação do instrumento;
- Deslocamento de alcance;
- Desgaste do instrumento. Por exemplo, a borda do orifício de um medidor de fluxo pode ser afetada por partículas de uma corrente com detritos.

A deriva é definida como uma variação na saída em um período específico de tempo para uma entrada constante. Os deslocamentos, por sua vez, são independentes do tempo e correspondem a erros na faixa de medição. Os deslocamentos do zero e de alcance estão ilustrados na Figura 2.6.

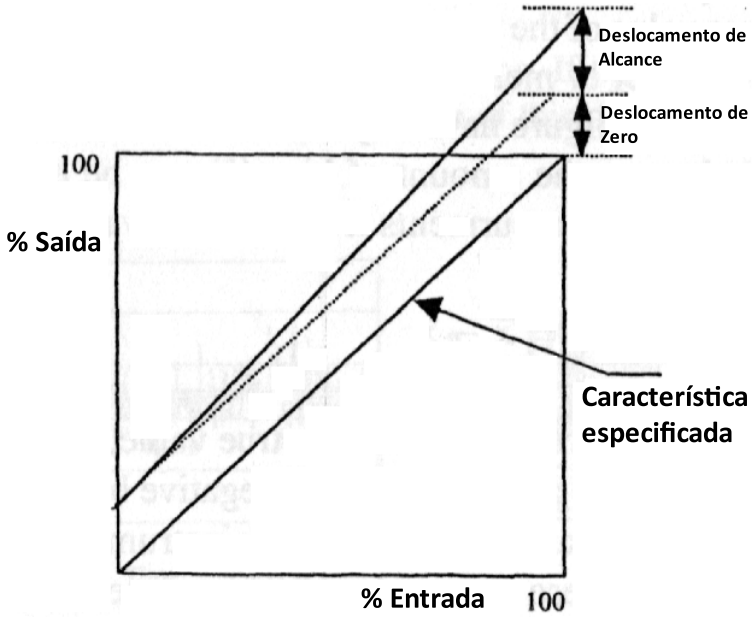


Figura 2.6: Deslocamento de alcance e de zero – adaptado a partir de Bagajewicz (2001)

2.6.6 Valores espúrios

Um **valor espúrio** (*outlier*) é definido como uma medida que não pode ser de forma alguma explicada, calculada, estimada ou antecipada. Um *outlier* é, desta forma, um ponto totalmente deslocado de um conjunto de medidas, de modo que é perceptível, por mera inspeção visual, sua discrepância em relação ao conjunto. As fontes típicas dos valores espúrios são os erros humanos, surtos de tensão elétrica e problemas na rede elétrica. A Figura 2.7 mostra um sinal sem valores espúrios presentes enquanto que na Figura 2.8 é mostrado um sinal com pequena variância e presença de valores espúrios.

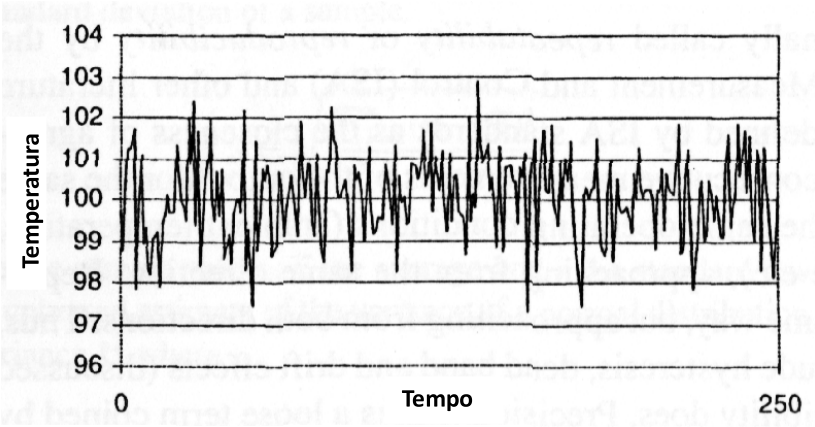


Figura 2.7: Sinal gaussiano – adaptado a partir de Bagajewicz (2001)

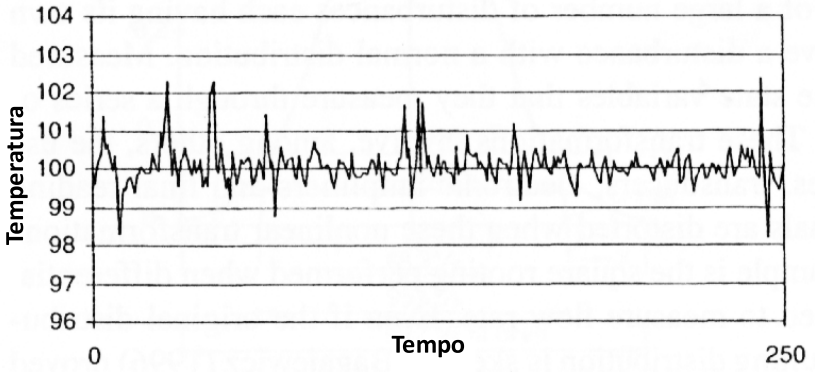


Figura 2.8: Sinal gaussiano com valores espúrios – adaptado a partir de Bagajewicz (2001)

A Figura 2.9 mostra o efeito de alta resistência ao sinal no cabeamento que causa comportamento errático e introduz um grande número de valores espúrios em uma dada variável. Por sua vez, na Figura 2.10 mostra um instrumento bem cabeado com valores espúrios ocasionais.

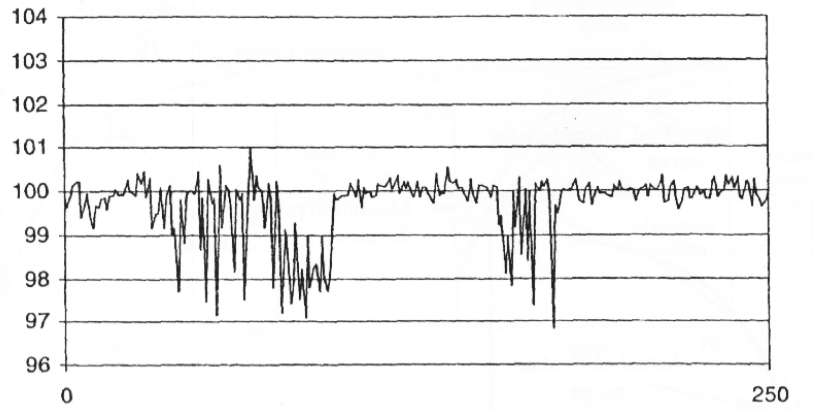


Figura 2.9: Efeito da resistência ao sinal no cabeamento (instrumento mal cabeado) – reproduzido a partir de Bagajewicz (2001)

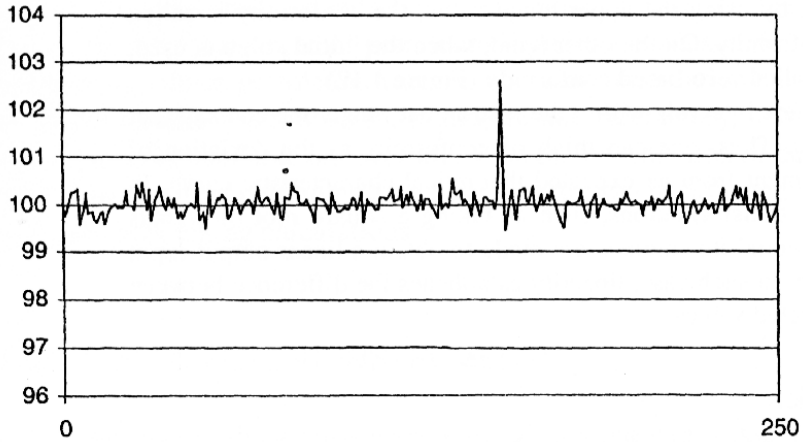


Figura 2.10: Efeito da resistência ao sinal no cabeamento (instrumento bem cabeado) – reproduzido a partir de Bagajewicz (2001)

2.6.7 Sensitividade e velocidade de resposta

A sensitividade de um instrumento é a menor variação na variável medida à qual o instrumento reage com uma variação na leitura. A velocidade de resposta, ou atraso, é uma característica dinâmica que descreve a reação de um instrumento a uma variável medida que varia com o tempo. Poucos problemas de sensitividade e atraso são encontrados porque a maioria dos instrumentos de medição tem respostas aceitavelmente boas. Uma sensitividade alta causa oscilações que resultam em um controle pobre. Por exemplo, os termopares são freqüentemente colocados em termopoços para protegê-los do fluido do processo. A presença dos termopoços cria um atraso de tempo na resposta do instrumento.

2.6.8 Histerese e banda morta

A histerese, ilustrada na Figura 2.11, é um fenômeno no qual uma saída correspondente a uma entrada crescente varia segundo um caminho que é distinto daquele mostrado pela saída quando a entrada é decrescente a partir do ponto máximo alcançado no caminho inicial. Já a banda morta é uma faixa na qual uma entrada não expressa valores na saída. Isto é tipicamente observado quando o sentido da variação na entrada é revertido bruscamente.

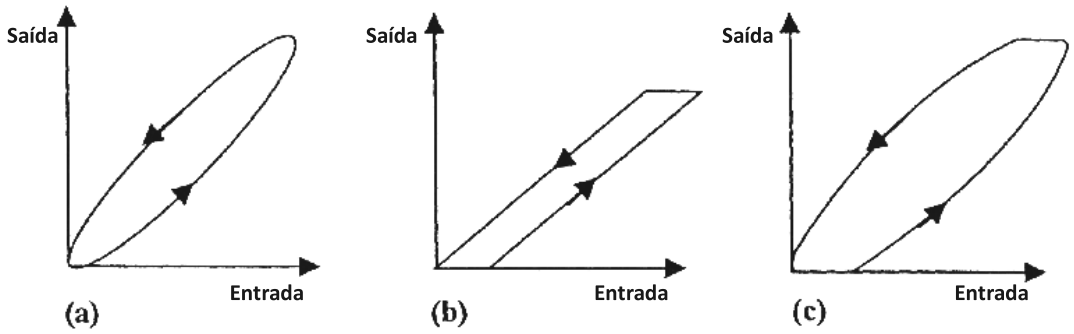


Figura 2.11: (a) Histerese, (b) Banda morta e (c) Histerese e banda morta – adaptado a partir de Bagajewicz (2001)

2.6.9 Linearidade

Quando o valor medido esperado da saída é a mesma variável da entrada, então a saída esperada é uma linha reta de 45^{o}. Este é o caso, por exemplo, quando o fluxo medido por um medidor de

fluxo é plotado contra o fluxo real depois que todas as compensações e conversões tenham sido feitas. Este diagrama é mostrado na Figura 2.12.

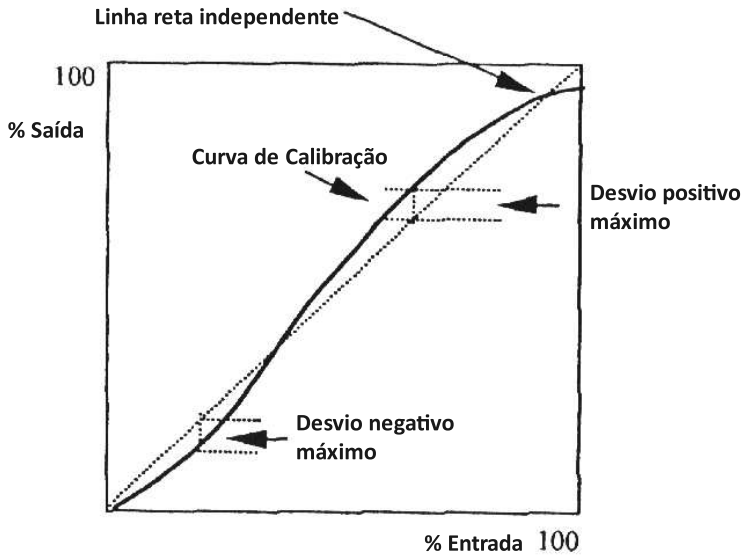


Figura 2.12: Linearidade independente – adaptado a partir de Bagajewicz (2001)

A conformidade entre o valor medido de fato e a linha característica é chamada linearidade independente e é obtida igualando os máximos desvios de sinais opostos. Estes desvios da linearidade são chamados de erros estáticos.

2.6.10 Exatidão

A exatidão de um instrumento é definida como o grau de conformidade com um valor padrão ou verdadeiro. Assim, a precisão e os erros sistemáticos contribuem para a exatidão de um instrumento de modo que um instrumento é dito exato se for não enviesado e preciso. A exatidão é reportada numericamente muitas vezes na forma

$$\sigma_a = \delta + \sigma_P \tag{2.5}$$

Segundo Bagajewicz (2001), em vários livros texto, a exatidão é definida do mesmo modo que a precisão, ignorando (ou assumindo) que não estão presentes viéses. A precisão, o viés e a exatidão são também freqüentemente reportados em termos relativos, ou seja, como um percentual de um valor medido. Algumas expressões típicas são listadas a seguir:

- A variável medida: $\pm 1^{\circ}C$
- Um percentual do *span*: $\pm 0.5\%$
- Um percentual do valor superior do *range*: $\pm 0.5\%$
- Um percentual da própria leitura de saída: $\pm 0.3\%$

A ISA e o Instituto Americano de Padrões (**ANSI** - *American National Standards Institute*) definem a exatidão como a soma da linearidade, histerese e repetibilidade comparadas com dois pontos fixos.

Uma nomenclatura mais moderna, ainda não refletida nos padrões ISA, usa o termo "inexatidão" ao invés de "exatidão" para indicar o grau de afastamento em relação a valores padrão ou ideais.

2.7 Trabalhos relevantes em reconciliação de dados e em detecção de erros grosseiros

Knepper e Gorman (1980) apresentaram as equações para o ajuste de dados e estimativa de parâmetros desenvolvidas a partir da teoria das inversas generalizadas que são usadas para analisar conjuntos de dados sujeitos a restrições lineares e não lineares. Examinaram também a convergência de problemas não lineares, para os quais sugeriram o uso de um algoritmo computacional. Desenvolveram a matriz covariância dos parâmetros calculados, a matriz dos dados ajustados e a covariância entre elas, que são usadas na decisão de validade de medidas experimentais adicionais para aumentar a precisão. Descreveram um teste χ^2 multivariado que permite verificar discrepâncias nas restrições inconsistentes com a matriz de covariância dos dados medidos. Descreveram, ainda, outros testes estatísticos para que as estimativas sejam livres de erros grosseiros.

Romagnoli e Stephanopoulos (1980a) estudaram a retificação de dados medidos em plantas químicas complexas em estado estacionário e a obtenção de estimativas das variáveis não-medidas. Propuseram o uso de um algoritmo de classificação das correntes medidas e não-medidas, baseado na **matriz de ocorrência**[7] para reduzir o tamanho do problema. Apresentaram também diferentes

[7] A matriz de ocorrência ou **matriz de incidência** no contexto da reconciliação de dados e detecção de erros grosseiros é uma descrição formal da topologia do sistema (processo) sendo analisado. Geralmente denota os nós do processo nas colunas e as correntes nas linhas. O sinal de positivo indica a corrente que entra em um nó e o de negativo denota a corrente que sai.

algoritmos para a solução de problemas práticos especiais, como a seleção de parâmetros corrigíveis, a seleção de medidas necessárias para a determinabilidade de um processo e de certos parâmetros. Propuseram ainda o uso de testes estocásticos para verificar a consistência de dados do processo e detectar erros grosseiros. Este estudo se aplica tanto ao caso de restrições lineares quanto para não lineares, considerando-se a situação na qual a composição de certas correntes seja medida para todos os componentes.

Romagnoli e Stephanopoulos (1981) estudaram a retificação de dados medidos na presença de erros grosseiros. Apresentaram um método para identificar a fonte de erros através da eliminação seriada de uma ou mais medidas de cada vez. Utilizaram o método dos mínimos quadrados para a obtenção dos ajustes, tendo como restrições os balanços de massa nodais. No método desenvolvido, faz-se o processamento seqüencial das equações de balanço, argumentando que isto simplifica a identificação de erros grosseiros durante o processo de reconciliação e facilita o tratamento de dados de processo de grandes dimensões, pois fornece um critério simples individualizado para cada balanço, ao contrário do teste global. O critério é baseado na distribuição χ^2 dos resíduos de balanço. Utilizam análise estrutural para identificar um conjunto reduzido de medidas suspeitas de erro. Este estudo é restrito ao caso de balanços lineares, que, na prática, são encontrados somente em balanços globais mássicos, mas que podem ser aplicados a balanços de componentes e de energia linearizáveis.

Mah (1981) apresentou uma revisão sobre estudos em diversas áreas relacionadas ao ajuste de dados de processo, envolvendo a reconciliação de dados, a classificação de variáveis, a detecção e identificação de erros grosseiros e a definição de pontos de tomada de medidas que formam uma base para o projeto e análise de sistemas de monitoramento. Observa que, inicialmente, dava-se ênfase aos procedimentos de cálculo para a reconciliação, que deu lugar à questão de classificação de variáveis e medidas e à detecção e identificação de erros grosseiros. Argumentou que um trabalho posterior era necessário para avaliar e generalizar os algoritmos desenvolvidos e para comparar os ganhos específicos em diferentes aspectos de sistemas de monitoramento do desempenho para que tenham uso industrial direto.

Stanley e Mah (1981a) apresentaram dois algoritmos para determinação da observabilidade local e global, redundância de variáveis individuais e medições de processos baseadas em técnicas de análise teórico-gráfica. O tratamento considera somente restrições de conservação de massa e de energia, porém observaram que pode ser estendido a processos com outras restrições adicionais, como frações de partição especificadas.

Stanley e Mah (1981b) desenvolveram uma teoria sobre observabilidade e redundância aplicável a sistemas no estado estacionário, através de analogias com sistemas dinâmicos. Demonstraram a importância desses conceitos na predição do desempenho qualitativo de estimadores, como os mínimos quadrados condicionados e outros. Mostraram como a observabilidade local está diretamente ligada à unicidade local da solução da equação de medida e como a não observabilidade local provoca a falha do estimador tanto no caso do estado estacionário quanto no quase-estacionário. Para restrições e medições lineares, as condições para observabilidade local também valem para a observabilidade global e o sistema é decomponível em subsistemas (redundante, estritamente observável e não observável).

Tamhane (1982), estudando o uso de resíduos (ajustes individuais das medidas) para detectar valores extremos em regressões lineares, mostra que um teste baseado no vetor transformado dos resíduos tem, na classe das transformações lineares do vetor dos resíduos, certas propriedades ótimas para revelar a presença de um único valor extremo, quando o observador não está ciente de que há exatamente um desses valores. Observa que é mais poderoso do que outros testes baseados no uso de resíduos, neste caso específico.

Mah e Tamhane (1982) tratam o problema de detecção e identificação da presença de um ou mais erros grosseiros em dados de processo, aplicado ao caso em que a reconciliação de dados é sujeita a restrições lineares. Consideram a situação em que todas as variáveis são medidas e que os erros grosseiros estão antes associados às medições do que a erros no modelo do processo. Observam que um teste simples, baseado nos ajustes ou resíduos (diferença entre o valor medido e o valor estimado) de cada variável, pode ser usado para detectar e identificar diretamente a fonte de erros grosseiros. Descrevem ainda um outro teste de medida baseado no vetor dos resíduos transformado (obtido pelo produto do inverso da matriz covariância pelo vetor resíduo), que conforme Tamhane (1982) possui máxima potência para detectar a presença de uma única medida com erro grosseiro. Este teste pode ser estendido ao caso de erros múltiplos, sem contudo ter assegurada suas propriedades de máxima potência.

Crowe *et al.* (1983) tratam da verificação e identificação de erros grosseiros na reconciliação de vazões, em balanços multicomponentes, no estado estacionário. Consideram que todas as medidas de concentração e vazão são independentes e normalmente distribuídas, com média desconhecida e variância conhecida. Observam que embora o produto de duas variáveis normais independentes não seja normalmente distribuído, este produto tem uma distribuição aproximadamente normal se as duas variáveis tiverem somente valores positivos e se seus coeficientes de variação (desvios padrão

relativos) forem suficientemente pequenos, isto é, menores que 5%. Então, consideram que o vetor das medidas é uma amostra de uma distribuição normal multivariada com média desconhecida e matriz covariância conhecida. Observam que medidas de diferentes correntes são, usualmente, tomadas como estatisticamente independentes, assim como medidas de concentração e de vazão total numa mesma corrente. Entretanto, as vazões dos componentes de uma mesma corrente não são independentes, pois a vazão total é comum. Utilizam três estatísticas-teste, sendo uma delas baseada na distribuição χ^2, a qual é usada para detectar erros grosseiros, coletivamente, em um conjunto de combinações lineares dos erros nas equações do balanço reduzido. Observam que qualquer combinação linear desses erros pode ser testada contra uma distribuição normal e utilizam uma que testa o desbalanço particular em cada nó e outra que testa o ajuste particular de cada medida. Aconselham que estas três estatísticas sejam empregadas em conjunto, para obter tantas evidências quanto possível e observam que, em um dos casos estudados, um dos testes não revelou a presença de erros grosseiros, enquanto os outros dois revelaram.

Madron (1985), estudando o problema de identificação de erros grosseiros em medidas, propõe uma abordagem que, ao complementar outras existentes, permite reduzir o número de elementos do conjunto de medidas suspeitas de portar erros, que deve ser testada na fase final da eliminação desse tipo de erro. Para isto, faz uso de informações disponíveis, a priori, sobre a credibilidade da medida, que representa o valor máximo possível do erro grosseiro de cada medida. Sugere o uso do teste global χ^2 para detectar erros grosseiros e de três métodos de identificação da fonte de erros descritos na literatura, os quais, restringindo-se ao caso de desvios de um modelo de erros pressuposto e não contemplando erros devidos à representação inadequada do modelo matemático (como vazamentos, por exemplo), considerando também a presença de um único erro grosseiro.

Jordache *et al.* (1985) tratam da avaliação do desempenho do teste de medida para detecção de erros grosseiros em dados de processo, proposto por Mah e Tamhane (1982), através da avaliação de sua potência, que é a probabilidade de detectar e identificar, corretamente a presença de erros grosseiros. Consideram que somente um erro grosseiro está presente nas medidas e exploram a influência de diferentes parâmetros na potência do teste. Observam que o efeito das restrições, da configuração do sistema e da posição da medida com erro grosseiro podem ser adequadamente considerados através da matriz transformada de restrições e que aqueles fatores que tornam as colunas (correntes) dessa matriz mais proporcionais tendem a reduzir as potências associadas a estas colunas e torná-las umas iguais às outras, enquanto fatores que as tornam menos proporcionais tendem a aumentar suas potências. Utilizam duas definições para a potência e observam que ela é

fortemente afetada pela razão entre magnitude do erro grosseiro e o desvio-padrão. Além disso, as potências também dependem da faixa e da distribuição dos desvios-padrão, quando eles são diferentes. A razão entre os desvios máximo e mínimo é uma medida do espalhamento das potências e quanto maior esta razão, maior a faixa das potências. Para uma mesma relação magnitude do erro grosseiro/desviopadrão, quanto menor a magnitude do erro, menor a potência associada àquela medida.

Narasimhan e Mah (1987) descrevem um método para detectar e identificar erros grosseiros devidos a desvios instrumentais e vazamentos em processos químicos no estado estacionário. Ele pode ser estendido a qualquer tipo de erro que possa ser modelado matematicamente. Como subproduto, ele permite a obtenção de estimativa da magnitude do erro grosseiro. Este método é chamado de método da razão de verossimilhança generalizada (GLR) – *Generalized Likelihood Ratio* – e se baseia no teste estatístico da razão da probabilidade condicional. Permite a identificação de vários erros grosseiros, para o que utiliza uma estratégia baseada na compensação seriada de erros grosseiros. Esta forma de compensação necessita menor tempo de cálculo do que outras existentes, baseadas na eliminação seriada. Observam que o método GLR apresenta vantagens sobre outros métodos para identificação de erros grosseiros em processos químicos no estado estacionário, pois somente ele possibilita diferençar os vários tipos de erros. Comentam que isto é importante, pois nem sempre se pode estar certo de que o processo opera estritamente no estado estacionário. Para avaliar o desempenho do teste relatam experiências de simulação, utilizando duas medidas de desempenho. Observam que quando somente erros grosseiros devidos a desvios instrumentais estão presentes, o teste GLR se reduz ao teste da medida e que ambos têm o mesmo desempenho, em condições idênticas.

Narasimhan e Mah (1989) observam que a maioria dos testes estatísticos para identificar erros grosseiros, propostos até então, trata de modelos simples de processos, no estado estacionário. Neles, as variáveis são consideradas diretamente medidas ou não medidas. Propõem técnicas para transformar adequadamente modelos gerais, no estado estacionário, em modelos simples nas mesmas condições de operação. Consideram os casos em que as variáveis de processo são medidas indiretamente e onde variáveis não medidas estão presentes nas restrições. Numa situação em que ambos os casos estejam presentes, o modelo geral do processo é transformado no modelo simples em duas etapas. Inicialmente, eliminam-se as variáveis não medidas e logo após são tratadas as variáveis indiretamente medidas. Assim, é possível aplicar qualquer um dos quatro testes estatísticos seguintes, previamente desenvolvidos: o teste global, o teste dos resíduos das restrições, o teste de

medida e o teste da razão de probabilidade condicional generalizada.

Kao *et al.* (1990) observam que uma das maiores desvantagens da maioria dos métodos propostos na literatura para detecção de erros grosseiros é que eles partem da hipótese de que as observações são não correlacionadas seriadamente (já que a independência implica em ausência de correlação), o que muitas vezes não é confirmado pelas evidências experimentais. Relatam que observações correlacionadas podem ser causadas por um número de fatores físicos associados com o tempo morto do processo, com sua dinâmica e controle, bem como devido a fatores relacionados a instrumentos de medição. Nesta investigação, tratam de erros de medida correlacionados seriadamente, em processos químicos no estado estacionário. Entretanto, consideram que variáveis diferentes medidas num mesmo tempo ou em diferentes tempos não são correlacionadas, isto é, que não há correlação cruzada. Usam como referência, o teste da medida e mostram que é bastante significativo o efeito de negligenciar a existência de correlação. Sugerem duas soluções para este problema, que são avaliadas analiticamente e através de simulações computacionais. Uma delas se baseia na obtenção da fórmula correta para a variância, considerando correlações seriadas dos dados do processo e a outra envolve a filtragem para remoção de correlações seriadas destes dados, obtendo resíduos não correlacionados, aos quais podem ser aplicados as técnicas desenvolvidas para a detecção de erros grosseiros em dados de processo independentes.

Keller (1992) propôs um algoritmo analítico para estimar a matriz de covariância dos erros de medida, que é empregada na reconciliação de dados de processo. Ele se baseia nos resíduos de restrições lineares, calculados a partir dos dados disponíveis. Observou que uma exigência indispensável para a sua aplicação é a redundância espacial suficiente e que este algoritmo pode ser aplicado a medidas correlacionadas se todos os sensores sujeitos a tais medidas forem conhecidos.

Karjala e Himmelblau (1994) apresentam uma forma alternativa para a reconciliação de dados dinâmica. Combinam o problema de reconciliação com a detecção de erros grosseiros, incluindo filtragem de dados e testes de hipóteses para as variáveis. Usam para a resolução do problema redes neurais artificiais, filtro de Kalman estendido e programação não linear com restrições. O ponto forte do uso da rede neural é o potencial de rápida resolução de problemas lineares baseados unicamente no histórico do processo. O filtro de Kalman requer o conhecimento anterior das matrizes de covariância, sendo computacionalmente eficiente. Já a retificação dinâmica de dados baseada em NLP (*non linear programming* – programação não linear), o tempo computacional pode ser grande e de modo geral sempre superior ao exigido pelo filtro de Kalman. Analisando os métodos, os autores mostram que o filtro de Kalman e as NLP são as que produzem as melhores estimativas,

tendo como inconveniente a necessidade de conhecimento a priori da matriz de covariância.

Zhang *et al.* (1995) utilizaram o pacote comercial **ASPEN PLUS** para realizar otimização e estimativa de parâmetros numa planta de ácido sulfúrico da Monsanto, utilizando interfaces em FORTRAN capazes de processar dados online. Para fazer a detecção de erros grosseiros e a reconciliação de dados utilizaram a linguagem de otimização **GAMS**, usando como método de modelagem as técnicas desenvolvidas por Tjoa e Biegler (1991a). Segundo os autores, o lucro teve um aumento de 9% após a adoção do procedimento.

Plácido (1995) em sua dissertação de mestrado estudou a aplicação da reconciliação à dados obtidos de várias unidades de uma planta industrial de produção de amônia. Foram desenvolvidos dois programas de reconciliação de dados em FORTRAN 77, com armazenagem de matrizes na forma esparsa. No programa não linear as restrições são linearizadas de modo a se utilizar o método da matriz de projeção de Crowe *et al.* (1983), em um esquema iterativo. Foram feitas várias simulações com geração aleatória de dados, com melhorias consideráveis na qualidade dos mesmos. A aplicação a casos reais, incluindo um balanço de energia com 106 equações e 145 variáveis, possibilitou a correta detecção de vários erros grosseiros.

Mendes (1995) em sua tese de doutorado desenvolveu procedimentos computacionais para reconciliação e retificação de dados de processo. Seus procedimentos utilizaram a técnica da projeção matricial para simplificação do problema de obtenção dos ajustes das quantidades medidas através do método dos mínimos quadrados ponderados e dos multiplicadores de Lagrange. Relatou os resultados da aplicação do programa de reconciliação de dados **RECON** a cinco processos distintos: um processo de síntese de amônia, uma rede de vapor com nove nós e vinte correntes, um processo de síntese de ácido nítrico e um circuito de moagem. Elaborou também um procedimento computacional para classificação de variáveis de processo chamado de **TCLASS**, aplicável a variáveis de balanços de massa multicomponentes no qual as restrições se referem essencialmente a quantidades extensivas, a menos das equações de normalização. Relatou a aplicação do **TCLASS** a oito casos do processo de síntese de amônia e a dois casos de um processo de evaporação-cristalização de nitrato de potássio.

Yang *et al.* (1995) revisaram brevemente os métodos de detecção de erros grosseiros, quais sejam:

Baseados na suposição de distribuição normal

- Teste Global (1965);

- Teste da Máxima Potência (1975);
- Teste Nodal ou de Restrição (1976);
- Teste da Medida (1982);
- Método do Critério de Informação de Akaike (1986);
- Razões de Verossimilhança Generalizada (1989);
- Análise de Componentes Principais (1994).

Baseados em distribuições não normais

- Distribuições de Verossimilhança Bivariada (1991);
- Função de Densidade de Probabilidade Não-central (1993);

e **Baseados em redes neurais**. Observaram que nenhum deles apresenta garantias consistentes de encontrar todos os erros grosseiros nos dados de processo. Um método combinado é então proposto baseado no teste da medida (*measurement test* ou **MT**) e no teste nodal (*nodal test* ou **NT**) e este método tem aplicação prática e factibilidade, com óbvias vantagens sobre as abordagens similares prévias.

A vantagem do MT está no fato de que a localização dos erros grosseiros pode ser apontada nos dados do processo diretamente mas com a desvantagem de tender a gerar outros erros grosseiros (erro **Tipo I** – apontar um erro quando este não ocorre realmente). O método NT não espalha os erros grosseiros por todos os dados mas tem como desvantagem a possibilidade de não detectar um erro grosseiro (erro **Tipo II** – não detectar a presença de um erro quando este de fato ocorre) no caso de haver dois erros grosseiros de mesma magnitude e sinais opostos na entrada ou saída do sistema ou dois erros grosseiros de mesma magnitude e sinal, um na entrada e outro na saída do sistema, cancelando-se mutuamente. Outra desvantagem é que este método não aponta qual corrente contém o erro grosseiro diretamente, sendo necessário o teste de diversas combinações de nós diferentes para rastrear o erro, o que torna o processo lento.

A abordagem combinada do MT e NT baseia-se em primeiramente encontrar os dois nós ligados à corrente detectada pelo MT, então checa-se estes dois nós pelo NT. Se nenhum deles apresenta

o erro, passa-se à próxima corrente candidata a portar o erro grosseiro que tenha sido apontada pelo MT e assim sucessivamente.

Barbosa Júnior (1996) em sua tese de doutorado tratou do problema de modelagem, reconciliação de dados e controle regulatório de um reator CSTR para polimerização em massa do estireno via radicais livres. A modelagem determinística do reator foi baseada no mecanismo cinético da reação e no princípio de conservação de massa e energia das quantidades envolvidas. Esta abordagem resultou num sistema de equações diferenciais ordinárias de primeira ordem não lineares. Para resolver o problema de reconciliação de dados dinâmica com restrições não lineares, foram usadas técnicas de programação não linear do tipo programação quadrática sucessiva, especificamente o método SQP. O autor resolveu o modelo do processo através do método da colocação ortogonal em elementos finitos.

Bagajewicz (1996) abordou o problema da distribuição de probabilidade dos erros na reconciliação de dados. Na quase totalidade dos trabalhos encontrados na literatura disponível, a formulação do problema de reconciliação recai na suposição da distribuição normal dos erros. O autor apontou que esta suposição é verdadeira para dispositivos de leitura que efetuem somente transformações lineares sobre as medidas. No caso de leituras como as de vazão que implicam em transformações não lineares sobre uma medida primária, mesmo que esta tenha uma distribuição normal, a leitura final a distorcerá e a formulação "clássica" torna-se ineficaz para captar a natureza dos erros nos procedimentos de reconciliação de dados.

Teixeira (1997) em sua dissertação de mestrado apresentou uma metodologia para reconciliação de dados linear e não linear de processo entre os quais não estejam presentes erros grosseiros. Foi usada a técnica da projeção de matrizes para simplificação das equações de balanço (restrições) de massa e/ou energia de processos complexos. Foi desenvolvido também um procedimento para detecção de erros grosseiros. Foram desenvolvidos e descritos dois programas computacionais, sendo que o primeiro faz a reconciliação de dados e o segundo detecta a existência de erros grosseiros.

Dempf e List (1998) procederam à reconciliação de dados com objetivos de otimização de processos em uma planta de acetato de vinila e outra de acetona. Os autores afirmam que o "ótimo" empírico está na maioria dos casos à 10% do ótimo global e que a abordagem do problema de reconciliação deve sempre partir de uma modelagem simples do balanço global de massa de toda a planta, com a gradual introdução de complexidade acompanhada de redução de escopo. Relataram,

ainda, os seguintes tópicos entre os resultados alcançados em sua aplicação prática:

- A RD compõe uma metodologia bastante sensível para detectar partidas do estado estacionário
- Suporte à manutenção preventiva, detectando vazamentos e/ou disfunção dos sensores ou das operações unitárias
- Alta credibilidade nas provas de conformidade com regras ambientais, de segurança e restrições econômicas (padrões de certificação)
- Várias análises laboratoriais puderam ser substituídas por medidas de temperatura e a taxa de amostragem foi reduzida em 50%
- A eficiência da unidade de craqueamento da acetona pôde ser calculada

Bagajewicz e Jiang (2000) compararam o desempenho entre a abordagem integral para a reconciliação de dados dinâmica e a reconciliação de dados estacionária. Foi mostrado que na ausência de viéses e vazamentos, o desempenho em ambas as abordagens é similar. Foi provado que uma vez que a variância apropriada seja escolhida, ambos os métodos são idênticos na ausência de termos de acúmulo. Finalmente, foi feita uma análise das discrepâncias como função do termo do acúmulo.

Vachhani *et al.* (2001) trataram do problema da detecção de erros grosseiros e reconciliação de dados em sistemas não lineares e dinâmicos. Baseiam o trabalho no tratamento dado por Liebman et al. (1992) que propuseram um *framework* de programação não linear para a resolução deste problema, mas com o pressuposto de se saber a priori qual o parâmetro que contém o viés, não sendo incluída a detecção dos parâmetros enviesados como parte da formulação. O trabalho de Vachhani *et al.* (2001) estende o trabalho de Liebman *et al.* (1992) para a inclusão da detecção dos parâmetros enviesados através do tratamento do problema de identificação como um problema de diagnóstico, tratável pelos métodos de diagnóstico de falhas na literatura, sendo que os autores optaram pelo tratamento por métodos quantitativos, mas precisamente uma rede neural elipsoidal. A abordagem foi demonstrada por um estudo de caso de um reator CSTR, não isotérmico com jaqueta.

Zhang *et al.* (2001) mostram que por vezes a eliminação de medidas detectadas como portadoras de erros grosseiros provoca perda de precisão na solução do problema de reconciliação de dados

subseqüente, de modo que propõem um método de análise de redundância, baseado na medida da precisão da reconciliação, que permite a eliminação criteriosa e seqüencial de medidas portadoras de erros grosseiros com o objetivo de preservar a solvabilidade do problema de reconciliação. Os autores aplicam o método proposto a uma unidade de síntese de metanol e na Fujian Petrochemical Ltd. Co., da República Popular da China.

Söderström *et al.* (2001) evidenciaram a estreita relação entre a reconciliação de dados e a detecção de erros grosseiros, indicando a busca de uma técnica que combine a solução dos dois problemas de forma que o resultado final seja um conjunto de dados com sua componente de erros aleatórios removida bem como com os erros grosseiros detectados. O artigo descreve o desenvolvimento de uma técnica deste tipo dentro de um *framework* de otimização inteira mista, para um sistema linear, invariante no tempo e em estado estacionário, como por exemplo redes de fluxos de um processo.

Özyurt e Pike (2004) mostraram a importância e a efetividade de procedimentos para reconciliação de dados e detecção de erros grosseiros simultâneos. Estes procedimentos, baseados em estatísticas robustas, reduzem o efeito de erros grosseiros, que na abordagem de detecção e eliminação iterativa tradicional, espalham erro por todas as estimativas. Além de derivar novos métodos robustos, são descritos novos critérios de detecção de erros grosseiros e os seus desempenhos são testados.

Benqlilou (2004) em sua tese de doutorado apresentou, discutiu e comparou um conjunto de metodologias e combinações de metodologias para prover estimativas com maior exatidão de variáveis de processos tanto para sistemas no estado estacionário quanto no dinâmico. Inicialmente, a exatidão das estimativas é aumentada através de novas técnicas de reconciliação de dados propostas que combinam filtragem baseada em dados e em modelos, considerando também a presença de atrasos entre as amostragens dos dados. Em seguida, se volta para o problema do projeto da rede de dispositivos de medição e seu uso ótimo.

Prata (2005) em sua dissertação de mestrado desenvolveu o monitoramento em tempo real da qualidade do produto e produtividade do processo de polimerização do polipropileno em massa, via catálise Ziegler-Natta, em um reator tanque de mistura contínua de propriedade da Polibrasil Resinas S.A., no pólo petroquímico de Duque de Caxias, RJ. A disponibilização de certos parâmetros em tempo real permite um operação mais eficiente quando comparada ao caso no qual são necessários demorados testes de laboratório. O autor propôs um modelo dinâmico simplificado do

processo, incluindo as variáveis mais significativas requeridas para determinar os parâmetros desejados e o utilizou em uma estratégia de monitoramento baseado na reconciliação de dados dinâmica e não linear simultânea com a estimativa de parâmetros. Isto resultou em um problema de otimização não linear baseada no critério da máxima verossimilhança. A autor usou uma abordagem de janela móvel para garantir estimativas em tempo real de modo eficiente.

ao (2005) em sua dissertação de mestrado avaliou os impactos da implantação de um sistema de reconciliação de dados em uma refinaria de petróleo (REFAP – Refinaria Alberto Pasqualini da Petrobras, RS), sendo uma das primeiras implementações desse tipo no Brasil. O autor utilizou entrevistas e questionários para as pessoas envolvidas na implantação e constatou que pelo menos dois dos objetivos propostos foram alcançados: houve um aumento de confiança nos dados do processo e melhoras nos índices de perdas. O sistema de reconciliação de dados reduziu o índice de perdas da refinaria de 1,09% (fev./2001 a jan./2002 – anterior à implantação) para 0,34% (jul./2002 a dez./2002 – logo após a implantação). A redução no índice equivaleu a uma redução de R$ 1.609.295,15 por mês nas perdas da refinaria.

Oliveira Júnior (2006) em sua tese de doutorado desenvolveu um novo código computacional para classificação de variáveis em sistemas dinâmicos que propiciou uma eficiente reconciliação dos dados de modo a possibilitar a estimativa dos parâmetros cinéticos da polimerização industrial do polipropileno com catalizadores Ziegler-Natta (Suzano-RJ). O algoritmo de classificação também foi testado para o balanço hídrico de uma fábrica de fertilizantes FAFEN (PETROBRAS), em Laranjeiras-SE.

Marques (2006) apresentou em sua dissertação de mestrado um novo algoritmo de classificação de variáveis desenvolvido para descrever modelos constituídos por equações algébricas, lineares ou não lineares, no sentido de dar suporte à identificação e localização de sensores em fluxogramas de processo e indicar o grau de observabilidade e redundância do processo analisado. foi analisado o fluxograma da FAFEN (PETROBRAS), unidade de Laranjeiras-SE. Esta planta produz fertilizantes nitrogenados, produzindo 1250 T/d de amônia e 1800 T/d de uréia. Foram demonstradas as vantagens econômicas da análise sendo possível otimizar a quantidade e a localização dos sensores na instalação.

2.8 Conclusões

Este capítulo introduziu a maior parte da terminologia usada no restante desta tese. Começou-se com o contexto no qual os temas da reconciliação de dados e da detecção de erros grosseiros se inserem, passando por uma descrição mais amiúde da natureza dos erros nas medições de dados de processo e do próprio conceito de qualidade de dados. Na seqüência foram apresentadas algumas justificativas para se utilizar as técnicas apresentadas neste trabalho, culminando com uma apresentação de trabalhos considerados mais relevantes para o escopo desta tese.

3 *Reconciliação de Dados em Estado Estacionário para Sistemas Lineares*

Neste capítulo são apresentas as bases conceituais e matemáticas da reconciliação de dados que serão ampliadas nos capítulos seguintes. Inicialmente são apresentados alguns conceitos básicos e definições que junto com uma base estatística para a reconciliação permitem a formulação do problema matemático de reconciliação de dados. Em seguida é apresentada a decomposição do problema geral de estimativa com a conseqüente classificação de variáveis do processo, onde são importantes a análise topológica e as abordagens para resolver o problema de classificação. Depois disso são apresentadas abordagens de decomposição usando transformações ortogonais para, finalmente, tratar da reconciliação de dados com todas as variáveis medidas e com algumas das variáveis não medidas.

3.1 Conceitos básicos

Quando um estimador é comparado com a observação, algumas questões podem surgir como, por exemplo, qual será o efeito do posicionamento da medida (observação) no desempenho do estimador e qual é o efeito de medidas enviesadas sobre esse estimador? Estas e outras questões são levadas em conta no problema de estimativa de parâmetros e na seleção da estrutura das medidas para o monitoramento ou controle de um dado processo. Está claro que os conceitos de redundância e a alocação das medidas têm um importante papel no problema da estimativa. Além disso a redundância é útil como segurança quando há viéses nas medidas ou imperfeições no modelo da situação física sob consideração (ROMAGNOLI; SÁNCHEZ, 2000).

Para se ter um problema de estimativa de parâmetros, qualquer que seja ele, em primeiro lugar deve haver um sistema com várias medidas disponíveis. O sistema é algum objeto físico e o seu comportamento, que pode ser dinâmico (discreto ou contínuo) ou estático, é descrito por equações

e seus dados associados. Define-se aqui "dado" como um conjunto de informações constituído de basicamente três elementos: a localização (a que corrente ou unidade de processo esse dado se relaciona), o momento no qual a medida foi colhida, chamado comumente de *time stamp* e o valor da própria medida em si. Daqui por diante, neste capítulo, será discutido um processo em estado estacionário.

Seja uma quantidade (ou vetor quantidade), associado com a operação de um sistema cujo valor se deseja a cada instante de tempo. Se esta quantidade não for diretamente mensurável ou só puder ser medida com erro, deve-se assumir que **y** medidas com **ruído**[1] estão disponíveis e que um experimento tenha sido projetado para medir ou estimar a referida quantidade associada do sistema. O conjunto dos valores reais das variáveis pode ser escrito como o vetor:

$$\mathbf{x}^{\mathrm{T}} = [x_1, x_2, ..., x_g] \tag{3.1}$$

A situação mais geral é aquela na qual as variáveis desejadas não podem ser observadas (medidas) diretamente e devem, portanto, ser indiretamente medidas como uma função de observações diretas. Assim, assume-se também que um conjunto de l medidas **y** possa ser expresso como uma função de g elementos de um vetor constante **x** mais um erro aleatório aditivo ε. Desta forma as medidas do processo podem ser modeladas como:

$$\mathbf{y} = \phi(\mathbf{x}) + \varepsilon, \quad \mathbf{y} \in \Re^l, \quad \mathbf{x} \in \Re^g \tag{3.2}$$

onde ϕ representa o modelo funcional da medida.

Se $\varepsilon = \mathbf{0}$, então $\mathbf{y} = \phi(\mathbf{x})$ e diz-se que as medidas são perfeitas (não contêm erros). Se $\varepsilon \neq \mathbf{0}$, então elas contêm ruído. Nos casos os quais ϕ é diferenciável no ponto $\mathbf{x}^0$, podemos definir a matriz **J**:

$$\mathbf{J}(\mathbf{x}^0) = \left.\frac{\partial \phi}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^0} \tag{3.3}$$

onde **J** é a versão linearizada das equações não lineares das medidas. Para sistemas lineares, a matriz **J** é constante e independente de **x**. Em geral, o sistema linear ou linearizado é expresso por:

[1] O ruído pode ser definido genericamente como tudo o que aparece sobreposto a um sinal e que não faça parte dele.

$$\mathbf{y} = \mathbf{J}\mathbf{x} + \boldsymbol{\varepsilon}, \quad \mathbf{J} \in \Re^{l \times g}, \tag{3.4}$$

onde $\mathbf{J}$ é a Jacobiana $(l \times g)$ de ϕ. Assim, quando as observações forem planejadas, deve ser especificado um modelo funcional geral sobre o sistema a ser avaliado (matriz $\mathbf{J}$). Tal modelo funcional, que se refere a um sistema finito fechado, é determinado por um certo número de variáveis e pelas relações entre elas.

Há sempre um número mínimo de variáveis independentes que determinam unicamente um modelo. Denota-se este número por g. A menos que as observações sejam suficientes para determinar as g variáveis, a situação é evidentemente deficiente. Estas observações devem ser linearmente independentes, isto é, nenhuma das l observações pode ser derivada de nenhuma outra das $(l-1)$ restantes.

Define-se um sistema como redundante quando a quantidade de dados disponíveis (informação) excede a quantidade mínima necessária para uma determinação singular das variáveis independentes do modelo.

Para o sistema na Equação 3.2, quando l é maior que g, diz-se que há redundância. Esta redundância, denotada por r, é dada por:

$$r = l - g \tag{3.5}$$

é igual ao conceito estatístico de graus de liberdade.

Um vez que os dados são geralmente obtidos de observações que estão sujeitas a flutuações probabilísticas, os dados redundantes são geralmente inconsistentes no sentido de que cada subconjunto com um número suficiente de variáveis para uma determinação singular do sistema provê um resultado diferente dos demais. Para se obter uma única solução, um critério adicional se faz necessário. Se o princípio dos mínimos quadrados for aplicado, entre todas as soluções que são consistentes com o modelo de medida, as estimativas que são as mais próximas o possível das medidas são consideradas como sendo a solução do problema de estimativa. Define-se o problema de estimativa por mínimos quadrados como:

$$\min N = (\mathbf{y} - \mathbf{J}\mathbf{x})^{\mathrm{T}}(\mathbf{y} - \mathbf{J}\mathbf{x}) \tag{3.6}$$

A solução dos mínimos quadrados é que minimiza da soma dos quadrados do resíduo $N = \varepsilon^{\mathrm{T}}\varepsilon$. A equação em **x**,

$$\mathbf{J}^{\mathrm{T}}\mathbf{J}\mathbf{x} = \mathbf{J}^{\mathrm{T}}\mathbf{y} \tag{3.7}$$

obtida pela diferenciação de N, é chamada de equação normal. Pode-se agora definir a propriedade da estimabilidade:

> **DEFINIÇÃO 3.1** Diz-se que um sistema é estimável se a equação normal admite uma única solução e, naturalmente, **x** é único.

Assim, as condições necessárias para a estimabilidade podem ser definidas. Para que as variáveis do processo, **x**, sejam estimáveis, o seguinte teorema tem que ser verdadeiro (RAO, 1973).

> **TEOREMA 3.1** O sistema descrito pela Equação 3.7 é globalmente estimável se e somente se
>
> $$\text{posto}\,\mathbf{J} = g \tag{3.8}$$
>
> onde g é a dimensão[a] do sistema.
> Reciprocamente, se
>
> $$\text{posto}\,\mathbf{J} < g \tag{3.9}$$
>
> o sistema é globalmente não-estimável.

[a] O posto (*rank*) de uma matriz **A** é a dimensão da maior matriz quadrada contida em **A** (formada pela eliminação de linhas e colunas) e que tenha o determinante não nulo (GELB, 1974).

Quando o sistema é não-estimável, um valor estimado de **x**, $\hat{\mathbf{x}}$, não é uma solução única do problema de mínimos quadrados. Neste caso a solução é possível somente com a adição de informação. Esta informação é introduzida através de equações do modelo do processo (equações de restrição). Elas ocorrem na prática quando algumas ou todas variáveis do sistema devem obedecer às relações que surgem das restrições físicas do processo.

Em alguns casos, a introdução de equações adicionais do modelo do processo pode aumentar o número de variáveis a serem estimadas, não diminuindo assim a deficiência de estimabilidade.

Com a introdução de restrições adicionais tem-se:

$$\begin{aligned} \mathbf{0} &= \varphi(\mathbf{x}) \qquad \mathbf{x} \in \Re^g \\ \mathbf{y} &= \phi(\mathbf{x}) + \varepsilon \quad \mathbf{y} \in \Re^l \end{aligned} \tag{3.10}$$

onde $\varphi \in \Re^m$, com m sendo o número de equações de restrição adicionais.

Estas relações funcionais que caracterizam o comportamento de processos reais não são nunca conhecidas exatamente. Uma maneira convencional de contabilizar a falta de exatidão gerada pelas aproximações é introduzir um ruído aditivo, o qual de certa forma reflete o grau de erro na modelagem, isto é,

$$\begin{aligned} \mathbf{0} &= \varphi(\mathbf{x}) + \mathbf{w} \quad \mathbf{x} \in \Re^g \\ \mathbf{y} &= \phi(\mathbf{x}) + \varepsilon \quad \mathbf{y} \in \Re^l \end{aligned} \tag{3.11}$$

Assumindo que $\varphi(\mathbf{x})$ e $\phi(\mathbf{x})$ sejam diferenciáveis em $\mathbf{x}^0$, e aplicando uma expansão em série de Taylor usando somente os termos zero e de primeira ordem (descartando os termos de segunda ordem em diante), chega-se ao sistema linear ou linearizado descrito por:

$$\begin{aligned} \mathbf{0} &= \mathbf{A}\mathbf{x} + \mathbf{w} \\ \mathbf{y} &= \mathbf{J}\mathbf{x} + \varepsilon \end{aligned} \tag{3.12}$$

onde $\mathbf{A}$ e $\mathbf{J}$ são as matrizes $(m \times g)$ e $(l \times g)$ das Jacobianas de φ e ϕ Neste caso a condição de redundância será satisfeita quando $(m+l) > g$. Pode-se agora definir o problema dos mínimos quadrados da seguinte forma:

$$\min N = (\mathbf{z} - \mathbf{M}\mathbf{x})^{\mathrm{T}}(\mathbf{z} - \mathbf{M}\mathbf{x}), \tag{3.13}$$

onde

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} \\ \mathbf{J} \end{bmatrix}, \ \mathbf{z} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{3.14}$$

A equação normal é dada por:

$$\mathbf{M}^{\mathrm{T}}\mathbf{M}\mathbf{x} = \mathbf{M}^{\mathrm{T}}\mathbf{z}. \tag{3.15}$$

De modo similar ao caso anterior, as condições gerais para a estimabilidade podem ser colocadas como segue:

TEOREMA 3.2 O sistema descrito pela equações 3.14 e 3.15 são globalmente estimáveis se, e somente se,

$$\text{posto } \mathbf{M} = \text{posto} \begin{bmatrix} \mathbf{A} \\ \mathbf{J} \end{bmatrix} = g. \tag{3.16}$$

Reciprocamente, se

$$\text{posto } \mathbf{M} = \text{posto} \begin{bmatrix} \mathbf{A} \\ \mathbf{J} \end{bmatrix} < g. \tag{3.17}$$

o sistema é globalmente não-estimável.

Define-se agora uma forma mais geral da função objetivo quadrática, a qual permite atribuir pesos pré-determinados aos componentes. Qual seja:

$$N = \begin{bmatrix} \mathbf{w} \\ \varepsilon \end{bmatrix} \Psi \begin{bmatrix} \mathbf{w} & \varepsilon \end{bmatrix} \tag{3.18}$$

onde Ψ é a matriz de ponderação (matriz de covariância), restrita a ser tanto simétrica quanto definida positivamente[2], ou seja, $\Psi = \Psi^{\mathrm{T}} > \mathbf{0}$. A introdução da matriz de ponderação define o problema como de mínimos quadrados ponderados e as mesmas condições estabelecidas pelos teoremas 3.1 e 3.2 também se aplicam nesta situação. (ROMAGNOLI; SÁNCHEZ, 2000).

3.2 Base estatística da reconciliação de dados

Tendo descrito a formulação do problema de reconciliação de dados a partir de um ponto de vista quase que puramente intuitivo, especialmente com respeito aos fatores ponderantes da função objetivo a serem usados em diferentes medidas, o problema de reconciliação de dados pode ser também explicado usando-se uma base teórica estatística que não somente ajuda na compreensão

[2] Seja uma matriz quadrada $\mathbf{Q}$ e λ_i seus autovalores, Diz-se que $\mathbf{Q}$ é ...

Definida positivamente se todos $\lambda_i > 0$;

Semidefinida positivamente se todos $\lambda_i \geq 0$;

Definida negativamente se todos $\lambda_i < 0$;

Semidefinida negativamente se todos $\lambda_i \leq 0$;

Indefinida não se pode estabelecer nenhum dos casos anteriores.

deste assunto, bem como dispõe informação quantitativa sobre o aumento de exatidão nos dados obtidos através da reconciliação e das propriedades estatísticas das estimativas resultantes. Estas propriedades podem ser usadas para identificar dados grosseiramente incorretos ou para projetar redes de sensores.

A base estatística para a reconciliação de dados vem das propriedades que são pressupostas para erros aleatórios nas medidas. Geralmente se assume que os erros aleatórios seguem uma distribuição normal multivariada com média zero e uma matriz de covariância, Ψ, conhecida. Contudo, deve ser levado em conta que algumas vezes o sinal primário mensurado é transformado na variável final de interesse. Se a transformação é não linear como na Equação 3.19

$$F = k\sqrt{\left(\frac{p_0 \Delta p}{T}\right)} \tag{3.19}$$

onde k é a constante do orifício da placa, Δp é a diferença de pressão no orifício, p_0 é a pressão de entrada no orifício e T é a temperatura do fluido. Então o erro na variável indicada não tem necessariamente uma distribuição normal.

Somente a forma linearizada pode ser aproximada por uma distribuição normal. Assim, se possível, as variáveis $\mathbf{x}$ no modelo de medida da Equação 3.4 devem representar as variáveis primárias medidas sendo que as relações entre a variável primária medida e as variáveis de interesse devem ser incluídas como restrições. No caso de restrições não lineares, então uma técnica de reconciliação de dados não linear deve ser usada.

A matriz Ψ contém informação sobre a exatidão das medidas e as correlações entre elas. Os elementos na diagonal de Ψ, σ_i^2, são a variância na i-ésima variável medida e os elementos fora da diagonal, σ_{ij}^2 são a covariância dos erros entre as variáveis i e j. Se os valores medidos são dados pelo vetor $\mathbf{y}$, então as estimativas mais prováveis para $\mathbf{x}$ são obtidas pela maximização da função de verossimilhança da distribuição normal multivariada.

$$\max_x \frac{1}{(2\pi)^{n/2}|\Psi|^{n/2}}\, e^{-0{,}5(y-x)^{\mathrm{T}}\Psi^{-1}(y-x)} \tag{3.20}$$

onde $|\Psi|$ é o determinante de Ψ. O problema de máxima verossimilhança acima é equivalente à minimização da função:

$$\min_{x} (\mathbf{y}-\mathbf{x})^{\mathrm{T}} \Psi^{-1} (\mathbf{y}-\mathbf{x}) \tag{3.21}$$

Como σ_i é o desvio padrão do erro na i-ésima medida, o fator de ponderação Ψ na Equação 3.21 é inversamente proporcional ao desvio padrão deste erro. Desta forma, um valor mais alto de desvio padrão implica em uma medida menos exata e assim tem-se pesos maiores para medidas mais exatas.

É possível também derivar propriedades estatísticas das estimativas obtidas através de reconciliação de dados. Considerando-se o caso quando todas as variáveis são medidas, as estimativas são dadas por

$$\hat{\mathbf{x}} = \mathbf{y} - \Psi \mathbf{A}^{\mathrm{T}} (\mathbf{A} \Psi \mathbf{A}^{\mathrm{T}})^{-1} \mathbf{A}\mathbf{y} = \left[I - \Psi \mathbf{A}^{\mathrm{T}} (\mathbf{A} \Psi \mathbf{A}^{\mathrm{T}}) \mathbf{A} \right] \mathbf{y} = \mathbf{B}\mathbf{y} \tag{3.22}$$

A Equação 3.22 mostra que as estimativas são obtidas se usando transformações lineares das medidas. As estimativas, portanto, são normalmente distribuídas, com valor esperado e matriz de covariância dados por:

$$E[\hat{\mathbf{x}}] = \mathbf{B}E(\mathbf{y}) = \mathbf{B}\mathbf{x} = \mathbf{x} \tag{3.23}$$

$$\mathrm{cov}[\hat{\mathbf{x}}] = E\left[(\mathbf{B}\mathbf{y})(\mathbf{B}\mathbf{y})^{\mathrm{T}} \right] = \mathbf{B} \Psi \mathbf{B}^{\mathrm{T}} \tag{3.24}$$

A Equação 3.23 implica em estimativas não enviesadas o que é uma propriedade de um estimador de máxima verossimilhança para sistemas lineares. A Equação 3.24 dá uma medida da exatidão das estimativas.

No caso em que algumas das variáveis não são medidas, também é possível se derivar propriedades estatísticas semelhantes. Estas propriedades são exploradas para a identificação de medidas portadoras de erros grosseiros bem como para projetar redes de sensores. (NARASIMHAN; JORDACHE, 2000).

3.3 Formulação do problema de reconciliação de dados

Como foi colocado nas seções anteriores, a reconciliação de dados aumenta a exatidão dos dados do processo através do ajuste dos valores mensurados de modo que estes satisfaçam a restrições do processo. A quantidade do ajuste feito sobre as medidas é minimizada pois os erros nas medidas são tidos como pequenos. No caso geral, não são todas as variáveis que são medidas devido a limitações de ordem técnica ou econômica.

As estimativas das variáveis não-medidas bem como os parâmetros do modelo são também obtidos como parte do problema de reconciliação. A estimativa de valores baseada em medidas reconciliadas é também conhecida como coaptação de dados. De um modo geral, a reconciliação de dados pode ser formulada pelo seguinte problema de otimização por mínimos quadrados sujeitos a restrições:

$$\min_{x_i,u_j} \sum_{i=1}^{n} \omega_i (y_i - x_i)^2 \tag{3.25}$$

sujeito a:

$$g_k(x_i,u_i) = 0, \qquad k = 1,\ldots,m \tag{3.26}$$

A função objetivo 3.25 define a soma total dos quadrados ponderados dos ajustes feitos sobre as medidas, onde os ω_i são os fatores ponderantes, y_i é a medida e x_i é a estimativa reconciliada para a variável i e os u_j são as estimativas das variáveis não medidas. A Equação 3.26 define o conjunto das restrições do modelo. Os fatores ponderantes ω_i são escolhidos dependendo da exatidão das diferentes medidas.

As restrições do modelo são geralmente balanços de massa e energia, mas também podem incluir relações de desigualdade impostas pela factibilidade das operações no processo. As leis determinísticas naturais da conservação de massa e energia são tipicamente usadas como restrições para a reconciliação de dados pois elas em geral são conhecidas. Equações empíricas e de outros tipos envolvendo vários parâmetros não medidos não são recomendadas, pois elas são, na melhor das hipóteses, conhecidas somente de modo aproximado. Forçar as variáveis medidas a obedecer relações inexatas pode levar a uma reconciliação de dados inexata e ao diagnóstico incorreto de erros grosseiros.

Qualquer lei de conservação de massa ou energia pode ser expressa pela seguinte fórmula geral:

entrada - saída + geração - consumo - acúmulo = 0

A quantidade para qual a Equação acima é escrita pode ser o fluxo material global, o fluxo de componentes individuais ou o fluxo de energia. Se não há acúmulo de qualquer destas quantidades, então estas restrições são algébricas e definem uma operação em estado estacionário.

Para um processo dinâmico, contudo, o termo do acúmulo não pode ser desprezado e as restrições são equações diferenciais. Para a maioria das unidades de processo, não há geração ou depleção de material. No caso dos reatores, porém, a geração ou consumo de componentes individuais devido à reação deve ser levada em consideração.

Para algumas unidades mais simples como *splitters*, não há nenhuma variação nem na composição nem na temperatura das correntes. Para tais unidades, os balanços de componente e de energia se reduzem à forma simples como:

$$x_i = x_j \tag{3.27}$$

onde a variável x_i representa tanto a temperatura quanto a composição da corrente i. A Equação acima é útil também quando dois ou mais sensores são usados para medir a mesma variável – uma taxa de fluxo ou a temperatura de uma corrente, por exemplo.

O tipo das restrições que são impostas na reconciliação de dados dependem do escopo do problema e do tipo da unidade. Além disso, a complexidade das técnicas de solução usadas depende fortemente das restrições impostas. Por exemplo, quando se está interessado em reconciliar somente taxas de fluxos de todas as correntes, então as restrições dos balanços materiais são lineares nas variáveis de fluxo e daí resulta um problema de reconciliação de dados linear. Por outro lado, se é desejado reconciliar composição, temperaturas ou pressão em conjunto com os fluxos, então tem-se um problema de reconciliação de dados não linear.

Um quesito importante é o tipo de restrições que podem ser legitimamente impostas na aplicação da reconciliação de dados. Uma vez que a reconciliação de dados força as estimativas de todas as variáveis a satisfazerem as restrições impostas, esta questão tem grande importância. Geralmente, as restrições de balanço de massa e energia são incluídas porque estas são leis físicas válidas. Deve ser notado, contudo, que estas equações são geralmente escritas assumindo que não

há perda de material ou de energia do processo para o meio ambiente. Enquanto isto pode ser válido para o fluxo material, uma perda significante de energia pode ocorrer fruto, por exemplo, do isolamento impróprio de trocadores de calor. Em tais casos, é melhor não impor os balanços de energia ou, de forma alternativa, incluir um termo não conhecido de perda na equação do balanço e este pode ser estimado como parte do problema de reconciliação.

Outra restrição além da conservação de massa e energia pode ser um modelo de uma unidade de processo que contenha equações envolvendo parâmetros da unidade. Por exemplo, o modelo de um trocador de calor pode incluir a equação relacionando a carga térmica ao coeficiente global de troca térmica, a área de troca e os fluxos e temperaturas das correntes. A Equação 3.28 descreve esta relação.

$$Q - UA\Delta T_{ln} = 0 \tag{3.28}$$

onde Q é a carga térmica, U é o coeficiente global de troca térmica e ΔT_{ln} é a diferença média logarítmica de temperatura.

Geralmente, uma equação como esta pode ser incluída como restrição em uma reconciliação sobre um trocador de calor se o coeficiente de troca for desconhecido e tenha que ser estimado a partir de dados medidos. Se não houver nenhuma informação prévia sobre U e nenhuma restrição de factibilidade, então a inclusão desta restrição não fornece nenhuma informação adicional e as estimativas de todas as outras variáveis serão as mesmas sendo esta restrição incluída ou não. Desta forma, o problema de reconciliação de dados pode ser resolvido sem esta restrição e U pode ser subseqüentemente estimado pela Equação anterior usando-se os valores reconciliados de fluxos e temperaturas.

Por outro lado, se U tem que estar dentro de limites específicos ou se há uma boa estimativa de U vinda de um ciclo anterior de reconciliação, então a restrição deve ser incluída junto com a informação adicional sobre U como parte do problema de reconciliação. O coeficiente global de troca térmica pode também ser relacionado às propriedades físicas das correntes, seus fluxos, temperaturas e as características do trocador de calor usando-se correlações. Não é aconselhável usar uma equação desse tipo no modelo de reconciliação, pois tais correlações podem conter muitos erros nelas mesmas, forçando os fluxos e temperaturas a se ajustarem à equação, o que pode aumentar a inexatidão das estimativas.

Outra questão importante é quanto ao procedimento de reconciliação de dados usar um modelo estacionário ou dinâmico de processo. Na prática, um processo nunca está verdadeiramente em estado estacionário. Apesar disso, uma planta é operada por várias horas ou dias em uma região em torno de um ponto de operação nominalmente em estado estacionário.

Para aplicações como otimização *on line*, nas quais a reconciliação é realizada uma vez em algumas horas, é apropriado se empregar reconciliação de dados em estado estacionário sobre a média das medidas em um intervalo de tempo de interesse.

Durante condições transientes, quando a saída do estado estacionário é significativa, a reconciliação de dados estacionária não deve ser aplicada porque resultará em grandes ajustes aos valores médios. Medidas tomadas durante tais períodos transientes podem ser reconciliadas, se necessário, usado-se um modelo dinâmico do processo. De um modo similar, para aplicações de controle de processo, nas quais a reconciliação precisa ser realizada a cada poucos minutos, a reconciliação de dados dinâmica é mais apropriada.

A reconciliação de dados é baseada na hipótese que somente erros aleatórios estão presentes nas medidas, os quais seguem uma distribuição normal (Gaussiana). Se um erro grosseiro devido a um viés na medida está presente em alguma medida ou se há um vazamento significante no processo que não tenha sido contabilizado no modelo das restrições, então os dados reconciliados podem ser bastante inexatos. É portanto necessário identificar e remover tais erros grosseiros. Isto é conhecido como o Problema de detecção de erros grosseiros.

Erros grosseiros podem ser detectados baseando-se na extensão na qual as medidas violam as restrições ou na magnitude dos ajustes feitos às medidas em uma reconciliação preliminar. Ainda que as técnicas de detecção de erros grosseiros tenham sido desenvolvidas primariamente para aumentar a exatidão de estimativas reconciliadas, elas também são úteis na identificação de instrumentos de medição que precisam ser trocados ou recalibrados. (NARASIMHAN; JORDACHE, 2000).

3.4 Decomposição do problema geral de estimativa

As seções anteriores discutiram a formulação matemática do problema de estimativa de parâmetros e as condições gerais de redundância e estimabilidade. Em algumas situações, o modelo do processo contempla variáveis que não são medidas em campo, mas das quais se deseja obter o valor de modo que se faz necessária uma decomposição do problema geral de estimativa em um problema

de reconciliação e um problema de coaptação de dados. Como subproduto, nesse procedimento de decomposição se consegue também um alívio computacional. Será analisada agora a decomposição do problema geral de estimativa. A divisão de sistemas lineares em suas partes observáveis e não observáveis foi primeiro sugerida por Kalman (1960a). O mesmo tipo de argumento pode ser usado aqui para decompor um sistema considerado em estado estacionário.

Quando os resultados da teoria de matrizes são aplicados ao problema geral de estimativa de parâmetros, o seguinte teorema pode ser definido:

TEOREMA 3.3 Para o sistema descrito pela equações 3.14 e 3.15, se

$$\text{posto } \mathbf{M} = \text{posto} \begin{bmatrix} \mathbf{A} \\ \mathbf{J} \end{bmatrix} = j < g, \tag{3.29}$$

então existe uma matriz não singular $\mathbf{T}$ tal que

$$\mathbf{MT} = \begin{bmatrix} \mathbf{A_U} & \mathbf{0} \\ \mathbf{J_U} & \mathbf{0} \end{bmatrix}, \tag{3.30}$$

onde $\mathbf{A_U}$ e $\mathbf{J_U}$ têm j colunas e

$$\text{posto} \begin{bmatrix} \mathbf{A_U} \\ \mathbf{J_U} \end{bmatrix} = j. \tag{3.31}$$

O sistema de equações 3.12 pode ser escrito usando a forma escalonada em colunas da matriz $\mathbf{M}$ tal como segue (ROMAGNOLI; SÁNCHEZ, 2000):

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} = \mathbf{MTT^{-1}x} + \begin{bmatrix} \mathbf{w} \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \mathbf{A_U} & \mathbf{0} \\ \mathbf{J_U} & \mathbf{0} \end{bmatrix} \mathbf{T^{-1}x} + \begin{bmatrix} \mathbf{w} \\ \varepsilon \end{bmatrix} \tag{3.32}$$

Dependendo da estrutura de $\mathbf{T}^{-1}$, duas situações podem surgir:

i. Se cada linha de $\mathbf{T}^{-1}$ tem somente um elemento não-zero, isto significa fisicamente que nas novas coordenadas $\mathbf{x}_c = [x_r, x_{g-r}]$, onde $\mathbf{x}_r$ é um vetor j-dimensional, o subsistema

$$\begin{aligned} \mathbf{0} &= \mathbf{A_U x}_r + \mathbf{w} \\ \mathbf{y} &= \mathbf{J_U x}_r + \varepsilon \end{aligned} \tag{3.33}$$

é estimável. O sistema como um todo admite uma decomposição em dois sub-sistemas menores: um estimável, de dimensão j e outro não estimável, de dimensão $(g-j)$. O primeiro inclui as variáveis $\mathbf{x}_r$ e o último contém as variáveis em $\mathbf{x}_{g-r}$.

ii. Se algumas linhas de $\mathbf{T}^{-1}$ têm mais de um elemento não zero, existem combinações lineares entre as variáveis em $\mathbf{x}_r$ e as variáveis em $\mathbf{x}_{g-r}$. Assim, a porção estimável do sistema é de dimensão ob tal que $(ob < j)$ e a porção não estimável é de dimensão $(g-ob)$.

Uma medida é considerada redundante se a sua remoção não causa perda de estimabilidade. Se o posto de $\mathbf{M} = g$ e $(m+l) > g$, isto é, há mais informação disponível que o necessário pra uma determinação singular do sistema, então o seguinte teorema pode ser enunciado:

TEOREMA 3.4 Se o sistema de equações (3.14) e (3.15) é estimável e redundante, isto é, $(m+l > g)$, com $(l-i)$ medidas redundantes e se as linhas de $\mathbf{J}$ são permutadas de modo que as primeiras $(l-i)$ linhas correspondam às medidas redundantes $(\mathbf{y}_1)$, ou seja,

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{bmatrix} \; e \;\; i > \mathbf{0}, \tag{3.34}$$

então existe uma matriz $\mathbf{F}$ $(g \times g)$, não-singular tal que

$$\mathbf{MF} = \begin{bmatrix} \mathbf{A}_{\mathrm{U}} & \mathbf{0} \\ \mathbf{J}_{1\mathrm{U}} & \mathbf{0} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{bmatrix} \tag{3.35}$$

e

$$\text{posto } \mathbf{J}_{22} = i, \quad \text{posto } [\mathbf{A}_{1\mathrm{U}}] \; = \; \text{posto} \begin{bmatrix} \mathbf{A}_{\mathrm{U}} \\ \mathbf{J}_{1\mathrm{U}} \end{bmatrix} = g \; - \; i \tag{3.36}$$

com cada uma das medidas do sistema $\mathbf{A}_{1\mathrm{U}}$=$\begin{bmatrix} \mathbf{A}_{\mathrm{U}} \\ \cdots \\ \mathbf{J}_{1\mathrm{U}} \end{bmatrix}$ sendo redundante.

A partir dos resultados do teorema anterior, conclui-se que qualquer sistema que seja estimável e redundante, $(r > 0)$, admite a decomposição em suas partes redundante $(\mathbf{x}_1)$ e não redundante $(\mathbf{x}_2)$. Essa decomposição permite uma nova formulação em duas partes do problema geral dos mínimos quadrados.

PROBLEMA 1

Problema dos mínimos quadrados:

$$\min_{\mathbf{x}_1} (\mathbf{z}_1 - \mathbf{A}_{1U}\mathbf{x}_1)^T \mathbf{W}_1 (\mathbf{z}_1 - \mathbf{A}_{1U}\mathbf{x}_1) \tag{3.37}$$

onde

$$\mathbf{z} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y}_1 \end{bmatrix} \tag{3.38}$$

PROBLEMA 2 Como a decomposição permite que $\mathbf{x}_1$ seja determinado primeiro, o passo seguinte é calcular $\mathbf{x}_2$ usando os valores já conhecidos de $\mathbf{x}_1$ e $\mathbf{y}_2$.

Esta formulação dividida em dois problemas distintos resulta numa significante redução de dimensionalidade em relação ao problema original.

3.5 Classificação das variáveis de processo

Uma planta de processos químicos é um sistema físico contendo uma grande quantidade de unidades e correntes. Por exemplo, a contagem dos equipamentos no setor de processo e de utilidades (considerando os mixers e os divisores de corrente) de uma planta petroquímica pode revelar a existência de aproximadamente 1000 unidades interconectadas e cerca de 2500 correntes. Em cada corrente as variáveis de interesse podem ser vazão, composição, temperatura, pressão e entalpia, é evidente que o tratamento dos dados de uma planta típica envolve a solução de um problema em grande escala (ROMAGNOLI; SÁNCHEZ, 2000).

A idéia original de reduzir os sistemas de equações usados no problema da reconciliação é devida a Václavek (1969), o qual propôs um procedimento de correção baseado somente em um subconjunto reduzido de equações e medidas. A idéia consiste na exploração da topologia para classificar as variáveis do processo e eliminar do problema original as que não são medidas, resultando em um subconjunto de equações envolvendo somente variáveis medidas. Várias estratégias foram desenvolvidas desde então para alcançar o mesmo objetivo, qual seja, a decomposição do processo para reduzir a dimensionalidade do problema. Algumas destas estratégias são baseadas na teoria dos grafos (Mah *et al.* (1976); Kretsovalis e Mah (1988a, 1988b); Meyer *et al.* (1993)), e outras abordagens orientadas a equação (Crowe *et al.* (1983); Crowe (1986, 1989a); Romagnoli e

Stephanopoulos (1980a); Joris e Kalitventzeff (1987)).

Do que foi apresentado se torna evidente que a aplicação de técnicas de reconciliação de dados em grandes plantas, representadas por modelos não lineares complexos é um problema desafiador. A decomposição através da classificação das variáveis do processo é uma importante ferramenta no tratamento da dimensionalidade do problema. O que é mais importante é que a compreensão da estrutura topológica da planta não somente permite decompô-la mas também pode ser muito útil no projeto ou análise de um sistema completo de monitoramento.

Seja um processo contendo K unidades denotadas por $k = 1,...,K$, e J correntes orientadas, $j = 1,...,J$, com C componentes, $c = 1,...,C$. A topologia da planta pode ser representada pela **matriz de incidência** também chamada de **matriz de ocorrência**, **A**, com as linhas correspondendo às ***unidades*** e colunas correspondendo às ***correntes***. Assim:

$A_{jk} = 1$ se a corrente j entra no nó k
$A_{jk} = -1$ se a corrente j sai do nó k
$A_{jk} = 0$ se a corrente j não tem contato com o nó k

As restrições de balanço para uma unidade de processo sem reações químicas e transferência de calor podem ser expressas tal como segue (ROMAGNOLI; SÁNCHEZ, 2000):

Balanços de Massa:

$$\sum_j \mathbf{A}_{j,k} f_j = 0 \tag{3.39}$$

Balanços de Massa para os Componentes:

$$\sum_j \mathbf{A}_{j,k} f_j \mathbf{M}_{c,j} = 0 \tag{3.40}$$

Balanços de Entalpia:

$$\sum_j \mathbf{A}_{j,k} f_j h_j = 0 \tag{3.41}$$

Equações de Normalização:

$$\sum_{c} f_j \mathbf{M}_{c,j} - f_j = 0 \tag{3.42}$$

onde f_j é o fluxo total na corrente j, $\mathbf{M}_{c,j}$ é a fração do componente c na corrente j e h_j é a entalpia específica da corrente j.

De um modo geral, o modelo de uma planta operando em estado estacionário é constituído de um sistema de equações algébricas não lineares da forma

$$\varphi(\mathbf{x}, \mathbf{u}) = \mathbf{0}, \varphi \in \Re^m \tag{3.43}$$

onde φ é uma função não linear tendo como variáveis $\mathbf{x}$ e $\mathbf{u}$, vetores das variáveis do processo, medidas e não medidas, respectivamente. Para balanços lineares de massa, a Equação 3.43 se torna:

$$\mathbf{A}_1\mathbf{x} + \mathbf{A}_2\mathbf{u} = \mathbf{0}, \quad \mathbf{x} \in \Re^g, \quad \mathbf{u} \in \Re^n \tag{3.44}$$

onde $\mathbf{A}_1$, e $\mathbf{A}_2$ são matrizes compatíveis de dimensão $(m \times g)$ e $(m \times n)$, respectivamente.

Se o estado do sistema for diretamente medido, então o modelo da medida é representado por:

$$\mathbf{y} = \mathbf{x} + \varepsilon \tag{3.45}$$

Neste caso a Jacobiana das funções de medição $\mathbf{J}$ é igual à matriz identidade, e o vetor do erro aleatório das medidas é

$$\varepsilon = \mathbf{y} - \mathbf{x} \tag{3.46}$$

Fica disposto assim que a formulação da reconciliação de dados é apenas um caso especial do problema geral de estimativa de parâmetro.

3.5.1 Definições

As variáveis de um processo podem ser classificadas em **medidas**, **não medidas** e **constantes**. As variáveis medidas se dividem em duas categorias: **redundantes** e **não redundantes**. As variáveis não medidas se classificam em **determináveis** e **não determináveis** como mostrado na Figura 3.1.

Uma variável não medida é dada como determinável se esta pode ser avaliada a partir das medidas disponíveis com auxílio das equações de balanço. Por outro lado, se esta variável não puder ser avaliada desta forma, ela é dita não determinável.

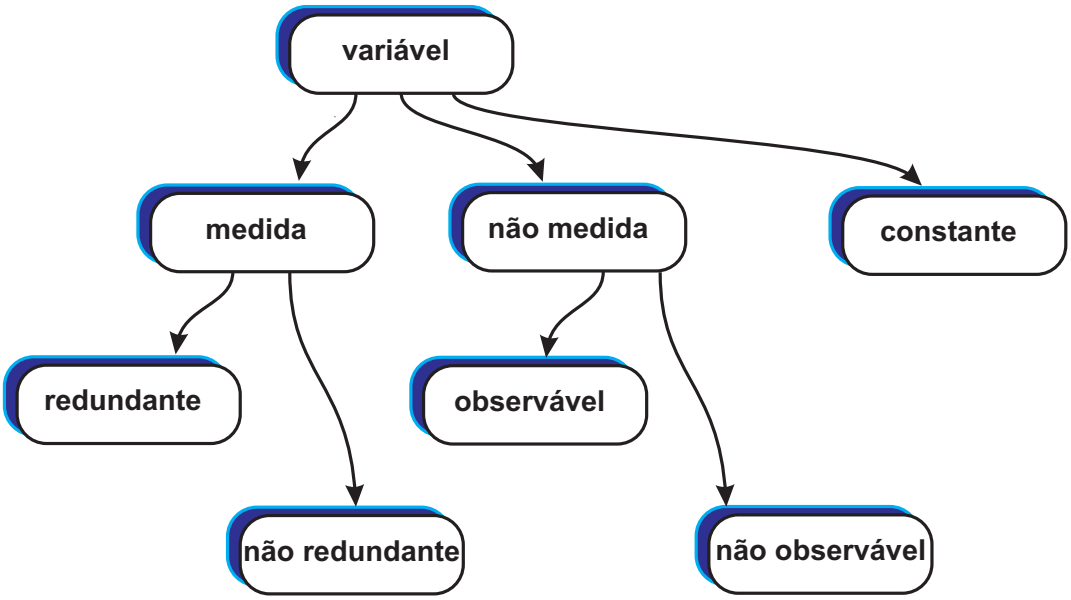


Figura 3.1: Classificação das variáveis não medidas

Uma medida é dita redundante (ou sobre-determinada) se esta puder ser encontrada através das equações de balanço aliadas às outras medições disponíveis. De modo contrário, se esta medida só pode ser alcançada através da própria medida, esta é dita não redundante.

Baseado na formulação anterior, os seguintes problemas podem ser definidos:

i. Classificar as variáveis não medidas;

ii. Definir o subconjunto de equações redundantes a serem usadas no ajuste das medidas;

iii. Classificar as variáveis medidas.

As ferramentas básicas para a avaliação estrutural das equações do processo serão brevemente discutidas. Estas ferramentas permitem analisar sistematicamente a estrutura topológica das equações de balanço e resolver os três problemas que foram colocados.

3.5.2 Análise da topologia de um processo

Seguindo a divisão das variáveis do processo nos vetores **x** e **u**, medidas e não medidas respectivamente, os sistemas de equações lineares ou linearizadas podem ser divididos nas matrizes compatíveis $\mathbf{A}_1$ e $\mathbf{A}_2$ através da Equação 3.44.

Esta divisão sugere uma representação estrutural do sistema onde as matrizes $\mathbf{A}_1$ e $\mathbf{A}_2$ consistem de alguns elementos que são geralmente não-nulos e outros que são sempre nulos.

O sistema de matrizes $\mathbf{A}_1$ e $\mathbf{A}_2$ descrevem a topologia estrutural das correntes e unidades em termos de variáveis e equações as quais podem ser associadas a um gráfico mostrando as suas influências mútuas.

Sejam os nós do gráfico as variáveis de processo e suas fronteiras as relações (equações de balanço) entre elas. Há uma fronteira em comum entre o nó *a* e o nó *i*, se *a* pertence ao intervalo de *i*, ou seja, se *a* é necessário para avaliar *i*.

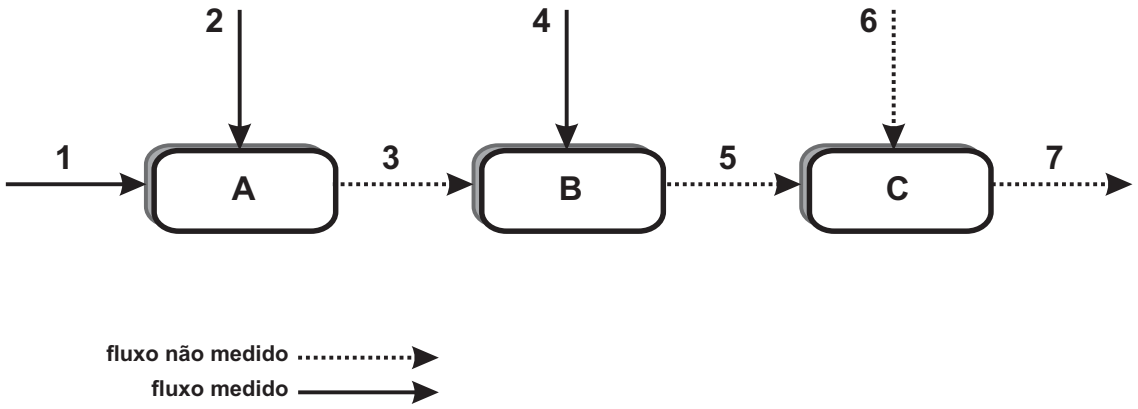


Figura 3.2: Diagrama de fluxo para um sistema simples em série – adaptado a partir de Romagnoli e Stephanopoulos (1980a)

Os conceitos expostos acima são ilustrados na Figura 3.2, onde três unidades são conectadas a sete correntes, sendo que destas, apenas as correntes 1, 2 e 4 são medidas. Procedendo a um balanço de massa total sobre cada unidade de acordo com a Equação 3.39, tem-se:

$$\begin{aligned} f_1 + f_2 - f_3 &= 0 \\ f_3 + f_4 - f_5 &= 0 \\ f_5 + f_6 - f_7 &= 0 \end{aligned}$$

A Resolução destas três equações para as variáveis f_3, f_5 e f_6 gera o grafo de fluxo de informação mostrado na Figura 3.3. No grafo mostra-se que a informação contida nas correntes 1 e 2 é propagada para a corrente 3 e esta última junto com a corrente 4 tem sua informação propagada para a corrente 5.

Introduz-se agora alguns conceitos em conexão com os sistemas estruturais e os seus gráficos associados

DEFINIÇÃO 3.2 (Inacessibilidade) Define-se um nó ***i*** como inacessível a partir do nó ***a*** se não houver possibilidade de alcançar ***i*** partindo de ***a*** (o qual corresponde a uma variável medida) e indo para o nó ***i*** na direção das setas ao longo de um caminho no gráfico de fluxo de informação.

DEFINIÇÃO 3.3 (Determinabilidade) Define-se um nó ***i*** como determinável se qualquer caminho para o nó ***i*** começar em um nó mensurado.

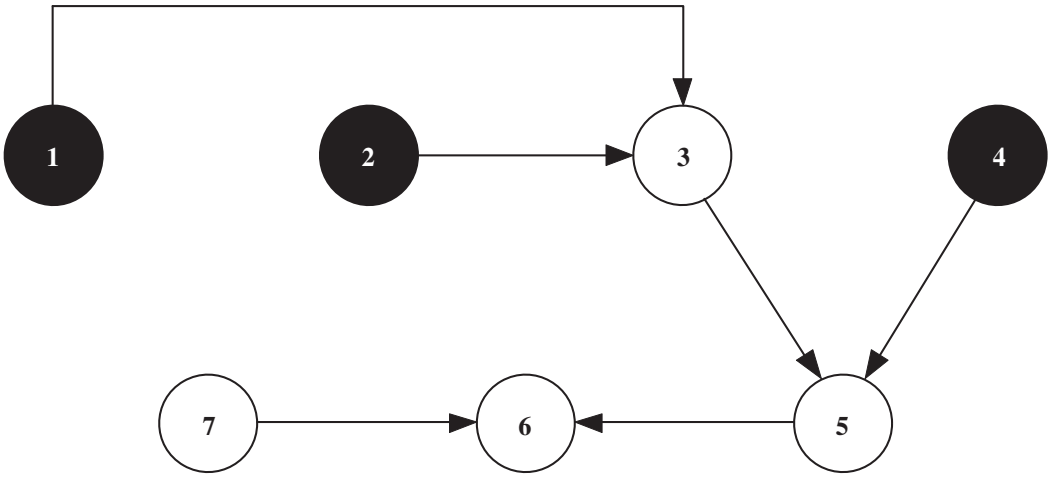


Figura 3.3: Grafo de fluxo de informação das correntes do sistema descrito na Figura 3.2 – adaptado a partir de Romagnoli e Stephanopoulos (1980a)

Na Figura 3.3, aplicando a definição de acessibilidade:

- Os nós 3, 5 e 6 são acessíveis;
- O nó 7 é inacessível.

Na mesma figura, aplicando agora o conceito de determinabilidade:

- Os nós 3 e 5 são determináveis;
- Os nós 6 e 7 são não determináveis.

3.5.3 Abordagens para solução do problema de classificação

Durante as três últimas décadas, várias estratégias foram formuladas para realizar a classificação de variáveis. Estas estratégias podem ser divididas em dois grandes grupos. Um dos grupos aplica os conceitos da teoria dos grafos[3] e o outro faz uso de técnicas de ordenação de matrizes e procedimentos computacionais. Está disposta a seguir uma breve revisão sobre o assunto que pode ser encontrada em Romagnoli e Sánchez (2000).

3.5.3.1 Técnicas orientadas a grafos

Dada a topologia de um processo, um grafo não orientado é construído onde os nós correspondem às unidades e os arcos às correntes do processo. O grafo do processo contém um nó para o meio-ambiente do qual o processo recebe as correntes de alimentação e para o qual vão as correntes de produtos finais.

As principais contribuições às técnicas orientadas a gráficos são devidas aos seguintes autores.

Václavek (1969)

Václavek (1969) foi quem primeiro definiu os conceitos de observabilidade e redundância. Ele formulou duas regras para se categorizar as variáveis em modelos lineares de plantas:

i. Agregar dois nós conectados com uma corrente não medida. O Esquema de Balanço Reduzido resultante contém somente medidas redundantes;

[3] A Teoria dos Grafos é o ramo da matemática que estuda as propriedades de grafos. Um grafo é um conjunto de pontos, chamados vértices (ou nós), conectados por linhas, chamadas de arestas (ou arcos). Dependendo da aplicação, arestas podem ou não ter direção, pode ser permitido ou não arestas ligarem um vértice a ele próprio e vértices e/ou arestas podem ter um peso (numérico) associado. Se as arestas têm uma direção associada (indicada por uma seta na representação gráfica) tem-se um grafo direcionado, ou digrafo.

ii. Retirar todas as correntes medidas e procurar por ciclos no grafo reduzido. Os ciclos no grafo resultante representam fluxos não determináveis.

Václavek e Loucka (1976) estenderam a abordagem a processos multicomponentes com a suposição que, em qualquer corrente, ou todas as frações mássicas são medidas ou nenhuma é. As reações químicas são contabilizadas pela adição de correntes fictícias no grafo. Os *splitters* (divisores de corrente) não são considerados em sua formulação.

Mah e colaboradores

Apresentaram uma abrangente teoria e algoritmos para a classificação de variáveis medidas e não medidas. Para processos com um único componente (somente balanços de massa), Mah *et al.* (1976) derivaram um procedimento simples de classificação baseado na teoria dos grafos. Em um trabalho posterior, Kretsovalis e Mah (1987) descreveram a categorização das variáveis para fluxos de multicomponentes sem pressuposições sobre a localização dos sensores. Não foram levadas em contas nem reações químicas nem *splitters*. Kretsovalis e Mah (1988a, 1988b) estenderam o seu tratamento para incluir reatores, *splitters* e unidades onde ocorrem fluxos de energia pura. As seguintes variáveis das correntes foram consideradas em sua análise: fluxo de massa, frações mássicas, fluxos de componentes e de energia e temperaturas. O conjunto de medidas é restrito a fluxos mássicos, frações mássicas e temperaturas. Foi pressuposto que existe uma correspondência unívoca entre temperatura e entalpia por unidade de massa.

A técnica requer uma análise extensa do grafo do processo e dos seus sub-grafos derivados (16+ número de componentes). Eles são testados por um conjunto de 19 teoremas de observabilidade e redundância. Estes subgrafos são atualizados durante a execução do procedimento. A classificação das variáveis não medidas é alcançada usando as regras derivadas somente da teoria dos grafos e da álgebra matricial.

(MEYER *et al.*, 1993)

Os autores (MEYER *et al.*, 1993) introduziram um método variante derivado do Kretsovalis e Mah (1987) que permite o tratamento de reações químicas e *splitters*. Ele leva à diminuição do tamanho do problema de reconciliação de dados bem como um particionamento das equações para classificação das variáveis não medidas.

3.5.3.2 Técnicas orientadas a equação

Dada a topologia do processo e um conjunto de medidas, estas estratégias geram primeiro o sistema de equações do modelo para a planta, procedendo posteriormente diferentes tipos de rearranjos e cálculos envolvendo matrizes e equações não lineares de modo a classificar as variáveis do processo. As principais contribuições para essa abordagem são dispostas a seguir:

Crowe *et al.* (1983)

Para modelos lineares, Crowe *et al.* (1983) usaram a projeção de matrizes para obter um conjunto reduzido de equações que permitem a classificação das variáveis medidas. Eles identificaram que as variáveis não medidas através da redução de colunas da sub-matriz correspondente a estas variáveis.

Crowe (1986) estendeu esta metodologia para a classificação de variáveis envolvidas em balanços bilineares de componentes. O modelo é modificado para uma forma linear usando o conhecimento da topologia do processo, da localização dos instrumentos e um conjunto de medidas que devem ser consistentes com as restrições do processo. Crowe (1989a) propôs um algoritmo de classificação de variáveis baseado numa série de lemas. Nesta formulação, balanços bilineares de energia são incluídos nas equações do modelo, assumindo uma correspondência unívoca entre a temperatura e a entalpia por unidade de massa.

O procedimento permite a inclusão de medições em lugares arbitrários, reações químicas, fluxos em *splitters* e fluxos de energia pura.

Joris e Kalitventzeff (1987)

O procedimento desenvolvido por Joris e Kalitventzeff (1987) objetiva classificar as variáveis e medições envolvidas em qualquer tipo de modelo de planta. O sistema de equações que representa a operação da planta envolve variáveis de estado (temperatura, pressão, taxas de fluxos parciais molares de componentes e extensão de reação), medidas e variáveis de ligação (as que relacionam certas medidas com as variáveis de estado). Este sistema é composto de balanços de massa e energia, relações de equilíbrio líquido-vapor, e etc. A classificação de variáveis não medidas e das medições é alcançado pela permuta de linhas e colunas da matriz de ocorrência correspondente à matriz Jacobiana do modelo.

Na maioria dos casos, o procedimento estrutural é capaz de determinar se as medições podem ser corrigidas e quando elas possibilitam o cômputo de todas as variáveis de estado do processo. Em

algumas configurações esta técnica, usada sozinha, falha na detecção de variáveis indetermináveis. Esta situação surge quando a Jacobiana usada na resolução é não inversível.

Madron (1992)

O procedimento de classificação desenvolvido por Madron (1992) é baseado na conversão da matriz com as equações do modelo da planta lineares ou linearizadas para a forma canônica. Inicialmente é formada uma matriz composta, contendo variáveis medidas e não medidas e um vetor de constantes. Então é realizada uma eliminação Gauss-Jordan, usada para pivotear as colunas pertencentes às quantidades não medidas. Na fase seguinte, o procedimento aplica a eliminação sobre uma sub-matriz resultante que contém variáveis medidas. A forma canônica final é obtida pelo rearranjo de linhas e colunas da macro-matriz, a qual permite a classificação de ambos os tipos de variáveis. O usuário deste procedimento deve prover estimativas inicias para todas as variáveis. Esta estratégia é extensamente descrita em Madron (1992).

3.6 Decomposição usando transformações ortogonais

Crowe *et al.* (1983) propuseram uma elegante estratégia para o desacoplamento de variáveis medidas a partir das equações lineares das restrições. Este procedimento permite tanto a redução do problema de reconciliação de dados quanto a classificação das variáveis do processo. Ele é baseado no uso de projeção de matrizes para eliminar as variáveis não medidas. Posteriormente Crowe estendeu essa metodologia (em Crowe (1986, 1989a)).

Uma decomposição equivalente pode ser realizada usando as transformações ortogonais QR (SÁNCHEZ *et al.*, 1992). As fatorações ortogonais foram primeiro usadas para por Swartz (1989) no contexto das técnicas de sucessivas linearizações para eliminar as variáveis não medidas das equações de restrição (ROMAGNOLI; SÁNCHEZ, 2000).

3.6.1 Abordagem da projeção de matrizes

Representando um processo em estado estacionário através de:

$$\mathbf{A}_1\mathbf{x}+\mathbf{A}_2\mathbf{u}=\mathbf{0}, \quad \mathbf{x} \in \Re^g; \, \mathbf{u} \in \Re^n, \tag{3.47}$$

onde $\mathbf{x}$ é o vetor $(g \times 1)$ das variáveis medidas e $\mathbf{u}$ é o vetor $(n \times 1)$ das variáveis não medidas. $\mathbf{A}_1$ e $\mathbf{A}_2$ são matrizes compatíveis de dimensão $(m \times g)$ e $(m \times n)$.

Uma matriz de projeção $\mathbf{P}$ foi definida por Crowe tal que $\mathbf{P}$ pré-multiplicando a matriz Jacobiana $\mathbf{A}_2$ resulta em:

$$\mathbf{P}\mathbf{A}_2 = \mathbf{0} \tag{3.48}$$

As colunas de $\mathbf{P}$ abarcam o espaço nulo de $\mathbf{A}_2$, e assim as variáveis não medidas são eliminadas. Para obter a matriz de projeção $\mathbf{P}$, Crowe propôs o seguinte procedimento:

i. Reduzir as colunas de $\mathbf{A}_2$ para obter a matriz $\mathbf{X}$ com colunas linearmente independentes

$$\mathbf{A}_2\mathbf{A}_3 = \begin{bmatrix} \mathbf{X} & \mathbf{0} \end{bmatrix} \tag{3.49}$$

onde $\mathbf{A}_3$ representa a matriz inversível que realiza as operações necessárias sobre $\mathbf{A}_2$.

ii. Particionar $\mathbf{X}$ tal que

$$\mathbf{A}_4\mathbf{A}_2\mathbf{A}_3 = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{0} \end{bmatrix} \tag{3.50}$$

com $\mathbf{X}_1$ quadrada e inversível. Então $\mathbf{P}$ é calculado pela seguinte expressão:

$$\mathbf{P} = \begin{bmatrix} -\mathbf{X}_2\mathbf{X}_1^{-1} & \mathbf{I} \end{bmatrix} \mathbf{A}_4 \tag{3.51}$$

Finalmente o problema reduzido é formulado como:

$$\mathbf{G}\mathbf{x} = \mathbf{0} \tag{3.52}$$

onde

$$\mathbf{G} = \mathbf{P}\mathbf{A}_1 \tag{3.53}$$

A reconciliação de dados pode agora ser realizada sobre um subsistema reduzido contendo somente variáveis medidas.

3.6.2 Abordagem da fatoração QR

Uma decomposição alternativa pode ser realizada usando a fatoração QR da matriz $\mathbf{A}_2$ para desacoplar as variáveis não medidas das medidas (SÁNCHEZ; ROMAGNOLI, 1996).

TEOREMA 3.5 (Teorema da Fatoração QR (DAHLQUIST; BJORK, 1974)) Seja $\mathbf{A}$ uma matriz $(m \times n)$ com $m \geq n$ e n colunas linearmente independentes. Então existe uma única matriz $\mathbf{Q}$ $(m \times m)$,

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{D}_i, \qquad \mathbf{D}_i = \begin{bmatrix} d_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_n \end{bmatrix}; \quad d_k > 0,\ k = 1, \dots, n \tag{3.54}$$

e uma única matriz triangular superior $\mathbf{R}$ $(m \times n)$, com $R_{kk} = 1, k = 1, \dots, n$ tal que

$$\mathbf{A} = \mathbf{QR} \tag{3.55}$$

Se $\mathbf{A}$ tem o seu posto deficiente a fatoração QR pode ser modificada de um modo simples para que fique da seguinte forma:

$$\mathbf{A}\Pi = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{3.56}$$

onde $r = \text{posto}(\mathbf{A})$, $\mathbf{Q}$ é ortogonal, $\mathbf{R}_{11}$ é triangular superior e Π é uma permutação. Se $\mathbf{A}\Pi = [a_{c1}, \dots, a_{cn}]$ e $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_m]$, então para $k = 1, \dots, n$, tem-se que

$$\mathbf{a}_{ck} = \sum_{i=1}^{min\{r,k\}} \mathbf{r}_{ik}\mathbf{q}_i \in span\{\mathbf{q}_1, \dots, \mathbf{q}_r\} \tag{3.57}$$

Segue que para qualquer vetor que satisfaça $\mathbf{A}\,\mathbf{x} = \mathbf{b}$,

$$\Pi^{\mathrm{T}}\mathbf{x} = \begin{bmatrix} \mathbf{s} \\ \mathbf{z} \end{bmatrix} \quad \text{e } \mathbf{Q}^{\mathrm{T}}\mathbf{b} = \begin{bmatrix} \mathbf{i} \\ \mathbf{l} \end{bmatrix} \tag{3.58}$$

onde **s** e **i** são vetores de dimensão r, **z** é um vetor de dimensão $(n-r)$ e **l** é um vetor de dimensão $(m-r)$.

De volta ao problema da reconciliação, a fatoração QR da matriz $\mathbf{A}_2$ permite obter as matrizes $\mathbf{Q}_\mathrm{u}$ e $\mathbf{R}_\mathrm{u}$ e a matriz de permuta Π_u tais que:

$$\mathbf{A}_2\Pi_\mathrm{u} \;=\; \mathbf{Q}_\mathrm{u}\mathbf{R}_\mathrm{u} \tag{3.59}$$

onde $\mathbf{Q}_\mathrm{u}$ e $\mathbf{R}_\mathrm{u}$ podem ser divididas em:

$$\mathbf{Q}_\mathrm{u} \;=\; \left[\begin{array}{cc} \mathbf{Q}_{\mathrm{u}1} & \mathbf{Q}_{\mathrm{u}2} \end{array}\right], \mathbf{R}_\mathrm{u} \;=\; \left[\begin{array}{cc} \mathbf{R}_{\mathrm{u}1} & \mathbf{R}_{\mathrm{u}2} \\ \mathbf{0} & \mathbf{0} \end{array}\right] \tag{3.60}$$

com r_u = posto($\mathbf{A}_2$) = posto($\mathbf{R}_{\mathrm{u}1}$). Tem-se que $\mathbf{Q}_\mathrm{u}$ é uma matriz ortogonal e $\mathbf{R}_{\mathrm{u}1}$ é uma matriz inversível triangular superior de dimensão r_u. Do mesmo modo as variáveis não medidas do processo podem ser particionadas em dois subconjuntos

$$\Pi_\mathrm{u}^\mathrm{T}\mathbf{u} \;=\; \left[\begin{array}{c} \mathbf{u}_{r_u} \\ \mathbf{u}_{n-r_u} \end{array}\right] \tag{3.61}$$

Pré-multiplicando as restrições linearizadas por $\mathbf{Q}_\mathrm{u}^\mathrm{T} \;=\; \mathbf{Q}_\mathrm{u}^{-1}$, obtém-se:

$$\left[\begin{array}{ccc} \mathbf{Q}_{\mathrm{u}1}^\mathrm{T}\mathbf{A}_1 & \mathbf{R}_{\mathrm{u}1} & \mathbf{R}_{\mathrm{u}2} \\ \mathbf{Q}_{\mathrm{u}2}^\mathrm{T}\mathbf{A}_1 & \mathbf{0} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \mathbf{x} \\ \mathbf{u}_{r_u} \\ \mathbf{u}_{n-r_u} \end{array}\right] \;=\; \mathbf{0} \tag{3.62}$$

as primeiras r_u equações para $\mathbf{u}_{r_u}$ podem ser escritas em termo das outras variáveis:

$$\mathbf{u}_{r_u} \;=\; \mathbf{R}_{\mathrm{u}1}^{-1}\mathbf{Q}_{\mathrm{u}1}^\mathrm{T}\mathbf{A}_1\mathbf{x} - \mathbf{R}_{\mathrm{u}1}^{-1}\mathbf{R}_{\mathrm{u}2}\mathbf{u}_{n-r_u} \tag{3.63}$$

como as variáveis não medidas não aparecem nas equações restantes, o primeiro sub-problema reduzido se torna:

PROBLEMA 1

$$\min_x (\mathbf{y}-\mathbf{x})^{\mathrm{T}} \Psi_x^{-1} (\mathbf{y}-\mathbf{x}) \tag{3.64}$$

$$\mathbf{G_x x} \;=\; \mathbf{0} \tag{3.65}$$

onde

$$\mathbf{G}_x = \mathbf{Q}_{u2}^{\mathrm{T}} \mathbf{A}_1$$

PROBLEMA 2

Estimar as variáveis não medidas, $\mathbf{u}$, resolvendo a Eq. 3.63 onde os componentes $\mathbf{u}_{n-r_u}$ são arbitrariamente colocados. A unicidade de $\mathbf{u}$ é relaciona a estimabilidade do sistema.

3.7 Reconciliação de dados linear com todas as variáveis medidas

Este problema pode ser formulado como:

$$\min_x J(\mathbf{y}-\mathbf{x})^{\mathrm{T}} \Psi^{-1} (\mathbf{y}-\mathbf{x}) \tag{3.66}$$

$$\mathbf{A_1 x} \;=\; \mathbf{0}$$

onde $\mathbf{A}_1$ é uma matriz $(m \times g)$ com constantes conhecidas. Neste caso todas as variáveis são redundantes.

3.7.1 Método dos multiplicadores de lagrange

Introduzindo o erro da medida nas restrições do processo tem-se que:

$$\mathbf{A}_1(\mathbf{y}-\varepsilon) = \mathbf{0} \tag{3.67}$$

conseqüentemente, o problema de otimização se torna agora

$$\min_{\varepsilon} \varepsilon^{\mathrm{T}}\Psi^{-1}\varepsilon \tag{3.68}$$
$$\mathbf{A}_1\mathbf{x} = \mathbf{0}$$

A solução é obtida pelo Método dos Multiplicadores de Lagrange. O Lagrangiano para este problema é:

$$L = \varepsilon^{\mathrm{T}}\Psi^{-1}\varepsilon - 2\lambda^{\mathrm{T}}(\mathbf{A}_1\mathbf{y} - \mathbf{A}_1\varepsilon) \tag{3.69}$$

como Ψ é positivamente definido e as restrições são lineares, as condições necessárias e suficientes para a minimização são:

$$\begin{aligned} \frac{\partial L}{\partial \varepsilon} &= 2\Psi^{-1}\varepsilon + 2\mathbf{A}_1^{\mathrm{T}}\lambda = \mathbf{0} \\ \frac{\partial L}{\partial \lambda} &= \mathbf{A}_1(\mathbf{y} - \varepsilon) = \mathbf{0} \end{aligned} \tag{3.70}$$

com

$$\varepsilon = -\Psi\mathbf{A}_1^{\mathrm{T}}\lambda \tag{3.71}$$

$$\lambda = -(\mathbf{A}_1\Psi\mathbf{A}_1^{\mathrm{T}})^{-1}\mathbf{A}_1\mathbf{y} \tag{3.72}$$

e finalmente, o estimador para as variáveis do processo, $\hat{\mathbf{x}}$, pode ser obtido como

$$\hat{\mathbf{x}} = \mathbf{y} - \Psi\mathbf{A}_1^{\mathrm{T}}(\mathbf{A}_1\Psi\mathbf{A}_1^{\mathrm{T}})^{-1}\mathbf{A}_1\mathbf{y} \tag{3.73}$$

3.7.2 Método da fatoração QR

Usando-se o método da fatoração QR, o problema de estimativa por mínimos quadrados ponderados sujeito a restrições é transformado em um problema não sujeito a restrições. Para tanto, os

seguintes passos devem ser dados (ROMAGNOLI; SÁNCHEZ, 2000):

Passo1: Calcular a solução geral do sistema indeterminado ($\mathbf{A}_1\mathbf{x} = \mathbf{0}$). Usando o procedimento indicado na Seção 3.6.2. A fatoração ortogonal QR de $\mathbf{A}_1$ produz as matrizes $\mathbf{Q}_x$, $\mathbf{R}_x$ e Π_x que permitem o cálculo de $\mathbf{Q}_{x1}$, $\mathbf{Q}_{x2}$, $\mathbf{R}_{x1}$, $\mathbf{R}_{x2}$, $\mathbf{x}_{r_x}$ e $\mathbf{x}_{g-r_x}$, tais que:

$$\mathbf{A}_1\Pi_x = \mathbf{Q}_x\mathbf{R}_x \tag{3.74}$$

$$\mathbf{Q}_x = \begin{bmatrix} \mathbf{Q}_{x1} & \mathbf{Q}_{x2} \end{bmatrix}, \quad \mathbf{R}_x = \begin{bmatrix} \mathbf{R}_{x1} & \mathbf{R}_{x2} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{3.75}$$

$$\Pi_x^T\mathbf{x} = \begin{bmatrix} \mathbf{x}_{r_x} \\ \mathbf{x}_{g-r_x} \end{bmatrix} \tag{3.76}$$

onde r_x = posto($\mathbf{R}_{x1}$) = posto($\mathbf{A}_1$). A solução geral do problema é:

$$\mathbf{x}_{r_x} = -\mathbf{R}_{x1}^{-1}\mathbf{R}_{x2}\mathbf{x}_{g-r_x} \tag{3.77}$$

onde $\mathbf{x}_{g-r_x}$ é um vetor arbitrário

Passo 2: Formulação do problema não restrito. Aplicando-se os resultados anteriores, o vetor $(\mathbf{y}-\mathbf{x})$ da função objetivo é modificado tal como segue:

$$\begin{aligned} (\mathbf{y}-\mathbf{x}) &= \\ \mathbf{y} - \begin{bmatrix} \mathbf{I}_{x1} & \mathbf{I}_{x2} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{rx} \\ \mathbf{x}_{g-rx} \end{bmatrix} &= \\ \mathbf{y} + \mathbf{I}_{x1}\mathbf{R}_{x1}^{-1}\mathbf{R}_{x2}\mathbf{x}_{g-r_x} - \mathbf{I}_{x2}\mathbf{x}_{g-r_x} &= \\ \mathbf{y} + (\mathbf{I}_{x1}\mathbf{R}_{x1}^{-1}\mathbf{R}_{x2} - \mathbf{I}_{x2})\mathbf{x}_{g-r_x} & \end{aligned} \tag{3.78}$$

onde

$$\mathbf{I}\Pi_x = \begin{bmatrix} \mathbf{I}_{x1} & \mathbf{I}_{x2} \end{bmatrix}, \quad \tilde{\mathbf{I}} = \mathbf{I}_{x1}\mathbf{R}_{x1}^{-1}\mathbf{R}_{x2} - \mathbf{I}_{x2} \tag{3.79}$$

$\mathbf{I}$ representa uma matriz identidade $(g \times g)$ e $\tilde{\mathbf{I}}$ é uma matriz $[g \times (g - r_x)]$ com colunas independentes.

A minimização não restrita pode agora ser declarada como:

$$\min (\mathbf{y} + \tilde{\mathbf{I}}\mathbf{x}_{g-r_x})^{\mathrm{T}} \Psi^{-1} (\mathbf{y} + \tilde{\mathbf{I}}\mathbf{x}_{g-r_x}) \tag{3.80}$$

Passo 3: Estimativa de $\mathbf{x}$. A solução do problema acima é:

$$\hat{\mathbf{x}}_{g-r_x} = -(\tilde{\mathbf{I}}^{\mathrm{T}} \Psi^{-1} \tilde{\mathbf{I}})^{-1} \tilde{\mathbf{I}}^{\mathrm{T}} \Psi^{-1} \mathbf{y} \tag{3.81}$$

usando-se o valor de $\mathbf{x}_{g-r_x}$, a Equação 3.77 é resolvida para se calcular $\hat{\mathbf{x}}_{r_x}$.

Nota-se também que a dimensão do problema de otimização não sujeito a restrições é menor que a do problema original.

3.8 Reconciliação de dados linear com variáveis não medidas

Segundo Romagnoli e Sánchez (2000), o pressuposto que todas as variáveis são medidas é geralmente falso, pois na prática algumas variáveis não são medidas e precisam ser estimadas. Na seção anterior, a decomposição de um problema de reconciliação de dados linear envolvendo somente variáveis medidas foi discutida, levando a um problema de mínimos quadrados reduzido. Nesta seção, são usados estes conceitos para prover uma solução geral do problema de reconciliação de dados linear quando algumas das variáveis não são medidas. A solução é baseada no desacoplamento entre as variáveis não medidas e as variáveis medidas, usando-se fatoração ortogonal QR. Desta forma, o problema global de estimativa é dividido em dois sub-problemas.

Considerando-se as equações de restrição

$$\mathbf{A}_1 \mathbf{x} + \mathbf{A}_2 \mathbf{u} = \mathbf{0} \tag{3.82}$$

Realizando-se uma decomposição QR sobre a matriz $\mathbf{A}_2$, as matrizes $\mathbf{Q}_u$, $\mathbf{R}_u$ e Π_u são obtidas tais que:

$$\mathbf{A}_2 \Pi_u = \mathbf{Q}_u \mathbf{R}_u \tag{3.83}$$

$$Qu = \begin{bmatrix} \mathbf{Q}_{u1} & \mathbf{Q}_{u2} \end{bmatrix}, \quad \mathbf{R}_u = \begin{bmatrix} \mathbf{R}_{u1} & \mathbf{R}_{u2} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{3.84}$$

onde r_u=posto($\mathbf{A}_2$)=posto($\mathbf{R}_{u1}$). o vetor das variáveis não medidas é particionado em dois subconjuntos

$$\Pi_u^{\mathrm{T}}\mathbf{u} = \begin{bmatrix} \mathbf{u}_{r_u} \\ \mathbf{u}_{n-r_u} \end{bmatrix} \tag{3.85}$$

Pré-multiplicando as restrições lineares por $\mathbf{Q}^{\mathrm{T}}$, obtém-se:

$$\begin{aligned} \mathbf{Q}_{u1}^{\mathrm{T}}\mathbf{A}_1\mathbf{x} + \mathbf{R}_{u1}\mathbf{u}_{r_u} + \mathbf{R}_{u2}\mathbf{u}_{n-r_u} &= \mathbf{0} \\ \mathbf{Q}^{\mathrm{T}}\mathbf{A}_{1x} &= \mathbf{0} \end{aligned} \tag{3.86}$$

Realizando a reconciliação no subsistema desacoplado representado pelas variáveis medidas $\mathbf{x}$ e as restrições

$$\mathbf{Q}_{u2}^{\mathrm{T}}\mathbf{A}_1\mathbf{x} = \mathbf{G}_x\mathbf{x} = \mathbf{0} \tag{3.87}$$

a solução do sistema é:

$$\hat{\mathbf{x}} = \mathbf{y} - \Psi\mathbf{G}_x^{\mathrm{T}}(\mathbf{G}_x\Psi\mathbf{G}_x^{\mathrm{T}})^{-1}\mathbf{G}_x\mathbf{y} \tag{3.88}$$

Contudo, este problema pode ser reduzido ainda mais, usando-se os conceitos desenvolvidos anteriormente, para o caso no qual todas as variáveis são medidas, levando à solução de uma seqüência de sub-problemas menores.

Para as variáveis não medidas, tem-se em geral

$$\mathbf{u}_{r_u} = -\mathbf{R}_{u1}^{\mathrm{T}}\mathbf{Q}_{u1}^{\mathrm{T}}\mathbf{A}_1\hat{\mathbf{x}} - \mathbf{R}_{u1}^{-1}\mathbf{R}_{u2}\mathbf{u}_{n-r_u} \tag{3.89}$$

onde os componentes $\mathbf{u}_{n-r_u}$ são dispostos arbitrariamente. Pode-se ter, assim, dois casos:

i. Posto $(\mathbf{R}_{u1}) = n$

ii. Posto $(\mathbf{R}_{u1}) < n$

Caso 1 Todos os parâmetros não medidos são estimáveis (observáveis) e uma solução única para as variáveis não medidas é possível usando-se os valores mensurados ajustados e as equações de balanço;

Caso 2 Algumas variáveis não medidas do processo não são estimáveis e é possível um infinito número de soluções. Assim, a solução básica é:

$$\mathbf{u}_{r_u} = -\mathbf{R}_{u1}^{-1}\mathbf{Q}_{u1}^{\mathrm{T}}\mathbf{A}_1\hat{\mathbf{x}}; \quad \mathbf{u}_{n-r_u} = \mathbf{0} \tag{3.90}$$

3.9 Conclusões

Neste capítulo foram vistas as bases matemáticas da reconciliação de dados, tanto no que concerne à decomposição (e redução) do problema de estimativa quanto aos detalhes da própria solução. Os desenvolvimentos apresentados aqui se referem à solução do problema linear, mais comumente fruto de um balanço global de massa, gerado por um processo em estado estacionário. Contudo, essas soluções formam a base matemática que será estendida em escopo progressivamente nos capítulos subseqüentes.

Os pontos mais importantes deste capítulo são as técnicas para decomposição usando transformações ortogonais e as metodologias para solução do problema de reconciliação com todas as variáveis medidas e quando há variáveis não medidas a serem determinadas.

4 *Reconciliação de Dados em Estado Estacionário para Sistemas Bilineares*

Neste capítulo são apresentadas técnicas para classificação de variáveis e resolução do problema de reconciliação de dados estacionária em sistemas não lineares com certas peculiaridades que permitem uma abordagem que tem como principal vantagem a velocidade de resolução. Inicialmente é feita a formulação geral do problema para logo em seguida introduzir o método que foi especificamente desenvolvido para a reconciliação de dados de sistemas bilineares – o método de Crowe. O método é detalhado para o caso da presença de variáveis não medidas e para fluxos de entalpia. Apesar deste método ser mais eficiente que técnicas de programação não linear, há a desvantagem de não poder tratar rigorosamente restrições de desigualdades, como simples limites sobre variáveis.

4.1 Reconciliação de dados em sistemas bilineares

Em uma planta química, as correntes do processo podem conter várias espécies ou componentes. Além das taxas de fluxo das correntes, as composições de algumas das correntes também são medidas. Como os analisadores de composição são comparativamente mais caros, os analisadores *on line* não podem ser usados em vários casos e estas medidas são obtidas em laboratório, o que por sua vez pode aumentar os erros nos dados reportados por não estarem prontamente disponíveis a todo instante. Geralmente nem o balanço global, nem o balanço de componentes são satisfeitos pelas medidas. Portanto, se faz necessária a reconciliação tanto das medidas de fluxo global quanto de composição simultaneamente.

As restrições do problema de reconciliação de dados são lineares considerando-se somente os balanços globais do fluxo. Contudo, se é necessário reconciliar simultaneamente medidas de fluxo e composição, então os balanços de componentes também devem ser incluídos como restrições

do problema de reconciliação. Estas restrições contêm termos de taxa de fluxo de componentes que são produto das taxas de fluxo pelas composições. Como estas restrições são não lineares, a solução é obtida usando técnicas de reconciliação de dados não lineares. É possível também resolver o problema de reconciliação multi-componente de um modo mais eficiente, explorando o fato dos termos não lineares nas restrições serem no máximo produto de duas variáveis.

O termo **reconciliação de dados bilinear** é usado para se referir a problemas não lineares devido à restrições que são produto de duas variáveis. As razões para desenvolver técnicas especiais para a solução de problemas deste tipo são duas. Primeiro, estas técnicas serão mais eficientes que as técnicas não lineares. Isto se torna especialmente importante quando se realiza a reconciliação na escala completa de uma planta e ainda mais quando o objetivo é a reconciliação em tempo real com a finalidade de controle. Segundo, um número significativo de aplicações industriais de reconciliação de dados é voltado para sistemas multicomponentes.

Um exemplo típico é a reconciliação de fluxos e composições em torno de uma única coluna de destilação ou uma seqüencia de colunas, tal como o trem de destilação de um complexo petroquímico. Em vários casos, a reconciliação de fluxos e temperaturas de fluxos energéticos são também problemas bilineares se a entalpia específica é função somente da temperatura. Um trem de pré-aquecimento de crus de uma refinaria e a rede de distribuição de vapor de um processo químico são dois importantes exemplos. Deve-se acrescentar, contudo, que estas técnicas somente resolvem o problema de modo mais eficiente em relação ao consumo de tempo computacional, mas não trazem nenhum benefício adicional em termos de exatidão das estimativas.

> **EXEMPLO 4.1** Um problema comum de reconciliação de dados bilinear é a reconciliação de fluxos globais e composições de uma coluna de destilação binária, ilustrada na Figura 4.1. Considera-se que todos os fluxos e frações molares dos componentes da alimentação, corrente do destilado e corrente da base sejam medidos. Um conjunto típico de valores medidos é mostrado na última coluna da Tabela 4.1. As discrepâncias nos fluxos mássicos e as equações de normalização são mostradas na Tabela 4.2. Pode-se observar nesta tabela que os fluxos medidos e as composições não satisfazem aos balanços materiais nem às equações de normalização.

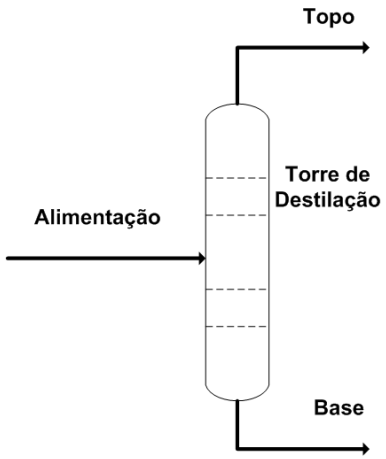


Figura 4.1: Coluna de destilação binária – adaptado de Narasimhan e Jordache (2000)

Tabela 4.1: Dados operacionais de uma coluna de destilação binária (NARASIMHAN; JORDACHE, 2000)

Corrente	**Variáveis**	**Valores Medidos**
Alimentação - F	Fluxo Mássico	1095,47
	Componente 1 (%)	48,22
	Componente 2 (%)	51,70
Destilado - D	Fluxo Mássico	478,4
	Componente 1 (%)	94,1
	Componente 2 (%)	5,01
Base - B	Fluxo Mássico	488,23
	Componente 1 (%)	1,97
	Componente 2 (%)	97,48

Tabela 4.2: Resíduos das restrições de balanço antes da reconciliação (NARASIMHAN; JORDACHE, 2000)

Tipo do Balanço	**Resíduo**
Balanço Global	-128,841
Balanço de Componentes	
1	-68,417
2	-66,406
Equações de Normalização	
Alimentação (F)	-0,089
Destilado (D)	-0,892
Base (B)	-0,551

A função objetivo da reconciliação é formulada como na Equação 3.25 e é dada por:

$$\begin{aligned}\min_{F,x} N &= W_F(\hat{F}-F)^2 + W_B(\hat{B}-B)^2 + W_D(\hat{D}-D)^2 \\ &+ \sum_{k=1}^{2}\Big[W_{x_{Fk}}(\hat{x}_{Fk}-x_{Fk})^2 + W_{x_{Bk}}(\hat{x}_{Bk}-x_{Bk})^2 \\ &+ W_{x_{Dk}}(\hat{x}_{Dk}-x_{Dk})^2\Big]\end{aligned} \tag{4.1}$$

onde os W são os fatores ponderantes e os x_{jk} são as frações molares dos componentes. Os três primeiros termos da função objetivo acima são a soma ponderada dos ajustes quadráticos feitos sobre os fluxos globais das correntes e os outros termos envolvem os ajustes feitos sobre as medidas das frações molares.

As estimativas reconciliadas têm que satisfazer os balanços materiais em torno da coluna. Os diferentes tipos de restrições que podem ser impostos são:

i. Balanço global em torno da coluna.

ii. Balanços de fluxos de componentes para todos os componentes.

iii. Equação de normalização para as frações molares de cada corrente.

Nem todas as restrições acima precisam ser impostas pois elas não são todas independentes. Para um separador, como na coluna de destilação considerada neste exemplo, um conjunto completo de restrições independentes é formado pelos balanços de componentes e as equações de normalização. O balanço global do fluxo material pode ser derivado usando estes dois tipos de equação e assim não precisa ser imposto. Um erro comum é assumir que através da imposição do balanço global e dos balanços dos componentes, as estimativas reconciliadas para as frações molares vão automaticamente satisfazer à restrição de normalização para todas as correntes. Este não é o caso, como pode ser demonstrado através do Exemplo 4.1, para o qual:

$$Fx_{F1} - Bx_{B1} - Dx_{D1} = 0 \tag{4.2}$$

$$Fx_{F2} - Bx_{B2} - Dx_{D2} = 0 \tag{4.3}$$

$$x_{F1} + x_{F2} = 1 \tag{4.4}$$

$$x_{B1} + x_{B2} = 1 \tag{4.5}$$

$$x_{D1} + x_{D2} = 1 \tag{4.6}$$

As restrições ao balanço de componentes nas Equações 4.2 e 4.3 contêm produtos da taxa de fluxo com a composição, o quê faz desse problema de reconciliação de dados mais difícil de resolver quando comparado com o caso linear considerado no Capítulo 3. A função objetivo 4.1, em conjunto com as restrições expressas nas Equações 4.2 até 4.6, podem ser tratadas como um problema de otimização não linear com restrições de igualdade e pode ser resolvido usando um programa de otimização não linear com restrições, como o método da programação quadrática sucessiva (SQP - *Successive Quadratic Programming*). Contudo, foram desenvolvidos métodos eficientes para este tipo de problema. Usando um destes métodos, os dados reconciliados para o Exemplo 4.1 foram obtidos e estão na Tabela 4.3.

Na Tabela 4.3, a segunda coluna mostra as estimativas reconciliadas quando os balanços de componentes e as equações de normalização são impostas, enquanto que na terceira coluna estão as estimativas quando o fluxo global e os balanços dos componentes são usados sem as restrições de normalização. Notar que, nesse último caso, os percentuais não fecham balanço. Os desbalanços nas restrições após a reconciliação para ambos os casos são dados na Tabela 4.4.

Tabela 4.3: Dados reconciliados de uma coluna de destilação binária (NARASIMHAN; JORDACHE, 2000)

	Valores Reconciliados	
Variáveis	**Com Restrições de Normalização**	**Sem Restrições de Normalização**
Fluxo Mássico	1009,51	1009,48
Componente 1 (%)	48,24	48,05
Componente 2 (%)	51,76	51,54
Fluxo Mássico	502,22	503,14
Componente 1 (%)	94,99	94,42
Componente 2 (%)	5,01	5,01
Fluxo Mássico	507,29	506,35
Componente 1 (%)	1,97	1,97
Componente 2 (%)	98,03	97,77

Os resultados nesta tabela demonstram claramente a necessidade da inclusão das restrições de

normalização nos problemas de reconciliação de dados multicomponentes.

Tabela 4.4: Resíduos das restrições de balanço depois da reconciliação (NARASIMHAN; JORDACHE, 2000)

	Valores dos Resíduos	
Tipo do Balanço	**Com Restrições de Normalização**	**Sem Restrições de Normalização**
Fluxo Mássico Global	0,0000E+00	0,0000E+00
Componente 1	8,5453E-13	9,9349E-10
Componente 2	$<$ 1,0E-13	4,8778E-10
Equações de Normalização		
Alimentação - F	0,0000E+00	0,41
Destilado - D	0,0000E+00	0,57
Base - B	0,0000E+00	0,26

4.2 Formulação geral do problema

O exemplo anterior mostra que a reconciliação de dados multi-componentes para uma coluna de destilação é um problema bilinear. De modo similar, os dados de uma seqüencia de colunas de separação podem ser reconciliados resultando também em um problema bilinear. Inicialmente é apresentada a formulação geral da reconciliação de dados multi-componente de tais processos típicos.

Segundo Narasimhan e Jordache (2000), dependendo do processo e do subsistema que é considerado, vários tipos diferentes de unidade de processo podem ser encontradas. Na indústria de processos químicos, as diferentes unidades onde o fluxo ou as composições das correntes passam por mudanças podem ser classificadas como misturadores (*mixers*), divisores de corrente (*splitters*), separadores e reatores. O tipo de restrições que podem ser impostas depende da natureza da unidade, portanto é importante se ter uma clara compreensão do conjunto completo de restrições independentes que podem ser impostas para cada unidade e assim, para o processo inteiro. Apesar de que para cada unidade de processo possam ser escritas diferentes combinações de restrições independentes, geralmente são escolhidas as equações a seguir, nas quais existem C componentes, S correntes, E espécies químicas e R reações independentes e onde o subscrito *in* denota entrada e o subscrito *out* denota saída.

4.2.1 Misturadores (*Mixers*)

Um misturador tem duas ou mais correntes de entrada e tem uma corrente de saída como é mostrado na Figura 4.2. Se as correntes são monofásicas[1], estão as restrições impostas para esta unidade são:

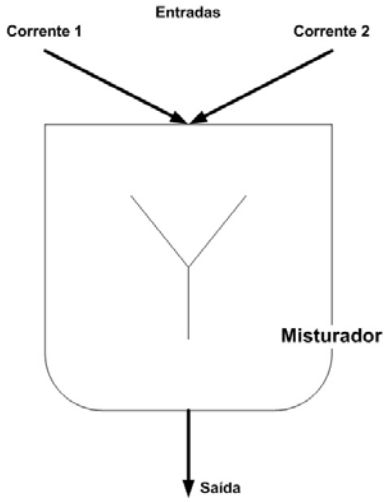


Figura 4.2: Unidade de mistura – adaptado de Narasimhan e Jordache (2000)

i. Balanços de Fluxo de Componente:

$$\sum_{j=1}^{S} F_j x_{jk} - F_{\text{out}} x_{\text{out}\,k} = 0 \quad k = 1 \ldots C$$

ii. Equações de Normalização

$$\sum_{k=1}^{C} x_{jk} = 1 \quad j = 1 \ldots S$$

$$\sum_{k=1}^{C} x_{\text{out}\,k} = 1$$

4.2.2 Divisores de Corrente (*Splitters*)

Um divisor de corrente divide uma corrente de entrada em duas ou mais correntes de saída, como é mostrado na Figura 4.3. As restrições que podem ser escritas para essa unidade são:

[1]Esta é uma condições importante, caso contrário seria necessário acrescentar à modelagem relações de equilíbrio de fases.

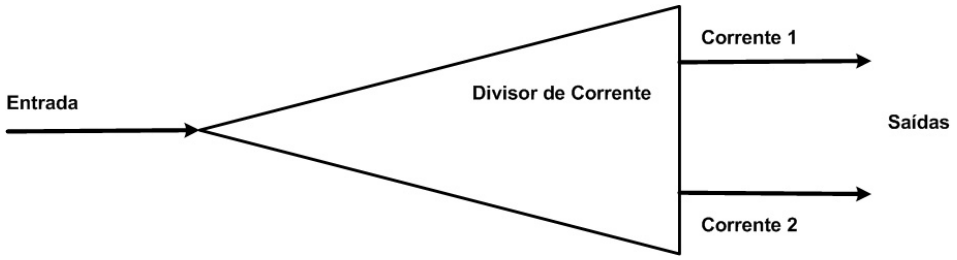


Figura 4.3: Unidade de divisão de corrente – adaptado de Narasimhan e Jordache (2000)

i. Balanços de fluxo de componente (igualdade de composição em cada corrente):

$$x_{\text{in}\,k} - x_{jk} = 0 \quad j = 1 \ldots S, \quad k = 1 \ldots C$$

ii. Balanço de fluxos globais

$$F_{\text{in}} - \sum_{j=1}^{S} F_j = 0$$

iii. Equação de normalização para a corrente de alimentação

$$\sum_{j=1}^{C} x_{\text{in}\,k} = 1$$

Todas as outras restrições, tais como balanços de componentes e restrições de normalização para as correntes de saída, podem ser derivadas por combinações apropriadas das equações acima. Uma observação é que se um divisor de corrente faz parte de um subsistema, então as equações de normalização devem ser escritas somente para a corrente de entrada e não para as de saída, haja vista que o divisor de corrente preserva as composições e a normalização das correntes de saída seria redundante.

Uma formulação alternativa que faz uso da definição das frações de partição é algumas vezes mais conveniente. Seja α_j a razão da taxa de fluxo da corrente de saída j sobre a corrente de entrada do divisor de corrente. Então as seguintes equações também constituem um conjunto completo e não redundante:

i. Balanços de fluxos de componente

$$F_j x_{jk} - \alpha_j F_{\text{in}} x_{\text{in}\,k} = 0 \quad j = 1 \ldots S; \quad k = 1 \ldots C$$

ii. Balanço de fluxo global

$$\sum_{j=1}^{S} \alpha_j = 1$$

iii. Equação de normalização para a corrente de entrada

$$\sum_{k=1}^{C} x_{\text{in}k} = 1$$

iv. Definição das frações de partição

$$F_j - \alpha_j F_{\text{in}} = 0 \quad j = 1 \ldots S$$

O uso da fração de divisão introduz um número de variáveis adicionais igual ao de correntes de saída e assim o número de equações independentes que devem ser escritas para um divisor de corrente usando frações de partição de corrente é igual a $CS + S + 2$. O uso das frações de partição de corrente também complica mais ainda o problema, pois os balanços dos componentes não são mais bilineares, mas sim trilineares (produto de três variáveis).

4.2.3 Separadores

Um separador, que é inverso de um misturador, toma uma corrente de entrada e a separa em duas ou mais correntes de diferentes composições como mostrado na Figura 4.4. Se todas as correntes são monofásicas (comentário em nota à página 89), as equações para estas unidades são similares àquelas para o misturador.

i. Balanços de fluxos de componentes

$$F_{\text{in}} x_{\text{in}k} - \sum_{j=1}^{S} F_j x_{jk} = 0 \quad k = 1 \ldots C$$

ii. Equações de normalização

$$\sum_{k=1}^{C} x_{\text{in}k} = 1$$

$$\sum_{k=1}^{C} x_{jk} = 1 \quad j = 1 \ldots S$$

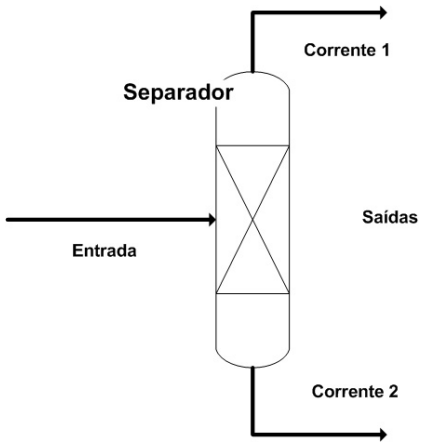


Figura 4.4: Unidade de separação – adaptado de Narasimhan e Jordache (2000)

4.2.4 Reatores

Considera-se um reator com uma corrente de alimentação e uma corrente de produtos tal como na Figura 4.5. Os reatores com múltiplas alimentações ou correntes de produtos podem ser modelados usando um misturador antes do reator e um separador depois. Devido às reações, nem o fluxo molar global, nem o fluxos molares dos componentes são conservados. Existem duas alternativas para as equações do modelo de um reator. Numa abordagem, assume-se que as reações independentes que ocorrem no reator são especificadas. Seja n_{kj} o coeficiente estequiométrico do componente k na reação j e sejam ξ_j, $j = 1...R$ as extensões[2] desconhecidas da reação, onde R é o número de reações independentes especificadas e *nu* o coeficiente estequiométrico. Usando as extensões, pode-se escrever as seguintes equações:

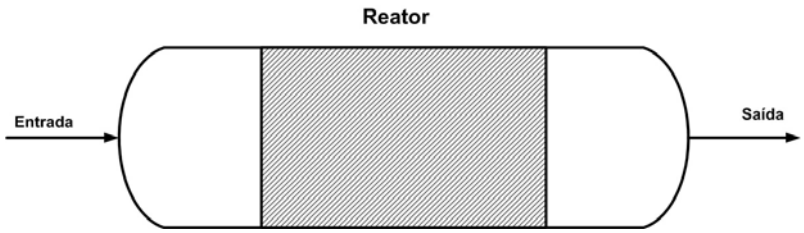


Figura 4.5: Reator – adaptado de Narasimhan e Jordache (2000)

[2]Extensão da reação ou grau de avanço é a quantidade extensiva que descreve o progresso de uma reação química como sendo igual ao número de transformações químicas, como indicado pela equação da reação numa escala molecular, dividido pela constante de Avogadro (é essencialmente a quantidade de transformações químicas). A variação na extensão da reação é dada por $d\xi = \frac{dn_B}{\nu_B}$, onde ν_B é o coeficiente estequiométrico de uma dada entidade B (reagente ou produto) e n_B é sua quantidade correspondente (IUPAC Compendium of Chemical Terminology, Electronic version, http://goldbook.iupac.org/E02283.html).

i. Balanço de componentes

$$F_{\text{in}} x_{\text{in}\,k} + \sum_{j=1}^{R} \nu_{kj} \xi_j - F_{\text{out}} x_{\text{out}\,k} = 0 \quad k = 1 \ldots C$$

ii. Equações de normalização

$$\sum_{k=1}^{C} x_{\text{in}\,k} = 1$$

$$\sum_{k=1}^{C} x_{\text{out}\,k} = 1$$

O conjunto alternativo de equações do modelo é obtido usando-se o fato de que cada espécie elementar é conservada. Denota-se o número de átomos do elemento j no componente k por a_{jk}, então pode-se escrever as seguintes equações para o reator:

i. Balanços elementares

$$\sum_{k=1}^{C} (a_{jk} x_{\text{in}\,k} F_{\text{in}} - a_{jk} x_{\text{out}\,k} F_{\text{out}}) = 0 \quad j = 1 \ldots E$$

ii. Equações de normalização

$$\sum_{k=1}^{C} x_{\text{in}\,k} = 1$$

$$\sum_{k=1}^{C} x_{\text{out}\,k} = 1$$

Foi mostrado em Reklaitis (apud NARASIMHAN; JORDACHE, 2000) que estes conjuntos de equações são equivalentes e dão resultados idênticos somente se o conjunto completo de reações independentes que podem ocorrer entre os componentes presentes é especificado. Na ausência de qualquer informação a respeito das reações que podem ocorrer, o modelo do balanço elementar pode ser usado. Contudo, se balanços de energia também têm que ser incluídos como parte da reconciliação, então o modelo de extensão de reação é mais conveniente.

4.2.5 Classificação das equações dos modelos

As equações dos modelos das várias unidades descritas podem ser classificadas tanto como do tipo **unidade de processo** ou do tipo **relações de correntes**. Os balanços de fluxo global e de

componentes são categorizados como equações do tipo unidade de processo enquanto que equações de normalização que relacionam variáveis de uma corrente são classificadas como relações de corrente. Essa classificação ajuda na rápida disposição das restrições para um processo ou subsistema que tenha combinações de algumas das unidades acima. Inicialmente, as equações do tipo unidade de processo para todas as unidades são escritas, seguidas pelas equações do tipo relações de corrente correspondentes às unidades de processo.

EXEMPLO 4.2 Considerando-se um processo simples de uma planta de suco orgânico sintético em Meyer *et al.* (apud NARASIMHAN; JORDACHE, 2000), mostrado na Figura 4.6, o qual consiste de um divisor de correntes e três separadores. As seguintes restrições são escritas para esse processo.

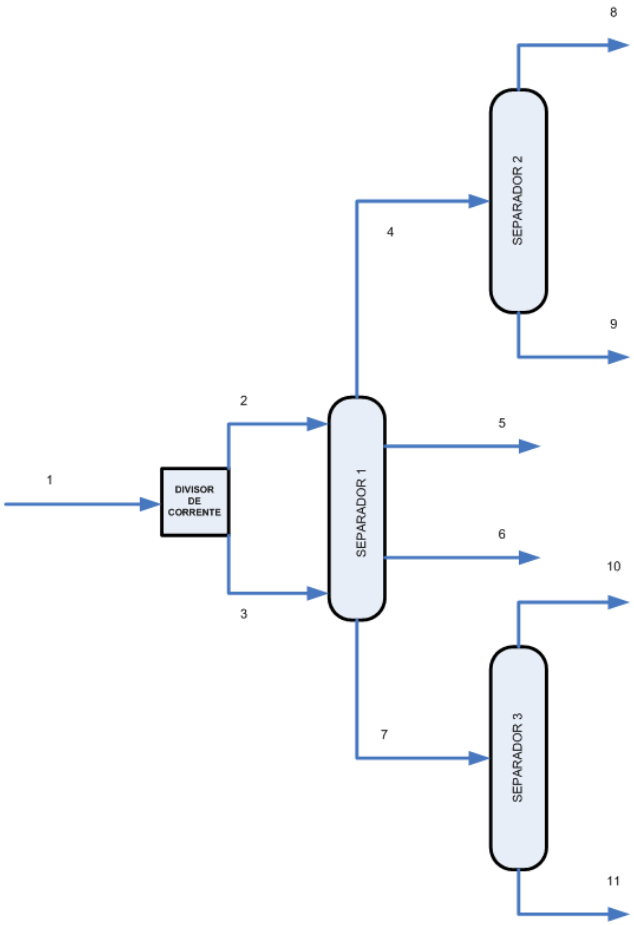


Figura 4.6: Processo de produção de suco orgânico sintético – adaptado de Narasimhan e Jordache (2000)

i. **Do Tipo Unidade de Processo** – Os balanços de componentes em torno de cada um

dos três separadores e igualdade de composição entre as correntes 1 e 2 e entre 1 e 3 (equações do divisor de corrente);

ii. **Do Tipo Relações de Corrente** – Equações de normalização para corrente 1 e correntes de 4 a 11.

As equações de normalização para as correntes 2 e 3 não são escritas pois são as correntes de saída do divisor de corrente.

4.3 Solução da reconciliação de dados bilinear

Um exame das equações para as diferentes unidades descritas no Exemplo 4.2 mostra que as únicas não linearidades que aparecem ocorrem como produtos de duas variáveis. A técnica proposta a seguir para resolver problemas de reconciliação de dados bilinear explora essa característica de uma maneira eficiente.

4.3.1 Método de Crowe

Para ilustrar o método desenvolvido por Crowe (1986) aplicado a problemas bilineares, considera-se inicialmente um processo multicomponentes consistindo somente de separadores e misturadores monofásicos e todos os fluxos de composições medidos. As restrições para esse processo consistem de balanços de componentes em torno de cada unidade e equações de normalização para cada corrente. Se existem m unidades, S correntes e C componentes, então as restrições do processo são:

$$\sum_{j=1}^{S} a_{ij} F_j x_{jk} = 0 \quad i = 1 \ldots m; \quad k = 1 \ldots C \tag{4.7}$$

$$\sum_{k=1}^{C} x_{jk} = 1 \quad j = 1 \ldots S \tag{4.8}$$

O objetivo é determinar estimativas de todos os fluxos e composições de modo que a soma quadrática ponderada dos ajustes feitos nos fluxos globais e nas composições seja minimizado. A função objetivo é dada por:

$$\min_{F_j,x_{jk}} N = \sum_{j=1}^{S} (\hat{F}_j - F_j)^2 W_{F_j} + \sum_{j=1}^{S} \sum_{k=1}^{C} (\hat{x}_{jk} - x_{jk})^2 W_{x_{jk}} \tag{4.9}$$

A formulação acima do problema de reconciliação de dados está em termos de taxa de fluxo e fração molar. Em uma formulação alternativa pode-se deixar o problema em termos de fluxo global e fluxo de componentes, onde o fluxo de componentes N_{jk} do componente k na corrente j é definido como:

$$N_{jk} = F_j x_{jk} \quad k = 1 \ldots C \quad j = 1 \ldots S \tag{4.10}$$

usando essas variáveis, os balanços dos componentes podem ser escrito como:

$$\sum_{j=1}^{S} a_{ij} N_{jk} = 0 \quad i = 1 \ldots m; \quad k = 1 \ldots C \tag{4.11}$$

as equações de normalização também podem ser escritas como:

$$F_j - \sum_{k=1}^{C} N_{jk} = 0 \quad j = 1 \ldots S \tag{4.12}$$

Pode-se observar das Equações 4.11 e 4.12 que as restrições são lineares nas variáveis de fluxos e esta característica pode ser explorada no procedimento da solução. Mesmo que as restrições sejam agora em termos de variáveis de fluxo, a função objetivo continua contendo frações molares pois estas são as quantidades medidas. Para superar este problema, Crowe (1986) propôs uma função objetivo modificada para o problema de reconciliação de dados, que é a minimização da soma quadrática ponderada dos ajustes feitos nos fluxos globais e nos fluxos dos componentes. Neste caso, a função objetivo modificada é da forma:

$$\min_{F_j,N_{jk}} N = \sum_{j=1}^{S} (\hat{F}_j - F_j)^2 W_{F_j} + \sum_{j=1}^{S} \sum_{k=1}^{C} (\hat{N}_{jk} - N_{jk})^2 W_{N_{jk}} \tag{4.13}$$

Como os fluxos dos componentes não são as quantidades medidas, é necessário esclarecer o conceito de valor medido de fluxo de componentes e os fatores de ponderação a serem usados para estas variáveis na função objetivo acima. Um fluxo de componente N_{jk} é tomado como uma quantidade medida se tanto o fluxo F_j e a composição x_{jk} são medidos.

Já foi visto no Capítulo 3 que o fator de ponderação da variável medida pode ser escolhido

como a inversa da variância do erro na medida. Uma estimativa de variância do erro no produto $N_j k$ é obtida por uma linearização em termos das medidas de taxa de fluxo e composição.

$$\hat{N}_{jk} \approx x^*_{jk}(\hat{F}_j - F^*_j) + F^*_j(\hat{x}_{jk} - x^*_{jk}) + F^*_j x^*_{jk} \tag{4.14}$$

A variância $\sigma^2_{\hat{N}_{jk}}$ do erro em $\hat{N}_{jk}$ pode ser obtida pela aplicação da regra para a soma linear de variáveis independentes normalmente distribuídas.

$$\sigma^2_{\hat{N}_{jk}} = (x^*_{jk})^2 \sigma^2_{\hat{F}_j} + (F^*_j)^2 \sigma^2_{\hat{x}_{jk}} \tag{4.15}$$

o fator ponderante $W_{N_{jk}}$ pode ser tomado como sendo igual a inversa da variância $\sigma^2_{\hat{N}_{jk}}$.

A escolha da função objetivo modificada para a reconciliação de dados e os fatores ponderantes para os fluxos de componentes "medidos" pode levar a ajustes maiores sobre as medidas. Contudo, a função objetivo continua indiretamente tentando minimizar o ajuste total feito sobre as variáveis medidas.

A função objetivo modificada 4.13 sujeita às Equações 4.11 e 4.12 leva a um problema de reconciliação de dados linear em variáveis de fluxos. Para o caso especial considerado aqui, todas as variáveis são medidas e as estimativas são imediatamente obtidas para todos os fluxos usando a solução analítica da Equação 3.22. A partir destas estimativas, os valores reconciliados das frações molares pode ser obtido como (NARASIMHAN; JORDACHE, 2000):

$$\hat{x}_{jk} = \frac{\hat{N}_{jk}}{\hat{F}_j} \tag{4.16}$$

EXEMPLO 4.3 O método de Crowe é aplicado para reconciliar os dados de uma destilação binária discutida no Exemplo 4.1. Os fluxos e composições medidos são dados na Tabela 4.1. Os valores verdadeiros e os reconciliados obtidos usando o método de Crowe são dados na Tabela 4.5. Para se obter os valores reconciliados, as variâncias nos erros das medidas de fluxos são tomadas como 5% dos valores verdadeiros e para as composições, a variância fica em 1% dos valores verdadeiros. Se comparado com os valores reconciliados mostrados na Tabela 4.3, que são obtidos usando uma técnica de otimização não linear, o método de Crowe provê estimativas de fluxo mais exatas às custas de uma maior inexatidão nas estimativas das composições. Isto se deve ao fato de que o método Crowe ajustar os fluxos dos componentes ao invés das composições.

Tabela 4.5: Dados reconciliados de uma coluna de destilação binária pelo método de Crowe (NARASIMHAN; JORDACHE, 2000)

Corrente	Variáveis	Valores Verdadeiros	Valores Reconciliados
Alimentação - F	Fluxo Mássico	1000	1002,7
	Componente 1 (%)	48,00	46,34
	Componente 2 (%)	52,00	50,33
Destilado - D	Fluxo Mássico	494,624	496,1
	Componente 1 (%)	95,00	91,22
	Componente 2 (%)	5,00	5,28
Base - B	Fluxo Mássico	505,376	506,6
	Componente 1 (%)	2,00	2,39
	Componente 2 (%)	98,00	94,43

4.3.2 Tratamento de variáveis não medidas

A presença de fluxos ou composições não medidas introduz complicações sutis no método de Crowe. Dependendo das medidas que são feitas, as correntes podem ser classificadas em duas categorias

i. Correntes com fluxos medidos e algumas ou todas composições não medidas;

ii. Correntes com fluxos não medidos e algumas ou todas as composições não medidas.

Não se pode obter o valor do fluxo de componentes nas situações onde a composição correspondente não é medida, o fluxo global não é medido, ou se ambos não são medidos. Como existe uma correspondência biunívoca entre as variáveis de composição e os fluxos dos componentes, é apropriado considerar o fluxo do componente como não medido se a variável de composição correspondente não é medida, não importando se o fluxo da corrente é medido ou não. Contudo se o fluxo global de uma corrente não é medido, então tratar todos os fluxos dos componente desta corrente como não medidos resultaria numa perda desnecessária de informação das composições medidas desta corrente. Para se evitar isso, o método de Crowe classifica os fluxos das correntes e os fluxos dos componentes nas seguintes categorias:

Categoria I Composta de todas as variáveis de fluxo de corrente medidas e os fluxos "medidos" de componentes. Assim essa categoria é composta somente das variáveis medidas;

Categoria II Composta de todos os fluxos de componentes correspondentes às composições medidas, mas com os fluxos globais não medidos e também todos os fluxos de corrente não

medidos. Assim essa categoria é composta de uma mistura de composições medidas e fluxos de correntes não medidos;

Categoria III Composta de todos os fluxos de componentes correspondentes às composições não medidas para os quais o fluxo da corrente seja ou não seja medido. Desta forma esta categoria é composta somente de variáveis não medidas.

Os fluxos globais e os fluxos de componentes nas diferentes categorias são denotadas por superescritos **I**, **II** e **III**. A função objetivo para o problema de reconciliação de dados pode agora ser formulada como

$$\begin{aligned} \min_{F^{\mathrm{I}},N_k^{\mathrm{I}},x_k^{\mathrm{II}}} & (\hat{\mathbf{F}}^{\mathrm{I}}-\mathbf{F}^{\mathrm{I}})^{\mathrm{T}}\mathbf{W}_{F^{\mathrm{I}}}(\hat{\mathbf{F}}-\mathbf{F}^{\mathrm{I}})^2 \quad + \\ & \sum_{k=1}^{C}(\hat{\mathbf{N}}_k^{\mathrm{I}}-\mathbf{N}_k^{\mathrm{I}})^{\mathrm{T}}\mathbf{W}_{N^{\mathrm{I}}}(\hat{\mathbf{N}}_k^{\mathrm{I}}-\mathbf{N}_k^{\mathrm{I}}) \quad + \\ & \sum_{k=1}^{C}(\hat{\mathbf{x}}_k^{\mathrm{II}}-\mathbf{x}_k^{\mathrm{II}})^{\mathrm{T}}\mathbf{W}_{x^{\mathrm{II}}}(\hat{\mathbf{x}}_k^{\mathrm{II}}-\mathbf{x}_k^{\mathrm{II}}) \end{aligned} \tag{4.17}$$

A função objetivo acima foi expressa de forma compacta usando os vetores $\mathbf{F}$, $\mathbf{N}_k$ e $\mathbf{x}_k$ correspondentes aos fluxos globais, aos fluxos dos k componentes e às composições de todas as correntes em cada categoria, respectivamente. As matrizes de ponderação $\mathbf{W}$ são matrizes diagonais, com as entradas na diagonal sendo os fatores ponderantes das variáveis apropriadas de todas as correntes em cada uma das categorias.

As restrições do problema de reconciliação de dados são os balanços materiais para cada unidade como descrito anteriormente. Estas equações podem ser colocadas em termos de variáveis nas três categorias. Para a solução deste problema Crowe propôs uma estratégia de decomposição em dois estágios para eliminação das variáveis não medidas das equações das restrições. No primeiro estágio, os fluxos de componentes não medidos na categoria III são eliminados usando uma matriz de projeção. Para tanto, pode ser seguido o procedimento usado na reconciliação de dados linear porque as restrições são lineares nos fluxos dos componentes. No segundo estágio, os fluxos globais não medidos na categoria II são eliminados usando uma segunda matriz de projeção. Isto requer algumas manipulações algébricas das equações das restrições, as quais são descritas em Crowe (1986).

O problema de reconciliação de dados reduzido requer ainda um procedimento iterativo para resolver as estimativas para as composições da categoria II e os fluxos dos componentes da categoria I, começando com as estimativas arbitrárias[3] iniciais dos fluxos globais de categoria II. Pode-se verificar que se as estimativas de fluxo de categoria II são dadas, então o problema de reconciliação de dados reduzido se torna linear e pode ser resolvido analiticamente. Estas estimativas reconciliadas são usadas para retro-calcular os fluxos não medidos da categoria II, usando-se um procedimento similar ao descrito no Capítulo 3, que são usadas como estimativas arbitrárias iniciais para a próxima iteração até a convergência.

Depois que as estimativas das variáveis das categorias I e II são obtidas, estas podem ser usadas para retro-calcular as estimativas para os fluxos de componentes não medidos de categoria III. Como o método de Crowe dá diretamente as estimativas dos fluxos dos componentes na categoria I e III, as estimativas de fração molar são obtidas através da Equação 4.16 (NARASIMHAN; JORDACHE, 2000).

EXEMPLO 4.4 Considerando-se o processo de flotação de minerais descrito por Smith e Ichiyen (apud NARASIMHAN; JORDACHE, 2000) mostrado na Figura 4.7. O processo consiste de três células de flotação (separadores), um misturador e de oito correntes, cada uma contendo dois minerais, cobre e zinco, em conjunto com o rejeito mineral. O fluxo da corrente 1 é tomado como uma unidade de massa (base de cálculo), enquanto que os outros fluxos das correntes não são medidos. As concentrações de mineral de todas as correntes, exceto a 8, são medidas. Estes valores são mostrados na primeira linha da Tabela 4.6.

Baseado nesta informação, as variáveis de fluxo global de componentes podem ser classificadas como.

Categoria I F_1, N_{11}, N_{12}

Categoria II $F_2, N_{21}, N_{22}, \ldots, F_7, N_{71}, N_{72}$

Categoria III F_8, N_{81}, N_{82}

As restrições impostas a este processo são os balanços dos fluxos globais e os balanços de fluxos de componentes em torno de cada unidade. As equações de normalização não são

[3]Cabe aqui um comentário quanto à terminologia empregada neste trabalho. Em todo este texto, o termo estimativa (*estimate*, na literatura em língua inglesa) é empregado como o resultado dos procedimentos de reconciliação de dados ou dos filtros, o quê entra em choque com o mesmo termo, estimativa (*guess* na literatura em língua inglesa), empregado como um valor necessário para iniciar um procedimento iterativo de qualquer natureza. Para não causar confusão no emprego dos termos, adota-se aqui **estimativa arbitrária** no sentido de um valor necessário para iniciar um procedimento iterativo.

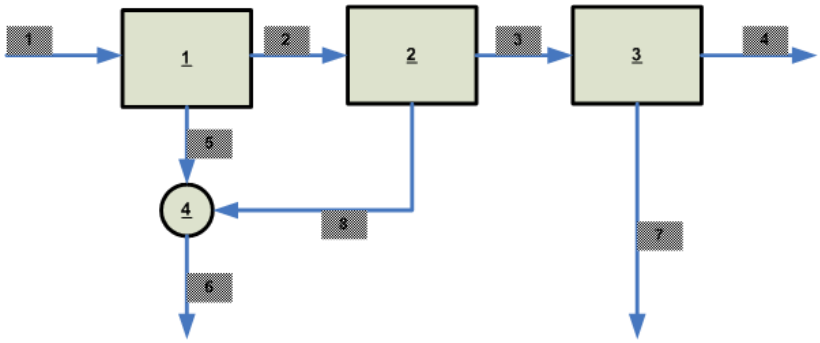


Figura 4.7: Processo de flotação de minério – adaptado de Narasimhan e Jordache (2000)

impostas porque foram eliminadas as variáveis de composição correspondentes ao rejeito não medido em cada corrente. Isto leva a um número reduzido de variáveis e restrições no problema de reconciliação de dados. Começa-se com a estimativa arbitrária inicial para os fluxos nas correntes 2 a 7, como mostrado na Tabela 4.6 e aplica-se o procedimento iterativo para se obter os valores reconciliados. A linha 2 da Tabela 4.6 mostra as estimativas reconciliadas dos fluxos e concentrações dos minerais obtidos pelo método de Crowe. Observa-se que a estimativa para a concentração de zinco na corrente 8 é negativa. Para efeito de comparação, as estimativas obtidas por técnica de programação não linear são também listadas na última linha da Tabela 4.6. Novamente, a estimativa para a concentração de zinco na corrente 8 é infactível. Isto indica a necessidade da imposição de limites restritivos do problema de reconciliação de dados. A máxima diferença entre as concentrações factíveis dos minerais nas duas soluções está em torno de 2-7 %. Como o método de Crowe usa uma função objetivo diferente do problema de reconciliação de dados padrão, as estimativas serão menos exatas do que aquelas obtidas pela abordagem da programação não linear.

4.3.3 Generalização das técnicas de reconciliação de dados bilinear

Na descrição da metodologia anterior, todos os valores medidos foram reconciliados. Nas aplicações industriais é mais comum manter algumas das variáveis medidas constantes durante a reconciliação. Uma forma simples de se conseguir isto é atribuir às medidas que devem ser mantidas constantes um fator ponderante muito alto (ou um desvio padrão muito baixo) na função objetivo. Isso vai forçar os ajustes feitos a estas variáveis serem suficientemente baixos para serem considerados nulos.

Tabela 4.6: Dados medidos e reconciliados de um processo de flotação de minerais (NARASIMHAN; JORDACHE, 2000)

Método	**Variável**	**Corrente**							
		1	2	3	4	5	6	7	8
Medido	F*	1	0,5	0,25	0,125	0,5	0,75	0,125	0,25
	y_{Cu}	1,928	0,450	0,128	0,090	19,88	21,43	0,513	35,36
	y_{Zn}	3,81	4,72	5,36	0,41	7,09	4,95	52,10	—
Crowe	F*	1	0,9229	0,9147	0,8324	0,0771	0,0853	0,0823	0,0081
	x_{Cu}	1,9451	0,4498	0,1285	0,0906	19,834	21,431	0,512	0,2976
	x_{Zn}	5,0356	4,8617	5,0461	0,4099	7,1167	4,9235	51,930	-15,91
PNL**	F*	1	0,9253	0,9164	0,8287	0,0747	0,0836	0,0877	0,0089
	x_{Cu}	1,9122	0,4509	0,1301	0,0899	20,00	21,44	0,5098	35,554
	x_{Zn}	4,2759	4,0584	5,3583	0,41	6,9694	4,95	52,116	-130,1

* Os valores iniciais dos fluxos globais das correntes de 1 a 8 estão nesta linha
** Programação não linear

O método de Crowe foi descrito a princípio para processos envolvendo misturadores e separadores. Se estiverem presentes no processo divisores de corrente e reatores, então este método tem que ser modificado adequadamente porque o tipo de equações impostas para estas unidades não está em conformidade com o daquelas para misturadores e separadores. Crowe (1986) descreveu as modificações necessárias para se levar os divisores de corrente em conta, de modo que as equações são formuladas usando frações de partição de corrente.

Como foi evidenciado anteriormente, o uso de variáveis de fração de partição leva a uma estrutura trilinear para o balanço de componentes. Para se poder usar o método de Crowe para o problema bilinear as variáveis de fração de divisão são estimadas em um *loop* interativo mais externo. Para cada estimativa arbitrária inicial das variáveis de fração de divisão, resulta um problema bilinear que pode ser resolvido usando o método de Crowe. Em geral, se faz necessária uma técnica de otimização sujeita a restrições para obter estimativas atualizadas das frações de partição a cada iteração, o quê diminui muito a eficiência do método (NARASIMHAN; JORDACHE, 2000).

4.3.4 Tratamento de fluxos de entalpia

Ainda que o método de Crowe tenha sido desenvolvido para resolver problemas de reconciliação de dados multi-componente, é possível estender a técnica para contabilizar balanços de entalpia e reconciliar variáveis de temperatura. De um modo geral a entalpia de uma corrente é uma função não linear da temperatura e da composição da corrente. Todavia, se a entalpia de uma corrente pode ser tida como uma função linear de temperatura e independente da composição, então a re-

conciliação simultânea dos balanços materiais e energéticos pode resultar em um problema bilinear. Mesmo a entalpia de uma corrente sendo uma função não linear da temperatura, mas independente da composição, os métodos aqui discutidos podem ser usados sem maiores modificações.

Um subsistema importante que satisfaz esta consideração é o do trem de pré-aquecimento de uma refinaria, no qual a entalpia de uma corrente de petróleo é relacionada à temperatura e a propriedades físicas tais como a densidade relativa e o ponto de ebulição normal da corrente. Para os propósitos deste trabalho, esta consideração será verdadeira e serão descritas as modificações necessárias para aplicar o método de Crowe para a reconciliação simultânea de balanços materiais e energéticos.

Do mesmo modo que foi feito anteriormente, são descritos a seguir os balanços energéticos para as diferentes unidades de processo.

Balanço de Entalpia no Misturador

$$\sum_{j=1}^{S} F_j H(T_j) - F_{\text{out}} H(T_{\text{out}}) = 0$$

onde $H(T)$ é a entalpia específica da corrente que é considerada como função exclusiva da temperatura.

Balanço de Entalpia do Divisor de Corrente

$$T_{\text{in}} - T_j = 0 \quad j = 1 \ldots S$$

Ou em termos de entalpia especificas das correntes

$$H(T_{\text{in}}) - H(T_j) = 0 \quad j = 1 \ldots S$$

Trocador de Calor Por definição, considera-se no trocador de calor os dados tanto do fluído frio quanto do fluído quente para serem reconciliados. Considera-se também que as correntes apresentem uma única fase. As equações nas quais o subscrito **h** significa **quente** e **c** significa **frio**, em referência às correntes desta unida demonstrada na Figura 4.8 são:

i. Balanços de Fluxo para os Fluidos Quentes e Frios

$$F_{\text{h in}} - F_{\text{h out}} = 0$$

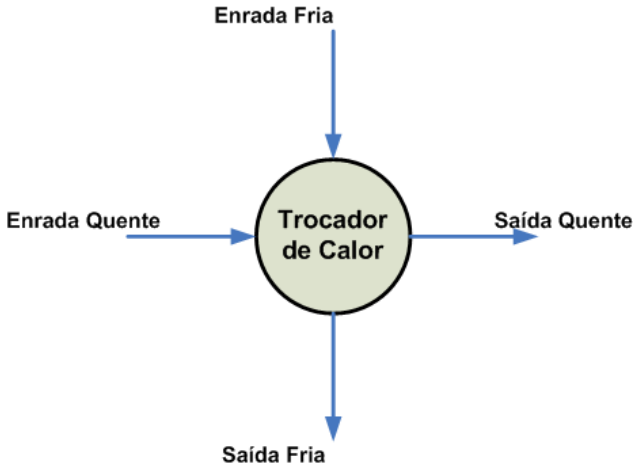


Figura 4.8: Trocador de calor – adaptado de Narasimhan e Jordache (2000)

$$F_{\mathrm{cin}} - F_{\mathrm{cout}} = 0$$

ii. Balanço de Entalpia

$$F_{\mathrm{hin}}H_{\mathrm{h}}(T_{\mathrm{hin}}) - F_{\mathrm{hout}}H_{\mathrm{h}}(T_{\mathrm{hout}}) + F_{\mathrm{cin}}H_{\mathrm{c}}(T_{\mathrm{cin}}) - F_{\mathrm{cout}}H_{\mathrm{c}}(T_{\mathrm{cout}}) = 0$$

iii. Balanços de fluxos de componentes

$$F_{\mathrm{hin}}x_{\mathrm{hin},k} - F_{\mathrm{hout}}x_{\mathrm{hout},k} = 0 \quad k = 1 \ldots C$$

$$F_{\mathrm{cin}}x_{\mathrm{cin},k} - F_{\mathrm{cout}}x_{\mathrm{cout},k} = 0 \quad k = 1 \ldots C$$

iv. Equações de normalização para as correntes de saída

$$\sum_{k=1}^{C} x_{\mathrm{hout},k} = 1$$

$$\sum_{k=1}^{C} x_{\mathrm{cout},k} = 1$$

Aquecedores ou resfriadores Um aquecedor ou resfriador é um trocador de calor para o qual somente os dados da corrente de processo é que são reconciliados, enquanto que os dados da corrente de utilidades são considerados como indisponíveis ou não importantes. As restrições para estas unidades são um subconjunto das restrições do trocador de calor. As equações para os balanços materiais e energéticos e as equações de normalização para as correntes de processo são as únicas escritas. O método de Crowe pode ser facilmente estendido para

incluir balanços entálpicos e variáveis de temperatura no problema de reconciliação como foi sugerido por Romagnoli e Sánchez (2000). As variáveis de entalpia específica podem ser tratadas de um modo similar às variáveis de composição. O fluxo entálpico de diferentes correntes pode ser classificado nas três categorias de um modo similar às variáveis de fluxo de componente. A função objetivo contém agora termos para os ajustes feitos aos fluxos entálpicos das correntes na categoria I e entalpias específicas em correntes de categoria II. A técnica de projeção de Crowe em dois estágios descrita anteriormente pode ser aplicada para obter também as entalpias específicas de todas as correntes. Se a entalpia específica é uma função não linear da temperatura, então a estimativa da temperatura para cada corrente pode ser recuperada a partir da entalpia específica. De um modo geral, isto pode requerer a solução de uma equação não linear unidimensional para cada corrente.

Uma desvantagem significativa deste método é ele não pode contabilizar limites simples sobre as variáveis e processo. Isso pode limitar bastante o uso desse método em aplicações industriais onde se requer a obtenção de estimativas factíveis para as variáveis de processo.

4.4 Conclusões

Um conjunto apropriado de soluções genéricas e *customizáveis* de reconciliação de dados tem que incluir soluções de reconciliação bilinear como alternativa aos métodos não lineares pois, quando o seu emprego é possível, pode-se efetivamente alcançar resultados com exatidão semelhante a dos métodos não lineares, mas com maior rapidez.

Dentre os pontos mais importantes levantados neste capítulo está a consideração de que um conjunto de restrições independentes tem que ser imposto para cada unidade de processo na formulação do problema de reconciliação de dados. Diferentes conjuntos podem ser impostos, mas tem que se levado em conta que alguns são mais convenientes do que outros. É importante incluir as restrições de normalização sobre as composições para garantir que as estimativas reconciliadas as satisfaçam.

É importante frisar também que os métodos especiais desenvolvidos para resolver os problemas de reconciliação de dados bilinear são eficientes mas não tratam todos os tipos de unidades nem lidam com restrições de factibilidade como limites sobre as variáveis. Além disso, as técnicas de reconciliação de dados não linear podem ser usadas para resolver problemas bilineares. Estas técnicas são menos eficientes mas não têm nenhuma outra limitação.

5 *Reconciliação de Dados em Estado Estacionário para Sistemas Não Lineares*

Neste capítulo são apresentadas as técnicas de reconciliação de dados não linear, iniciando com a formulação do problema para mostrar em seguida métodos que lidam com problemas sujeitos a restrições de igualdade e depois métodos de programação não linear que podem lidar com as restrições de desigualdade. Logo após, o problema da classificação de variáveis é tratado e, finalmente, é feita uma comparação de estratégias de otimização não linear para reconciliação de dados.

5.1 Formulação de problemas de reconciliação de dados não linear

As restrições de conservação em estado estacionário que são usadas para descrever a maioria dos processos químicos são não lineares por natureza. Quando o objetivo é somente a reconciliação dos balanços globais de fluxo então as técnicas de reconciliação linear, apresentadas no Capítulo 3 são suficientes. Quando o objetivo é a reconciliação de dados de fluxos globais e de componentes, onde a não linearidade é fruto do produto de fluxos pelas composições, é possível lançar mão às técnicas de reconciliação bilinear, descritas no Capítulo 4, mas se o objetivo é levar em consideração relações de equilíbrio termodinâmico e complexas correlações para as propriedades físicas e termodinâmicas, então as técnicas de reconciliação de dados não linear devem ser usadas (NARASIMHAN; JORDACHE, 2000). Para iniciar a discussão sobre as questões de reconciliação de dados não linear, será apresentado o exemplo da reconciliação de dados de um vaso *flash* isotérmico ilustrado na Figura 5.1, com uma corrente de alimentação composta de propano (1), n-butano (2) e n-pentano (3). As equações de balanço para o estado estacionário para esta unidade são as dadas

abaixo:

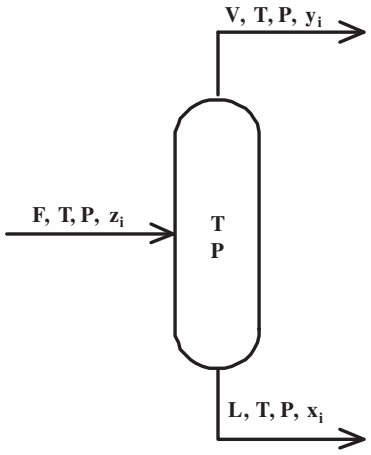


Figura 5.1: Vaso flash – adaptado de Narasimhan e Jordache (2000)

Balanço dos componentes:

$$Fz_i - Lx_i - Vy_i = 0 \quad i = 1,2,3 \tag{5.1}$$

Equações de normalização:

$$\sum_{i=1}^{3} x_i - 1 = 0 \tag{5.2}$$

$$\sum_{i=1}^{3} y_i - 1 = 0 \tag{5.3}$$

$$\sum_{i=1}^{3} z_i - 1 = 0 \tag{5.4}$$

Relações de equilíbrio:

$$y_i = \frac{P_i^{\text{sat}}(T)x_i}{P} \quad i = 1,2,3 \tag{5.5}$$

Por uma questão de simplicidade, a Lei de Raoult foi usada para descrever o equilíbrio. A pressão de saturação é obtida através da Equação de Antoine, a qual é dada por:

$$\ln P_i^{\text{sat}} = A_i + \frac{B_i}{(T + C_i)} \quad i = 1,2,3 \tag{5.6}$$

O problema de reconciliação não linear é reconciliar as medidas da taxa de fluxo, temperatura, pressão e composição da alimentação e das correntes de produto líquido e vapor de modo a satisfazer às restrições 5.1 até 5.6.

5.1.1 Formulação geral de problema

Tal como no caso linear, considera-se que os erros aleatórios nas medidas seguem uma distribuição normal com média zero e uma matriz covariância Ψ conhecida. O problema geral de reconciliação não linear pode ser formulado como o problema da minimização por mínimos quadrados, tal como segue:

$$\min_{\mathbf{x},\mathbf{u}} (\mathbf{y} - \mathbf{x})^{\mathrm{T}} \Psi^{-1} (\mathbf{y} - \mathbf{x}) \tag{5.7}$$

sujeito a:

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0} \tag{5.8}$$

$$\mathbf{g}(\mathbf{x}, \mathbf{u}) \leq \mathbf{0} \tag{5.9}$$

Onde:

f: vetor $m \times 1$ das restrições de igualdade

g: vetor $q \times 1$ das restrições de desigualdade

Ψ: matriz $n \times n$ de covariância

u: vetor $p \times 1$ das variáveis não medidas

x: vetor $n \times 1$ das variáveis medidas

y : vetor $n \times 1$ dos valores medidos das variáveis **x**

As restrições de igualdade definidas pela Equação 5.8 geralmente incluem todas as relações de

conservação de massa e energia, as restrições de equilíbrio termodinâmico e equações constitutivas similares às Equações 5.1 até 5.6 do exemplo do vaso *flash*. As restrições de desigualdade dadas na Equação 5.9 podem ser tão elementares quanto a atribuição de limites superiores e inferiores sobre as variáveis, ou complexas como restrições de factibilidade relacionada com a operação do equipamento.

Na formulação acima, assume-se implicitamente que as variáveis $\mathbf{x}$ são diretamente medidas, contudo, isto não impõe qualquer limitação. Se as medidas são funções lineares ou não lineares das variáveis (por ex.: pH é uma função da concentração de H^+) então sempre é possível definir uma nova variável de estado para o pH a qual é diretamente medida e a relação entre pH e a concentração de H^+ pode ser incluída como parte do conjunto de restrições de igualdade.

Serão consideradas aqui inicialmente técnicas para a solução de problemas de reconciliação de dados não linear contendo somente restrições de igualdade no modelo do processo. Duas técnicas de solução e suas variantes são discutidas a seguir:

5.2 Métodos para problemas com restrições de igualdade

A minimização da Equação 5.7 sujeita às restrições de igualdade da Equação 5.8 pode ser alcançada com uma técnica de otimização não linear de propósito geral, mas como a função objetivo é quadrática por natureza, toma-se partido disto usando-se técnicas mais eficientes desenvolvidas para esse tipo de problema. As estimativas obtidas pela solução deste problema de otimização são **estimadores de máxima verossimilhança**, mas deve-se notar, contudo, que estas estimativas podem ser enviesadas enquanto que no caso linear as estimativas eram não enviesadas.

5.2.1 Métodos usando multiplicadores de Lagrange

O problema de reconciliação de dados não linear sujeito à restrições de igualdade pode ser resolvido pelo método dos multiplicadores de Lagrange. O Lagrangiano para este problema é dado por:

$$L(\mathbf{x},\mathbf{u},\lambda) = (\mathbf{y}-\mathbf{x})^{\mathrm{T}}\Psi^{-1}(\mathbf{y}-\mathbf{x}) + 2\lambda^{\mathrm{T}}\mathbf{f}(\mathbf{x},\mathbf{u}) \tag{5.10}$$

A solução para o problema de reconciliação de dados pode ser obtida igualando as derivadas

parciais da Equação 5.10 com respeito às variáveis $\mathbf{x}$, $\mathbf{u}$ e λ a zero, que são as condições necessárias para uma solução ótima do problema definido pelas Equações 5.7 e 5.8, com a subseqüente solução das equações resultantes. Quais sejam:

$$\frac{\partial L}{\partial \mathbf{x}} = -\Psi^{-1}(\mathbf{y}-\mathbf{x}) + \mathbf{J_x}^{\mathrm{T}}\lambda = 0 \tag{5.11}$$

$$\frac{\partial L}{\partial \mathbf{u}} = \mathbf{J_u}^{\mathrm{T}}\lambda = 0 \tag{5.12}$$

$$\frac{\partial L}{\partial \lambda} = \mathbf{f}(\mathbf{x},\mathbf{u}) = 0 \tag{5.13}$$

onde:

$$\mathbf{J_x} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \tag{5.14}$$

$$\mathbf{J_u} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \tag{5.15}$$

são as matrizes jacobianas contendo as derivadas parciais das funções não lineares $\mathbf{f}$ com respeito a $\mathbf{x}$ e $\mathbf{u}$, respectivamente.

Como as restrições são não lineares, resolver para $\mathbf{x}$, $\mathbf{u}$ e λ envolve um procedimento numérico iterativo. O sistema das equações normais 5.11 até 5.13 pode ser resolvido por qualquer *solver* de equações simultâneas (DENNIS; SCHNABEL, 1983 apud NARASIMHAN; JORDACHE, 2000). Stephenson e Shewchuck (apud NARASIMHAN; JORDACHE, 2000) usaram um método Newton-Raphson iterativo baseado numa linearização quasi-newton[1] do modelo não linear. O algoritmo deles tira partido da esparsidade da matriz Jacobiana e da invariância das derivadas parciais dos termos lineares nas equações do modelo, o que torna o cálculo computacional mais eficiente para sistemas grandes. Serth *et al.* (apud NARASIMHAN; JORDACHE, 2000) reportaram uma abordagem semelhante, mas com um *solver* de equações não lineares diferente.

[1]Os métodos quasi-newton ou de métrica variável podem ser usados quando a a avaliação da matriz Hessiana é difícil ou demorada. Ao invés de se obter uma estimativa da matriz Hessiana em um ponto singular, estes métodos constroem gradualmente uma matriz Hessiana aproximada usando a informação do gradiente advinda de algumas ou mesmo de todas as iterações anteriores visitadas pelo algoritmo.

Madron (1992) sugeriu uma abordagem iterativa para a solução das equações normais 5.11 até 5.13 baseada em linearizações sucessivas. Sejam $\hat{\mathbf{x}}_k$ e $\hat{\mathbf{u}}_k$ representações das estimativas das variáveis obtidas no inicio da iteração k. Uma aproximação linear pode ser obtida para as restrições não lineares a partir da expansão em série de Taylor da função $\mathbf{f}(\mathbf{x},\mathbf{u})$ na Equação 5.8 com a retenção somente do termo constante e do termo de primeira ordem :

$$\mathbf{f}(\mathbf{x},\mathbf{u}) = \mathbf{f}(\hat{\mathbf{x}}_k,\hat{\mathbf{u}}_k) + \mathbf{J_x}^k(\mathbf{x}-\hat{\mathbf{x}}_k) + \mathbf{J_u}^k(\mathbf{u}-\hat{\mathbf{u}}_k) = 0 \tag{5.16}$$

onde as matrizes Jacobianas $\mathbf{J_x}^k$ e $\mathbf{J_u}^k$ são as definidas nas Equações 5.14 e 5.15 com o superescrito k indicando que são avaliadas nas estimativas $\hat{\mathbf{x}}_k$ e $\hat{\mathbf{u}}_k$. As matrizes Jacobianas que aparecem nas Equações 5.11 e 5.12 são substituídas pelos seus valores estimados na iteração k. O conjunto de equações semelhante é, desta forma, linear. No procedimento de Madron, estas equações lineares são desacopladas, eliminando o vetor $\mathbf{x}$ das Equações 5.12 e 5.13, usando-se a Equação 5.11.

A partir da Equação 5.11, tem-se:

$$\mathbf{x} = \mathbf{y} - \Psi(\mathbf{J}_x^k)^{\mathrm{T}}\lambda. \tag{5.17}$$

Usando-se a Equação 5.16 e 5.17 nas Equações 5.12 e 5.13 e rearranjando, obtém-se as seguintes equações lineares envolvendo $\mathbf{u}$ e λ

$$\begin{bmatrix} \mathbf{J_x}^k\Psi(\mathbf{J_x}^k)^{\mathrm{T}} & \mathbf{J_u}^k \\ (\mathbf{J_u}^k)^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \lambda \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} -\mathbf{f}(\hat{\mathbf{x}}_k,\hat{\mathbf{u}}_k) + \mathbf{J_x}^k\hat{\mathbf{x}}_k + \mathbf{J_u}^k\hat{\mathbf{u}}_k - \mathbf{J_x}^k\mathbf{y} \\ \mathbf{0} \end{bmatrix} \tag{5.18}$$

A Equação 5.18 pode ser resolvida para obter-se as novas estimativas para $\mathbf{u}$ e λ. As estimativas para λ são usadas na Equação 5.17 para obter as novas estimativas para $\mathbf{x}$. Este procedimento é repetido usando-se as estimativas obtidas em uma iteração como estimativas arbitrárias iniciais para a próxima iteração (ver nota de rodapé na página 100). Uma desvantagem de todos estes métodos é que a inclusão dos multiplicadores de Lagrange λ na solução aumenta o tamanho do problema e isto reflete no tempo computacional requerido.

Para reduzir o tamanho do problema, Madron (1992) propôs um processo de eliminação Gauss-Jordan das matrizes de restrições lineares/linearizadas ($\mathbf{J_x}|\mathbf{J_u}$ para o caso não linear). A estrutura da matriz resultante oferece informação útil para classificação de variáveis.

5.2.2 Método da reconciliação de dados linear sucessiva

Uma maneira mais simples de lidar com a reconciliação de dados não linear é resolver sucessivamente uma série de problemas de reconciliação de dados linear através da linearização das restrições não lineares. Uma aproximação linear às restrições não lineares é obtida na Equação 5.16. Obtém-se assim um problema de reconciliação de dados linear com a minimização de 5.7 sujeita às restrições de igualdade lineares da Equação 5.16 que podem ser resolvidas usando a técnica demonstrada no Capítulo 3.

O método desenvolvido para restrições lineares é estendido aos problemas sujeitos a restrições não lineares considerando-se que as restrições não lineares $\varphi(\mathbf{x},\mathbf{u}) = \mathbf{0}$ podem ser linearizadas por uma expansão em série de Taylor em torno de uma estimativa da solução (x_i, u_i). Em geral, os valores das medidas são usados como estimativas arbitrárias iniciais. Assim, o seguinte sistema de equações lineares é obtido:

$$\mathbf{J_x x} + \mathbf{J_u u} = \mathbf{c}_1 \tag{5.19}$$

onde:

$$\mathbf{J_x} = \left.\frac{\partial \varphi}{\partial \mathbf{x}}\right|_{x_i,u_i}, \mathbf{J_u} = \left.\frac{\partial \varphi}{\partial \mathbf{u}}\right|_{x_i,u_i} \tag{5.20}$$

$$\mathbf{c}_1 = \mathbf{J_x} x_i + \mathbf{J_u} u_i - \varphi(x_i, u_i) \tag{5.21}$$

As variáveis não medidas são então eliminadas usando, por exemplo, uma fatoração ortogonal como mostrado na Seção 3.6. Uma vez que o subconjunto de equações contendo somente variáveis medidas tenha sido identificado, o problema colocado por Swartz (1989) é resolvido.

$$\min_x = (\mathbf{y} - \mathbf{x})^{\mathrm{T}} \psi^{-1} (\mathbf{y} - \mathbf{x}) \tag{5.22}$$

sujeito a:

$$\mathbf{G_x x} = \mathbf{b} \tag{5.23}$$

onde:

$$\mathbf{G_x} = \mathbf{Q}_{u2}^{\mathrm{T}}\mathbf{J_x},\ \mathbf{b} = \mathbf{Q}_{u_2}^{\mathrm{T}}\mathbf{c}_1 \tag{5.24}$$

E sua solução pode ser colocada na forma:

$$\hat{\mathbf{x}} = \mathbf{y} - \psi\mathbf{G_x}^{\mathrm{T}}(\mathbf{G_x}\psi\mathbf{G_x}^{\mathrm{T}})^{-1}(\mathbf{G_x}\mathbf{y} - \mathbf{b}) \tag{5.25}$$

$$\mathbf{u}_{r_u} = \mathbf{R}_{u_1}^{-1}\mathbf{Q}_{u_1}^{\mathrm{T}}\mathbf{c}_1 - \mathbf{R}_{u_1}^{-1}\mathbf{Q}_{u_1}^{\mathrm{T}}\mathbf{J_x}\hat{\mathbf{x}} - \mathbf{R}_{u_1}^{-1}\mathbf{R}_{u_2}\mathbf{u}_{n-r_u} \tag{5.26}$$

Esta solução é o ponto ótimo para as restrições lineares. Uma série de iterações são realizadas pela linearização das restrições em torno da iteração anterior até que seja obtida uma solução que satisfaça às restrições não lineares.

O método das linearizações sucessivas tem como vantagem sua relativa simplicidade e rapidez de cálculo. Além disso, ele pode ser modificado de modo a escolher um tamanho de passo que minimize uma **função de penalidade**[2] pré-determinada. O tamanho do passo é escolhido pelo método da bisseção (PAI; FISHER, 1988). Entretanto, os limites sobre as variáveis não podem ser tratados e ele pode falhar na convergência para um determinado mínimo e ainda apresentar oscilações se houver múltiplos mínimos.

Britt e Luecke (apud NARASIMHAN; JORDACHE, 2000) propuseram um procedimento de solução alternativo para o problema linearizado. A solução deles para as estimativas que devem ser usadas na próxima iteração é dada por:

$$\hat{\mathbf{u}}_{k+1} = \hat{\mathbf{u}}_k \left[(\mathbf{J_u}^k)^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{J_u}^k \right]^{-1} (\mathbf{J_u}^k)^{\mathrm{T}}\mathbf{R}^{-1} \left\{ \mathbf{f}(\hat{\mathbf{x}}_k, \hat{\mathbf{u}}_k) + \mathbf{J_x}^k(\mathbf{y} - \hat{\mathbf{x}}_k) \right\} \tag{5.27}$$

$$\hat{\mathbf{x}}_{k+1} = \mathbf{y} - \Psi(\mathbf{J_x}^k)^{\mathrm{T}}\mathbf{R}^{-1} \left\{ \mathbf{f}(\hat{\mathbf{x}}_k, \hat{\mathbf{u}}_k) + \mathbf{J_x}^k(\mathbf{y} - \hat{\mathbf{x}}_k) + \mathbf{J_u}^k(\hat{\mathbf{u}}_{k+1} - \hat{\mathbf{u}}_k) \right\} \tag{5.28}$$

onde:

[2]Um método de penalidade substitui um problema de otimização com restrições através de uma série de problemas de otimização sem restrições com a função objetivo modificada com um termo aditivo de penalidade. O termo cresce com a proximidade de violação das restrições e é zero nas regiões onde as restrições não são violadas. O termo da penalidade é normalmente o produto de uma função penalidade e um coeficiente positivo de penalidade.

$$\mathbf{R} = \mathbf{J_x}^k \Psi (\mathbf{J_x}^k)^{\mathrm{T}} \tag{5.29}$$

As Equações 5.27 e 5.28 foram derivadas por Britt e Luecke (apud NARASIMHAN; JORDACHE, 2000) para a estimativa de parâmetros em regressão não linear e adaptadas por Knepper e Gorman (apud NARASIMHAN; JORDACHE, 2000) para reconciliação de dados não linear e também usadas por MacDonald e Howat (apud NARASIMHAN; JORDACHE, 2000). O algoritmo requer estimativas arbitrárias iniciais para todas as variáveis contidas nos vetores $\mathbf{x}$ e $\mathbf{u}$. Os valores medidos $\mathbf{y}$ podem ser usados para inicializar as variáveis $\mathbf{x}$. Britt e Luecke (apud NARASIMHAN; JORDACHE, 2000) projetaram também um algoritmo simplificado que pode ser usado para inicializar os parâmetros não medidos $\mathbf{u}$. A cada iteração, a função $\mathbf{f}(\mathbf{x},\mathbf{u})$ e as matrizes Jacobianas $\mathbf{J_x}$ e $\mathbf{J_u}$ são reavaliadas com as novas estimativas. As iterações continuam até que $||\mathbf{u}_{k+1} - \mathbf{u}_k||$ e $||\mathbf{x}_{k+1} - \mathbf{x}_k||$ satisfaçam a um critério de tolerância. Se a convergência é alcançada, a solução pode não ser um mínimo global. Esta dificuldade é comum à maioria dos problemas de estimativas por mínimos quadrados não linear.

Uma variante do algoritmo acima foi sugerida por Knepper e Gorman (apud NARASIMHAN; JORDACHE, 2000) no sentido de reduzir o tempo computacional e consiste em manter as matrizes Jacobianas constantes nas estimativas iniciais e somente calculá-las novamente depois que as restrições são satisfeitas (abordagem de direção constante). Esta abordagem contudo é caracterizada por uma convergência lenta.

Uma outra variação proposta por MacDonald e Howat (apud NARASIMHAN; JORDACHE, 2000) é um procedimento desacoplado no qual as estimativas para $\mathbf{u}$ são mantidas constantes e a Equação 5.28 é repetidamente usada até que as estimativas para $\mathbf{x}$ convirjam. A Equação 5.27 é agora usada para obter as novas estimativas para $\mathbf{u}$ e o procedimento é repetido até que todas as estimativas convirjam. MacDonald e Howat (apud NARASIMHAN; JORDACHE, 2000) demonstram através da aplicação a um vaso *flash* de não equilíbrio que o algoritmo acoplado oferece estimativas marginalmente mais exatas às custas de um tempo computacional maior. O procedimento desacoplado pode ser um esquema computacional útil quando as equações não lineares são implícitas nos parâmetros.

O método de Britt e Luecke (apud NARASIMHAN; JORDACHE, 2000) e suas variantes descritas acima tem algumas limitações. As Equações 5.27 e 5.28 envolvem a inversa do produto de duas matrizes, $\mathbf{R}$ e $(\mathbf{J_u}^k)^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{J_u}^k)$. Para que a inversa dos produtos das matrizes possa existir, as seguintes condições devem ser satisfeitas:

i. A matriz $\mathbf{J_x}^k$ deve ter o posto da linha completo;

ii. A matriz $\mathbf{J_u}^k$ deve ter o posto de coluna completo.

A segunda das condições acima implica que todas as variáveis não medidas devem ser observáveis. Isto é idêntico à condição observada no caso linear no Capítulo 3, onde foi mostrado que para todas variáveis não medidas serem observáveis, as colunas das matrizes da restrições correspondentes a estas variáveis devem ser linearmente independentes (posto completo). Mesmo que essa condição seja satisfeita, a primeira das condições pode não ser satisfeita em alguns processos, dependendo de quais das variáveis são medidas. Desta forma os métodos descritos acima podem não ser aplicados em geral para todos os processos.

Uma abordagem que pode ser usada de um modo geral é baseada na técnica da projeção de matrizes de Crowe (CROWE *et al.*, 1983) para resolver o problema de reconciliação de dados linear em cada iteração. Os passos básicos envolvidos nesta abordagem são os seguintes:

Passo 1 Iniciar com os valores medidos como sendo estimativas iniciais arbitrárias para as variáveis **x** e com estimativas iniciais arbitrárias para **u** providas pelo usuário;

Passo 2 Avaliar as matrizes Jacobianas das restrições não lineares com respeito às variáveis **x** e **u** com as estimativas correntes;

Passo 3 Calcular a matriz de projeção $\mathbf{P}^k$ tal que esta satisfaça

$$\mathbf{P}^k \mathbf{J_u}^k = \mathbf{0} \tag{5.30}$$

A matriz de projeção pode ser obtida também usando-se uma fatoração QR da matriz $\mathbf{J_u}^k$, tal como descrito no Capítulo 3;

Passo 4 Calcular as novas estimativas para **x** usando

$$\hat{\mathbf{x}} = \mathbf{y} - \Psi(\mathbf{P}^k \mathbf{J_x}^k)^{\mathrm{T}} \left[\mathbf{P}^k \mathbf{J_x}^k \Psi(\mathbf{P}^k \mathbf{J_x}^k)^{\mathrm{T}}) \right]^{-1} \mathbf{P}^k \left[\mathbf{J_x}^k \mathbf{y} + \mathbf{f}(\hat{\mathbf{x}}_k, \hat{\mathbf{u}}_k) - \mathbf{J_x}^k \hat{\mathbf{x}}_k \right] \tag{5.31}$$

Passo 5 Calcular as novas estimativas para **u** através da Equação 3.90 utilizando a fatoração QR da matriz $\mathbf{J_u}^k$;

Passo 6 Parar se as novas estimativas não são significativamente diferentes daquelas obtidas na iteração anterior. Caso contrário, usando estas novas estimativas, repetir o procedimento começando pelo **Passo 2**.

Pai e Fisher (1988) usaram um procedimento semelhante ao algoritmo descrito acima. As modificações adicionais em seu algoritmo são:

i. Um procedimento de atualização de Broyden[3] para a Jacobiana ao invés de seu re-cálculo em cada iteração com o objetivo de reduzir o esforço computacional envolvido;

ii. Um procedimento de busca em linha após o **Passo 5**, baseado em um método de função penalidade (ver nota à página 114) para calcular as estimativas a serem usadas para iniciar a próxima iteração. A função penalidade $||\mathbf{f}(\mathbf{x},\mathbf{u})|| + \alpha(\mathbf{y}-\mathbf{x})^{\mathrm{T}}\boldsymbol{\Psi}^{-1}(\mathbf{y}-\mathbf{x})$ foi usada, onde α é um número arbitrário em $0 \leq \alpha \leq 1$.

A modificação do ítem **i** pode melhorar a eficiência computacional do algoritmo e foi demonstrado para pequenos problemas, contudo a modificação em **ii** é de utilidade questionável devido ao fato da função objetivo da reconciliação de dados ser quadrática e a solução para as estimativas obtidas nos **Passos 4** e **5** serem ótimas, ainda que não satisfaçam às restrições não lineares. Um procedimento de busca em linha para a modificação destas estimativas pode facilitar a factibilidade com respeito às restrições não lineares mas com o sacrifício da otimalidade. Isto pode não levar a uma redução global do esforço computacional. A escolha correta para o parâmetro α é um fator preponderantemente subjetivo para a função objetivo dos mínimos quadrados e é de difícil determinação. Pai e Fisher (1988) usaram $\alpha = 0,1$ (NARASIMHAN; JORDACHE, 2000).

Swartz (1989) recomendou o uso de uma fatoração QR para separação das estimativas das variáveis medidas das não medidas a cada iteração. Se o problema é altamente não linear e o tamanho do problema é grande, isto pode se tornar computacionalmente ineficiente. Ramamurthi e Bequette (1990) reportaram um aumento no tempo computacional com o nível de ruído (magnitude dos erros grosseiros) entre as medidas. Devido ao processo de linearização sucessiva, mais iterações são geralmente necessárias para a convergência de um problema com erros de grande magnitude entre os dados. Este procedimento é bastante utilizado, mas para problemas de reconciliação de dados envolvendo balanços globais de massa e energia.

No Capítulo 4, os valores reconciliados para a coluna de destilação binária reportados na Tabela 4.6 foram obtidos usando o algoritmo de reconciliação com linearizações sucessivas descrito acima. Pode-se notar, a partir dos resultados da Tabela 4.6 que a estimativa reconciliada para a

[3] O procedimento de Broyden é uma extensão do método da secante para encontrar raízes de dimensões maiores (BROYDEN, 1965).

concentração de zinco na **Corrente 8** tem um alto valor negativo, o que é um resultado espúrio, e assim este método não pode garantir o resultado com valores factíveis para as estimativas em todos os casos.

5.3 Métodos de programação não linear (*NLP - nonlinear programming*)

A maior limitação dos métodos descritos na seção anterior é a sua incapacidade de lidar com restrições de desigualdade. Em várias situações, em especial na presença de erros grosseiros significativos, a reconciliação de dados padrão pode causar o espalhamento dos erros grosseiros por todas as estimativas. Se não houver redundância suficiente, as estimativas para as variáveis que têm valores pequenos são fortemente prejudicadas por erros grosseiros. São obtidas, em alguns casos, estimativas infactíveis tais como valores negativos para taxa de fluxo ou composições. Para enfrentar este problema se faz necessário impor limites ou contornos sobre as variáveis. Estas restrições de desigualdade sobre as variáveis medidas e não medidas tomam a forma (NARASIMHAN; JORDACHE, 2000):

$$\mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max} \tag{5.32}$$

$$\mathbf{u}_{\min} \leq \mathbf{u} \leq \mathbf{u}_{\max} \tag{5.33}$$

Mais raramente, outros tipos de restrições de factibilidade têm que ser impostas. Por exemplo, quando a reconciliação de dados é aplicado a uma rede de trocadores de calor para a reconciliação de fluxos e temperaturas, é possível que as estimativas violem a termodinâmica, como por exemplo em uma estimativa da temperatura de uma corrente quente sendo inferior à estimativa da corrente fria correspondente. Para que isto não ocorra, uma restrição de desigualdade deve ser imposta, de modo que force a temperatura da corrente quente a ser maior do que a temperatura da corrente fria correspondente em ambas as extremidades de cada trocador de calor da rede. Este tipo de restrição de factibilidade pode ser disposto na forma da Equação 5.9.

A solução do problema de reconciliação de dados quando restrições como a Equação 5.2 ou 5.32 e 5.33 são impostas pode ser obtida com o uso de técnicas de programação não linear de

propósito geral[4]. São discutidas a seguir duas técnicas mais comuns de programação não linear no tocante ao seu uso para a solução de problemas de reconciliação de dados não linear.

5.3.1 Programação Quadrática Sucessiva (*SQP – successive quadratic programming*)

A técnica da programação quadrática sucessiva (HAN; POWELL; CHEN; STADTHERR, 1977a, 1977, 1984 apud NARASIMHAN; JORDACHE, 2000) resolve o problema de otimização não linear através da solução sucessiva de uma série de problemas de programação quadrática. A cada iteração, uma aproximação do problema geral de otimização é obtida por uma aproximação quadrática da função objetivo e uma aproximação linear das restrições, ambas usando uma expansão em série de Taylor em torno das estimativa correntes. No caso de problemas de reconciliação de dados definido pelas Equações 5.7 até 5.9, a função objetivo já é quadrática e somente as restrições têm que ser linearizadas. O problema quadrático resultante na iteração $k+1$ é formulado como:

$$\min_{\mathbf{s}} (\nabla\phi)^{\mathrm{T}}\mathbf{s} + \mathbf{s}^{\mathrm{T}}\mathbf{B}\mathbf{s} \tag{5.34}$$

sujeito a:

$$f_i(\hat{\mathbf{z}}_k) + [\nabla f_i(\hat{\mathbf{z}}_k)]^{\mathrm{T}}\mathbf{s} = \mathbf{0} \quad i = 1,\ldots,m \tag{5.35}$$

$$g_j(\hat{\mathbf{z}}_k) + \left[\nabla g_j(\hat{\mathbf{z}}_k)\right]^{\mathrm{T}}\mathbf{s} \leq \mathbf{0} \quad j = 1,\ldots,q \tag{5.36}$$

onde $\mathbf{z}$ é o vetor das variáveis originais $(\mathbf{x},\mathbf{u})$, $\mathbf{s} = \mathbf{z} - \hat{\mathbf{z}}_k$ é a direção de busca para iteração $k+1$; $\nabla\phi$, ∇f_i, ∇g_j são, respectivamente, os gradientes (derivadas com respeito às variáveis $\mathbf{z}$) da função objetivo, da restrição de igualdade i e da restrição de desigualdade j, todas avaliadas na estimativa corrente $\hat{\mathbf{z}}_k$. $\mathbf{B}$ é a Hessiana (matriz com as derivadas de segunda ordem da função objetivo com respeito às variáveis $\mathbf{z}$) avaliada na estimativa corrente $\hat{\mathbf{z}}_k$.

Na formulação do programa quadrático, todas as variáveis são incluídas na função objetivo. Quando comparada com a função objetivo da reconciliação de dados na Equação 5.7, pode-se dizer que:

[4]Uma descrição mais profunda destes métodos pode ser encontrada em Edgar e Himmelblau (1988)

$$\nabla\phi = \begin{bmatrix} -2\Psi^{-1}(\mathbf{y} - \hat{\mathbf{x}}_k) \\ \mathbf{0} \end{bmatrix} \tag{5.37}$$

$$\mathbf{B} = 2\begin{bmatrix} \Psi^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{5.38}$$

Nota-se que a matriz Hessiana é constante e singular se houver variáveis não medidas no processo.

A solução para o programa quadrático dá a direção de busca para obtenção das estimativas. Uma busca unidimensional é realizada na direção $\mathbf{s}_k$ a cada iteração k, de modo que o novo valor de $\hat{\mathbf{z}}$ na próxima iteração seja:

$$\hat{\mathbf{z}}_{k+1} = \hat{\mathbf{z}}_k + \alpha_k \mathbf{s}_k \tag{5.39}$$

onde α é parâmetro para o tamanho do passo e está entre 0 e 1. O tamanho do passo é obtido pela minimização de uma função de penalidade (similar ao Lagrangiano). O procedimento é repetido usando as novas estimativas até que seja atingida a convergência.

Há uma série de questões de particular interesse na solução de problemas de reconciliação de dados usando SQP. De um modo geral, no SQP a matriz Hessiana exata (ou sua inversa) não é calculada a cada iteração devido ao peso computacional que isto representaria e, ao invés disso, uma inversa aproximada da matriz Hessiana (ou sua raiz quadrada) é obtida por uma técnica de atualização simétrica de Broyden. No caso da reconciliação de dados, a Equação 5.38 mostra que a matriz Hessiana é constante e portanto não há necessidade de atualizá-la. Em segundo lugar, a Equação 5.38 também mostra que a matriz Hessiana é semidefinida positivamente (ver nota de rodapé à Página 54) se houver a presença de variáveis não medidas e portanto o *solver* QP (*quadratic programming* – programação quadrática) que é usado deve ser capaz de tratar com matrizes Hessianas semidefinidas positivamente. Em terceiro lugar, a solução obtida usando a QP é usada como uma direção de busca e o comprimento de um passo ótimo nesta direção é obtido através de minimização de uma função penalidade.

Se a função objetivo contém termos não lineares de ordem maiores que a quadrática, então a minimização desta linha dá estimativas que são mais próximas ao ótimo e menos infactíveis. No caso de reconciliação de dados, contudo, como a função objetivo é quadrática, o comprimento

do passo de uma unidade dá as estimativas ótimas que satisfazem a aproximação linearizada das restrições. Neste caso, a minimização desta linha vai melhorar a factibilidade com respeito às restrições não lineares através do sacrifício da otimalidade. Assim, mesmo que uma técnica SQP de propósito geral possa ser usada, a exploração de características peculiares discutas acima torna possível o desenvolvimento de uma técnica SQP mais eficiente e feita sob medida para a solução de problemas de reconciliação de dados não lineares (NARASIMHAN; JORDACHE, 2000).

EXEMPLO 5.1 Para ilustrar a necessidade da inclusão de limites na reconciliação de dados para obtenção de estimativas factíveis usando o processo de flotação mineral, considerado no Exemplo 4.4, no qual os dados medidos são listados na primeira linha da Tabela 4.6. A última linha nesta tabela dá também as estimativas reconciliadas obtidas usando um procedimento de solução linear sucessivo de reconciliação de dados em conjunto com a projeção de Crowe (método de Pai e Fisher, discutido na seção anterior). Como este método não pode lidar com limites, estes não foram impostos. Estas estimativas reconciliadas mostram que é obtido um valor negativo absurdamente grande para a estimativa da concentração do zinco na **Corrente 8**. O mesmo problema foi resolvido usando o SQP com a imposição de um limite inferior de 0,1% e um limite superior a 100% em todas as concentrações e um limite inferior de 0 e um superior de 1 para todos os fluxos. As estimativas reconciliadas obtidas são mostradas na Tabela 5.1. Assim, mostra-se que é possível obter estimativas factíveis através da inclusão de limites no problema de reconciliação de dado não linear. As concentrações de *Cu* na **Corrente 4** e *Zn* na **Corrente 8** estão no limite inferior na solução reconciliada. Comparando estes resultados com os da Tabela 4.6, nota-se que as estimativas reconciliadas para todas outras variáveis não são significativamente diferentes.

Tabela 5.1: Dados reconciliados com SQP para o processo de flotação mineral (NARASIMHAN; JORDACHE, 2000)

Método	**Variável**	**Corrente**							
		1	2	3	4	5	6	7	8
	F	1	0,9267	0,9157	0,8448	0,0733	0,0843	0,0709	0,0110
SQP	$\hat{\mathbf{x}}_{Cu}$%	1,9042	0,4526	0,1316	0,1	20,267	21,164	0,5080	27,126
	$\hat{\mathbf{x}}_{Zn}$%	4,5580	4,3728	4,4242	0,4101	6,9002	4,95	52,266	0,1

5.3.2 Gradiente Reduzido Generalizado (GRG)

A técnica de otimização **GRG** resolve um problema de otimização não linear essencialmente através da solução de uma série de problemas de programação linear. A cada iteração, uma aproximação por um programa linear (*linear programming* – LP) é obtida pela linearização da função objetivo e das restrições. O subproblema LP é formulado como nas Eqs. 5.34 a 5.36 com a diferença que o segundo termo (quadrático) na Eq. 5.34 não está presente. A técnica GRG difere do SQP por um aspecto: a cada iteração, o método GRG requer estimativas que satisfaçam às restrições não lineares, enquanto que no SQP isto não ocorre necessariamente. Ao invés da minimização de uma linha, como no SQP, o subproblema LP é ajustado usando um procedimento iterativo como Newton-Raphson no sentido de obter estimativas que satisfaçam às restrições não lineares.

O subproblema LP é resolvido usando um algoritmo padrão pelo particionamento das variáveis entre *dependentes* (básicas) e *independentes* (não básicas). As variáveis dependentes são implicitamente determinadas pelas variáveis independentes, tornando a função objetivo uma função exclusiva de variáveis não básicas. As variáveis não básicas são dividas ainda em variáveis superbásicas, as quais estão entre seus limites e as variáveis não básicas, as quais estão sobre seus limites. Uma busca unidimensional é realizada na direção do gradiente das variáveis superbásicas (portanto o termo "gradiente reduzido"). Vários algoritmos comerciais para o GRG diferem nos métodos que usam para fazer a busca e recuperar um ponto factível com respeito às restrições não lineares (ABADIE; LASDON; WAREN, 1978, apud NARASIMHAN; JORDACHE, 2000).

Uma questão interessante está relacionada com o modo no qual as variáveis não medidas são tratadas no SQP e no GRG. Em ambas abordagens não é feita nenhuma distinção entre as variáveis medidas e não medidas. Vale então a pena considerar qual seria o proveito do uso do método de projeção de matrizes de Crowe para desacoplar as variáveis medidas das não medidas. O problema é que a técnica de projeção de matrizes de Crowe não pode ser utilizada para eliminação de variáveis não medidas se houver a imposição de limites sobre estas variáveis porque as estimativas para as variáveis não medidas obtidas com esta técnica podem violar os limites impostos. Além disso, tanto o SQP quanto o GRG empregam uma técnica de projeção para eliminar não somente variáveis não medidas mas um conjunto de variáveis dependentes (igual ao número de restrições de igualdade que é geralmente maior do que o número de variáveis não medidas).

5.4 Classificação de variáveis para a reconciliação de dados não linear

No Capítulo 3 foram vistos os métodos de classificação de variáveis quanto à observabilidade e redundância para a reconciliação de dados linear. Alguns destes métodos (SWARTZ; CROWE; KRETSOVALIS; MAH; MEYER *et al.*; ROMAGNOLI; STEPHANOPOULOS; SÁNCHEZ *et al.*; SÁNCHEZ; ROMAGNOLI, 1989, 1989a, 1988a, 1988b, 1993, 1980a, 1980b, 1992, 1996 apud NARASIMHAN; JORDACHE, 2000) são também aplicáveis à reconciliação de dados não linear (principalmente para problemas com restrições bilineares).

Para problemas com níveis maiores de não linearidade, um procedimento comum é a realização primeiro de uma linearização do modelo e a aplicação de métodos de classificação de variáveis para modelos lineares. Albuquerque e Biegler (1996) descrevem um procedimento como este. Ainda que projetado para classificação de variáveis em problemas de reconciliação de dados em sistemas dinâmicos, o método deles pode ser usado para reconciliação de dados não linear em estado estacionário. Nesta abordagem uma **decomposição LU** foi usada para construir uma matriz de projeção com o objetivo de separar as variáveis não medidas das medidas. As regras de classificação de variáveis são bem semelhantes às descritas por Swartz (1989) com o algoritmo de decomposição QR. Como a decomposição LU é parte de alguns métodos de solução de programação não linear para a reconciliação de dados, é descrito aqui o algoritmo de Albuquerque e Biegler para a classificação de variáveis.

A Equação 5.16 descrevendo um modelo linearizado, pode também ser escrita de forma abreviada como:

$$\mathbf{J_x x} + \mathbf{J_u u} = \mathbf{c} \tag{5.40}$$

onde foram agrupados todos os termos constantes vindos da linearização do modelo em um vetor global de constantes $\mathbf{c}$. Foi também omitido o subscrito k com base na consideração de que a linearização é realizada em torno do ponto de solução final. Para eliminar as variáveis não medidas $\mathbf{u}$, deve ser construída uma matriz de projeção $\mathbf{P}$, tal que $\mathbf{PJ_u} = \mathbf{0}$. Seja uma decomposição LU de tal modo que:

$$\mathbf{E}\mathbf{J}_u\Pi = \mathbf{L}\begin{bmatrix}\mathbf{U}_1 & \mathbf{U}_2\\ \mathbf{0} & \mathbf{0}\end{bmatrix} \tag{5.41}$$

onde $\mathbf{E}$ e Π são matrizes de permutação, $\mathbf{L}$ é uma matriz triangular inferior, $\mathbf{U}_1$ é uma matriz triangular superior de posto r (o posto das colunas da matriz $\mathbf{J_u}$) e $\mathbf{U}_2$ é uma matriz retangular. Se $\mathbf{J_u}$ tem o posto das linhas completo, as linhas nulas na matriz triangular superior não devem existir. Além disso, se $\mathbf{J_u}$ tem o posto das colunas completo, $\mathbf{U}_2$ também não deve existir e por conseguinte não deve haver nenhuma variável não observável. Uma matriz de projeção $\mathbf{P}$ para a matriz $\mathbf{J_u}$ pode ser criada como:

$$\mathbf{P} = \begin{bmatrix}\mathbf{0} & | & \mathbf{I}\end{bmatrix}\mathbf{L}^{-1} \tag{5.42}$$

De uma maneira semelhante às regras derivadas por Swartz (1989) ou Crowe (1989a), a observabilidade requer uma linha nula em $\mathbf{U}_1^{-1}\mathbf{U}_2$ e as variáveis medidas redundantes terão colunas nulas na matriz $\mathbf{P}\mathbf{J_x}$. Todas as outras variáveis medidas são, assim, declaradas redundantes. Observa-se que estes métodos dependem dos valores das medidas e podem dar lugar a uma classificação incorreta devido a problemas numéricos.

Várias questões concernentes à classificação de variáveis em conexão com o algoritmo de solução do SQP são importantes e dignas de nota. Em primeiro lugar o SQP precisa de estimativas iniciais arbitrárias para todas as variáveis (medidas e não medidas). Se houver variáveis não observáveis, o SQP será ainda capaz de fornecer estimativas para todas as variáveis. Ele usa as estimativas arbitrárias iniciais das variáveis não medidas (apenas em número suficiente para tornar todas as outras variáveis não medidas observáveis) como "especificações" para as não observáveis e realiza então a reconciliação de dados.

Quais das variáveis não medidas serão escolhidas é algo implícito ao método numérico (basicamente quando a escolha de variáveis independentes e dependentes é feita baseada nas colunas da matriz de restrições linearizada a cada iteração). A única forma de saber qual foi a variável escolhida é observando os resultados finais. Se a estimativa reconciliada de uma variável não medida é igual a estimativa inicial fornecida pelo usuário, então a variável não medida é não observável (podem haver outras variáveis não observáveis que eventualmente tenham sido recalculadas, o que torna sua identificação impossível por uma mera inspeção dos resultados).

De modo similar, uma medida redundante pode ser identificada pelo exame dos valores recon-

ciliados. Se o valor reconciliado de uma variável medida é igual ao valor medido, então a medida é não redundante. Em alguns casos, a estimativa inicial e a estimativa final reconciliada podem ser iguais devido a valores numéricos (pequenas variâncias, etc.). Novamente não é possível apontar precisamente qual a causa para o ajuste nulo, assim a única forma de obter alguma classificação confiável de variáveis é através de algoritmos abrangentes citados acima.

O algoritmo RND-SQP (VASANTHARAJAN; BIEGLER, 1988) gera automaticamente um programa quadrático reduzido pela eliminação de todas as restrições de igualdade (balanços de massa, energia e componentes) e um número igual de variáveis do problema original. Portanto isto dá o menor problema de reconciliação possível e não há necessidade de identificar variáveis redundantes porque o RND-SQP usa uma técnica de programação linear para separar as variáveis em dependentes/independentes e assim eliminar todas as variáveis dependentes (uma mistura de variáveis medidas e não medidas) do problema para construir um programa quadrático reduzido a cada iteração, mas se uma análise de redundância é necessária para a disposição de sensores ou outros motivos, deve ser realizada uma análise de redundância em separado.

5.5 Comparação das estratégias de otimização não linear para reconciliação de dados

Códigos de programação não linear são disponíveis comercialmente e já provaram sua robustez numérica e confiabilidade para problemas de grande escala na indústria. Eles atuam melhor quando modelos rigorosos são usados (NAIR; JORDACHE, 1990, 1991 apud NARASIMHAN; JORDACHE, 2000). A programação não linear permite uma formulação completa para o problema de reconciliação de dados, como descrito nas Equações 5.7 até 5.9.

Tjoa e Biegler (1991b) desenvolveram um método eficiente de SQP híbrido especificamente talhado para solução de problemas de reconciliação de dados não lineares. O pacote de *software* para reconciliação de dados **RAGE**, desenvolvido por Ravikumar *et al.* (1994) também usa um *solver* SQP que foi especialmente adaptado para problemas de reconciliação de dados. Liebman e Edgar (1988) compararam o gradiente reduzido generalizado (a versão GRG2 de Lasdon e Waren (1978)) com a solução da reconciliação de dados pelo método das linearizações sucessivas e encontraram que o método da programação não linear era mais robusto às custas de um tempo computacional maior. Enquanto o GRG2, um método de **caminho provável**, precisa da conver-

gência das restrições a cada iteração, o SQP – um método do **caminho improvável**[5] – satisfaz às restrições somente no fim quando a convergência é alcançada. O SQP e outros métodos de caminho improvável (como o MINOS, um outro método de gradiente reduzido generalizado, desenvolvido por Murtagh e Saunders (apud NARASIMHAN; JORDACHE, 2000)) geralmente precisam de menos tempo computacional que outros métodos de caminho provável. Ramamurthi e Bequette (apud NARASIMHAN; JORDACHE, 2000) comparam o SQP, GRG e SL para propósitos de reconciliação de dados e apontam o seguinte:

i. Linearizações sucessivas geram viéses significativos, particularmente entre as variáveis não medidas, enquanto que as abordagens NLP geram pequenos viéses tanto entre as variáveis medidas quanto as não medidas;

ii. O tempo computacional aumenta com a magnitude dos erros nas medidas para SL, mas não para SQP e GRG;

iii. O tempo computacional é uma função estrita da exatidão desejada para SL, mas não para SQP e GRG;

iv. Os algoritmos NLP são mais eficientes e mais robustos para problemas altamente não lineares. O SQP é mais eficiente enquanto que o GRG é mais confiável.

5.6 Conclusões

Neste capítulo foi visto que as restrições de um problema de reconciliação de dados não linear podem abarcar restrições de igualdade (balanços materiais, balanços de energia, restrições de equilíbrio e correlações) e de desigualdade (limites nas variáveis e restrições de factibilidade termodinâmica). Os problemas de reconciliação de dados não linear que contenham somente restrições de igualdade podem ser resolvidos usando técnicas iterativas baseadas em linearizações sucessivas

[5] No contexto da otimização em larga escala usando simuladores seqüenciais modulares, o meio mais eficiente de combinar o simulador com o otimizador é diretamente unir o algoritmo de otimização com o código de *flowsheeting*. Existem duas classes extremas de estratégias para tanto: as de caminho provável e de caminho improvável. As estratégias de caminho provável resolvem as restrições de igualdade a cada iteração (buscam pela convergência de cada módulo) para valores fixos das variáveis de projeto e então ajustam as variáveis de projeto através do procedimento de otimização. Os resultados de cada iteração, portanto, provêem um projeto candidato para a planta, ainda que o projeto seja sub-ótimo. Estratégias de caminho improvável, por outro lado, não demandam uma solução exata dos módulos a cada passagem do simulador, assim, se um um método de caminho improvável falha, a última solução é de pouco valor (EDGAR; HIMMELBLAU, 1988, p. 574).

e a solução analítica do problema de reconciliação de dados linear. Os problemas de reconciliação de dados não linear contendo restrições de desigualdade somente podem ser resolvidos usando técnicas de otimização não linear sujeita a restrições. Se são impostos limites sobre as variáveis não medidas, então as variáveis não podem ser eliminadas por nenhuma técnica de fatoração ou projeção para obter o problema reduzido e é necessário às vezes impor limites nas variáveis de modo a se obter estimativas factíveis.

Foi visto também que os métodos GRG e SQP são duas técnicas de otimização não linear usadas para resolver problemas de reconciliação de dados não linear. Dentre os métodos apresentados destaca-se pela abrangência e aplicabilidade o SQP. Contudo, devem ser tomadas algumas precauções, apontadas na Seção 5.3.1, no sentido de adaptar o *solver* SQP às necessidades específicas dos problemas de reconciliação de dados. É desejável que o SQP seja associado a algum outro método, de modo a reduzir o espaço de busca e garantir o ótimo global.

6 *Reconciliação de Dados em Sistemas Dinâmicos*

Neste capítulo são apresentados conceitos e abordagens disponíveis para realizar a reconciliação de dados (e estimativas de estado) em sistemas transientes. Inicialmente são dadas justificativas para se tratar desse problema, depois é iniciada uma modelagem dos sistemas de interesse objetivando o uso das técnicas apresentadas. Em seguida é tratada a estimativa ótima de estado usando filtro de Kalman, sendo feita uma analogia entre este e a reconciliação de dados. Na seqüência, é tratado o controle ótimo e o filtro de Kalman, com a sua implementação para finalmente ser abordada a reconciliação dinâmica de sistemas não lineares tanto por estimativas de estado não lineares quanto por métodos de reconciliação de dados não lineares.

6.1 Justificativas para a reconciliação dinâmica de processos

Nos capítulos anteriores, a reconciliação de dados foi aplicada a um único vetor de medidas de variáveis do processo. Este vetor pode ser de medidas realizadas em qualquer instante de tempo e corresponde a um instantâneo do processo. É mais provável, contudo, que a reconciliação em estado estacionário seja realizada sobre um vetor contendo os valores médios de medidas feitas em um dado período de tempo, por exemplo 2 horas. Esta abordagem é satisfatória se os dados reconciliados se destinam a aplicações como simulações em estado estacionário ou otimização *on line* nas quais os *set points* ótimos são calculados uma vez dentro de um período de algumas horas.

Se forem consideradas aplicações tais como controle regulatório, que requerem estimativas com grande exatidão de variáveis do processo freqüentemente, então a reconciliação de dados tem que ser aplicada às medidas feitas a cada instante de amostragem. Neste caso não se pode mais considerar que as variáveis obedecem às relações de balanço material e energético em estado estacionário. Devem ser levadas em consideração as capacidades de armazenamento e atrasos de

transporte e devem ser usados os balanços materiais e de energia que relacionam as variáveis.

A estimativa de variáveis de processo que usam medidas e relações dinâmicas entre as variáveis foram desenvolvidas muito antes do termo "reconciliação de dados" surgir. São discutidas a seguir algumas dessas importantes técnicas de estimativa em conjunto com os recentes avanços na reconciliação dinâmica de dados.

Inicialmente é necessário descrever o sentido utilizado aqui para o termo "estado dinâmico" de um processo, através de duas características determinantes:

i. Os valores verdadeiros das variáveis de processo variam com o tempo e assim as medidas dessas variáveis são também função do tempo mesmo considerando-se a possibilidade extrema da ausência total de erros;

ii. Devido às entradas continuamente variantes, o acúmulo dentro de uma unidade de processo também varia continuamente e tem que ser levado em consideração.

As características acima definem tanto operações em torno de um estado estacionário nominal bem como transientes de processo que levam de um estado estacionário nominal a outro.

Diferentes técnicas estão disponíveis para o desenvolvimento do modelo dinâmico de um processo. Estas técnicas são descritas sob o tópico de **identificação de modelos** em vários livros como Aguirre (2000), Ljung (1999) e Söderström e Stoica (1989). Neste capítulo serão considerados **modelos discretos no tempo**[1] em oposição aos modelos contínuos, porque se lidará com medidas que são feitas em instantes discretos de tempo, que são convenientemente tratados por computadores. Além disso, serão considerados **modelos de espaço de estados**[2] em oposição aos modelos entrada-saída devido às suas vantagens inerentes. Inicia-se a descrição com um sistema linear discreto antes de se passar a sistemas não lineares.

[1] Segundo Aguirre (2000) os **modelos contínuos** são descritos por equações diferenciais e representam a evolução do sistema continuamente no tempo. Em oposição a isso, os **modelos dinâmicos discretos no tempo** representam a evolução do sistema em instantes discretos e são descritos por equações a diferenças.

[2] Segundo Ogata (apud GARCIA, 2005) o estado de um sistema dinâmico é o menor conjunto de variáveis (chamadas de variáveis de estado) tal que o conhecimento destas variáveis em $t = t_0$, junto com o conhecimento da entrada para $t \geq t_0$, determina completamente o comportamento do sistema para qualquer instante $t \geq t_0$. As variáveis de estado de um sistema dinâmico são as variáveis que constituem o menor conjunto de variáveis determinantes do estado do sistema. O vetor de estados é composto pelas variáveis de estado necessárias para descrever completamente o comportamento de um dados sistema. O espaço n-dimensional cujos eixos coordenados consistem nos eixos formados pelas variáveis de estado é chamado de **espaço de estados** (ou espaço de fase).

6.2 Modelagem do problema dinâmico

Um modelo discreto, dinâmico, linear e de espaço de estados de um processo é comumente descrito pela seguintes equações:

$$\mathbf{x}_k = \mathbf{A}_k\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \tag{6.1}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{6.2}$$

onde:

$\mathbf{x}_k$: vetor $n \times 1$ de variáveis de estado;

$\mathbf{u}_k$: vetor $p \times 1$ de entradas manipuladas;

$\mathbf{w}_k$: vetor $s \times 1$ de perturbações aleatórias;

$\mathbf{y}_k$: vetor $m \times 1$ de medidas;

$\mathbf{v}_k$: vetor $m \times 1$ de erros aleatórios nas medidas.

O subscrito k representa o instante de tempo $t = kT$ quando as variáveis são amostradas ou medidas, sendo T o período de amostragem. As matrizes $\mathbf{A}_k$, $\mathbf{B}_k$ e $\mathbf{H}_k$ são matrizes de dimensões apropriadas cujos coeficientes são conhecidos em qualquer tempo. Se os coeficientes dessas matrizes não variam com o tempo, então o modelo resultante é conhecido como linear e invariante com o tempo. É também comum se usar variáveis de desvio ao invés das próprias variáveis nas equações do modelo. Desta forma as variáveis de estado, $\mathbf{x}_k$, representam as diferenças entre os valores verdadeiros das variáveis e os seus valores no estado estacionário nominal. De modo similar as variáveis $\mathbf{u}_k$ e $\mathbf{y}_k$ também representam variáveis de desvio. Doravante, neste capítulo, assume-se implicitamente que todas as variáveis são de desvio.

A Equação 6.1 descreve a evolução dinâmica das variáveis de estado enquanto que a Equação 6.2 é o modelo de medição que descreve o relacionamento entre as medidas e as variáveis de estado. As hipóteses padrão feitas a respeito das perturbações aleatórias, $\mathbf{w}_k$, e sobre os erros aleatórios, $\mathbf{v}_k$, são de que estes são normalmente distribuídos com propriedades estatísticas dadas por:

$$E[\mathbf{w}_k] = E[\mathbf{v}_k] = 0 \tag{6.3}$$

$$cov[\mathbf{w}_k] = \mathbf{R}_k \tag{6.4}$$

$$cov[\mathbf{v}_k] = \mathbf{Q}_k \tag{6.5}$$

$$cov[\mathbf{w}_k, \mathbf{w}_j] = cov[\mathbf{v}_k, \mathbf{v}_j] = 0 \quad j \neq k \tag{6.6}$$

$$cov[\mathbf{w}_k, \mathbf{v}_j] = \mathbf{0} \tag{6.7}$$

As Equações 6.4 e 6.5 implicam que as variáveis aleatórias $\mathbf{w}_k$ e $\mathbf{v}_k$ têm média zero e matrizes de covariância dadas por $\mathbf{R}_k$ e $\mathbf{Q}_k$, respectivamente. A Equação 6.6 significa que as perturbações em diferentes instantes de tempo não são correlacionadas e o mesmo vale para os erros na medições. Além disso, a Equação 6.7 estipula que as perturbações e os erros não apresentam nenhuma correlação.

Os erros aleatórios nas medidas, $\mathbf{v}_k$, surgem pelas diversas razões mostradas no Capítulo 2. Por outro lado, as causas das perturbações aleatórias, $\mathbf{w}_k$, na equação da evolução de estado pode ser melhor explicada se for considerado um modelo determinístico derivado dos balanços diferenciais de massa e energia do processo. Neste caso, flutuações aleatórias nas características da alimentação do processo como fluxo, temperatura, pressão e composição podem ser modeladas como perturbações.

Quaisquer erros aleatórios nas entradas do controle fruto de ruídos elétricos nas linhas de transmissão do controlador ou devido ao posicionamento impreciso do atuador podem também ser modeladas como perturbações aleatórias. Por outro lado, se um modelo de entrada-saída é levantado a partir dos dados do processo, então pode não ser possível separar os efeitos ocasionados por erros aleatórios dos causados por perturbações aleatórias sobre as medidas. Neste caso, as diferenças entre as predições do modelo e as verdadeiras medições podem ser atribuídas ao efeito combinado dos erros nas medidas, perturbações na alimentação do processo e erros entre as entradas manipuladas reais e as calculadas.

Um modelo linear do sistema dado pelas Equações 6.1 e 6.2 pode ser derivado para qualquer processo a partir das equações diferenciais que descrevem as relações de conservação de massa e energia. Como alternativa, técnicas de identificação de modelos podem ser usadas para obter um modelo dinâmico a partir das saídas ou respostas de um processo para um conjunto de entradas. O desenvolvimento de um modelo determinístico para um processo simples de controle de nível é

mostrado a seguir no Exemplo 6.1.

EXEMPLO 6.1 Um processo simples de controle de nível é mostrado na Figura 6.1 com uma alimentação (F_1) e duas saídas (F_2 e F_3). A válvula V_1 é mantida aberta em uma posição fixa, enquanto que a válvula V_2 é manipulada para manter o controle do nível do tanque (ao invés de se calcular diretamente a nova posição da válvula, se assume que o ajuste, a, à posição x da válvula é calculado a cada intervalo de tempo).

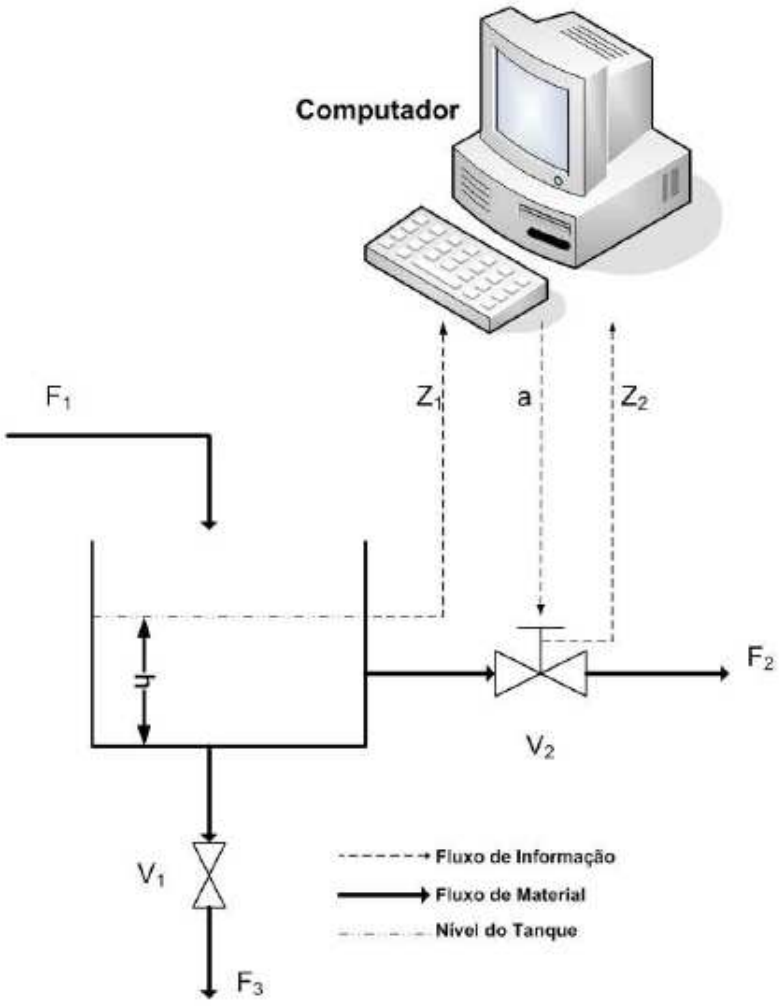


Figura 6.1: Processo de controle de nível – adaptado de Narasimhan e Jordache (2000)

O nível do tanque e a posição da válvula V_2 são medidos e denotados por Z_1 e Z_2, respectivamente. A equação diferencial descrevendo o balanço mássico para esse processo é dada por:

$$A\frac{dh}{dt} = F_1 - F_2 - F_3$$

As taxas de fluxo de saída são relacionadas ao nível do tanque e às posições da válvula por:

$$F_3 = K_{01}h$$

$$F_2 = K_{02}h + K_{03}x$$

Substituindo as relações acima na equação de balanço de massa, tem-se que:

$$A\frac{dh}{dt} = F_1 - (K_{01} + K_{02})h - K_{03}x$$

Assumindo um intervalo de amostragem uniforme T entre as medidas e usando o subscrito k para representar as variáveis no instante kT de amostragem, então o equivalente discreto da equação diferencial acima pode ser obtido usando o método descrito em Franklin *et al.* (1980).

$$h_{k+1} = \alpha h_k - \frac{(1-\alpha)K_{03}}{(K_{01}+K_{02})}x_k + \frac{(1-\alpha)}{(K_{01}+K_{02})}F_1$$

onde

$$\alpha = \exp\frac{-(K_{01}+K_{02})T}{A}$$

Quando se deriva a representação discreta acima, é implicitamente assumido que a posição x da válvula é constante no valor x_k durante o intervalo de tempo kT até $(k+1)T$ e também que as perturbações aleatórias na alimentação F_1 têm uma magnitude constante dentro de cada intervalo mas têm uma magnitude aleatória de intervalo para intervalo. Se o ajuste a_k à posição da válvula (calculado pelo controlador depois das medidas serem feitas no instante k de amostragem) é implementado no início do próximo intervalo de amostragem, então a posição da válvula em cada instante de amostragem é dada por:

$$x_{k+1} = x_k + a_k + e_{k+1}$$

onde e_{k+1} é o erro aleatório no posicionamento da válvula. A Tabela 6.1 dá os valores para as diferentes constantes usadas neste processo. Usando estes valores, o modelo de espaço de estados do processo de controle de nível é obtido como:

$$\begin{bmatrix} h_{k+1} \\ x_{k+1} \end{bmatrix} = \begin{bmatrix} 0,995 & -0,1373 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} h_k \\ x_k \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}a_k + \begin{bmatrix} 0,00012 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} F_{1,k+1} \\ e_{k+1} \end{bmatrix}$$

$$\begin{bmatrix} z_{1,k+1} \\ z_{2,k+1} \end{bmatrix} = \begin{bmatrix} 0,631 & 0 \\ 0 & 1,57 \end{bmatrix}\begin{bmatrix} h_{k+1} \\ x_{k+1} \end{bmatrix} + \begin{bmatrix} v_{1,k+1} \\ v_{2,k+1} \end{bmatrix}$$

onde $v_{1,k+1}$ e $v_{2,k+1}$, são os erros aleatórios nas medidas do nível e no posicionamento da

válvula (em volts), respectivamente.

Tabela 6.1: Valores dos parâmetros para o processo de controle de nível (NARASIMHAN; JORDACHE, 2000)

Parâmetro	Valor	Unidade
K_{01}	7,2	cm^2/min
K_{02}	34,78	cm^2/min
K_{03}	1156,0	cm^2/min
A	280,0	cm^2
T	2,0	s

As entradas manipuladas $\mathbf{u}_k$ na Equação 6.1 podem ser obtidas usando uma lei de controle que é geralmente uma função das medidas quando as variáveis que precisam ser controladas são também medidas. Uma lei de controle proporcional linear pode ser descrita como:

$$\mathbf{u}_k = \bar{\mathbf{C}}_k(\mathbf{y}_k - \mathbf{y}_{sp,k}) \tag{6.8}$$

onde $\mathbf{y}_{sp,k}$ representa o desvio ou variação entre os *set points* de operação correntes e é igual a $\mathbf{0}$ se não há mudança a partir dos *set points* correntes. Em alguns casos quando é difícil ou caro medir as variáveis controladas, uma estratégia de controle inferencial é usada. As entradas manipuladas neste caso são uma função das estimativas de estado. Mesmo no caso quando as variáveis controladas são medidas, pode ser melhor basear a lei de controle nas estimativas destas variáveis porque estas têm maior probabilidade de serem mais exatas se o estimador é projetado apropriadamente. Como foi mencionado na seção anterior, uma razão primária para a reconciliação dinâmica de dados é gerar estimativas que possam ser usadas para um controle mais eficiente. Desta forma se assume uma lei de controle da seguinte forma:

$$\mathbf{u}_k = \bar{\mathbf{C}}_k(\hat{\mathbf{x}}_k - \mathbf{x}_{sp,k}) \tag{6.9}$$

onde $\hat{\mathbf{x}}_k$ são estimativas para os valores reais das variáveis de estado e $\mathbf{x}_{sp,k}$ são as variações nos *set points* das variáveis de estado a partir dos *set points* correntes. Com o objetivo de se obter um bom controle é portanto necessário estimar as variáveis de estado com a maior exatidão possível.

6.3 Estimativa ótima de estado usando filtro de Kalman

Lidou-se primeiro com o problema da estimativa ótima para as variáveis de estado de um processo que possa ser descrito por um modelo linear de forma dada pela Equação 6.1 e 6.2 e que satisfaça às hipóteses das Equações 6.3 até 6.7. Assume-se também que as entradas manipuladas a cada intervalo de tempo são valores constantes conhecidos e desconsidera-se temporariamente o fato de que estas são estimativas de funções de estado.

O estimador linear ótimo chamado de **filtro de Kalman** que é descrito nesta seção pode ser derivado usando diferentes formulações teóricas. Será usada uma abordagem baseada na formulação dos mínimos quadrados porque esta ajuda a prontamente comparar o filtro de Kalman com a reconciliação de dados.

As estimativas inicias das variáveis de estado, pressupostamente disponíveis, possuem as seguintes propriedades estatísticas:

$$\hat{\mathbf{x}}_0 = E\left[\mathbf{x}_0\right] \tag{6.10}$$

$$cov[\hat{\mathbf{x}}_0] = \mathbf{P}_0 \tag{6.11}$$

Dado um conjunto de medias, $\mathbf{Y}_k = (y_1, y_2, \ldots, y_k)$, deseja-se obter estimativas das variáveis de estado $\mathbf{x}_k$ que sejam as melhores sob determinado aspecto. Estas estimativas são denotadas por $\hat{\mathbf{x}}_{k|k}$ interpretadas como as estimativas de estado no tempo k obtidas usando todas medidas a partir do tempo $t = 1$ até o tempo $t = k$. Nota-se que o uso de todas as medidas a partir do tempo inicial para se derivar as estimativas explora automaticamente a redundância temporal nos dados medidos.

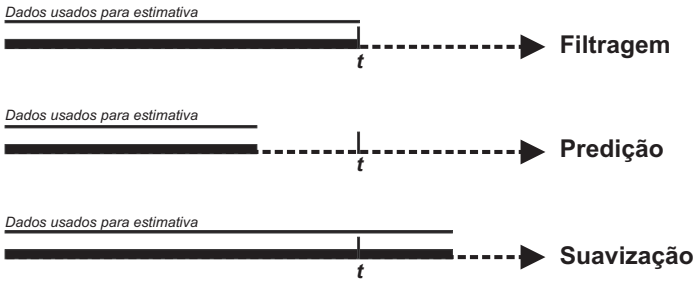


Figura 6.2: Problemas de estimativa de estado – adaptado de Bagajewicz (2001)

O problema de estimativa considerado aqui é um caso específico de um problema mais geral no

qual se deseja obter estimativas de variáveis de estado $\mathbf{x}_j$ para um tempo j, usando todas as medidas feitas a partir do tempo inicial até o tempo k. As estimativas assim derivadas são denotadas como $\hat{\mathbf{x}}_{j|k}$. O problema de estimativa é dito de predição se $j > k$, é chamado de filtro se $j = k$ e de suavização se $j < k$ (Figura 6.2). Aqui é trabalhado o problema do filtro.

Devido à presença de perturbações aleatórias na Equação 6.1, os valores verdadeiros das variáveis de estado a cada instante de tempo são eles próprios variáveis aleatórias. Portanto, uma medida probabilística tem que ser usada para determinar as melhores estimativas das variáveis de estado. As melhores estimativas das variáveis de estado no tempo k são obtidas pela minimização da seguinte função:

$$\mathbf{J}_k = E\left[(\hat{\mathbf{x}}_{k|k} - \mathbf{x}_k)^{\mathrm{T}}(\hat{\mathbf{x}}_{k|k} - \mathbf{x}_k)\right] \tag{6.12}$$

A Equação 6.12 é a esperança da soma dos quadrados das diferenças entre as estimativas e os valores verdadeiros das variáveis de estado e é assim uma extensão da bem conhecida função objetivo determinística de mínimos quadrados. A solução do problema foi primeiro obtida por Kalman (KALMAN, 1960b, 1961 apud NARASIMHAN; JORDACHE, 2000) em uma forma recursiva conveniente e é agora geralmente referida como o filtro de Kalman:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}) \tag{6.13}$$

onde

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_k\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_k\mathbf{u}_{k-1} \tag{6.14}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}} + \mathbf{Q}_k)^{-1} \tag{6.15}$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}_k\mathbf{P}_{k-1|k-1}\mathbf{A}_k^{\mathrm{T}} + \mathbf{R}_k \tag{6.16}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} \tag{6.17}$$

As matrizes $\mathbf{P}_{k|k}$ e $\mathbf{P}_{k|k-1}$ são as matrizes de covariância das estimativas $\hat{\mathbf{x}}_{k|k}$ e $\hat{\mathbf{x}}_{k|k-1}$, respectivamente. Iniciando com as estimativas $\hat{\mathbf{x}}_0$ e $\mathbf{P}_0$, as equações acima podem ser aplicadas em ordem reversa da Equação 6.16 para a Equação 6.13 para obter as estimativas de estado a cada tempo k.

A forma recursiva das equações do estimador reduz consideravelmente o esforço computacional envolvido na obtenção das estimativas. Pode ser observado a partir destas equações que o esforço gasto na obtenção das estimativas de estado em um dado tempo é efetivamente utilizado para obter as estimativas no próximo instante de tempo. A Equação 6.13 pode também ser interpretada como um método preditor-corretor na obtenção das estimativas. As estimativas $\hat{\mathbf{x}}_{k|k-1}$ são as estimativas preditas das variáveis de estado no tempo k e baseadas em todas as medidas até o tempo $k-1$.

O segundo termo nessa equação é a correção a estas estimativas baseada nas medidas no tempo k. A matriz $\mathbf{K}_k$ é conhecida como o ganho do filtro de Kalman e a diferença $(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1})$ é conhecida como "inovações". As inovações são equivalentes aos resíduos nas medidas de um processo no estado estacionário (Capítulo 7) e representam um papel importante na detecção de erros grosseiros. As estimativas de filtro de Kalman possuem a desejável propriedade estatística de serem não enviesadas, ou seja:

$$E[\hat{\mathbf{x}}_{k|k}] = E[\hat{\mathbf{x}}_k] \tag{6.18}$$

e têm também a mínima variância entre todos os estimadores não-enviesados. Além disso as estimativas do filtro de Kalman de máxima verossimilhança (SAGE; MELSA, 1971). Para um processo linear e invariante no tempo, o ganho de Kalman se torna constante depois de algum tempo, o que é conhecido como o ganho de Kalman em estado estacionário.

As aplicações do filtro de Kalman na Engenharia Química foram discutidas por diversos autores. Fisher e Seborg (apud NARASIMHAN; JORDACHE, 2000) aplicaram o filtro de Kalman a um evaporador de múltiplo efeito em escala piloto no contexto da investigação de vários tipos de estratégias de controle. Stanley e Mah (apud NARASIMHAN; JORDACHE, 2000) aplicaram-no a uma subseção de uma refinaria para a estimativa de fluxos globais e temperaturas. O modelo dinâmico usado nesta aplicação é um modelo heurístico de caminhada aleatória[3] para as variáveis de es-

[3] A caminha aleatória é a formalização matemática de uma trajetória que consiste numa série de passos aleatórios sucessivos. Os resultados da caminhada aleatória foram aplicados à ciência da computação, física, ecologia e várias outras áreas como um modelo fundamental para processos aleatórios no tempo. Um exemplo é o caminho traçado por uma molécula que atravessa um fluido.

tado, o quê é apropriado para descrever processos que operam por longos períodos em torno de um estado estacionário nominal com transições lentas ocasionais para um novo estado estacionário nominal. As variáveis de estado foram também forçadas a satisfazer os balanços de massa e energia em estado estacionário. Através dessa abordagem, é explorada tanto a redundância espacial quanto a temporal na reconciliação de dados.

EXEMPLO 6.2 Ilustra-se a aplicação do filtro de Kalman na obtenção de estimativas ótimas para processos de controle de nível descrito no Exemplo 6.1. Neste processo, as flutuações do fluxo de alimentação e o erro aleatório no posicionamento da válvula de controle são tomados como perturbações de estado. Os desvios padrão dessas perturbações aleatórias são 250 cm^3/min e 0,05; respectivamente. 0s desvios padrão dos erros nas medidas de nível e na posição da válvula são tomados como 0,01 volt cada um. As medidas correspondentes ao comportamento em malha fechada do processo são simuladas baseadas no modelo de espaço de estados derivado no Exemplo 6.1. A lei de controle usada nessa simulação é dada por:

$$a_k = 0,02057\tilde{z}_{1k} - 0,6369\tilde{z}_{2k}$$

onde $\tilde{z}_{1k}$ e $\tilde{z}_{2k}$ são os valores medidos do nível e da posição da válvula (em cm) obtidos pela divisão da leitura em volts por 0,631 e 1,57, respectivamente (ver Exemplo 6.1). Um filtro de Kalman é usado para estimar os estados usando o ganho de Kalman em estado estacionário porque se usou um modelo invariante no tempo. O ganho de Kalman no estado estacionário é obtido pela solução da equação da matriz de Ricatti (SORENSON, 1985 apud NARASIMHAN; JORDACHE, 2000) e é obtido como:

$$K_{EE} = \begin{bmatrix} 0,2758 & -0,0291 \\ -0,0177 & 0,3406 \end{bmatrix}$$

A Figura 6.3 mostra os valores verdadeiros, medidos e estimados para o nível. Pode-se observar na Figura 6.3 que as estimativas são mais próximas aos valores verdadeiros que as medidas. A variância do erro na medida, calculada a partir de amostras de dados sobre um período de 200 segundos, é de 0,0294 cm^2 enquanto que a variância do erro nas estimativas é 0,0023 cm^2. A variância da diferença entre os valores verdadeiros e o *set point*, que neste caso é zero, é um indicador do desempenho do controlador. Para este caso foi calculado como sendo 0,0068 cm^2. A variância no posicionamento da válvula para alcançar este controle é 0,1057 cm^2.

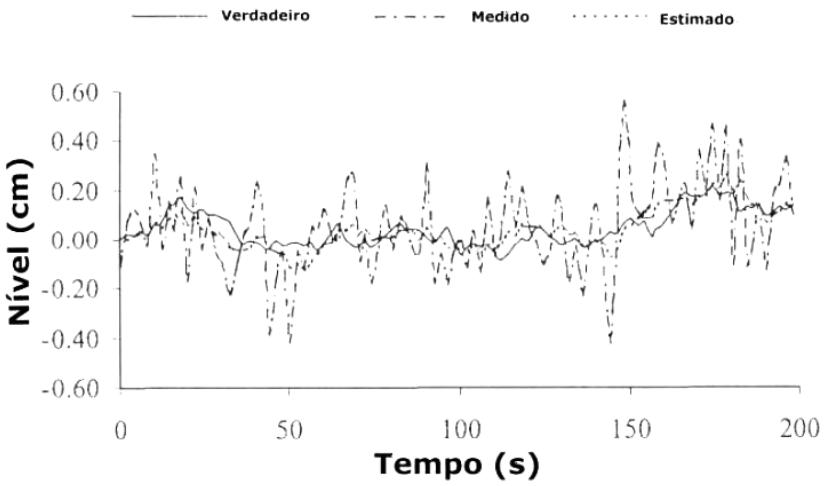


Figura 6.3: Valores medidos, verdadeiros e estimados para o nível em um processo de controle de nível usando uma lei de controle baseada em medidas – adaptado de Narasimhan e Jordache (2000)

6.4 Analogia entre o filtro de Kalman e a reconciliação de dados em estado estacionário

As técnicas de reconciliação de dados foram desenvolvidas inicialmente para processos em estado estacionário enquanto que o filtro de Kalman foi desenvolvido independentemente para um processo linear dinâmico. Ambas as técnicas podem ser derivadas usando um procedimento de estimativa de mínimos quadrados ponderados. Para proceder com a ligação entre estas duas abordagens, prova-se que a reconciliação de dados em estado estacionário pode ser vista como um caso especial do filtro de Kalman. Já foi visto no Capítulo 3 que para processos em estado estacionário, as relações de conservação de massa e energia são escritas como restrições algébricas. A forma dinâmica diferencial dessas relações de conservação pode também ser usada para derivar um modelo discreto e linear de espaço de estados da forma da Equação 6.1, como mostrado no Exemplo 6.1. Como um caso especial considera-se uma forma livre de perturbações dessa equação fazendo $\mathbf{w}_k$ ser identicamente nula para todos os tempos para obter:

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_{k-1} \tag{6.19}$$

Define-se um novo vetor de variáveis de estado $\mathbf{x}$ composto de $\mathbf{x}_k$ e $\mathbf{x}_{k-1}$, no qual se omite o índice de tempo k para uma comparação direta com a reconciliação de dados em estado estacionário. A Equação 6.19 pode ser reescrita como:

$$\mathbf{A}\mathbf{x} = \mathbf{c} \tag{6.20}$$

onde

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{x}_{k-1} \end{bmatrix} \tag{6.21}$$

$$\mathbf{A} = [\mathbf{I} \quad -\mathbf{A}_k] \tag{6.22}$$

$$\mathbf{c} = \mathbf{B}_k \mathbf{u}_{k-1} \tag{6.23}$$

Se todas as variáveis de estado são medidas, então a Equação 6.2 pode ser escrita como:

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{v}_k \tag{6.24}$$

Assume-se também que se tenha estimativas não enviesadas $\hat{\mathbf{x}}_{k-1|k-1}$ das variáveis $\mathbf{x}_{k-1}$ de estado obtidas do instante de tempo anterior e que a matriz de covariância seja $\mathbf{P}_{k-1|k-1}$. Estas estimativas podem ser tratadas como medidas adicionais e podem ser escritas como:

$$\hat{\mathbf{x}}_{k-1|k-1} = \mathbf{x}_{k-1} + \varepsilon_{k-1} \tag{6.25}$$

onde ε_{k-1} são os erros aleatórios na estimativa de $\mathbf{x}_{k-1}$ com média zero e matriz de covariância $\mathbf{P}_{k-1|k-1}$. A partir das propriedades pressupostas para $\mathbf{v}_k$ (Equações 6.5 a 6.7) pode-se facilmente provar que elas não são correlacionadas com ε_{k-1}. Combinando as Equações 6.24 e 6.25 pode-se escrever o modelo de medição modificado como:

$$\mathbf{y} = \mathbf{x} + \mathbf{e} \tag{6.26}$$

onde

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_k \\ \hat{\mathbf{x}}_{k-1|k-1} \end{bmatrix} \tag{6.27}$$

$$\mathbf{e} = \begin{bmatrix} \mathbf{v}_k \\ \varepsilon_{k-1} \end{bmatrix} \tag{6.28}$$

onde $\mathbf{v}$ é o vetor de erros aleatórios com média zero e matriz de covariância Ψ definida por:

$$\Psi = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{k-1|k-1} \end{bmatrix} \tag{6.29}$$

As Equações 6.20 e 6.26 são similares ao modelo de medida e restrições nas Equações 3.10. Aplicando-se a solução de reconciliação de dados em estado estacionário a esse modelo pode-se obter as estimativas de $\mathbf{x}$ usando a Equação 3.22. Substituindo as diferentes variáveis definidas pelas Equações 6.21 até 6.23, 6.27 e 6.29 nesta solução tem-se:

$$\begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\mathbf{x}}_{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_k \\ \hat{\mathbf{x}}_{k-1|k-1} \end{bmatrix} - \begin{bmatrix} \mathbf{Q}_k \\ -\mathbf{P}_{k-1|k-1}\mathbf{A}_k^{\mathrm{T}} \end{bmatrix} \times (\mathbf{Q}_k + \mathbf{A}_k\mathbf{P}_{k-1|k-1}\mathbf{A}_k^{\mathrm{T}})^{-1} \times (\mathbf{y}_k - \mathbf{A}_k\mathbf{x}_{k-1|k-1} - \mathbf{B}_k\mathbf{u}_{k-1}) \tag{6.30}$$

Considerando somente as estimativas de $\mathbf{x}_k$ na Equação 6.30 e usando as estimativas preditas pela Equação 6.14, obtém-se:

$$\hat{\mathbf{x}}_k = \mathbf{y}_k - \mathbf{Q}_k(\mathbf{Q}_k + \mathbf{A}_k\mathbf{P}_{k-1|k-1}\mathbf{A}_k^{\mathrm{T}})^{-1}(\mathbf{y}_k - \hat{\mathbf{x}}_{k|k-1}) \tag{6.31}$$

As estimativas dadas pela Equação 6.31 podem ser demonstradas como idênticas às estimativas de filtro de Kalman dadas pela Equação 6.13 com $\mathbf{R}_k = \mathbf{0}$ e $\mathbf{H}_k = \mathbf{I}$ como no caso do modelo simplificado considerado nesta seção. Devido ao filtro de Kalman também contabilizar perturbações aleatórias no modelo do processo, pode ser considerado como uma extensão da técnica de reconciliação de dados linear em estado estacionário para processos dinâmicos. Um outro resultado interessante da análise apresentada nesta seção é que se pode provar que as estimativas $\hat{\mathbf{x}}_{k-1}$ dadas na Equação 6.30 são idênticas às estimativas ótimas suavizadas $\hat{\mathbf{x}}_{k-1|k}$ do modelo simplifi-

cado. As técnicas de reconciliação de dados linear dinâmica foram aplicadas nas estimativas de fluxo de um processo por Darouach e Zasadzinski (apud NARASIMHAN; JORDACHE, 2000) e por Rollins e Devanathan (apud NARASIMHAN; JORDACHE, 2000). Estes autores converteram as equações diferenciais em equações algébricas trocando o termo derivativo por uma diferença avançada. O problema pôde então ser resolvido usando técnicas de solução linear de reconciliação de dados similares ao procedimento descrito anteriormente. Devido ao constante aumento da dimensão do problema com o tempo, foram propostas por estes autores técnicas eficientes para obtenção das estimativas. Bagajewicz e Jiang (1997) consideraram também o problema de reconciliação de dados dinâmica de fluxo de processos e *holdups* de tanques como uma função polinomial do tempo e converteram as equações diferenciais em equações algébricas. Os coeficientes dos polinômios foram estimados usando uma janela de medições.

6.5 Controle ótimo e filtro de Kalman

Considerando o problema de controle ótimo para um processo linear determinístico que evolua segundo a Equação 6.1, mas sem qualquer perturbação de estado. Assume-se também que as variáveis de estado são medidas diretamente sem quaisquer erros, ou seja, os valores verdadeiros das variáveis de estado estão disponíveis. Deseja-se determinar os valores ótimos das entradas manipuladas que minimizem o índice de desempenho.

$$\min_{\mathbf{u}_i} N_k^c = \sum_{i=1}^{n} (\mathbf{x}_i^{\mathrm{T}} \mathbf{E}_i \mathbf{x}_i + \mathbf{u}_i^{\mathrm{T}} \mathbf{F}_i \mathbf{u}_i) \tag{6.32}$$

onde $\mathbf{E}_i$ e $\mathbf{F}_i$ são matrizes de fatores ponderantes especificados. O primeiro termo na Equação 6.32 tenta manter as variáveis de estado (os desvios das variáveis de estado em relação a seus *set points* atuais) no valor alvo de zero enquanto que o segundo termo tenta minimizar valores altos das entradas manipuladas. As matrizes de ponderação são escolhidas baseadas na importância relativa entre as variáveis de estado e as entradas manipuladas.

A solução do problema acima leva a uma lei de controle linear da forma:

$$\mathbf{u}_i = \bar{\mathbf{K}}_i \mathbf{x}_i \tag{6.33}$$

onde $\bar{\mathbf{K}}_i$ é dependente das matrizes $\mathbf{E}_i$ e $\mathbf{F}_i$ e das matrizes do sistema.

Considerando-se agora o problema de controle ótimo para um sistema estocástico linear descrito pela Equação 6.1. Neste caso o índice de desempenho para o problema de controle ótimo pode ser escrito como:

$$\min_{\mathbf{u}_i} N_k^c = E\left[\sum_{i=1}^{n}(\mathbf{x}_i^{\mathrm{T}}\mathbf{E}_i\mathbf{x}_i + \mathbf{u}_i^{\mathrm{T}}\mathbf{F}_i\mathbf{u}_i)\right] \tag{6.34}$$

Os valores ótimos das entradas manipuladas que minimizam a Equação 6.34 podem ser obtidos da seguinte forma:

i. Calcular $\hat{\mathbf{x}}_{k|k-1}$ que são as estimativas preditas das variáveis de estado no tempo k, usando as equações do filtro de Kalman tratando as entradas manipuladas anteriores ao tempo k como sendo entradas determinísticas e conhecidas.

ii. Calcular as entradas manipuladas no tempo k usando as estimativas $\hat{\mathbf{x}}_{k|k-1}$ no lugar de $\mathbf{x}_k$ na Equação 6.33.

Apesar do fato das entradas manipuladas serem elas próprias funções de estimativas de estado, elas são consideradas como sendo entradas determinísticas e conhecidas quando se derivam as estimativas de estado usando o filtro de Kalman. Por outro lado a lei de controle ótimo foi obtida para um sistema determinístico, para o qual os valores verdadeiros das variáveis de estados são considerados disponíveis, mas também é usada para sistemas estocásticos. Essencialmente isso implica que o problema de estimativa ótima e o problema de controle ótimo foram separados. A prova de que esses procedimentos fornecem as entradas manipuladas ótimas para a minimização da Equação 6.34 segue do **teorema da separação** ou do **princípio da equivalência da certeza**[4] (ANDERSON; MOORE, 1989 apud NARASIMHAN; JORDACHE, 2000).

EXEMPLO 6.3 Para investigar o efeito do uso de uma lei de controle baseada em estimativas de estado, o processo de controle de nível descrito nos exemplos precedentes foi simulado com uma lei de controle similar àquela usada no Exemplo 6.2. A lei de controle,

[4] O princípio da equivalência da certeza foi formulado para problemas de controle no início dos anos 1960 e diz que a lei de controle ótimo para um problema de controle estocástico tem a mesma estrutura que a lei de controle ótimo para o problema determinístico associado. A única diferença seria que na lei de controle estocástico as variáveis de estado verdadeiras são trocadas por seus valores estimados. A validade deste princípio leva à conclusão que o projeto do estimador e do controlador podem ser otimizados separadamente.

contudo, foi baseada nas estimativas do nível e da posição da válvula obtidas usando um filtro de Kalman. A lei de controle, assim, é dada por:

$$a_k = 0,02057\hat{h}_k - 0,6369\hat{x}_k$$

Os valores verdadeiros, os medidos e os simulados do nível para este caso são mostrados na Figura 6.4. Pode-se comparar os valores verdadeiros obtidos neste caso com aqueles obtidos no Exemplo 6.2 e mostrar que há um ganho marginal no desempenho do controle. A variância do erro entre os valores verdadeiros e o *set point* é $0,0065cm^2$ o que é marginalmente mais baixo que o obtido quando a lei de controle é baseada em valores medidos. Contudo, a variância no posicionamento da válvula neste caso é somente 0,0523 cm^2, um valor cerca de 5% do obtido no exemplo anterior. Isto implica que é possível alcançar um nível de controle tão bom quanto a do exemplo anterior, mas com menos variação na variável manipulada.

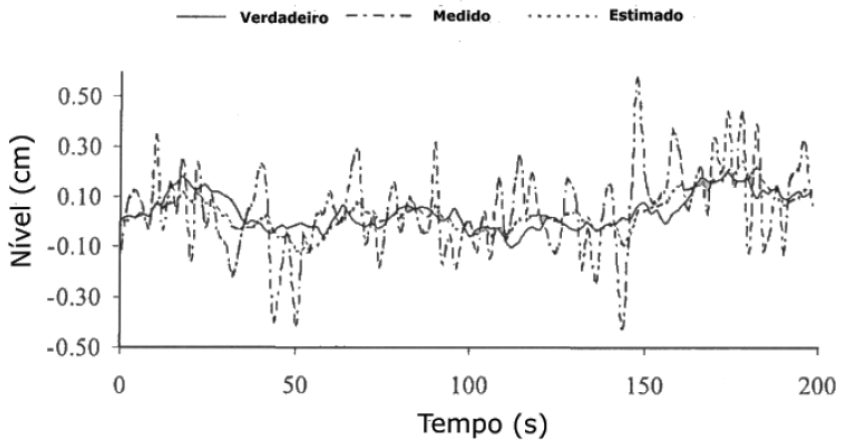


Figura 6.4: Valores medidos, verdadeiros e estimados para o nível em um processo de controle de nível usando uma lei de controle baseada em estimativas – adaptado de Narasimhan e Jordache (2000)

6.5.1 Implementação do filtro de Kalman

As matrizes $\mathbf{P}_{k|k-1}$ e $\mathbf{P}_{k|k}$, que estão nas equações do filtro de Kalman, por serem as matrizes de covariância, geralmente devem ser positivamente definidas ou seja, os seus autovalores devem ser positivos. Se isso é assegurado, então o filtro de Kalman será estável. Contudo se o filtro de Kalman é implementado como dado nas Equações 6.13 até 6.17, estas matrizes tendem a perder

sua definição de não-negatividade característica e as estimativas tendem a divergir devido à falta de exatidão nos cálculos.

Uma forma conhecida como filtro da raiz quadrada da covariância pode ser usada para implementar o filtro de Kalman. A Equação 6.16 que é usada para obter $\mathbf{P}_{k|k-1}$, preserva as características de simetria e de definição positiva da matriz, mas a Equação 6.17 usada para obter a matriz de covariância atualizada pode causar problemas numéricos porque envolve uma inversão matricial no cálculo da matriz do ganho do filtro. É essa equação que é remodelada em termos das raízes quadradas das matrizes de covariância.

Além disso, a eficiência computacional é também aumentada pelo processamento das medidas seqüencialmente ao invés de simultaneamente, evitando assim a necessidade de calcular a inversa de uma matriz na Equação 6.15. Os passos envolvidos na implementação são descritos a seguir (BORRIE; BAGCHI, 1992, 1993 apud NARASIMHAN; JORDACHE, 2000).

Passo 1 Começando com as estimativas de $\hat{\mathbf{x}}_{k-1|k-1}$ e $\mathbf{P}_{k-1|k-1}$, aplicar às Equações 6.14 e 6.16 para calcular as predições um passo adiante, $\hat{\mathbf{x}}_{k|k-1}$ e $\mathbf{P}_{k|k-1}$;

Passo 2 Obter as raízes quadradas $\mathbf{S}_{k|k-1}$ e Ψ_k das matrizes de covariância $\mathbf{P}_{k|k-1}$ e $\mathbf{Q}_k$ respectivamente definidas por:

$$\mathbf{P}_{k|k-1} = \mathbf{S}_{k|k-1}\mathbf{S}_{k|k-1}^{\mathrm{T}} \tag{6.35}$$

$$\mathbf{Q}_k = \Psi_k \Psi_k^{\mathrm{T}} \tag{6.36}$$

As raízes quadradas das matrizes podem ser obtidas usando a **fatoração de Cholesky**[5];

Passo 3 Calcular as medidas transformadas e a matriz das medidas transformadas definidas por:

$$\Psi_k \mathbf{y}_k^* = \mathbf{y}_k \tag{6.37}$$

$$\Psi_k \mathbf{H}_k^* = \mathbf{H}_k \tag{6.38}$$

Como Ψ_k é triangular superior, $\mathbf{y}_k^*$ e $\mathbf{H}_k^*$ podem ser calculadas sem a necessidade de inverter Ψ_k.

Passo 4 Como as medidas transformadas não são correlacionadas elas podem ser processadas

[5] A fatoração ou decomposição de Cholesky decompõe uma matriz simétrica positivamente definida em uma matriz triangular inferior e a transposta da matriz triangular inferior. A fatoração de Cholesky é usada principalmente na solução numérica de sistemas de equações lineares.

seqüencialmente usando as seguintes equações

$$\mathbf{t}_{ki} = \mathbf{S}_{k|k,i-1}\mathbf{h}^*_{k,i} \tag{6.39}$$

$$\beta_{ki} = \frac{1}{\mathbf{t}^{\mathrm{T}}_{ki}\mathbf{t}_{ki} + 1} \tag{6.40}$$

$$\mathbf{S}_{k|k,i} = \mathbf{S}_{k|k,i-1} - \frac{\beta_{ki}\mathbf{S}_{k|k,i-1}\mathbf{t}_{ki}\mathbf{t}^{\mathrm{T}}_{ki}}{1 + \sqrt{\beta_{ki}}} \tag{6.41}$$

onde $\mathbf{S}_{k|k-1}$ é a raiz quadrada da matriz de covariância $\mathbf{P}_{k|k}$ atualizada após o processamento das i primeiras medidas e $\mathbf{h}^*_{k,i}$ é a i-ésima linha da matriz de medidas transformada, $\mathbf{H}^*_k$. Inicializa-se os cálculos deste passo usando:

$$\mathbf{S}_{k|k,0} = \mathbf{S}_{k|k-1} \tag{6.42}$$

Desta forma, após todos as n medições serem processadas, a matriz da covariância atualizada é obtida como:

$$\mathbf{P}_{k|k} = \mathbf{S}_{k|k,n}\mathbf{S}^{\mathrm{T}}_{k|k,n} \tag{6.43}$$

Mesmo que a matriz do ganho do filtro de Kalman não seja explicitamente calculada no procedimento seqüencial acima, isto pode ser feito, se necessário, usando-se:

$$\mathbf{K}_{k,i} = \beta_{ki}\mathbf{S}_{k|k,i}\mathbf{t}_{ki} \tag{6.44}$$

onde $\mathbf{K}_{k,i}$ é a i-ésima coluna da matriz de ganho.

6.6 Reconciliação de dados dinâmica de sistemas não lineares

Foi mostrado anteriormente que a técnica do filtro de Kalman pode ser usada com sucesso sobre dados de um processo dinâmico para suavizá-los recursivamente e estimar parâmetros para sistemas lineares. A seguir são apresentadas modificações para lidar com sistemas não lineares. Estas modificações envolvem tipicamente a substituição das equações não lineares que representam o sistema por aproximações de primeira ordem. Para processos operando em regiões de alta não linearidade, estas aproximações lineares podem não ser satisfatórias.

6.6.1 Estimativas de estado não lineares

O tratamento de processos não lineares apresenta várias dificuldades que não são encontradas em sistemas lineares. Inicialmente, em geral não é possível obter analiticamente uma representação da forma discreta do processo análoga à Equação 6.1 iniciando com um conjunto de equações diferenciais não lineares. Em segundo lugar, é matematicamente difícil tratar o ruído aleatório se as equações de transição de estado ou as equações das medidas são funções não lineares do ruído.

Desta forma, o efeito do ruído em um processo não linear é modelado como um termo aditivo linear e mesmo se o ruído aleatório é tomado como normalmente distribuído, nem as variáveis de estado nem as próprias medidas seguem uma distribuição gaussiana devido à não linearidade das equações. Assim, um *framework* probabilístico pode ser usado somente sob algumas aproximações. Contudo, uma formulação em mínimos quadrados pode ser usada para derivar as estimativas.

Sob as limitações apontadas acima, o comportamento das variáveis de estado para um processo não linear geral é modelado pela seguinte equação diferencial:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{t}, \mathbf{x}, \mathbf{u}) + \dot{\beta} \tag{6.45}$$

$$\dot{\beta} = \mathbf{w}(\mathbf{t}) \tag{6.46}$$

onde $\mathbf{w}(\mathbf{t})$ é um ruído branco de processo com média da função zero e função de matriz de covariância $\mathbf{R}(t)\delta(t-\tau)$, onde $\delta(t-\tau)$ é a função delta de Dirac.

As variáveis são amostradas em tempos discretos $t = kT$ e a relação entre as medidas e as variáveis de estado são representadas como:

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \tag{6.47}$$

onde $\mathbf{v}_k$ são os erros aleatórios das medidas aos quais se atribui uma distribuição gaussiana com média zero e matriz de covariância $\mathbf{Q}_k$. Como no caso linear, assume-se que $\mathbf{w}(\mathbf{t})$ e $\mathbf{v}_k$ não são correlacionadas ente si. As Equações 6.45 e 6.47 descrevem um processo estocástico não linear e contínuo com medidas discretas.

Um filtro de Kalman estendido, que é a versão não linear do filtro de Kalman, baseado em uma aproximação linear das Equações 6.45 e 6.47, a cada tempo em torno das estimativas atuais

de estado, pode ser usado para obter as estimativas de estado recursivamente usando as seguintes equações que são análogas às Equações 6.45 e 6.47:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left[\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1})\right] \tag{6.48}$$

$$\dot{\hat{\mathbf{x}}}_{\tau|k-1} = \mathbf{f}(\tau, \hat{\mathbf{x}}_{\tau|k-1}, \mathbf{u}_{k-1}), \quad \hat{\mathbf{x}}_{k|k-1} = \hat{\mathbf{x}}_{\tau=kT|k} \tag{6.49}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}} + \mathbf{Q}_k)^{-1} \tag{6.50}$$

$$\dot{\mathbf{P}}_{\tau|k-1} = \mathbf{F}_{\tau|k-1}\mathbf{P}_{\tau|k-1} + \mathbf{P}_{\tau|k-1}\mathbf{F}_{\tau|k-1}^{\mathrm{T}} \tag{6.51}$$

$$\mathbf{P}_{k|k-1} = \mathbf{P}_{\tau=kT|k-1} + \mathbf{R}(t = kT) \tag{6.52}$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} \tag{6.53}$$

onde:

$$\mathbf{F}_{\tau|k-1} = \left.\frac{\partial \mathbf{f}(t, \mathbf{x}, \mathbf{u})}{\partial \mathbf{x}}\right|_{\tau, \hat{\mathbf{x}}_{k-1|k-1}, u_{k-1}} \tag{6.54}$$

$$\mathbf{H}_k = \left.\frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k|k-1}} \tag{6.55}$$

As Equações 6.49 e 6.51 são equações diferenciais não lineares que têm que ser numericamente integradas para obter as estimativas preditas das variáveis de estado e a matriz de covariância predita das estimativas. A Equação 6.51 que envolve a solução de n^2 equações diferenciais acopladas, pode ser computacionalmente custosa, mas pode ser evitada pelo cálculo de uma matriz de transição de estado, $\mathbf{A}_k$, que a cada tempo é baseada numa aproximação linear das funções não lineares, considerando-as constantes durante cada período de amostragem. Com essa aproximação adicional, a Equação 6.16 pode ser usada para obter a matriz de covariância predita. O método discreto aqui representa uma das várias diferentes abordagem para desenvolver técnicas de estimativas recursivas que podem ser encontradas em Muske e Edgar (1997).

EXEMPLO 6.4 Um reator tanque de mistura contínua (CSTR) com trocador de calor externo (LIEBMAN *et al.*, 1992), e no qual ocorre uma reação exotérmica de primeira ordem (decomposição do reagente *A*) é usado para ilustrar a aplicação da estimativa de estado para um processo não linear. As equações diferenciais descrevendo a variação na concentração

do reagente A e a temperatura do reator são dados por:

$$\frac{dA}{dt} = \frac{q}{V}(A_0 - A) - kA$$

$$\frac{dT}{dt} = \frac{q}{V}(T_0 - T) - \frac{\Delta H_R A_r}{\rho C_p T_r} kA - \frac{UA_R}{\rho C_p V}(T - T_c)$$

onde A_0 e T_0 são a concentração e a temperatura na alimentação e A e T são a concentração e a temperatura no reator, respectivamente. As variáveis da concentração e da temperatura são escalonadas usando os fatores A_r e T_r. A constante da taxa de reação é dada por:

$$k = k_0 \exp \frac{-E_A}{TT_r}$$

Os valores de todos os parâmetros são listados na Tabela 6.2. A concentração e a temperatura do reator no estado estacionário são 0,1531 e 4,6091, respectivamente. A concentração e a temperatura no reator para esse processo são amostradas em um período de 2,5 s e os desvios padrão dos erros aleatórios nestas medidas são 0,0077 e 0,2305 (5% nos valores nos estados estacionários), respectivamente.

Tabela 6.2: Valores dos parâmetros para o CSTR (NARASIMHAN; JORDACHE, 2000)

Parâmetro	**Valor**	**Unidade**
q	10,0	cm^3/s
V	1000,0	cm^3
ΔH_R	-27000	cal/mol
ρ	0,001	g/cm^3
C_p	1,0	$cal/(molK)$
U	$5{,}0 \times 10^{-4}$	$cal/(cm^2 sK)$
A_R	10,0	cm^3
T_c	340,0	K
k_0	$7{,}86 \times 10^{12}$	s^{-1}
E_a/R	14090,0	K
A_0	6,5	–
T_0	3,5	–
A_r	$1{,}0 \times 10^{-6}$	mol/cm^3
T_r	100	K

A resposta em malha aberta deste processo é simulada para uma variação em degrau na concentração da entrada de 6,5 para 7,5 e um filtro de Kalman estendido é usado para estimar a concentração e a temperatura do reator. Na implementação do filtro de Kalman estendido, as estimativas de estado preditas e a matriz de covariância predita dos erros em cada instante de amostragem são obtidos pela integração das equações diferenciais 6.49 e 6.51 usando o método de Runge-Kutta de quarta ordem. Os valores verdadeiros medidos e estimados da

concentração e da temperatura no reator são mostradas nas Figuras 6.5 e 6.6.

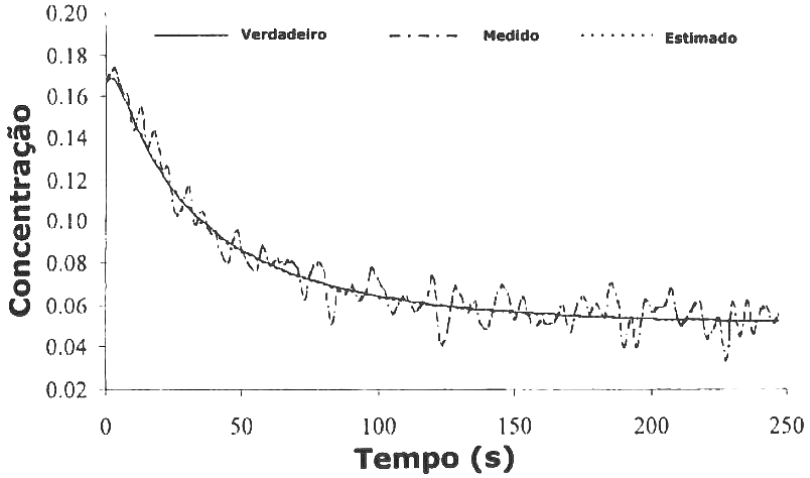


Figura 6.5: Concentração estimada de um CSTR usando um filtro de Kalman Estendido – adaptado de Narasimhan e Jordache (2000)

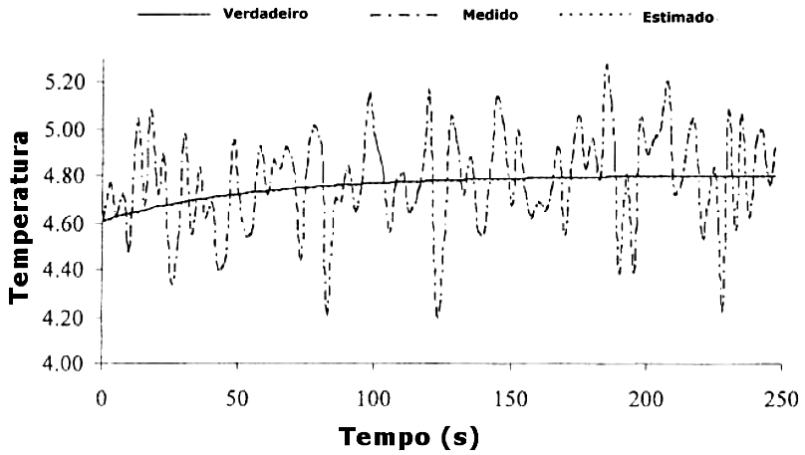


Figura 6.6: Temperatura estimada de um CSTR usando um filtro de Kalman Estendido – adaptado de Narasimhan e Jordache (2000)

Pode-se observar que os valores estimados estão bem próximos ao verdadeiro (nas figuras, os valores estimados praticamente coincidem com os valores verdadeiros). As variâncias dos erros nas medidas de concentração e temperatura do reator, calculadas a partir dos dados amostrados em um período total de tempo de 250 s são $5,9 \times 10^{-5}$ e 0,0534 respectivamente. Em comparação, as variâncias dos erros na concentração e temperatura estimadas são somente $1,36 \times 10^{-7}$ e $2,52 \times 10^{-8}$ respectivamente.

6.6.2 Métodos de reconciliação de dados não linear

Foi demonstrado anteriormente que o filtro de Kalman é equivalente à reconciliação de dados se as equações de transição de estado não são corrompidas por ruídos ou perturbações aleatórias. Uma progressão similar do filtro não linear para a reconciliação de dados é possível se for retirado o termo do ruído aleatório na Equação 6.45. Contudo, existem outras diferenças importantes na formulação e solução de problemas de reconciliação de dados não linear dinâmica quando comparada com os problemas de filtro não linear. Liebman *et al.* (1992) e posteriormente Ramamurthi *et al.* (1993) formularam o problema de reconciliação de dados dinâmica não linear e também propuseram estratégias de solução. A definição geral do problema tal como proposto por Liebman *et al.* (1992) pode ser formulada como:

$$\min_{\mathbf{x}} N = \sum_{j=t_0}^{t_N} (\mathbf{y}_j - \mathbf{x}_j)^{\mathrm{T}} \mathbf{Q}_j^{-1} (\mathbf{y}_j - \mathbf{x}_j) + \sum_{j=t_0}^{t_N} (\mathbf{u}_{cj} - \mathbf{u}_j)^{\mathrm{T}} \mathbf{Q}_{uj}^{-1} (\mathbf{u}_{cj} - \mathbf{u}_j) \tag{6.56}$$

sujeito a:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}); \quad \mathbf{x}(t_0) = \hat{\mathbf{x}}_0 \tag{6.57}$$

$$\mathbf{h}(\mathbf{x}) = 0 \tag{6.58}$$

$$\mathbf{g}(\mathbf{x}) \leq 0 \tag{6.59}$$

Na formulação acima, as variáveis manipuladas $\mathbf{u}$ são incluídas como parte da função objetivo e são estimadas em cada passo de tempo, mesmo sendo tomadas como constantes dentro de cada intervalo de amostragem. Os valores calculados das entradas manipuladas, $\mathbf{u}_{cj}$, a cada tempo j, usando a Equação 6.9 ou qualquer outra lei de controle, são diferentes das entradas manipuladas reais do processo devido a erros inerentes aos atuadores. Desta forma, os valores calculados das entradas manipuladas servem como medidas e os valores verdadeiros destas variáveis têm que ser estimados.

Esta formulação é mais geral quando comparada com o modelo usado no filtro, no qual as entradas manipuladas são tomadas como sendo conhecidas exatamente e as variáveis de estado são pressupostamente medidas de forma direta (ou de modo equivalente, a matriz $\mathbf{H}_k$ é tomada como a matriz identidade). Isto não impõe qualquer limitação porque usando-se uma simples transformação, o problema pode continuar a ser formulado como colocado acima.

Se uma medida é uma função não linear de variáveis de estado, pode-se então introduzir uma nova variável de estado artificial correspondente a essa medida e a relação não linear entre a variável de estado artificial e a verdadeira variável de estado pode ser incluída como parte das restrições de igualdade na Equação 6.58. Esta transformação é similar ao tratamento de variáveis indiretamente medidas na reconciliação de dados em estado estacionário. Por último, as restrições de desigualdade (Equação 6.59) permitem a inclusão de limites nas variáveis de estado ou outras restrições de factibilidade. Deve-se notar que métodos de filtragem não podem tratar restrições de desigualdade e podem dar lugar, portanto, a estimativas infactíveis com uma probabilidade de ocorrência muito maior que os métodos de reconciliação de dados linear em estado estacionário. Assim, a formulação dada pela Equação 6.56 até 6.59 é extremamente geral e de uso prático.

A formulação geral do problema de reconciliação de dados dinâmica não linear tem uma contrapartida negativa: não é mais possível desenvolver uma técnica de solução recursiva como na abordagem por filtro. Além disso, a análise da função objetivo revela que todas as variáveis de estado a partir do tempo inicial até o tempo corrente estão sendo simultaneamente estimadas a cada instante de amostragem. Isto leva a um número sempre crescente de variáveis que têm que ser estimadas com o tempo, o quê não é aceitável na prática.

Com o objetivo de reduzir o esforço computacional, uma abordagem de **janela móvel** foi adotada, (LIEBMAN *et al.*, 1992; RAMAMURTHI *et al.*, 1993). Nesta abordagem, a cada tempo t, somente uma janela de medidas a partir do tempo $t - N$ até o tempo t é usada para estimar todas as variáveis de estado dentro desta janela de tempo de dimensão N. A função objetivo a ser minimizada é a soma ponderada do quadrado das diferenças entre as estimativas de estado e as medidas dentro desta janela de tempo. As estimativas obtidas para as variáveis de estado no tempo t a partir desta otimização são usadas para calcular as entradas manipuladas. O procedimento é repetido no próximo instante de amostragem, o que dá origem do termo "janela móvel".

A estratégia para a solução do problema de estimativa a cada tempo na abordagem da janela móvel requer algumas explicações devido à presença de equações diferenciais não lineares (Equações 6.57 em conjunto com as Equações algébricas 6.58 e 6.59. Liebman *et al.* (1992) converteram as equações diferenciais em restrições algébricas de igualdade através da discretização por *colocação ortogonal*. Nesta técnica as funções das variáveis de estado (no tempo) dentro de cada período de amostragem são expressas com a soma ponderada dos valores das variáveis de estado em diferentes instantes de tempo, dentro deste período de amostragem, representando os pontos nodais de colocação. Os pesos usados nesta representação são os polinômios ortogonais. Ainda que um

período de amostragem possa ser subdividido em vários elementos, por conveniência é usado um elemento por intervalo de amostragem. Com esta escolha, as funções das variáveis de estado dentro de cada intervalo de amostragem j podem ser escritas como:

$$\mathbf{x}(t) = \sum_{i=1}^{n_c} l_i(t)\mathbf{x}_i^j \tag{6.60}$$

onde $l_i(t)$ são os polinômios da base ortogonal, n_c é a ordem e $\mathbf{x}_i^j$ são os valores das variáveis de estado no i-ésimo ponto de colocação no intervalo de amostragem j. Os pontos extremos do intervalo, 1 e n_c, correspondem aos instantes de amostragem. Usando a Equação 6.60, as derivadas também podem ser expressas em termos dos valores das variáveis de estado em diferentes instantes de tempo. As Equações 6.57 podem agora ser forçadas a satisfazer a todos os pontos de colocação resultando nas seguintes equações algébricas para cada intervalo de amostragem j.

$$\mathbf{D}\mathbf{x}^j - Ts(\mathbf{x}^j, \mathbf{u}_{j-1}) \tag{6.61}$$

onde $\mathbf{x}^j$ é o vetor de todas as variáveis de estado em todos os pontos de colocação no intervalo de amostragem j. A Equação 6.61 pode ser escrita para cada um dos N intervalos de amostragem na janela escolhida com a consideração adicional de que os valores das variáveis no fim de um intervalo de amostragem sejam iguais àqueles no início do próximo intervalo. Uma técnica de otimização não linear como GRG ou SQP pode ser usada para minimizar a Equação 6.56 sujeita às Equações 6.58, 6.59 e 6.61. Deve-se notar que o número de variáveis neste problema de otimização é maior que o número de variáveis de estado e de entrada nos N instantes de amostragem dentro da janela pois também se está estimando simultaneamente as variáveis de estado nos pontos de colocação dentro de cada intervalo (NARASIMHAN; JORDACHE, 2000). Liebman *et al.* (1992) dá detalhes sobre o polinômio ortogonal usado, a dimensão do problema e a estrutura da matriz de derivadas **D**.

Para reduzir o esforço computacional necessário pela estratégia de programação não linear descrita acima, Ramamurthi *et al.* (1993) propuseram um método do horizonte de estimativa sucessivamente linearizado no qual as Equações 6.57 e 6.58 são linearizadas em torno de uma trajetória de referência dada para as variáveis de estado. Os valores de referência a cada instante de amostragem j são usados para obter a forma linearizada destas equações para o período j de amostragem. Se não são incluídas restrições de desigualdade, então pode ser obtida uma solução analítica para as

estimativas das variáveis de estado no início da janela de tempo, a qual é então usada para integrar numericamente as equações diferenciais para obter as estimativas de estado em outros instantes de amostragem dentro da janela. Mesmo sendo este método eficiente, ele pode dar lugar a estimativas infactíveis porque não pode tratar restrições de desigualdade.

Na discussão acima, não foram explicitamente incluídas as variáveis não medidas ou os parâmetros como parte das equações do modelo. Os métodos de programação não linear podem também ser usados para estimar simultaneamente tanto a variável de estado medida quanto os parâmetros não medidos. As estimativas simultâneas do estado e dos parâmetros em processos dinâmicos foi considerada por Kim *et al.* (1990, 1991), que se referem a isto como estimativa pelo método do erro nas variáveis (*EVM – Error in the variables method*).

Resumidamente, as estratégias de reconciliação de dados dinâmica têm várias vantagens sobre as técnicas clássicas de filtro, como foi discutido aqui, mas não lidam com o problema do ruído aleatório nas equações de estado, os quais podem ser causados por perturbações não medidas no processo. Estas técnicas são também computacionalmente mais pesadas porque não apresentam uma forma recursiva do estimador.

EXEMPLO 6.5 O CSTR não isotérmico descrito no Exemplo 6.4 é usado para ilustrar a aplicação da técnica de reconciliação de dados não linear dinâmica. Foram simuladas as medidas correspondentes à resposta em malha aberta deste processo para uma variação em degrau na concentração de alimentação de 6,5 para 7,5 no tempo inicial, assim como no Exemplo 6.4. Usando uma janela de 10 períodos de amostragem, é aplicada uma técnica de reconciliação de dados não linear dinâmica para estimar a concentração e a temperatura no reator. Foram impostos limites inferiores e superiores para a concentração (0,01 e 0,2) e para a temperatura (4,0 e 5,0). Como é realizada uma simulação em malha aberta neste exemplo, a função objetivo de reconciliação de dados é a soma ponderada quadrática das diferenças entre os valores medidos e estimativas sobre os últimos 10 períodos de amostragem, ou seja, o segundo termo na Equação 6.56 não está presente. A otimização a cada instante t de amostragem é feita através das estimativas iniciais arbitrárias da concentração e temperatura no tempo $t-10$. As equações diferenciais que descrevem o CSTR são integradas do tempo $t-10$ até o tempo t usando o método de Runge-Kutta de quarta ordem para obter as estimativas das variáveis de estado em todos os instantes de amostragem dentro deste período de tempo. O valor da função objetivo é calculado para estas estimativas e as estimativas de estado no tempo $t-10$ são iteradas até que um valor mínimo da função objetivo seja obtido sujeito às restrições de limites inferiores e superiores sobre as estimati-

vas iniciais. Esta abordagem difere do método Ramamurthi *et al.* (1993) no sentido de que as equações diferenciais não lineares não são linearizadas, mas explicitamente integradas. Deve-se notar que nesta abordagem são impostos limites somente sobre as estimativas de estado no início da janela de tempo e é possível que as estimativas de estados em outros instantes de amostragem obtidos pela integração explícita possam violar os limites. Contudo, esta abordagem tem a vantagem de ser mais eficiente que o método de Liebman *et al.* (1992). A concentração e as temperaturas estimadas obtidas desta forma são mostradas nas Figuras 6.7 e 6.8, respectivamente. Pode-se observar que os estados estimados estão bem próximos dos valores verdadeiros. Para assegurar a convergência do problema de otimização em cada tempo, verificou-se que foi necessária a imposição de limites sobre as variáveis. Quando se compara estas estimativas com as obtidas usando o filtro de Kalman Estendido, no Exemplo 6.4, percebe-se que o filtro de Kalman dá resultados melhores e é também mais eficiente computacionalmente. Assim é melhor usar a reconciliação de dados não linear dinâmica somente quando as técnicas de filtro de Kalman estendido não dão estimativas que satisfaçam aos limites sobre as variáveis.

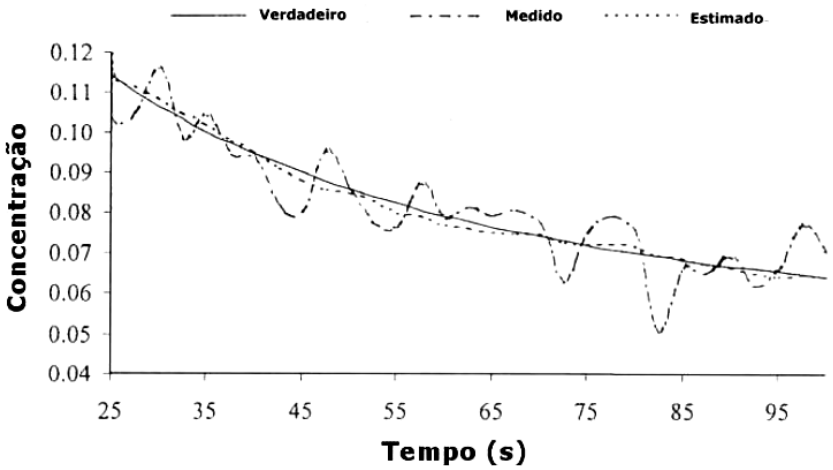


Figura 6.7: Concentração estimada do CSTR usando reconciliação de dados dinâmica com um comprimento de janela móvel de 10 pontos – adaptado de Narasimhan e Jordache (2000)

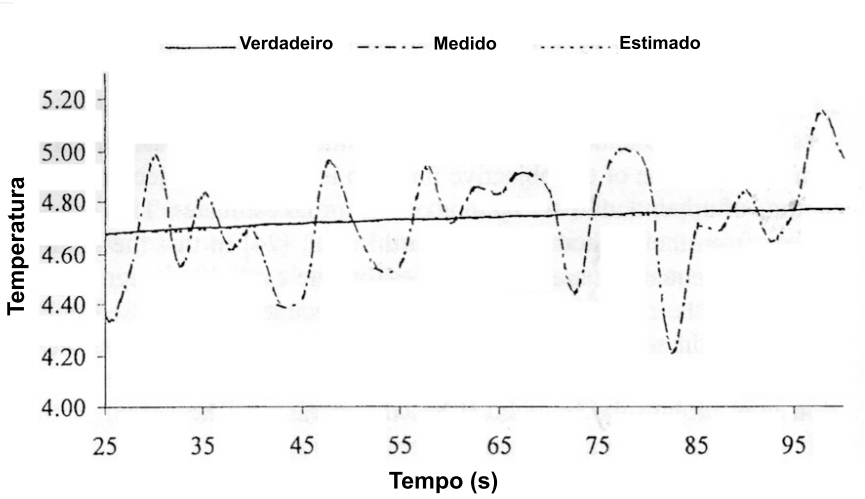


Figura 6.8: Temperatura estimada do CSTR usando reconciliação de dados dinâmica com um comprimento de janela móvel de 10 pontos – adaptado de Narasimhan e Jordache (2000)

6.7 Conclusões

Dentre os pontos que foram vistos neste capítulo, se destaca a importância que a reconciliação de dados tem para aplicações de controle de processos, pois o uso de estados estimados no lugar das medidas pode levar a um controle mais eficiente. Foi visto também que para explorar a redundância temporal dos dados, foram usados modelos dinâmicos que descrevem o comportamento das variáveis de estado em conjunto com as medições.

A técnica mais abrangente apresentada foi a do filtro de Kalman, que pode ser usado para estimar variáveis de estado em sistemas dinâmicos não lineares. Se perturbações nas variáveis de estado são ignoradas, então o filtro de Kalman é equivalente à reconciliação de dados. Além disso, a estimativa de estados em sistemas dinâmicos não lineares pode ser realizada usando um filtro de Kalman estendido ou suas variantes, mas esses métodos não tratam restrições de factibilidade sobre as variáveis. Por outro lado, métodos de otimização não linear podem ser usados para a reconciliação dinâmica de dados em processos não lineares. Estes métodos podem contabilizar restrições de desigualdade, mas são menos eficientes que os filtros de Kalman estendidos.

7 *Detecção de Erros Grosseiros*

Neste capítulo são apresentadas técnicas de detecção e identificação de erros grosseiros. Estas técnicas têm uma importância tal, que dependendo do ambiente industrial em que se apliquem, podem trazer resultados que são usados de uma forma mais direta e sem restrições que os resultados da reconciliação e coaptação de dados. O capítulo inicia com a definição do problema e sua contextualização para em seguida apresentar os testes estatísticos básicos usados na detecção de erros grosseiros, fazendo uma comparação entre eles. Depois, é apresentado o teste baseado em componentes principais aplicado tanto aos resíduos do balanço quanto aos ajustes às medidas, fazendo no fim uma comparação entre o teste baseado em componentes principais e os testes básicos. Na seqüência, é tratado o problema do desacoplamento entre variáveis medidas e não medidas. Logo após, são apresentadas técnicas de identificação de erros grosseiros como a estratégia da eliminação serial, as estratégias combinatoriais e a identificação por componentes principais. Finalmente, é abordado o problema de detectabilidade e identificabilidade.

7.1 Definição do problema

A técnica de reconciliação de dados se apóia na hipótese da presença exclusiva de erros aleatórios nos dados e da ausência de erros sistemáticos tanto nas medidas quanto nas equações do modelo. Se essa hipótese é inválida, a reconciliação pode levar a grandes ajustes sobre os valores medidos e as estimativas resultantes podem ser bastante inexatas e por vezes infactíveis. Assim, é importante identificar tais erros sistemáticos ou grosseiros antes das estimativas reconciliadas finais serem obtidas. Parte das técnicas de detecção de erros grosseiros estão intimamente relacionadas com as técnicas de reconciliação de dados, sendo que algumas delas pressupõem uma dada seqüencia de aplicação.

Nos Capítulos 2 e 3, foi dito que a reconciliação pode ser realizada somente se restrições estão

presentes. O mesmo é válido para a detecção de erros grosseiros. Sem a disponibilidade de restrições para a verificação das medidas, a detecção de erros grosseiros não pode ser realizada. Portanto, a reconciliação de dados e a detecção de erros grosseiros exploram a mesma informação disponível a partir das medidas e das restrições, estando relacionadas e interagindo no processamento de dados.

Existem dois tipos principais de erros grosseiros, como indicado no Capítulo 2. Um está relacionado ao desempenho do instrumento e inclui viéses de medida, tendências, descalibração e falha total de instrumentação. O outro é relacionado ao modelo de restrições e inclui perdas não contabilizadas de material e energia resultante de vazamentos em equipamentos de processo ou inexatidões de modelagem devidas a parâmetros igualmente inexatos. Várias técnicas foram propostas e desenvolvidas para a detecção e eliminação destes dois tipos de erros grosseiros.

Segundo Narasimhan e Jordache (2000), qualquer estratégia abrangente de detecção de erros grosseiros deve possuir as seguintes capacidades, que estão também apresentadas de forma condensada na Figura 7.1:

- Habilidade de detectar a presença de um ou mais erros grosseiros entre os dados (problema de detecção);
- Habilidade de identificar o tipo e a localização do erro grosseiro (problema de identificação);
- Habilidade de localizar e identificar múltiplos erros grosseiros que possam estar presentes simultaneamente entre os dados (problema da identificação de múltiplos erros grosseiros);
- Habilidade de estimar a magnitude dos erros grosseiros (problema de estimativa).

Nem todas as técnicas de detecção de erros grosseiros podem preencher os requesitos propostos acima. O último dos requesitos, ainda que útil, não é absolutamente necessário. Uma estratégia de detecção de erros grosseiros pode ser analisada em termos dos métodos componentes que esta usa para atacar os três principais problemas: detecção, identificação de erros grosseiros singulares e múltiplos, sendo que o desempenho de uma estratégia é função direta destes métodos componentes.

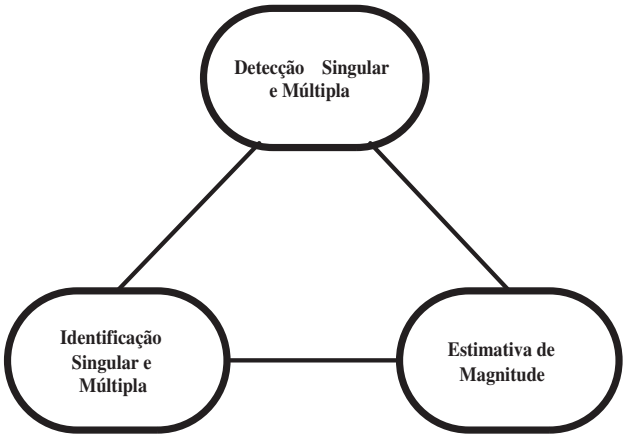


Figura 7.1: Estratégia de detecção e identificação de erros grosseiros

7.2 Testes estatísticos básicos para a detecção de erros grosseiros

Este componente da estratégia de detecção de erros grosseiros é devotado à questão da presença ou ausência de erros grosseiros entre os dados. Como todos os métodos de detecção se utilizam direta ou indiretamente do fato que a presença de erros grosseiros entre as medidas causa a violação dos balanços do modelo, então se as medidas não contêm quaisquer erros aleatórios, a violação de qualquer restrição do modelo por um valor medido pode ser imediatamente interpretada como devida à presença de erros grosseiros. Este é um método puramente determinístico. Já foi pressuposto, contudo, que todas as medidas contêm erros aleatórios, devido aos quais não se pode esperar que as medidas satisfaçam estritamente qualquer restrição do modelo, mesmo se erros grosseiros estão ausentes. Assim, é atribuída uma tolerância à violação das restrições devido a erros aleatórios. Uma abordagem probabilística é usada para resolver este problema sob uma distribuição de probabilidades pré-definida para os erros aleatórios.

O princípio básico na detecção de erros grosseiros é derivado da detecção de valores espúrios (*outliers*) em aplicações estatísticas. O erro aleatório inerente a qualquer medida segue por hipótese uma distribuição normal com média zero e variância conhecida. O *erro normalizado*, que é a diferença entre o valor medido e o valor médio esperado, dividido pelo seu desvio padrão, segue uma distribuição normal padrão. A maioria dos erros normalizados está dentro de um intervalo de confiança $(1-\alpha)$ a um nível de significância arbitrário α. Qualquer valor (um erro normalizado) que recaia fora deste intervalo de confiança é declarado como um *outlier* ou um erro grosseiro.

Um grande número de testes estatísticos é derivado deste princípio básico e estes são capazes de detectar erros grosseiros, mas nem todos testes são capazes de identificar os diferentes tipos e a localização destes erros. Alguns testes básicos são capazes de detectar somente erros de medidas (viéses). Outros testes só podem detectar vazamentos ou erros no modelo do processo. Por outro lado, o teste da razão de verossimilhança generalizada, o qual é derivado do princípio da estimativa por máxima verossimilhança em estatística, pode ser usado tanto para detectar problemas de instrumentação como vazamentos no processo.

As técnicas estatísticas mais comumente usadas para detecção de erros grosseiros são baseadas no teste de hipótese. Em um caso de detecção de erro grosseiro, a hipótese nula, H_0, é a de que não há erros grosseiros presentes e a hipótese alternativa, H_1, é a de que há um ou mais erros grosseiros entre os dados. Todas as técnicas estatísticas para a escolha entre estas duas hipóteses fazem uso de uma estatística de teste que é função das medidas e do modelo de restrições. A estatística de teste é comparada com um valor pré-determinado de referência e a hipótese nula é rejeitada ou aceita, dependendo se esta estatística excede o valor de referência ou não. O valor de referência é também conhecido como critério de teste, valor crítico ou valor crítico de teste (NARASIMHAN; JORDACHE, 2000).

O resultado de um teste de hipótese não é perfeito. Um teste estatístico pode declarar a presença de erros grosseiros quando, de fato, não há nenhum erro grosseiro presente (H_0 é verdadeiro). Neste caso, o teste comete um **Erro do Tipo I**, ou levanta um *alarme falso*. Por outro lado, o teste pode detectar a medida como livre de erros, quando de fato esta contém um ou mais erros. Este é um **Erro do Tipo II**. A probabilidade deste tipo de erro é igual a β. A **potência de um teste estatístico**, que corresponde à probabilidade da correta detecção, é igual a probabilidade de $1-\beta$. A potência e a probabilidade do Erro do Tipo I em qualquer teste estatístico estão intimamente relacionadas. Enquanto que a probabilidade de um Erro do Tipo I de um teste é descrita por um único número, o nível de significância α, a potência depende do quanto a hipótese nula desvia da realidade. Se é possível medir esse desvio de alguma forma, a dependência de β do desvio é chamada de característica operacional do teste. Se as hipóteses podem ser testadas por dois ou mais testes com o mesmo nível de significância, o teste que dá o menor valor de probabilidade de um Erro do Tipo II é declarado o teste mais forte ou de **Máxima Potência** (*MP – Maximum Power Test*).

A potência de um teste estatístico pode ser aumentada através da permissão para uma probabilidade maior para o Erro do Tipo I. Portanto, quando se projeta um teste estatístico, a potência do

teste deve ser balanceada face à probabilidade de falsa detecção. Se a distribuição de probabilidade da estatística de teste pode ser obtida sob a consideração da hipótese nula, então o critério de teste pode ser selecionado de modo que a probabilidade do Erro do Tipo I seja menor ou igual a um valor especificado α. O parâmetro α é também chamado de nível de significância para o teste estatístico.

Doravante, neste capítulo, serão considerados testes estatísticos para a detecção de erros grosseiros considerando-se modelos lineares e operação em estado estacionário.

Assumindo-se um modelo de restrições lineares dado por:

$$\mathbf{Ax} = \mathbf{c} \tag{7.1}$$

onde $\mathbf{A}$ é a matriz das restrições lineares e o vetor $\mathbf{c}$ contém coeficientes conhecidos. Em geral, para fluxos lineares, $\mathbf{c}$ é o vetor nulo, a menos que algumas das variáveis sejam conhecidas exatamente. Como nos capítulos anteriores, os erros nas medidas seguem por hipótese uma distribuição normal e uma matriz de covariância, Ψ, conhecida.

Quatro testes estatísticos básicos foram desenvolvidos e amplamente aplicados para a detecção de erros grosseiros. Para simplificar a descrição destes testes, um modelo linear com todas as variáveis medidas será inicialmente proposto.

Os dois primeiros testes são baseados no vetor dos resíduos de balanço, $\mathbf{r}$, o qual é descrito por:

$$\mathbf{r} = \mathbf{Ay} - \mathbf{c} \tag{7.2}$$

Na ausência de erros grosseiros, o vetor $\mathbf{r}$ segue uma distribuição normal multivariada com média zero e uma matriz de covariância, $\mathbf{V}$, dada por:

$$\mathbf{V} = \mathrm{cov}(\mathbf{r}) = \mathbf{A}\Psi\mathbf{A}^{\mathrm{T}} \tag{7.3}$$

Portanto, sob H_0, $\mathbf{r} \sim N(\mathbf{0}, \mathbf{V})$. Na presença de erros grosseiros, os elementos do vetor residual $\mathbf{r}$ refletem o grau de violação das restrições do processo (leis de conservação de massa e energia). Por outro lado, a matriz $\mathbf{V}$ retém informações sobre a estrutura do processo vidas da matriz $\mathbf{A}$, e sobre a matriz de covariância, Ψ. As duas quantidades, $\mathbf{r}$ e $\mathbf{V}$, podem ser usadas para construir

estatísticas para detectar a presença de erros grosseiros.

7.2.1 O teste global (*GT – global test*)

O teste global usa o teste descrito por:

$$\gamma = \mathbf{r}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{r} \tag{7.4}$$

Sob H_0, a estatística acima segue uma distribuição χ^2 com ν graus de liberdade, com ν igual ao posto da matriz **A**. Se o critério de teste é escolhido como $\chi^2_{1-\alpha,\nu}$, onde $\chi^2_{1-\alpha,\nu}$ é o valor crítico da distribuição χ^2 a um dado nível de significância α, então H_0 é rejeitado e um erro grosseiro é detectado se $\gamma \geq \chi^2_{1-\alpha,\nu}$. Esta escolha do critério de teste assegura que a probabilidade de Erro do Tipo I para este teste é menor ou igual a α. O teste global combina todos os resíduos dos balanços na obtenção da estatística de teste e assim dá lugar a um teste multivariado ou coletivo.

Para esclarecer a implementação deste teste é considerado um sistema de troca de calor com *bypass* (NARASIMHAN; JORDACHE, 2000), mostrado na Figura 7.2.

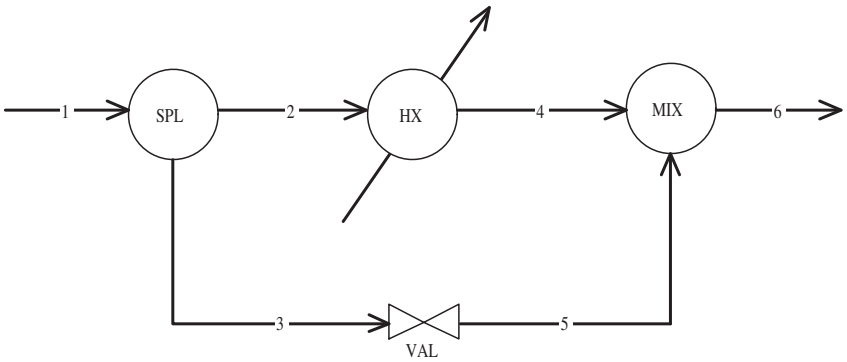


Figura 7.2: Sistema de troca de calor com *bypass* – adaptado de Narasimhan e Jordache (2000)

EXEMPLO 7.1 Considerando-se a reconciliação sobre os fluxos do trocador de calor ilustrado na Figura 7.2, assume-se que todos os fluxos sejam medidos e os valores verdadeiros, os medidos e os reconciliados (desconsiderando-se a presença de erros grosseiros) estão

Tabela 7.1: Reconciliação de dados com a presença de erros grosseiros para o processo da Figura 7.2 – primeiro caso

Número da Corrente	Valores Verdadeiros dos Fluxos	Valores Medidos dos Fluxos	Valores Reconciliados
1	100	100,91	100,89
2	64	68,45	65,83
3	36	34,65	35,05
4	64	64,20	65,83
5	36	36,44	35,05
6	100	98,88	100,89

dispostos na Tabela 7.1. A matriz das restrições para este processo é dada por:

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}$$

onde as linhas correspondem aos balanços materiais (fluxos) para o *splitter*, o trocador de calor, a válvula de *bypass* e o *mixer* e as colunas correspondem às seis correntes. Os resíduos das restrições, dado pela Equação 7.2 para as dadas medidas são iguais a:

$$\mathbf{r} = \begin{bmatrix} -2,19 & 4,25 & -1,79 & 1,76 \end{bmatrix}$$

e a matriz covariância dos resíduos das restrições, através de 7.3, é igual a

$$\mathbf{V} = \begin{bmatrix} 3 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 3 \end{bmatrix}$$

Usando-se a Equação 7.4, a estatística do teste global é calculada como 15,2942. Isto pode ser verificado como sendo igual à soma dos quadrados das diferenças entre os valores medidos e reconciliados. O critério de teste, a um nível de significância de 5%, para uma distribuição χ^2 com 4 graus de liberdade é igual a 9,4877. Desta forma o teste global rejeita a hipótese nula, H_0, e um erro grosseiros é detectado (NARASIMHAN; JORDACHE, 2000).

7.2.2 Teste nodal ou teste da restrição (*NT – nodal test*)

O vetor **r** também pode ser usado para derivar estatísticas de teste, uma para cada restrição *i*, na forma:

$$z_{r,i} = \frac{|r_i|}{\sqrt{V_{ii}}}, \quad i = 1,2,\ldots,m \tag{7.5}$$

ou escrevendo na forma vetorial:

$$\mathbf{z}_r = [diag(\mathbf{V})]^{-1/2}\,\mathbf{r} \tag{7.6}$$

onde diag(**V**) é a matriz diagonal cujos elementos são V_{ii}. O teste nodal ou teste da restrição usa as estatísticas de teste $z_{r,i}$ para a detecção de erros grosseiros. $z_{r,i}$ segue uma distribuição normal padrão $N(0,1)$ sob H_0. Se qualquer uma das estatísticas $z_{r,i}$ excede o critério de teste $z_{1-\alpha/2}$, onde $z_{1-\alpha/2}$ é o valor crítico da distribuição normal padrão para um nível α de significância, então um erro grosseiro é detectado.

Ao contrário do teste global, o teste nodal processa o resíduo de cada restrição separadamente e dispõe *m* testes univariados. Como múltiplos testes são realizados usando o mesmo valor crítico, isto aumenta a probabilidade que um destes testes possa ser rejeitado mesmo se nenhum erro grosseiro estiver presente. Em outras palavras, a probabilidade do Erro do Tipo I será maior que o valor especificado de α. Se é desejável controlar a probabilidade de Erro do Tipo I, pode ser usado o nível de significância modificado β (Equação de Sidak), proposto por Mah e Tamhane (1982):

$$\beta = 1 - (1-\alpha)^{1/m} \tag{7.7}$$

Para qualquer valor especificado de α, o valor modificado de β pode ser calculado a partir da Equação 7.7 e o critério de teste para todos os testes nodais pode ser escolhido como $z_{1-\beta/2}$. Isto assegura que a probabilidade de que qualquer um dos testes nodais ser rejeitado sob H_0 é menor ou igual a α. Deve-se notar que α é apenas o limite superior da probabilidade de Erro do Tipo I e no sentido de se assegurar que a probabilidade de Erro do Tipo I seja exatamente igual a α, pode-se adotar um critério de teste por tentativa e erro através de simulação. De um modo alternativo, Rollins e Davis (1992) propuseram o uso de um valor crítico baseado no intervalo de confiança de Bonferroni, o qual é dado por:

$$\beta = \alpha / m \tag{7.8}$$

Para valores grandes de m, a Equação 7.7 se reduz à Equação 7.8.

É possível se obter outras formas do teste nodal através do uso de transformações lineares dos resíduos das restrições. Contudo, nem todas estas formas possuem a mesma potência de detecção de erros grosseiros. Crowe (1989b) obteve uma forma em particular do teste nodal que detém a propriedade da máxima potência (ver página 162). As estatísticas do teste nodal de máxima potência são dadas por:

$$z^*_{r,i} = \frac{\left| \left[\mathbf{V}^{-1} \mathbf{r} \right]_i \right|}{\sqrt{\left[\mathbf{V}^{-1} \right]_{ii}}} \tag{7.9}$$

ou, se escrevendo na forma vetorial:

$$\mathbf{z}^*_r = \left[diag(\mathbf{V}^{-1}) \right]^{-1/2} \mathbf{V}^{-1} \mathbf{r} \tag{7.10}$$

O critério de teste é escolhido como sendo o mesmo do teste nodal padrão. Se existir um erro grosseiro no processo, então o valor esperado do máximo entre as estatísticas do teste dadas pela Equação 7.9 é maior que o valor esperado do máximo entre as estatísticas do teste dadas pela Equação 7.5. Isto implica que se houver um erro grosseiro, então o teste nodal baseado na estatística da Equação 7.9 tem uma maior probabilidade de detectá-lo do que o teste baseado nas estatísticas da Equação 7.5. Se as estatísticas do teste nodal são derivadas usando qualquer outra transformação linear sobre os resíduos, estas não possuem esta propriedade. Assim, o teste nodal baseado nas estatísticas da Equação 7.9 tem a propriedade da **máxima potência - (MP)**.

EXEMPLO 7.2 Para o processo considerado no Exemplo 7.1, os resíduos das restrições e sua matriz de covariância foram calculados. A partir destes dados, as estatísticas do teste nodal são obtidas: $\begin{bmatrix} 0,687 & 3,0052 & 1,2657 & 1,0161 \end{bmatrix}$. O critério de teste em uma distribuição normal padrão, a um nível de significância de 5% é de 1,96, assim, somente o teste para o resíduo da restrição 2 é rejeitado, significando que um erro grosseiro foi detectado entre as medidas relacionadas com o nó 2 do processo (o trocador de calor) (NARASIMHAN; JORDACHE, 2000).

7.2.3 Teste da medida (*MT – measurement test*)

O terceiro teste é baseado no vetor dos ajustes às medidas:

$$\mathbf{a} = \mathbf{y} - \hat{\mathbf{x}} \tag{7.11}$$

onde $\hat{\mathbf{x}}$ é o vetor das estimativas reconciliadas obtidas a partir da Equação 3.73, por exemplo. Usando esta solução, os ajustes também podem ser escritos como:

$$\mathbf{a} = \Psi \mathbf{A}^{\mathrm{T}} \mathbf{V}^{-\mathbf{1}} \mathbf{r} \tag{7.12}$$

o qual segue, sob H_0, uma distribuição normal multivariada: $N(\mathbf{0}, \overline{\mathbf{W}})$, onde

$$\overline{\mathbf{W}} = \mathrm{cov}(\mathbf{a}) = \Psi \mathbf{A}^{\mathrm{T}} \mathbf{V}^{-\mathbf{1}} \Psi \tag{7.13}$$

A estatística

$$z_{a,j} = \frac{|a_j|}{\sqrt{\overline{W}_{jj}}}, \quad j = 1, 2, \ldots, n \tag{7.14}$$

conhecida como a estatística do teste da medida, segue uma distribuição normal padrão, $N(0,1)$ sob H_0. Tamhane (1982) mostrou que para uma matriz de covariância Ψ não diagonal, um vetor de estatísticas de teste com máxima potência para a detecção de um único erro grosseiro é obtido pela pré-multiplicação de $\mathbf{a}$ por Ψ^{-1}, o que resulta em:

$$\mathbf{d} = \Psi^{-1} \mathbf{a} \tag{7.15}$$

Sob H_0, $\mathbf{d}$ é também normalmente distribuído com média zero e uma matriz de covariância

$$\mathbf{W} = \mathrm{cov}(\mathbf{d}) = \mathbf{A}^{\mathrm{T}} (\mathbf{A} \Psi \mathbf{A}^{\mathrm{T}})^{-1} \mathbf{A}^{\mathrm{T}} \tag{7.16}$$

Mah e Tamhane (1982) propuseram a seguinte estatística:

$$z_{d,j} = \frac{|d_j|}{\sqrt{W_{jj}}}, \quad j = 1, 2, \ldots, n \tag{7.17}$$

conhecida como teste da medida de máxima potência, o qual segue uma distribuição normal padrão, $N(0,1)$ sob H_0. De modo similar ao teste nodal, o teste da medida também envolve múltiplos testes univariados. A probabilidade do Erro do Tipo I será menor ou igual a α se o critério do teste estatístico é dado por $Z_{1-\beta/2}$ onde β é dado pela Equação 7.7 ou 7.8 com m sendo substituído por n que é o número de testes da medida univariados.

EXEMPLO 7.3 A partir dos valores medidos e reconciliados listados na Tabela 7.1, os ajustes são calculados:

$$\begin{bmatrix} 1,0233 & 2,6167 & -0,4033 & -1,6333 & 1,3867 & -2,0067 \end{bmatrix}$$

A matriz de covariância dos ajustes às medidas é dada por:

$$\overline{\mathbf{W}} = \begin{bmatrix} 0,6667 & -0,1667 & -0,1667 & -0,1667 & -0,1667 & -0,3333 \\ -0,1667 & 0,6667 & 0,1667 & -0,3333 & 0,1667 & -0,1667 \\ -0,1667 & 0,1667 & 0,6667 & 0,1667 & -0,3333 & -0,1667 \\ -0,1667 & -0,3333 & 0,1667 & 0,6667 & 0,1667 & -0,1667 \\ -0,1667 & 0,1667 & -0,3333 & 0,1667 & 0,6667 & -0,1667 \\ -0,3333 & -0,1667 & -0,1667 & -0,1667 & -0,1667 & 0,6667 \end{bmatrix}$$

As estatísticas do teste da medida são então obtidas:

$$\begin{bmatrix} 1,2533 & 3,2047 & 0,494 & 2,0004 & 1,6983 & 2,4577 \end{bmatrix}$$

Devido à matriz de covariância dos erros nas medidas ser diagonal neste exemplo, as estatísticas de máxima potência para o tese da medida são também as mesmas. Para um nível de significância de 5%, o critério de teste normal padrão é 1,96. A partir destas informações os testes para as medidas 2, 4 e 6 são rejeitados. O nível de significância dado pela Equação 7.7 é igual a 0,0085, enquanto o baseado no intervalo de confiança de Bonferroni (Equação 7.8) é igual a 0,0083. Correspondendo a estes níveis de significância modificados, os critérios de teste são, respectivamente, 2,6315 e 2,6396. Assim, se forem usados os níveis de significância modificados, somente a medida 2 é rejeitada.

7.2.4 Teste da razão de verossimilhança generalizada (*GLR – generalized likelihood ratio test*)

O quarto teste para a detecção de erros grosseiros em estado estacionário é o teste da razão de verossimilhança generalizado, baseado no princípio estatístico da máxima verossimilhança. Em contraste com outros testes, a formulação deste teste requer um modelo do processo na presença de erros grosseiros, também conhecido como *modelo de erros grosseiros*. Este teste pode identificar diferentes tipos de erros grosseiros para os quais sejam fornecidos modelos.

O modelo de erros grosseiros para um viés de magnitude desconhecida b na medida j é dado por:

$$\mathbf{y} = \mathbf{x} + \boldsymbol{\varepsilon} + b\mathbf{e}_j \tag{7.18}$$

onde $\mathbf{e}_j$ é o vetor unidade com valor 1 na j-ésima posição e zero nas demais. Por outro lado, vazamentos de material devem ser modelados como parte das restrições, pois podem ser consideradas como "correntes". Um vazamento de fluxo mássico em um nó i do processo com uma magnitude b desconhecida, pode ser modelado como:

$$\mathbf{A}\mathbf{y} - b\mathbf{m}_i = \mathbf{c} \tag{7.19}$$

Os elementos do vetor $\mathbf{m}_i$ são relativamente fáceis de definir quando somente balanços com fluxos totais estão envolvidos. Se o vazamento vem de uma unidade i, então somente a restrição de fluxo para este vetor da unidade é afetado e assim $\mathbf{m}_i$ é idêntico a $\mathbf{e}_i$. Contudo, se as restrições também incluem balanços de componentes e balanços de energia (com valores conhecidos de composição e temperatura), então o vetor $\mathbf{m}_i$ só pode ser definido aproximadamente a partir de uma decisão arbitrária. Uma recomendação de Narasimhan e Mah (1987) é escolher os elementos de $\mathbf{m}_i$ da seguinte forma:

i. Correspondendo à restrição de fluxo de massa total da unidade i, $\mathbf{m}_i$ tem o valor de 1 na i-ésima posição

ii. Correspondendo à restrição de fluxo de energia associada ao nó i, o valor do i-ésimo elemento de $\mathbf{m}_i$ pode ser escolhido como a entalpia específica média das correntes que incidem no nó i.

O mesmo pode ser aplicado à restrição do fluxo de um componente para o nó *i*, substituindo a entalpia específica pela concentração

iii. Os elementos em $\mathbf{m}_i$ não associados com restrições sobre o nó *i*, são escolhidos como iguais a zero.

Perdas de energia ou de fluxos de componentes no nó *i* também podem ser modeladas pela Equação 7.19, escolhendo-se o elemento correspondente em $\mathbf{m}_i$ igual a 1 e todos os outros elementos iguais a zero.

É possível se derivar a distribuição estatística dos resíduos das restrições sob H_1 utilizando-se modelos de erros grosseiros quando um erro deste tipo está presente entre as medidas ou nas restrições. Os resíduos das restrições seguem uma distribuição normal sob H_0, com média zero e matriz de covariância dada pela Equação 7.3. Sob H_1, os resíduos das restrições continuam seguindo uma distribuição normal com matriz de covariância dada pela Equação 7.3, mas o valor esperado depende do tipo de erro grosseiro presente. Se um erro grosseiro devido a um viés de magnitude *b* está presente na medida *j*, então:

$$E[\mathbf{r}] = b\mathbf{A}\mathbf{e}_j \tag{7.20}$$

Por outro lado, se um erro grosseiro devido a um vazamento no processo está presente nas medidas no nó *i*, então:

$$E[\mathbf{r}] = b\mathbf{m}_i \tag{7.21}$$

Portanto, quando um erro grosseiro devido a um viés ou vazamento no processo estiver presente, pode-se definir:

$$E[\mathbf{r}] = b\mathbf{f}_k \tag{7.22}$$

onde

$$\mathbf{f}_k \begin{cases} \mathbf{A}\mathbf{e}_j & \text{para um viés na } i\text{-ésima medida} \\ \mathbf{m}_i & \text{para um vazamento no nó } i \end{cases} \tag{7.23}$$

Os vetores $\mathbf{f}_k$ são também conhecidos como **vetores de assinatura de erro grosseiro**. Definindo-se μ como o valor esperado desconhecido de $\mathbf{r}$, pode-se formular as hipóteses de detecção de erros grosseiros como:

$$\begin{aligned} H_0 &: \mu = \mathbf{0} \\ H_1 &: \mu = b\mathbf{f}_k \end{aligned} \tag{7.24}$$

onde H_0 é a hipótese nula que nenhum erro grosseiro esteja presente e H_1 é a hipótese alternativa de que esteja presente entre as medidas um viés ou um vazamento. A hipótese alternativa tem duas incógnitas, b e $\mathbf{f}_k$. O parâmetro b pode ser qualquer número real e $\mathbf{f}_k$ pode ser qualquer vetor do conjunto F dado por:

$$F = \{\mathbf{A}\mathbf{e}_j, \mathbf{m}_i : \ i = 1, \ldots, m \ : j = 1, \ldots, n\} \tag{7.25}$$

onde m é o número de nós ou unidades do processo e n é o número de variáveis medidas.

No sentido de testar as duas hipóteses dadas pela Equação 7.24, pode-se usar o teste da razão de verossimilhança. Este teste, para este caso é dado por:

$$\lambda = \sup \frac{P\{\mathbf{r}|H_1\}}{P\{\mathbf{r}|H_0\}} \tag{7.26}$$

onde $P\{\mathbf{r}|H_0\}$ e $P\{\mathbf{r}|H_1\}$ são as probabilidades de se obter o vetor residual $\mathbf{r}$ sob as hipóteses H_0 e H_1. O supremo[1] (*sup* na Equação 7.26) é calculado sobre todos os valores possíveis dos parâmetros presentes nas hipóteses. Usando-se uma função de densidade de probabilidade normal para $\mathbf{r}$, pode-se reescrever a Equação 7.26 como:

$$\lambda = \sup_{b,\mathbf{f}_k} \frac{e^{-0,5(\mathbf{r}-b\mathbf{f}_k)^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{r}-b\mathbf{f}_k)}}{e^{-0,5\mathbf{r}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{r}}} \tag{7.27}$$

como o lado direito da Equação 7.27 é sempre positivo, pode-se simplificar o cálculo através da escolha da estatística:

[1] O supremo (supremum) é o menor valor do limite superior de um dado conjunto, tal que nenhum membro do conjunto excede este valor.

$$T = 2\ln\lambda = \sup_{b,\mathbf{f}_k} \left[\mathbf{r}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{r} - (\mathbf{r} - b\mathbf{f}_k)^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{r} - b\mathbf{f}_k)\right] \tag{7.28}$$

O cálculo de T é realizado da seguinte forma: para qualquer vetor $\mathbf{f}_k$, calcula-se a estimativa b^* de b, que dá o supremo na Equação 7.28. Assim, obtém-se uma estimativa de máxima verossimilhança

$$b^* = (\mathbf{f}_k^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{f}_k)^{-1}(\mathbf{f}_k^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{f}_k) \tag{7.29}$$

substituindo b^* na Equação 7.28 e denotando o valor correspondente de T por T_k, tem-se:

$$T_k = d_k^2/C_k \tag{7.30}$$

onde

$$d_k = \mathbf{f}_k\mathbf{V}^{-1}\mathbf{r} \tag{7.31}$$

$$C_k = \mathbf{f}_k\mathbf{V}^{-1}\mathbf{f}_k \tag{7.32}$$

Este procedimento é realizado para cada vetor $\mathbf{f}_k$ no conjunto F e a estatística T é portanto obtida como:

$$T = \sup_k T_k \quad k = 1,\ldots,m+n \tag{7.33}$$

Seja $\mathbf{f}^*$ o vetor que leva ao supremo na Equação 7.33. A estatística T é comparada com um valor de referência T_{cr} pré-determinado e um erro grosseiro é detectado se T excede T_{cr}. Pode-se interpretar T_k como uma estatística para testar a presença do erro grosseiro k. Como T é o máximo entre os T_k, o teste *GLR* detecta um erro grosseiro se qualquer das estatísticas T_k exceder o valor crítico. Assim o teste *GLR*, como o teste da medida e o teste nodal, realiza múltiplos testes univariados para detectar um erro grosseiro. A distribuição de T_k, sob H_0, é uma distribuição χ^2 central com um grau de liberdade. Desta forma, com o objetivo de manter a probabilidade do Erro do Tipo I do teste *GLR* menor ou igual a um dado valor α, pode-se escolher um critério de teste

como $\chi^2_{1-\beta,1}$, o quantil $1-\beta$ superior da distribuição χ^2 com um grau de liberdade, onde β é dado por:

$$\beta = 1 - (1-\alpha)^{\frac{1}{m+n}} \tag{7.34}$$

EXEMPLO 7.4 Considerando-se erros grosseiros causados por viéses nas medidas para o processo analisado nos exemplos anteriores, o vetor de assinatura dos erros grosseiros para um viés numa medida *i* é a *i*-ésima coluna da matriz de restrições mostrada no Exemplo 7.1. As estatísticas do teste *GLR* calculadas a partir dos resíduos das restrições e sua matriz de covariância dada no Exemplo 7.2 são $\begin{bmatrix} 1,5708 & 10,2704 & 0,244 & 4,0017 & 2,8843 & 6,0401 \end{bmatrix}$. As estatísticas do teste *GLR* são o quadrado dos testes da medida de máxima potência, calculados no Exemplo 7.3. O critério de teste a um nível de significância de 5% e a níveis modificados de significância – Sidak e Bonferroni, são simplesmente o quadrado do critério de teste normal padrão (3,8415; 6,925 e 6,9676), respectivamente. Assim os testes *GLR* para as medidas 2, 4 e 6 são rejeitadas a um nível de significância de 5% enquanto que somente o teste para a medida 2 é rejeitado sob os níveis de significância modificados. Para testar a presença de vazamentos em todos os quatro nós os vetores de assinatura para estes quatro erros grosseiros são simplesmente os vetores unidade. As estatísticas do teste *GLR* para estes quatro erros grosseiros são dadas por: $\begin{bmatrix} 1,5708 & 13,2496 & 0,3844 & 6,0401 \end{bmatrix}$. Os testes *GLR* para vazamentos nos nós 2 e 4 são rejeitados a um nível de significância de 5% enquanto o teste para vazamento sob os níveis de significância modificados rejeitam somente o nó 2.

7.2.5 Comparação de potência entre os testes básicos para detecção de erros grosseiros

Como foi descrito nas seções anteriores, vários testes foram desenvolvidos para detecção de erros grosseiros entre as medidas causados por viéses nos instrumentos de medição ou erros grosseiros nas restrições de conservação em estado estacionário devido a vazamentos desconhecidos. Para que se obtenha o melhor desempenho, é importante aplicar o teste que tenha a máxima potência (a máxima probabilidade de detectar a presença de um erro grosseiro quando de fato há um) sem aumentar a probabilidade de aumentar o Erro do Tipo I.

Assim, uma questão importante que pode se feita é qual, dentre os quatro testes demonstrados

anteriormente, o que confere a máxima potência para detecção de um único erro grosseiro entre os dados. A maioria dos trabalhos que comparam o desempenho de diferentes estratégias de detecção de erros grosseiros considera somente o desempenho global, que inclui todos os componentes da detecção, identificação e da detecção de múltiplos erros, mas não compara os componentes da detecção na estratégia de maneira isolada. São apresentados a seguir, alguns resultados que respondem parcialmente a esta questão (NARASIMHAN; JORDACHE, 2000).

Fazendo esta comparação, é necessário considerar somente os testes nodal de máxima potência, o teste da medida de máxima potência, o teste global e o teste da razão de verossimilhança generalizada. Pode-se ainda simplificar a tarefa fazendo uso dos resultados teóricos que foram derivados por Crowe (1989b) e por Narasimhan (1990) para mostrar que o teste *GLR* é mais poderoso que um teste da medida de máxima potência, baseado em qualquer transformação linear singular ou não singular dos resíduos das restrições, para a detecção de um único erro grosseiro.A prova deste resultado segue:

Pode ser observado, a partir das Equações 7.12, 7.15 e 7.17 que as estatísticas do teste das medidas de máxima potência são obtidos usando uma transformação linear dos resíduos da restrição. Considerando-se a raiz quadrada positiva das estatísticas dos testes *GLR*, pode-se então mostrar que as estatísticas do teste *GLR* são também obtidas usando uma transformação linear dos resíduos da restrição.

Portanto o teste nodal de máxima potência, o teste da medida e o teste *GLR* derivam todos eles estatísticas baseadas em uma transformação linear dos resíduos da restrição. Para mostrar qual transformação linear dos resíduos da restrição resulta nos testes mais poderosos, considera-se uma transformação linear arbitrária sobre os resíduos da restrição dada por:

$$\mathbf{r}^* = \mathbf{Y}\mathbf{r} \tag{7.35}$$

Seja um erro grosseiro de magnitude b, quer seja devido a um viés na medida ou a um vazamento presente, com o vetor de erros grosseiros correspondente $\mathbf{f}_k$. Usando a Equação 7.22, o valor esperado dos resíduos da restrição transformados é obtido como:

$$\mathrm{E}[\mathbf{r}^*] = b\mathbf{Y}\mathbf{f}_k \tag{7.36}$$

A matriz de covariância dos resíduos restritos é dado por:

$$\text{cov}(\mathbf{r}^*) = \mathbf{V}^* = \mathbf{Y}\mathbf{V}\mathbf{Y}^{\mathrm{T}} \tag{7.37}$$

Um teste pode ser então observado baseado nos resíduos restritos transformados com uma estatística dada por:

$$z_i^* = \frac{|r_i^*|}{\sqrt{V_{ii}^*}} \tag{7.38}$$

A Equação 7.38 também pode ser escrita como:

$$z_i^* = \frac{\left|\mathbf{e}_i^{\mathrm{T}}\mathbf{r}^*\right|}{\sqrt{\mathbf{e}_i^{\mathrm{T}}\mathbf{V}^*\mathbf{e}_i}} \tag{7.39}$$

Pode-se verificar que fazendo-se $\mathbf{Y}$ como $\mathbf{V}^{-1}$, $\mathbf{A}^{\mathrm{T}}\mathbf{V}^{-1}$, ou $\mathbf{F}^{\mathrm{T}}\mathbf{V}^{-1}$ (onde $\mathbf{F}$ é uma matriz de vetores $\mathbf{f}_k$ definidos pela Equação 7.23), são obtidas, respectivamente, as estatísticas do teste nodal de máxima potência, do teste da medida da máxima potência ou o teste *GLR*. Para mostrar que o teste *GLR* tem a máxima potência, precisa-se provar que dentre os valores esperados para a estatística do teste *GLR*, o máximo é maior ou igual ao valor esperado de qualquer das estatísticas de testes dadas pela Equação 7.39. Para isto, primeiro se prova que o máximo entre os valores esperados para a estatística do teste *GLR* é alcançado por T_k, ou seja:

$$\mathrm{E}\left[\sqrt{T_k}\right] \geq \mathrm{E}\left[\sqrt{T_i}\right] \tag{7.40}$$

Onde a partir das Equações 7.30 e 7.35 através de 7.37 com $\mathbf{Y} = \mathbf{F}^{\mathrm{T}}\mathbf{V}^{-1}$, os valores esperados de $\sqrt{T_k}$ e $\sqrt{T_i}$ são respectivamente:

$$\mathrm{E}\left[\sqrt{T_k}\right] = b\sqrt{\mathbf{f}_k^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{f}_k} \tag{7.41}$$

$$\mathrm{E}\left[\sqrt{T_i}\right] = b\sqrt{\mathbf{f}_i^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{f}_i} \tag{7.42}$$

Os resultados acima podem ser estabelecidos usando a desigualdade de Cauchy-Schwartz:

$$\left|\mathbf{v}^{\mathrm{T}}\mathbf{w}\right| \leq (\mathbf{v}^{\mathrm{T}}\mathbf{v})^{\frac{1}{2}}(\mathbf{w}^{\mathrm{T}}\mathbf{w})^{\frac{1}{2}} \tag{7.43}$$

definindo os vetores **v** e **w** como:

$$\mathbf{v} = \mathbf{R}\mathbf{f}_i, \quad \text{e} \quad \mathbf{w} = \mathbf{R}\mathbf{f}_k \tag{7.44}$$

onde **R** é uma matriz tal que $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{V}^{-1}$

Para provar que $\mathrm{E}\left[\sqrt{T_k}\right] \geq \mathrm{E}[z_i^*]$ para qualquer i, primeiro é necessário definir $\mathrm{E}[z_i^*]$, o qual de acordo com as Equações 7.39, 7.36 e 7.37 é:

$$\mathrm{E}[z_i^*] = \frac{b\mathbf{e}_i^{\mathrm{T}}\mathbf{Y}\mathbf{f}_k}{\mathbf{e}_i^{\mathrm{T}}\mathbf{Y}\mathbf{V}\mathbf{Y}^{\mathrm{T}}\mathbf{e}_i} \tag{7.45}$$

Então pode-se novamente fazer uso da desigualdade de Cauchy-Schwartz pela definição das matrizes.

$$\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{V}^{-1} \quad \text{e} \quad \mathbf{P} = \mathbf{Y}\mathbf{R}^{-1} \tag{7.46}$$

e identificando os vetores **v** e **w** como:

$$\mathbf{v} = \mathbf{R}\mathbf{f}_k \quad \text{e} \quad \mathbf{w} = \mathbf{P}^{\mathrm{T}}\mathbf{e}_i \tag{7.47}$$

Como os testes nodal de máxima potência e da medida de máxima potência são obtidos usando uma transformação linear particular dos resíduos da restrição, pode-se afirmar, baseados nos resultados acima, que na média o teste *GLR* tem a máxima potência em detectar um único erro grosseiro. Considerando-se erros grosseiros oriundos somente de viéses nas medidas, o teste *GLR* torna-se idêntico ao teste da medida de máxima potência.

Por outro lado, se estão presentes apenas erros grosseiros que afetam uma só restrição (por exemplo, vazamento no balanço global de fluxo), então o teste *GLR* torna-se idêntico ao teste nodal de máxima potência. Contudo, quando há a possibilidade dos dois tipos de erros grosseiros estarem presentes no sistema, então o teste *GLR* é mais poderoso que os outros dois testes. Apesar destes resultados, deve-se notar que há uma consideração implícita nesta derivação de que se conhece

precisamente os vetores $\mathbf{f}_k$ dos erros grosseiros para diferentes tipos de erros que podem ocorrer no processo. Esta consideração pode não ser válida se houver incertezas no modelo de distribuição ou no modelo de erros grosseiros. Além disso, todos estes resultados somente são válidos sob a hipótese de presença de no máximo um erro grosseiro.

Agora, é necessário somente examinar, dentre o teste global e o *GLR*, qual detém a máxima potência na detecção da presença de um único erro grosseiro. Nota-se a partir da Equação 7.4 que o teste global é também baseado nos resíduos das restrições e assim usa a mesma informação que o teste *GLR* na detecção dos erros grosseiros, mas há uma diferença fundamental na forma como esta informação é processada. O teste global realiza um único teste multivariado na detecção, enquanto que o teste *GLR* realiza vários testes univariados, um para cada possível erro grosseiro, para a detecção de qualquer um deles que esteja presente. A questão é qual desses esquemas de processamento confere a máxima potência. Este problema tem sido objeto de estudo na literatura de estatística, mas é difícil obter uma resposta singular a esta questão de maneira teórica.

Podem ser feitos estudos de simulações em processos selecionados para avaliar a potência dos dois testes. Antes de se tentar uma comparação como esta, no entanto, deve-se assegurar que ambos os testes têm a mesma probabilidade para o Erro Tipo I. Isto implica que o critério para cada teste tem que ser escolhido para dar um valor específico para a probabilidade do Erro do Tipo I. Isto é possível somente no caso do teste global. Como já foi explicado como o teste *GLR* realiza múltiplos testes univariados, o critério de teste tem que ser escolhido por tentativa e erro usando simulação.

Os resultados obtidos desta simulação podem ser usados, na melhor das hipóteses, para se obter algumas conclusões de sentido amplo mas não podem se generalizadas para todos os processos.

Foi observado, a partir da discussão anterior, que um teste não tem uma potência uniformemente maior que outro, mas é recomendável que para o propósito da detecção da presença de um ou mais erros grosseiros, que o teste global seja usado baseado nas seguintes considerações

i. O cálculo da estatística do teste global é mais eficiente, pois é calculado somente um teste estatístico;

ii. Na prática, para incutir confiança entre os operadores do processo, é preciso manter a probabilidade do alarme falso abaixo de um determinado limite. O critério de teste para o teste global pode ser escolhido para manter este limite. Qualquer valor mais alto para o critério de teste pode satisfazer este limite, mas isto resulta numa potência menor para o teste. No caso

do teste *GLR*, o valor mais baixo que satisfaz este limite pode ser escolhido somente através de simulações;

iii. O teste *GLR* requer conhecimento sobre os vetores de erros grosseiros para os diferentes erros que podem ocorrer no processo para obtenção da estatística do teste. O teste global não precisa de qualquer informação acerca dos erros grosseiros para a detecção. Isto pode ser uma consideração de ordem prática importante, pois o conhecimento completo sobre a sorte de erros grosseiros possíveis de ocorrer em um processo não é, de um modo geral, uma informação disponível.

Pode ser questionada a qualidade do teste global, pois este apenas indica a presença ou ausência de erros grosseiros, sendo necessária uma estratégia para a detectar a natureza e a localização do erro (No caso de teste *GLR*, a estatística do teste pode ser diretamente usada para identificação como será descrito na Seção 7.5). Esta questão pode ser esclarecida por duas considerações.

O uso do teste global para a detecção não exclui o uso do teste *GLR* na identificação do tipo e localização do erro grosseiro. Neste caso é necessário construir a estatística de teste do *GLR* somente se o erro grosseiro foi detectado pelo teste Global. Em segundo lugar, será demonstrada que a estratégia de identificação inerente ao teste *GLR* é a técnica da eliminação serial padrão que foi proposta e primeiro usada por Ripps (apud NARASIMHAN; JORDACHE, 2000), em combinação com o teste global, para a identificação de viéses entre as medidas. Assim, o teste *GLR* para a detecção e identificação de um erro grosseiro singular pode ser visto como uma estratégia de detecção que tem como um dos seus componentes o teste global para detecção e uma eliminação serial para identificação.

7.3 Detecção de erros grosseiros usando teste de componentes principais

As matrizes de covariância dos resíduos de balanço $\mathbf{V}$ e dos ajustes às medidas ($\overline{\mathbf{W}}$ ou $\mathbf{W}$) são sempre densas. Isto implica que, mesmo que as medidas sejam independentes ou fracamente correlacionadas, os dados reconciliados são sempre fortemente correlacionados. Os valores reconciliados e, portanto, os ajustes às medidas são correlacionados porque são amarrados pelo modelo do processo e o mesmo é verdadeiro para os resíduos dos balanços.

Contudo nem todos os testes básicos exploram a informação completa contida nas matrizes $\mathbf{V}$, $\overline{\mathbf{W}}$ ou $\mathbf{W}$. O teste nodal simples e o teste da medida univariado (de máxima potência ou não) descritos anteriormente usam somente os termos diagonais destas matrizes. Como alternativa, os testes de componentes principais usam as matrizes inteiras. A análise de componentes principais (PCA – *principal component analysis*) é uma ferramenta efetiva na análise de dados multivariados. Ela transforma um conjunto de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas, conhecidas como componentes principais. Cada componente principal é uma combinação linear das variáveis originais. Os coeficientes de cada combinação linear são obtidos a partir de um autovetor da matriz de covariância das variáveis originais.

Espera-se que estes testes sejam capazes de detectar erros grosseiros mais sutis pois são testes multivariados. Testes multivariados como o teste global detectam erros grosseiros que não são detectados por testes univariados. Este aspecto é muito importante porque a falha na detecção de todos os erros grosseiros pode resultar no fracasso da reconciliação de dados (as estimativas reconciliadas são infactíveis ou questionáveis). Os testes de componentes principais são relacionados com os testes nodal e da medida univariados, porque usam uma transformação linear do vetor dos resíduos do balanço ou das medidas.

A seguir tem-se uma breve descrição dos testes de componentes principais dada por Tong e Crowe (1995). Como nos testes anteriores, esta análise é restrita a modelos lineares com todas as variáveis medidas. O caso com variáveis não medidas pode ser resolvido com uma matriz de projeção. Dois tipos principais de testes de componentes principais podem ser derivados, tal como se segue.

7.3.1 Teste de componentes principais para resíduos de balanço do processo

Considerando um conjunto de combinações lineares do vetor $\mathbf{r}$ (resíduos do balanço, dado pela Equação 7.2)

$$\mathbf{p_r} = \mathbf{W}_r^{\mathrm{T}}\mathbf{r} \tag{7.48}$$

onde as colunas de $\mathbf{W}_r$ são os autovetores de $\mathbf{V}$ que satisfazem

$$\mathbf{W}_r = \mathbf{U}_r\Lambda_r^{-1/2} \tag{7.49}$$

A matriz Λ_r é diagonal e consiste nos autovalores de $\mathbf{V}$, λ_{ri}, $i = 1 \ldots q$, na sua diagonal e satisfaz

$$\Lambda_r = \mathbf{U}_r^{\mathrm{T}} \mathbf{V} \mathbf{U}_r \tag{7.50}$$

A matriz $\mathbf{U}_r$ consiste dos autovetores ortonormalizados de $\mathbf{V}$ de modo que

$$\mathbf{U}_r \mathbf{U}_r^{\mathrm{T}} = \mathbf{I} \tag{7.51}$$

O vetor $\mathbf{p}_r$ consiste dos componentes principais dos resíduos de balanço e os seus elementos são os *scores* dos componentes principais (TONG; CROWE, 1995).

Se nenhum erro grosseiro estiver presente, então $\mathbf{r} \sim N(\mathbf{0}, \mathbf{V})$ e $\mathbf{p}_r \sim N(\mathbf{0}, \mathbf{I})$. Portanto, um conjunto de variáveis correlacionadas, $\mathbf{r}$, é transformado em um novo conjunto de variáveis não correlacionadas, $\mathbf{p}_r$. Os componentes principais são numerados em ordem descendente das magnitudes dos autovalores correspondentes.

Por outro lado, as Equações 7.48 e 7.49 podem ser combinadas e reescritas como:

$$\mathbf{r} = \mathbf{U}_\mathbf{r} \Lambda_r^{1/2} \mathbf{p}_r \tag{7.52}$$

o que significa que o vetor residual $\mathbf{r}$ pode ser unicamente reconstruído a partir de seus componentes principais se todos eles forem retidos, ou seja, $\mathbf{p}_\mathbf{r} \in \Re^m$, onde m é o número de equações (resíduos do balanço). Contudo, se menos que m componentes são retidos, tem-se que:

$$\mathbf{r} = \mathbf{U}_\mathbf{r} \Lambda_r^{1/2} \mathbf{p}_\mathbf{r} + (\mathbf{r} - \hat{\mathbf{r}}) \tag{7.53}$$

onde:

$$\hat{\mathbf{r}} = \mathbf{U}_r \Lambda_r^{1/2} \mathbf{p}_r \tag{7.54}$$

Com $\mathbf{p}_r \in \Re^k$ e $k < m$. A Equação 7.54 é chamada de modelo de componentes principais do vetor $\mathbf{r}$. A Equação 7.53 indica que os resíduos no vetor $\mathbf{r}$ podem ser decompostos nas contribuições vindas do termo dos componentes principais e os resíduos do modelo de componentes principais, $\mathbf{r} - \hat{\mathbf{r}}$. Isto significa que para a detecção de erros grosseiros, ao invés de usar testes estatísticos

para $\mathbf{r}$, pode-se realizar um teste de hipótese sobre $\mathbf{p_r}$ e $\mathbf{r} - \hat{\mathbf{r}}$. Como cada elemento do vetor $\mathbf{p}_r$ é distribuído segundo uma normal padrão, uma regra de detecção semelhante ao teste nodal univariado pode ser usada e o teste para o i-ésimo resíduo do balanço nodal é rejeitado se $p_{r,i}$ excede $Z_{1-\beta/2}$. Como ocorre nos testes univariados, para limitar o Erro do Tipo I a um nível α, o β pode ser escolhido como na Equação 7.7, na qual o expoente m seja substituído pelo número de componentes principais retidos, k.

7.3.2 Teste de componentes principais sobre ajustes às medidas

De um modo semelhante à estatística de teste de componentes principais baseada nos resíduos dos balanços, a estatística de teste de componentes principais das medidas pode ser definida como:

$$p_{ai} = \left(\mathbf{W}_a^{\mathrm{T}}\mathbf{a}\right)_i \quad i = 1 \ldots k \tag{7.55}$$

onde as colunas de $\mathbf{W}_a$ são os autovetores de $\overline{\mathbf{W}}$ e k é o número de componentes principais retidos. De um modo geral, $k < n$, onde n é o número de medidas.

Se não houver erros grosseiros presentes entre as medidas, então $\mathbf{p_a} \sim N(\mathbf{0}, \mathbf{I})$ e portanto os componentes principais dos ajustes às medidas são também não correlacionados. Como no teste da medida, pode-se então conduzir um teste sobre cada p_{ai} comparando-o com um valor limite $Z_{1-\beta/2}$.

EXEMPLO 7.5 Aplicando o teste de componentes principais baseado nos ajustes às medidas sobre o processo considerado nos exemplos anteriores. Os autovalores não nulos da matriz $\mathbf{W}$ calculados no Exemplo 7.3 são todos iguais a unidade. A matriz $\mathbf{U}_a$ cujas colunas são os autovetores normalizados correspondentes é dada por:

$$\mathbf{U}_a = \begin{bmatrix} -0,7475 & 0,1152 & -0,0067 & 0,0287 \\ 0,4077 & -0,6881 & 0,0500 & -0,4232 \\ 0,3843 & -0,4449 & -0,5698 & 0,2554 \\ -0,0167 & 0,4956 & -0,6011 & -0,0597 \\ 0,0067 & 0,2523 & 0,0188 & -0,7382 \\ 0,3564 & 0,0773 & 0,5578 & 0,4542 \end{bmatrix}$$

Quatro componentes principais foram retidos: $\begin{bmatrix} -0,5317 & -2,1181 & 0,2422 & -3,0187 \end{bmatrix}$. A um nível de significância de 5%, os testes para os componentes principais 2 e 4 são re-

jeitados enquanto que nos níveis de significância modificados somente o teste para o último componente principal foi rejeitado.

7.3.3 Relação entre testes de componentes principais e os testes estatísticos básicos

Os testes de componentes principais são também baseados em uma transformação linear dos resíduos de balanço como na Equação 7.35. Pode-se verificar que a matriz de transformação usada na derivação do teste dos componentes principais das restrições é $\mathbf{Y} = \mathbf{W}_r^{\mathrm{T}}$ e que para o teste dos componentes principais das medidas, a matriz de transformação é $\mathbf{Y} = \mathbf{W}_a^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{V}^{-1}$. Tong e Crowe (1995) indicaram que como o número de componentes principais retidos é geralmente menor que o número de componentes principais, o nível modificado de significância para os testes de componentes principias é menor. Portanto, seria esperado que o Erro do Tipo I seja reduzido, mas não há fundamento nesta expectativa, porque a probabilidade do Erro do Tipo I para qualquer teste pode sempre ser reduzida pela escolha de um valor menor de α. Além disso, porque os testes de componentes principias não identificam diretamente os erros grosseiros, é possível que a estratégia usada na identificação cometa Erros do Tipo I adicionais.

De modo análogo ao teste global, um teste global coletivo, baseado em componentes principais pode também ser proposto, para o qual a estatística de teste é definida por:

$$\gamma_k = \mathbf{p_r}^{\mathrm{T}}\mathbf{p_r} \tag{7.56}$$

Outro teste coletivo importante é definido por:

$$Q_r = (\mathbf{r} - \hat{\mathbf{r}})^{\mathrm{T}}(\mathbf{r} - \hat{\mathbf{r}}) \tag{7.57}$$

Conhecida como a estatística Q ou erro de predição quadrático ou ainda como a estatística de Rao. Pode ser demonstrado que Q_r é a soma ponderada dos quadrados dos últimos $m-k$ componentes principais

$$Q_r = \sum_{i=k+1}^{m} \lambda_i \mathbf{p}_i^2 \tag{7.58}$$

As duas quantidades, γ_k, e Q_r são complementares. A primeira examina os componentes principais retidos e a última os não retidos de modo coletivo. γ_k contabiliza a variância explicada pelo modelo de componentes principias, enquanto que Q_r contabiliza a variância não explicada. Os testes baseados nestas quantidades podem ser conduzidos para verificação da presença de erros grosseiros entre os componentes principais retidos ou não retidos.

Como já foi colocado anteriormente, a maior diferença entre os testes univariados e os testes multivariados χ^2 é que os primeiros não contabilizam a correlação entre os resíduos e assim tornam-se menos confiáveis quando esta correlação cresce. Contudo o teste *GLR* (ou teste da medida de máxima potência) e o teste nodal de máxima potência incorporam a correlação pela transformação dos resíduos usando a inversa da matriz de covariância. Isto conduz à máxima potência para a correta detecção de um erro grosseiro sobre todos os outros testes, mas apenas quando houver um erro grosseiro singular.

Quando múltiplos erros grosseiros estão presentes, estes testes não detêm a máxima potência. Tong e Crowe (1995) indicam que os testes multivariados de componentes principais não apenas fornecem uma melhor detecção de erros grosseiros sutis, mas têm maior potência na correta identificação das variáveis com erros sobre os outros testes.

7.4 Testes estatísticos para modelos gerais em estado estacionário

Nas seções anteriores, diferentes testes estatísticos para detecção de erros grosseiros foram descritos e implementados para o caso mais simples no qual todas as variáveis são medidas diretamente. De um modo geral, como já foi colocado no Capítulo 3, podem estar presentes variáveis não medidas e/ou as medidas podem ser indiretamente relacionadas às variáveis. Narasimhan e Mah (1989) descreveram transformações simples através das quais os modelos gerais em estado estacionário podem ser convertidos no modelo simples acima mencionado. Usando-se estas transformações, todos os testes estatísticos podem ser derivados como descrito a seguir.

Se variáveis não medidas estão presentes, o modelo de restrições é descrito por:

$$\mathbf{A}_1\mathbf{x} + \mathbf{A}_2\mathbf{u} = \mathbf{c} \tag{7.59}$$

onde $\mathbf{x} : n \times 1$ é o vetor das variáveis medidas, $\mathbf{u} : p \times 1$ é o vetor das variáveis não medidas e $\mathbf{A}_2$ tem o posto das colunas completo, p.

Como mostrado na Seção 3.6, as variáveis não medidas podem ser eliminadas pela pré-multiplicação das restrições por uma matriz de projeção $\mathbf{P} : (m-p) \times m$, de posto $m-p$, onde m é o número de restrições, de modo a gerar as restrições reduzidas:

$$\mathbf{PA}_1\mathbf{x} = \mathbf{Pc} \tag{7.60}$$

Os resíduos das restrições para o conjunto reduzido de restrições é definido em uma forma exatamente análoga à Equação 7.2:

$$\rho = \mathbf{P}(\mathbf{A}_1\mathbf{y} - \mathbf{c}) \tag{7.61}$$

A matriz de covariância do vetor ρ é

$$\mathbf{V}_\rho = \mathrm{cov}(\rho) = \mathbf{PA}_1\Psi(\mathbf{PA}_1)^{\mathrm{T}} \tag{7.62}$$

As estatísticas para os testes global, nodal e da medida podem ser obtidas usando $\mathbf{PA}_1$, ρ e $\mathbf{V}_\rho$ no lugar de $\mathbf{A}$, $\mathbf{r}$ e $\mathbf{V}$, respectivamente, nas equações apropriadas. Para se derivar estatísticas para os testes *GLR*, os vetores de assinatura de erros grosseiros para viéses e vazamentos são também transformados pela matriz de projeção.

Este vetor de assinatura transformado é dado por:

$$\mathbf{f}_{\rho k} = \mathbf{Pf}_k \tag{7.63}$$

onde $\mathbf{f}_k$ é dado pela Equação 7.23. As estatísticas do teste *GLR* são então obtidas usando as Equações 7.30 até a 7.33, substituindo-se $\mathbf{f}_{\rho k}$, $\mathbf{V}_\rho$ e ρ por $\mathbf{f}_k$, $\mathbf{V}$ e $\mathbf{r}$, respectivamente. O teste *GLR* é um teste de máxima potência para a detecção de erros grosseiros únicos mesmo em meio à variáveis não medidas.

Em alguns casos, as medidas não podem ser diretamente relacionadas às variáveis como na Equação 3.2. Por exemplo, na Equação 3.19, onde a queda de pressão é relacionada ao quadrado da taxa de fluxo. Outro exemplo é a relação entre uma medida de pH e a concentração de íons

de hidrogênio e a temperatura do processo. Estas relações são tipicamente não lineares, mas por simplicidade serão representadas aqui por equações lineares

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \varepsilon \tag{7.64}$$

Assumindo-se as restrições sejam dadas por

$$\mathbf{A}\mathbf{x} = \mathbf{c} \tag{7.65}$$

Define-se as variáveis artificiais $\mathbf{x_a}$ como

$$\mathbf{x_a} = \mathbf{D}\mathbf{x} \tag{7.66}$$

Então a Equação 7.64 se torna

$$\mathbf{y} = \mathbf{x_a} + \varepsilon \tag{7.67}$$

As Equações 7.65 e 7.66 podem ser escritas em conjunto como:

$$\begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{x}_a + \begin{bmatrix} \mathbf{A} \\ -\mathbf{D} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0} \end{bmatrix} \tag{7.68}$$

As Equações 7.67 e 7.68 representam um modelo alternativo equivalente do processo no qual as variáveis $\mathbf{x}$ seriam "variáveis não medidas" e as variáveis $\mathbf{x}_a$ seriam as "variáveis diretamente medidas". Portanto, o método descrito para o tratamento de variáveis não medidas pode ser usado para derivar estatísticas para todos os testes.

A técnica descrita acima pode ser aplicada mesmo quando as medidas são relacionadas às variáveis através de equações não lineares. Contudo, as equações de restrição modificadas serão não lineares e assim técnicas de reconciliação e de detecção de erros grosseiros não lineares terão que ser usadas para resolução destes processos.

7.5 Técnicas para identificação de erros grosseiros

O segundo componente de uma estratégia de detecção de erros grosseiros trata do problema de identificar corretamente o tipo e a localização de um erro grosseiro detectado por um determinado teste. O problema de identificação somente surge a partir da rejeição da hipótese nula. Nem todos os testes de detecção descritos nas seções anteriores são projetados para distinguir os diferentes tipos de erros grosseiros. Somente o teste *GLR* é adequado para tanto, porque usa informação concernente ao efeito de cada tipo de erro grosseiro sobre o modelo do processo.

Para comparar as diferentes técnicas desenvolvidas em conjunto com os diferentes testes para identificação de erros grosseiros e depreender o relacionamento entre elas, esta seção se restringirá a considerações sobre erros causados por viéses nas medidas com a presença de somente um único erro grosseiro. Neste caso, o problema de identificação se reduz a simplesmente identificar corretamente a medida que contém o erro grosseiro.

As técnicas para identificação da medida contendo o erro vão de uma regra simples até estratégias complexas, dependendo do teste que é usado. O teste da medida e o teste *GLR*, em virtude do modo como são derivadas suas estatísticas, usam uma regra simples para identificar o erro grosseiro. O teste *GLR* e o teste da medida de máxima potência são idênticos considerando-se erros grosseiros causados somente por viéses. Neste caso, há uma estatística correspondente a cada medida. A regra de identificação usada nestes testes pode ser formulada como:

> *Identificar o erro grosseiro na medida que corresponda à máxima estatística que exceda ao critério de teste*

Devido à simplicidade da regra acima, diz-se que os testes *GLR* e da medida não necessitam de uma estratégia propriamente dita para identificação de erros grosseiros.

7.5.1 Estratégia da eliminação serial

Quando se usa o teste global para a detecção de erros grosseiros, uma estratégia comparativamente mais complexa tem que ser aplicada para a identificação da medida com o erro. Ripps (apud NARASIMHAN; JORDACHE, 2000) foi quem primeiro delineou um procedimento que foi posteriormente estudado e refinado por Serth e Heenan (apud NARASIMHAN; JORDACHE, 2000) e Rosenberg

et al. (apud NARASIMHAN; JORDACHE, 2000) e que é conhecido como **procedimento de eliminação serial**. Nele, cada medida é eliminada por vez, e a estatística do teste global é recalculada. Quando se elimina uma medida, faz-se a variável correspondente passar por não medida, assim o teste global tem que ser recalculado usando o conjunto deduzido dos resíduos das restrições (Seção 7.4).

Devido ao aumento no número de variáveis não medidas, o valor da função objetivo, e por conseguinte a estatística do teste global, vai diminuir. Ripps (apud NARASIMHAN; JORDACHE, 2000) sugeriu que o erro grosseiro pode ser identificado naquela medida cuja eliminação leva à maior redução na função objetivo.

Ao invés de resolver o problema de reconciliação de dados repetidamente ou calcular a projeção de matrizes para a eliminação de cada medida por vez, Crowe (apud NARASIMHAN; JORDACHE, 2000) derivou expressões simplificadas para a redução do valor da função objetivo da reconciliação de dados devido à eliminação da i-ésima medida. Esta expressão é dada por:

$$\Delta N_i = N - N_i = \frac{(\mathbf{e}_i \mathbf{A}^{\mathrm{T}} \mathbf{V}^{-1} \mathbf{r})^2}{\mathbf{e}_i \mathbf{A}^{\mathrm{T}} \mathbf{V}^{-1} \mathbf{e}_i} \tag{7.69}$$

A redução no valor da função objetivo através da eliminação da i-ésima medida é igual à estatística do teste *GLR* (ou o quadrado da estatística do teste da medida) para a variável i. Isto implica que se a regra usada em conjunto com o teste global for "*Identificar o erro grosseiro na medida que dá o máximo* ΔJ_i", então esta é precisamente a mesma regra usada no teste *GLR* (ou no teste da medida de máxima potência) para identificação do erro grosseiro na medida correspondente à máxima estatística do teste. Ou seja, o teste global combinado com a estratégia de eliminação serial é equivalente ao teste *GLR*.

Um outro resultado útil é obtido da interpretação do princípio do teste *GLR* quanto ao valor da função objetivo da reconciliação de dados. Considerando-se a Equação 7.28, que define a estatística do teste *GLR*, os dois termos dentro dos parênteses, no lado direito da equação, podem ser interpretados como os valores ótimos para os problemas de reconciliação de dados. Já foi observado que o primeiro termo nesta expressão é o valor ótimo da função objetivo para o problema da reconciliação de dados padrão. O segundo termo é o valor ótimo da função objetivo do problema de reconciliação, no qual a estimativa do erro grosseiro na i-ésima medida é obtida como parte da solução, qual seja:

Problema I:

$$\min_{x,b}(\mathbf{y}-\mathbf{x}-b\mathbf{e}_i)^{\mathrm{T}}\boldsymbol{\Psi}^{-1}(\mathbf{y}-\mathbf{x}-b\mathbf{e}_i)$$

sujeito a

$$\mathbf{A}\mathbf{x}=\mathbf{c}$$

Assim, a estatística do teste *GLR* também pode se interpretada como a máxima diferença entre o ótimo da função objetivo da reconciliação de dados, sob a hipótese da ausência de erros, e o ótimo da função objetivo obtido pela solução do Problema I. No Problema I, todas as medidas são retidas, mesmo que a *i*-ésima medida seja apontada como contendo um erro grosseiro. Ao invés da eliminação desta *i*-ésima medida, todas as medidas são usadas para obter uma estimativa do erro grosseiro. Baseado nesta interpretação e no resultado da Equação 7.69, pode-se concluir que os valores ótimos da função objetivo são iguais, independente da eliminação da *i*-ésima medida ou de sua manutenção e obtenção de uma estimativa para o erro grosseiro.

Tabela 7.2: Redução na estatística do teste global com a eliminação da *i*-ésima medida

Medida Eliminada	Redução na Estatística do Teste Global
1	1,571
2	10,27
3	0,244
4	4,002
5	2,884
6	6,040

EXEMPLO 7.6 A partir dos resultados dos Exemplos 7.3 e 7.4, observa-se que quando se escolhe identificar um erro grosseiro na medida correspondente à máxima estatística do teste, então o erro grosseiro na medida 2 é corretamente identificada tanto pelo teste da medida de máxima potência quanto pelo teste *GLR*. De fato, mesmo considerando-se erros grosseiros devidos a vazamentos, a maior estatística do teste *GLR* corresponde a um viés na medida 2. O teste global também rejeita a hipótese nula e então pode se utilizar a estratégia de eliminação de variáveis para identificar a localização do erro grosseiro. A Tabela 7.2 mostra a redução na estatística do teste global quando diferentes medidas são eliminadas. Como a maior redução na estatística do teste global é obtida quando a medida 2 é eliminada, o procedimento de eliminação em conjunto com o teste global identificou o erro corretamente.

7.5.2 Estratégias combinatórias

Similar ao teste global, o teste nodal não identifica completamente o tipo ou localização do erro grosseiro. Ainda que contenha mais informação, pois identifica o nó (ou equação de balanço) com o erro, o teste nodal também requer uma estratégia para proceder à identificação. Mah *et al.* (apud NARASIMHAN; JORDACHE, 2000) desenvolveram um algoritmo (para o caso de uma rede de fluxos de massa) para a identificação das medidas contribuindo para os desbalanços nodais. Se nenhuma medida é encontrada com erro, qualquer desbalanço nodal significativo é atribuído a um vazamento ou um erro de modelagem. O maior problema do teste nodal é a possibilidade de cancelamentos mútuos de erros que tornam difícil a localização do erro grosseiro.

Existem várias estratégias de identificação de erros grosseiros que fazem uso do teste nodal, mas em sua maioria são projetadas para problemas lineares (na prática, balanços de massa).

O princípio básico da estratégia para identificação de erros grosseiros baseado no teste nodal foi proposto por Mah *et al.* (apud NARASIMHAN; JORDACHE, 2000). Se houver um erro grosseiro em qualquer medida de fluxo, isto afeta o resíduo do balanço na medida onde o erro ocorre e assim espera-se que o teste nodal para os dois nós que são ligados por essa corrente rejeite a hipótese nula. Se esses dois nós são agrupados, a corrente que os interconecta é eliminada e o teste nodal para este pseudo-nó provavelmente não detectará um erro grosseiro. Para explorar este princípio, os testes nodais são conduzidos sobre os resíduos em torno de nós únicos bem como em combinações de dois ou mais nós conectados por correntes. Se o teste nodal para qualquer combinação não é rejeitado, então as medidas de fluxo de todas as correntes que incidem sobre o pseudo-nó são declaradas livres de erros.

Não se pode tomar nenhuma decisão a respeito das medidas de fluxo das correntes que ligam quaisquer dois nós formadores de um pseudo-nó. Por outro lado, se o teste nodal é rejeitado, então uma ou mais medidas incidentes ao pseudo-nó podem conter erros grosseiros, mas não se pode fazer nenhuma afirmação direta sobre estas medidas.

Através da seleção adequada dos nós nos quais o teste nodal será realizado, é possível identificar um conjunto de medidas prováveis de estarem livres de erros grosseiros. As medidas restantes são suspeitas de conter erros. Uma posterior sondagem sobre este último conjunto pode ser feita usando outras estratégias para identificação.

7.5.2.1 Técnica da combinação linear

No método proposto por Rollins *et al.* (1996), os testes nodais são conduzidos sobre os resíduos de balanço em torno de nós simples bem como certas combinações de nós. De um modo geral, para quaisquer destes testes nodais, as hipóteses podem ser expressas como:

$$H_{0i} : \mathbf{l}_i^{\mathrm{T}} \mu_r = 0 \quad \times \quad H_{li} : \mathbf{l}_i^{\mathrm{T}} \mu_r \neq 0 \tag{7.70}$$

onde μ_r é o resultado esperado do resíduo de balanço e $\mathbf{l}_i$ é o vetor de zeros e uns representando a combinação linear do i-ésimo teste. A um nível de significância α, H_{0i} é rejeitada se:

$$\frac{\mathbf{l}_i^{\mathrm{T}} \mathbf{r}}{\sqrt{\mathbf{l}_i^{\mathrm{T}} \mathbf{V} \mathbf{l}_i}} \geq Z_{1-\alpha/2} \tag{7.71}$$

Se H_{0i} não é rejeitada, todas as medidas das correntes incidentes sobre o pseudo-nó são declaradas livres de erros. Depois de realizar todas as combinações lineares de testes, dois conjuntos de variáveis medidas são obtidos. O conjunto **SET1** contém variáveis cujas medidas não são suspeitas e o outro, **SET2**, com as medidas suspeitas de conterem erros grosseiros. Obviamente, o algoritmo pode resultar na classificação incorreta das medidas dando lugar a Erros do Tipo I (medidas aceitáveis dentro do conjunto **SET2**) ou do Tipo II (medidas com erros no conjunto **SET1**). O nível de significância α, escolhido para estes testes, desempenha um papel importante no balanço entre os dois tipos de erros.

Para reduzir o número de combinações lineares para o teste de hipótese, Rollins *et al.* (1996) adotaram as seguintes regras:

i. Conduzir m testes nodais sobre os nós individuais. Se H_{ok} para o nó $k(k = 1, \ldots, m)$ não é rejeitado, nenhum teste nodal sobre combinações de nós contendo o nó k é realizado;

ii. Um erro grosseiro em uma corrente com um baixo valor de taxa de fluxo é geralmente difícil de se detectar. Assumindo implicitamente que nenhuma corrente com baixa taxa de fluxo contenha um erro grosseiro, os nós conectados a estas correntes também não são avaliados pelo teste nodal;

iii. Não é realizado o teste nodal sobre combinações de nós que não sejam diretamente conectados;

iv. Não é realizado o teste nodal sobre combinações de nós conectadas por correntes classificadas como livres de erros grosseiros.

As regras acima são usadas para evitar a realização de testes nodais sobre combinações de nós que não forneçam informação adicional para identificação de boas medidas. Mah *et al.* (apud NARASIMHAN; JORDACHE, 2000) usaram uma estratégia similar, fazendo uso das regras acima, exceto da regra **ii**. Além disso, o procedimento deles também tentou como último recurso identificar vazamentos nos nós, quando o teste nodal para um nó é rejeitado, mas todas as correntes incidentes neste nó são classificadas como livres de erros (listadas em **SET1**). Serth e Heenan (1986) propuseram três variantes diferentes da estratégia anterior e em uma delas, as informações sobre limites nas variáveis também era explorada. O Exemplo 7.7 ilustra a aplicação da técnica da combinação linear.

EXEMPLO 7.7 A técnica da combinação linear é aplicada sobre os dados do trocador de calor com *bypass* considerada nos exemplos anteriores, mas para o caso onde são induzidos três erros grosseiros nas correntes 1, 2 e 5. Os valores verdadeiros, medidos e reconciliados para estas condições estão listados na Tabela 7.3. São realizados testes nodais com $\alpha = 0,05$ sobre nós singulares e combinações apropriadas de nós de modo a gerar a Tabela 7.4 com a classificação das correntes. A menos das correntes 1 e 2, o restante é classificado como bom e assim dois dos três erros são corretamente identificados pelo algoritmo.

Tabela 7.3: Reconciliação de dados com a presença de erros grosseiros para o processo da Figura 7.2 – segundo caso

Número da Corrente	Valores Verdadeiros dos Fluxos	Valores Medidos dos Fluxos	Valores Reconciliados
1	100	106,91	102,0533
2	64	68,45	67,1667
3	36	34,65	34,8867
4	64	64,20	67,1667
5	36	33,44	34,8867
6	100	98,88	102,0533

Tabela 7.4: Detecção de erros grosseiros usando a técnica da combinação linear

Combinação de nós	Estatística do Teste Nodal	Status	Medidas classificadas livres de erros grosseiros
1	2,200	rejeitado	–
2	3,005	rejeitado	–
3	0,856	não rejeitado	3 e 5
4	0,716	não rejeitado	4 e 6
1+2	4,653	rejeitado	–
1+3	–	não rejeitado – corrente 3 ok	
1+4	–	não rejeitado – desconectado	
2+3	–	não rejeitado – desconectado	
2+4	–	não rejeitado – corrente 4 ok	
3+4	–	não rejeitado – corrente 5 ok	

7.5.2.2 Técnica combinatória MT-NT

Segundo Yang *et al.* (1995), a idéia por trás de sua técnica combinatória é encontrar um meio de evitar qualquer manipulação artificial dos dados, como o re-cálculo no MIMT (*modified iterative measurement test – teste da medida iterativo modificado*) com imposição de limites sobre as variáveis (HEENAN; SERTH, 1986), tirando vantagem tanto do teste da medida quanto do teste nodal.

A vantagem do teste da medida (MT) é que as localizações dos erros grosseiros nos dados do processo podem ser apontadas diretamente, mas ele tende a gerar erros extra (Erro do Tipo I). O teste nodal (NT) não espalha os erros grosseiros pelos dados, mas é propenso ao Erro do Tipo II. A técnica MT-NT propõe então combinar os dois testes de modo que compensem as falhas mútuas.

Desta forma, é necessário encontrar um meio de identificar qual dos erros apontados pelo MT é um verdadeiro erro grosseiro e qual é falso. No MIMT, somente um erro grosseiro pode ser encontrado em cada iteração, mas não há garantias que assegurem que o primeiro a ser encontrado é um erro grosseiro verdadeiro depois da primeira iteração. Para distingui-los, é introduzido o teste nodal para fazer a checagem. O procedimento pode ser descrito assim: primeiro encontrar dois nós que ligam a corrente que tem o maior z_a para então checar estes nós por meio do teste nodal verificando se as estatísticas do teste, z_r, são maiores que o valor crítico, $Z_{1-\beta/2}$ (com β calculado por 7.7 ou 7.8). Se não forem, quer dizer que não há um erro grosseiro nesta corrente. Retorna-se então ao MT e toma-se o segundo maior z_a. Se um ou ambos os z_r (dos nós ligados à corrente) excederem $Z_{1-\beta/2}$, isso quer dizer que ao menos uma das correntes que ligam este nó tem um erro grosseiro. O principal suspeito é a corrente com o maior ajuste relativo, a_i/y_i, checa-se então

o outro nó que se liga a esta corrente. Se o z_r deste nó é inferior ao critério, então a corrente é declarada livre de erros e passa-se à corrente com o segundo maior ajuste. Caso contrário, o valor exceda $Z_{1-\beta/2}$, então o erro grosseiro foi encontrado, devendo ser retirado do conjunto das medidas e tratado como uma variável não medida, para em seguida recomeçar com uma segunda iteração do MT e assim por diante.

Esta técnica de localizar com MT e checar com NT agrega a vantagem de ambos os métodos e evita o problema de busca combinatorial de grandes proporções, apresentando uma quantidade bem menor de cálculos e nenhuma manipulação artificial.

EXEMPLO 7.8 O processo considerado neste exemplo está ilustrado na Figura 7.3 e consiste de quatro unidades (de **A** a **D**) com oito correntes, sendo as duas últimas não medidas. Os valores verdadeiros estão listados na coluna x da Tabela 7.5, onde a coluna y guarda os valores medidos (corrompidos com erros aleatórios e grosseiros), a coluna $\hat{x}$ mostra o resultado do primeiro procedimento de reconciliação de dados, a coluna a, os ajustes feitos às medidas, a coluna z_a, a estatística do teste de medida e a coluna a/y, os ajustes relativos.

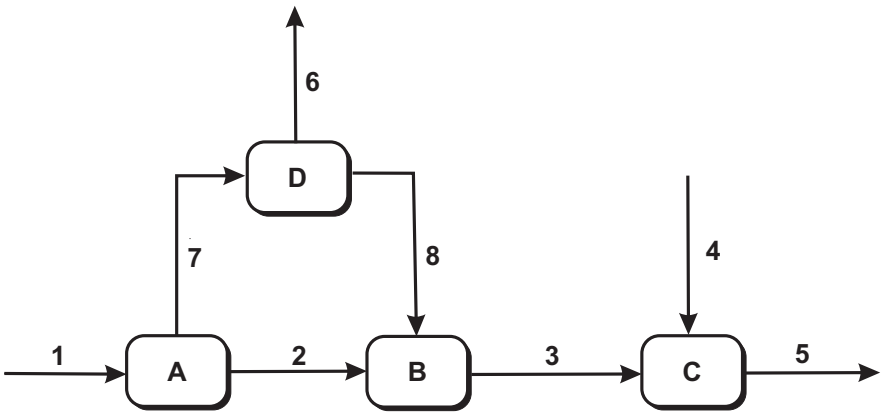


Figura 7.3: Exemplo da aplicação da técnica MT-NT – adaptado de Yang *et al.* (1995)

A Tabela 7.6 mostra o teste nodal aplicado aos nós do processo, incluindo o nó das vizinhanças (ambiente). Com $\alpha = 0,05$ os critérios de corte para o teste da medida e nodal são $Z_a = 2,63$ e $Z_r = 1,96$, respectivamente. Um erro grosseiro foi induzido na corrente 1, sendo medido o valor de 110,1 no lugar do valor verdadeiro 98,7. Os valores de z_a que excedem o valor crítico de 2,63 são os das correntes 1, 2 e 6. Já os valores de z_r que excedem o valor crítico de 1,96 são correspondentes ao nó 1 e ao nó das vizinhanças. Seguindo a técnica apresentada e baseando-se nos dados das Tabelas 7.5 e 7.6, segue que as correntes 2 e 6 não passam pela checagem do teste nodal, deixando de ser suspeitas. Os resultados finais,

Tabela 7.5: Resultado da primeira iteração do método MT-NT (resultados do teste da medida) – reproduzido de Yang *et al.* (1995)

Corrente	x	y	$\hat{x}$	a	z_a	a/y
1	98,700000	110,100000	101,583577	-8,516423	4,538834	0,077352
2	41,100000	41,100000	41,100000	0,000000	4,538834	0,000000
3	78,900000	79,000000	81,508146	2,508146	2,249785	0,031749
4	30,200000	30,600000	30,318458	-0,281542	1,950316	0,009201
5	109,100000	108,300000	111,826604	3,526604	1,950316	0,032563
6	19,800000	19,800000	20,075431	0,275431	4,538834	0,013911
7	57,600000	–	60,483597	–	–	–
8	37,800000	–	40,408165	–	–	–

depois de mover a corrente 1 para o grupo das não medidas e repetir o procedimento de reconciliação de dados e estimativa de parâmetros, são apresentados nas Tabelas 7.7 e 7.8, onde se pode perceber que nenhum dos valores de z_a e z_r excede os valores críticos. Assim a corrente 1 foi corretamente identificada como contendo o erro grosseiro.

Tabela 7.6: Resultado da primeira iteração do método MT-NT (resultados do teste nodal) – reproduzido de Yang *et al.* (1995)

Nó	z_r
A	3,221708
B	1,282414
C	0,472727
D	0,182680
Ambiente	5,434266

Tabela 7.7: Resultado final do método MT-NT (correntes)– reproduzido de Yang *et al.* (1995)

Corrente	$\hat{x}$	a	z_a	a/y
1	98,370824	–	–	–
2	41,100000	0,000000	0,000000	0,000000
3	78,570867	-0,429133	0,472727	0,005432
4	30,535616	-0,064384	0,472727	0,002104
5	109,106482	0,806482	0,472727	0,007447
6	19,800000	0,000000	0,000000	0,000000
7	57,270809	–	–	–
8	37,470812	–	–	–

Tabela 7.8: Resultado final do método MT-NT (nós)– reproduzido de Yang *et al.* (1995)

Nó	z_r
A	0,000006
B	0,222115
C	0,472727
D	0,000002
Ambiente	0,384051

Ainda que as estratégias baseadas em testes nodais reduzam o número de predições equivocadas (erros do Tipo I) em comparação às estratégias seriais, elas sofrem dos seguintes problemas:

- Se múltiplos erros estão presentes entre os dados, devido a cancelamentos parciais ou completos dos erros, os testes nodais para as combinações de nós nos quais estas correntes incidam podem não ser rejeitados, resultando numa classificação incorreta (Erro do Tipo II);
- Estes testes são projetados para processos lineares de fluxo e são difíceis de estender a processos não lineares.

7.5.3 Identificação por componentes principais

Os balanços com erros grosseiros podem ser identificados pela inspeção da contribuição do j-ésimo resíduo em $\mathbf{r}$, r_j, para um componente principal suspeito, $p_{r,i}$, que pode ser calculado por:

$$g_j = \left(\mathbf{w}_{r,i}\right)_j r_j \quad j = 1, \ldots, m \tag{7.72}$$

onde $\mathbf{w}_{r,i}$ é o i-ésimo vetor coluna da matriz $\mathbf{W}_r$.

Definindo $\mathbf{g} = (g_1, \ldots, g_m)^{\mathrm{T}}$, e fazendo $\mathbf{g}'$ ser o mesmo que $\mathbf{g}$, exceto que seus elementos são ordenados de modo decrescente pelos seus valores absolutos. As contribuições dos resíduos ao componente principal suspeito são diferentes e dominadas pelos primeiros elementos. Estes são os maiores contribuintes para o componente principal suspeito. Os maiores contribuintes são diretamente relacionados às restrições que estão também sob suspeição. A quantidade de grandes contribuintes, k, pode ser definida de modo que:

$$\left| \frac{\left(\sum_{j=1}^{k} g'_j\right) - p_{r,i}}{p_{r,i}} \right| < \varepsilon_1 \tag{7.73}$$

onde ε_1 é uma tolerância preestabelecida de, por exemplo, 0,1.

A Equação 7.73 leva em consideração o efeito de mútuo cancelamento dos sinais nos elementos de $\mathbf{w}_{r,i}$ e $\mathbf{r}$.

De modo similar ao teste nodal, o teste de componentes principais sobre os resíduos do balanço indica apenas quais dos resíduos das restrições são os maiores contribuintes para o componente principal suspeito. Uma estratégia adicional é necessária para identificar a fonte do erro (vazamento ou viés de medida) e qual das medidas contém os erros grosseiros. Pode-se, por outro lado, usar o teste dos componentes principais sobre os ajustes às medidas para identificar a medida com o erro grosseiro pela inspeção da contribuição do j-ésimo ajuste em $\mathbf{a}$, a_j, para um componente principal suspeito i.

A j-ésima contribuição ao ajuste pode ser calculada por

$$g_j = (\mathbf{w}_{a,i})_j a_j \quad j = 1, \ldots, n \tag{7.74}$$

onde $\mathbf{w}_{a,i}$ é o i-ésimo autovetor de $\mathbf{W}_a$ e n é o número total de medidas. As contribuições podem ser avaliadas pela checagem dos sinais e magnitudes dos elementos em $\mathbf{g}$. De um modo geral, como no teste de componentes principais para os resíduos de balanço, as contribuições variam e são dominadas por poucos elementos. A regra de identificação para o teste de componentes principais nas medidas é a seguinte:

Identificar o erro grosseiro na medida que corresponde ao maior contribuinte para o máximo componente principal que exceda o critério de teste

EXEMPLO 7.9 Para identificar o erro grosseiro usando testes de componentes principais, tem-se que examinar as contribuições para o componente principal rejeitado. Considerando somente o último componente principal que é rejeitado a um nível de significância modificado (Exemplo 7.6). Os contribuintes (ajustes às medidas) para este componente principal podem ser analisados pelo cálculo do vetor $\mathbf{g}$ (Equação 7.74). Este vetor é dado por

$$\begin{bmatrix} 0,0293 & -1,1073 & -0,1030 & 0,0975 & -1,0237 & -0,9114 \end{bmatrix}$$

O maior contribuinte para o componente principal suspeito é o ajuste 2, e portanto um erro

grosseiro é identificado nesta medida

Tong e Crowe (1995) fizeram uma extensa analise dos testes de componentes principais e propuseram algumas regras práticas para implementação de um estratégia de detecção e identificação. A maior parte de suas recomendações, como fazer uso de testes χ^2 coletivos primeiro, usando uma matriz de covariância dos erros nas medidas de grande exatidão e uma distribuição de erros apropriada, são válidas para todas estratégias envolvendo testes univariados.

Eles recomendam que os testes de componentes principais devem ser usados em combinação com outros testes estatísticos, pois não há garantias de detecção para todos os erros grosseiros. Eles também alertam sobre o tempo computacional aumentado para o cálculo de autovalores e autovetores e também para a análise da contribuição de estatísticas dos componentes principais para identificação de erros. Para resumir, os testes de componentes principais são efetivos em certas situações, mas não são superiores, de modo geral, aos testes estatísticos básicos.

7.6 Detectabilidade e identificabilidade de erros grosseiros

A detectabilidade e a identificabilidade são duas características importantes na detecção de erros grosseiros. A primeira se refere à possibilidade de detectar erros grosseiros em meio às medições e a segunda, à possibilidade de distinção entre dois ou mais erros grosseiros. Os conceitos de detectabilidade propostos por Madron (1992) e de identificabilidade discutidos por vários pesquisadores como Bagajewicz e Jiang (1998), Jordache *et al.*; Charpentier *et al.* (apud NARASIMHAN; JORDACHE, 2000) são mostrados a seguir.

7.6.1 A detectabilidade de erros grosseiros

Semelhante à reconciliação de dados, um pré-requisito fundamental para a detecção de erros grosseiros é a redundância de medições. Teoricamente, somente é possível detectar erros grosseiros entre medições redundantes. Isto se deve ao fato de que uma medida não redundante é eliminada junto com variáveis não medidas e não participa do problema reduzido de reconciliação e assim nenhum teste estatístico pode ser derivado.

Na Seção 3.5, foram apresentados métodos para classificação de variáveis de processo quanto

à observabilidade e redundância. Somente medições **redundantes** são passíveis de ajustes pela reconciliação de dados e somente as variáveis não medidas **observáveis** podem ser estimadas. Através da adição de novos sensores ou pela inclusão de restrições adicionais, que é possível eliminar/mitigar a não observabilidade e a deficiência de redundância.

Tanto a abordagem por grafos quanto a matricial podem ser usadas para classificação quanto à observabilidade e redundância. Na prática, contudo, várias variáveis redundantes se comportam como não redundantes. Tais variáveis podem ser chamadas de medições **praticamente não redundantes**. Jordache *et al.*; Crowe (apud NARASIMHAN; JORDACHE, 2000), Madron (1992) e Charpentier *et al.* (apud NARASIMHAN; JORDACHE, 2000) reportaram dificuldades na reconciliação e detecção de erros grosseiros em tais variáveis. De modo semelhante, mesmo se algumas variáveis não medidas são observáveis, as suas estimativas podem ter desvios padrão tão altos que podem ser consideradas variáveis **praticamente não observáveis**.

Se uma medida de uma variável redundante contém um erro grosseiro, então a reconciliação de dados deve teoricamente fazer um grande ajuste nesta medida para obter uma estimativa tão próxima quanto possível do valor verdadeiro da variável. Em alguns casos, no entanto, devido à natureza das restrições e aos desvios padrão das variáveis, a reconciliação pode fazer um ajuste insignificante na medida redundante errada e, em seu lugar, fazer um ajuste maior em uma variável que seja na verdade livre de erros de modo a satisfazer às restrições. Uma medida deste tipo não é verdadeiramente redundante mesmo se teoricamente classificada como tal. É um procedimento mais difícil a identificação de erros grosseiros em tais medições.

Madron (1992) define uma medida praticamente redundante como aquele cuja **ajustabilidade** seja maior que um dado valor limite. Essa condição pode ser colocada da seguinte forma:

$$a_i = 1 - \frac{\sigma_{\hat{x}_i}}{\sigma_{y_i}} > a_{cr} \tag{7.75}$$

onde a_i é a ajustabilidade, $\sigma_{\hat{x}_i}$ é o desvio padrão do valor reconciliado i, e σ_{y_i} é o desvio padrão do erro na medida. O limite crítico, a_{cr}, é um valor no intervalo $0 \leq a_{cr} \leq 1$. Por exemplo, se a_{cr} é escolhido igual a $0,1$, todas as medidas i tais que $a_i < 0,1$ são consideradas praticamente não redundantes. Para tais medidas $\frac{\sigma_{\hat{x}_i}}{\sigma_{y_i}} > 0,9$ e portanto o ajuste feito ao valor medido é insignificante. A ajustabilidade a_i é também uma medida da melhora na exatidão de um valor medido que pode ser alcançada através da reconciliação de dados.

Charpentier *et al.* (1991) sugeriram usar a razão:

$$d_i = \sqrt{\left(1 - \frac{\sigma_{\hat{x}i}^2}{\sigma_{yi}^2}\right)} \tag{7.76}$$

para identificar medidas com fraca redundância. Este fator é uma medida da **detectabilidade** de um erro. Como desbalanços nas restrições indicam a existência de erros grosseiros, a detectabilidade de um erro grosseiros depende da sua contribuição aos desbalanços nas restrições. A contribuição de uma medida no resíduo de uma restrição depende da própria restrição e da precisão relativa das medidas (desvios padrão relativos). A contribuição de um erro ao desbalanço de uma restrição é proporcional ao fator de detectabilidade. Quanto maior o fator, maior é a probabilidade do erro grosseiro ser detectado. Isto também indica que se o fator de detectabilidade é alto, então erros grosseiros de menor magnitudes nas medições correspondentes podem ser detectados com relativa facilidade.

Uma análise completa da redundância prática é útil na identificação de todas variáveis medidas com fraca redundância. Para processos lineares, o desvio padrão das estimativas reconciliadas pode ser calculado analiticamente e as medidas de ajustabilidade ou detectabilidade podem ser calculadas. Para casos não lineares, no entanto, estas medidas podem ser calculadas somente depois da solução do problema de reconciliação para um dado conjunto de medições e através da linearização das restrições em torno das estimativas reconciliadas.

Narasimhan e Jordache (2000) mostram que estudos por simulação foram feitos e foram identificadas como praticamente não redundantes as variáveis com as seguintes características:

i. Variáveis com desvios padrão relativamente pequenos em comparação com o desvio padrão das outras medições pertencentes ao mesmo balanço. Geralmente, este é o caso para as medições cuja ordem de magnitude é também pequena em relação às outras variáveis na mesma equação de balanço, (a exemplo de fluxo de correntes pequenas que aparecem em balanços com fluxos de grandes correntes). A razão necessária entre erro/desvio padrão para detecção de erros grosseiros é muito maior para as variáveis com um desvio padrão pequeno do que para aquelas com um grande desvio padrão (JORDACHE *et al.*, 1985 apud NARASIMHAN; JORDACHE, 2000);

ii. Correntes paralelas, a exemplo dos fluxos de saída de um divisor de corrente que não estejam

restritos por nenhum outro balanço (JORDACHE *et al.*, 1985 apud NARASIMHAN; JORDACHE, 2000);

iii. Fluxos que ocorrem no balanço de entalpia, mas não no de massa (CHARPENTIER *et al.*, 1991 apud NARASIMHAN; JORDACHE, 2000). Estes são tipicamente refluxos usados no balanço entálpico da coluna de destilação principal e balanços dos trocadores de calor associados, mas não incluídos no balanço mássico da torre. Um balanço global de massa usando as taxas medidas de alimentação e de produtos para a fracionadora inteira é comumente escolhido para evitar o grande número de fluxos não medidos em torno da própria coluna;

iv. Temperaturas de correntes pequenas no mesmo balanço com temperaturas de correntes grades, mesmo que a ordem de magnitude e desvio padrão de tais temperaturas sejam similares;

v. Temperatura de entrada do primeiro trocador de calor no trem de pré-aquecimento (CHARPENTIER *et al.*, 1991 apud NARASIMHAN; JORDACHE, 2000). Tipicamente, ela só aparece em um balanço entálpico enquanto que as temperaturas seguintes experimentam uma redundância extra por fazer parte de ao menos dois balanços de energia;

vi. Variáveis medidas que aparecem em somente uma equação com uma variável não medida que não é restrita por qualquer outra equação de balanço ou limitação. Um erro grosseiro nesta variável medida é comumente transferido para uma variável não medida que tem mais liberdade para ajustes.

Segundo Narasimhan e Jordache (2000), não há uma solução simples para a reconciliação de dados e detecção de erros grosseiros em tais variáveis. Restrições e instrumentação extra certamente são úteis, mas não são sempre possíveis. Em algumas situações variáveis "medidas" artificiais podem ser criadas a partir de valores calculados para aumentar a redundância (CHARPENTIER *et al.*; KNEILE, 1991, 1995 apud NARASIMHAN; JORDACHE, 2000). A informação a respeito dos pontos de fraca redundância pode ser útil para os usuários de um pacote de reconciliação de dados no sentido de possibilitar o reconhecimento das limitações na exatidão dos métodos de detecção de erros grosseiros a ajudar nas melhorias de instrumentação.

O conhecimento sobre a classificação das variáveis práticas é uma informação importante que pode ser incluída nos algoritmos de detecção de erros grosseiros. Por exemplo, o fator de detectabilidade de um erro grosseiro pode ser usado como um critério de desempate quando mais de

uma medida apresentam o mesmo valor de teste estatístico (JORDACHE; TILTON, 1999 apud NARASIMHAN; JORDACHE, 2000).

Tabela 7.9: Valores de ajustabilidade e detectabilidade para o processo descrito na Figura 7.2

Medida	Variância do Erro	Ajustabilidade	Detectabilidade
1	1,0000	0,5025	0,8675
2	0,9801	0,4975	0,8646
3	0,0001	0,2929	0,7071
4	0,9801	0,4975	0,8646
5	0,0001	0,2929	0,7071
6	1,0000	0,5025	0,8675

EXEMPLO 7.10 Para o processo considerado nos exemplos anteriores, a matriz de covariância dos erros nas medidas foi dada como a matriz identidade. A matriz de covariância para as estimativas de todas estas variáveis pode ser calculada usando a Equação 3.24. Os elementos da diagonal desta matriz são as variâncias das estimativas. Para este processo, as variâncias para todas as estimativas acabam sendo iguais a 0,3333. Usando estes valores nas Equações 7.75 e 7.76, obtém-se a ajustabilidade igual a 0,4226 e a detectabilidade igual a 0,8165 para todas variáveis. Isto implica que os erros grosseiros em todas as medidas têm iguais chances de serem detectados. Por outro lado, tomando-se os valores verdadeiros das variáveis de fluxo como [100, 99, 1, 99, 1, 100] e assumindo-se que os desvios padrão nos erros das medidas como 1% dos valores verdadeiros, tem-se então a ajustabilidade e a detectabilidade para as diferentes variáveis (Tabela 7.9). A partir dos resultados dados desta tabela, pode-se concluir que é relativamente mais difícil identificar erros grosseiros nas medições das correntes 3 e 5 quando comparadas com as demais. Para verificar este resultado Narasimhan e Jordache (2000) fizeram cerca de 20 simulações e em cada uma delas um erro grosseiro de magnitude 5 a 15 vezes o desvio padrão foi simulado na medida do fluxo da corrente 1, com a aplicação do teste *GLR* para identificação do erro grosseiro. De modo semelhante, 20 simulações foram feitas com um erro grosseiro presente na medida 2 e assim por diante para cada uma das correntes. Os resultados mostraram que enquanto os erros grosseiros nas correntes 1, 2 e 6 foram identificados corretamente em todas as simulações, somente 60% dos erros grosseiros na corrente 3 e 30% dos erros na corrente 5 foram identificados corretamente. Ainda que o número de simulações tenham sido pequeno, a tendência dos resultados corrobora as observações.

7.6.2 Identificabilidade de erros grosseiros

Mesmo que uma medida tenha alta detectabilidade, é importante determinar se um erro grosseiro nesta medida pode ser identificado ou distinguido de um erro grosseiro presente em uma outra medida. Para processos lineares, esta questão pode ser respondida de diferentes formas.

Jordache *et al.* (apud NARASIMHAN; JORDACHE, 2000) mostraram que as estatísticas de teste de duas medidas diferentes são idênticas se as colunas da matriz **A** que correspondem às duas variáveis medidas são proporcionais entre si. Um caso especial disto ocorre quando duas correntes paralelas ligam os mesmos dois nós de um processo. Isto implica que não é possível distinguir erros grosseiros presentes nestas medidas. No contexto do teste *GLR*, Narasimhan e Mah (1987) indicaram que se os vetores de assinatura de dois erros grosseiros são proporcionais, então estes não podem ser distinguidos entre si. Quando se restringe a considerar os viéses das medidas somente, então esta observação é a mesma feita por Jordache *et al.* (apud NARASIMHAN; JORDACHE, 2000). Podem ser descobertos problemas de identificabilidade entre diferentes tipos de erros grosseiros através do uso dos vetores de assinatura.

Bagajewicz e Jiang (1998) propuseram o conceito de **conjuntos equivalentes de erros grosseiros**. Um conjunto de erros grosseiros é equivalente a outro se não poderem ser distinguidos um do outro. Para o caso dos viéses nas medidas, os autores provaram que se um conjunto de medidas de k variáveis forma um ciclo no grafo do processo, então erros grosseiros em qualquer combinação de $k-1$ medidas deste conjunto não podem ser distinguidos de erros grosseiros em qualquer outra combinação. Isto pode ser verificado se for considerado que na eliminação serial uma medida supostamente contendo um erro grosseiro é eliminada, tornando a variável correspondente não medida.

Eliminar qualquer conjunto de $k-1$ medidas de um ciclo de k medidas vai automaticamente tornar o restante das medidas não redundantes e eliminá-las do problema de reconciliação. Isto implica que a solução do problema do problema reduzido de reconciliação será a mesma, não importando qual combinação de $k-1$ medidas seja eliminada. Desta forma, não é possível identificar qual conjunto de $k-1$ medidas deste ciclo contém erros grosseiros e todos estes conjuntos são declarados equivalentes. Pela mesma razão, não é possível distinguir erros grosseiros nas medidas de todas as k variáveis de um ciclo a partir de qualquer conjunto de erros grosseiros nas medidas de $k-1$ variáveis deste ciclo.

Um caso especial é o ciclo formado por duas correntes paralelas no qual não é possível dis-

tinguir entre um erro grosseiro numa corrente ou na outra. Também não é possível distinguir se ambas as correntes contém eros grosseiros ou se somente uma delas. Ainda mais, se o número de restrições independentes é igual a m, então todos os conjuntos de erros grosseiros linearmente independentes são também equivalentes e isto se deve ao fato que o problema de reconciliação reduzido não terá mais nenhum redundância e não há informação disponível para fazer qualquer distinção entre eles.

Narasimhan e Jordache (2000) se referem aos conjuntos equivalentes como pertencentes a uma **classe de equivalência**. Classes de equivalência podem ser obtidas em termos dos vetores de assinatura dos erros grosseiros o quê permite outros tipos de erros grosseiros, como vazamentos, serem considerados também. O seguinte princípio pode então ser colocado: se os vetores de assinatura para um conjunto de k erros grosseiros formam um conjunto de posto $k-1$ linearmente dependente, então não é possível distinguir teoricamente entre uma combinação de $k-1$ erros grosseiros a partir de qualquer outra combinação de $k-1$ erros grosseiros escolhidos deste mesmo conjunto. Também não é possível distinguir se k erros grosseiros ou $k-1$ erros grosseiros deste conjunto estão presentes no processo. No entanto é possível distinguir uma combinação de menos que $k-1$ erros grosseiros de outras combinações. Como um caso especial, se o número máximo de vetores de assinatura independentes é m, então qualquer conjunto de m erros grosseiros com vetores de assinatura linearmente independentes é equivalente a qualquer outro conjunto definido da mesma forma.

EXEMPLO 7.11 O grafo do processo considerado nos exemplos anteriores é mostrado na Figura 7.2. Os três seguintes ciclos podem ser identificados neste grafo:

- Ciclo 1 - correntes 2, 3, 4 e 5
- Ciclo 2 - correntes 1, 2, 4 e 6
- Ciclo 3 - correntes 1, 3, 5 e 6

Assim, as seguintes classes de equivalência de conjuntos de viéses são obtidas:

- Classe 1: [2, 3, 4]; [2, 3, 5]; [2, 4, 5]; [3, 4, 5] e [2, 3, 4, 5];
- Classe 2: [1, 2, 4]; [1, 2, 6]; [1, 4, 6]; [2, 4, 6] e [1, 2, 4, 6];
- Classe 3: [1, 3, 5]; [1, 3, 6]; [1, 5, 6]; [3, 5, 6] e [1, 3, 5, 6];

Considerando-se por exemplo um vazamento no nó 1 (divisor de correntes), então os vetores de assinatura são usados para identificar os conjuntos equivalentes. Os vetores de assinatura

para os viéses nas medidas são as colunas da matriz **A**. O vetor de assinatura para um vazamento no nó 1 é o primeira coluna da matriz **A**, que é idêntico a um viés na medida 1. Deste modo, um vazamento no nó 1 não pode ser distinguido de um viés na medida 1 e, além disso, são obtidos os mesmos conjuntos equivalentes que os obtidos usando os ciclos de um grafo porque os vetores de assinatura para os viéses nas medidas das correntes 2, 3, 4 e 5 são linearmente dependentes como posto igual a 3 e assim por diante. Se um conjunto G de erros grosseiros contém um sub-conjunto de erros grosseiros, g_{ci}, pertencente a uma classe de equivalência C, então outros conjuntos de equivalência de G podem ser obtidos pela substituição de g_{ci} por outros conjuntos pertencentes a C. Assim, por exemplo, podem ser derivados os conjuntos de equivalência para a combinação [1, 2, 3, 4] pela substituição do subconjunto [2, 3, 4] por outros conjuntos da Classe 1. De modo similar, pode se substituir [1, 2, 4] por outros conjuntos da Classe 2. Deste modo, seria obtida uma outra classe de equivalência dada por:

- Classe 4: [1, 2, 3, 4]; [1, 2, 3, 5]; [1, 2, 4, 5]; [1, 3, 4, 5]; [1, 2, 3, 6]; [1, 3, 4, 6]; [2, 3, 4, 6]; [1, 2, 5, 6]; [2, 3, 4, 6]; [1, 4, 5, 6]; [2, 4, 5, 6]; [3, 4, 5, 6]

Os últimos cinco conjuntos foram adicionados à Classe 4 porque são equivalentes aos conjuntos [1, 2, 3, 5]; [1, 2, 4, 5] e [1, 3, 4, 5] que pertencem à classe 4. A Classe de Equivalência 4 pode também ser gerada usando o fato que este processo tem quatro restrições independentes e todos os conjuntos de quatro erros grosseiros com vetores de assinatura linearmente independentes são equivalentes.

Se problemas de identificabilidade podem ocorrer em processos lineares, de um modo geral, esse não é um problema em processos não lineares. Se restrições não lineares são linearizadas em torno das estimativas reconciliadas, é bastante improvável que as colunas da matriz das restrições linearizadas se tornem dependentes e mesmo que isto ocorra, isto teria que ser interpretado mais como um problema numérico do que de identificabilidade.

7.7 Conclusões

Alguns resultados e afirmações vistos neste capítulo se destacam como por exemplo a definição dos dois tipos de erros associados com qualquer teste estatístico: O Erro do Tipo I (quando o teste

detecta um erro que não existe de fato) e o Erro do Tipo II (quando o teste falha na detecção de um erro que de fato está presente) e também o fato de que qualquer estratégia de abordagem aos erros grosseiros precisa detectar e também identificar a localização do erro grosseiro.

Quanto aos testes básicos, somente o teste da medida e o teste GLR podem diretamente identificar a localização de um erro grosseiro (por uma simples regra de identificação). O teste GLR é o único teste que pode identificar tanto viéses nas medidas quanto vazamentos pelo mesmo tipo de teste. A estratégia de detecção de erros grosseiros pelo teste GLR envolve também a estimativa da magnitude dos erros grosseiros.

Os testes de máxima potência podem ser derivados para o teste da medida e para o nodal, mas o teste GLR é mais poderoso que ambos para o caso da presença de um único erro grosseiro. A potência do teste GLR é a mesma que a do teste da medida para um único viés nas medidas. Por outro lado, a estatística do teste GLR é equivalente à estatística do teste da medida para um único viés entre as medidas.

Em relação aos testes baseados em componentes principais, estes não podem identificar diretamente a localização do erro grosseiro pois requerem uma análise adicional para encontrar a restrição ou medida que contribua majoritariamente com o componente principal que falhou no teste.

Já em relação às técnicas de identificação, a redução na estatística do teste global depois da eliminação de uma medida é igual à estatística do teste GLR. A eliminação serial pode ser usada para identificar os erros grosseiros detectados pelo teste global e abordagem combinatória MT-NT se mostra excelente do ponto de vista computacional, agregando vantagens e cancelando fraquezas dos testes MT e NT.

A detectabilidade de um erro grosseiro depende principalmente de sua magnitude e de sua localização. Alguns erros grosseiros podem ser detectados, mas nem sempre identificado apropriadamente.

8 *Apresentação dos Softwares Desenvolvidos*

Este capítulo apresenta o principal resultado deste trabalho: o desenvolvimento de um *software* genérico de suporte a análises e monitoramento da qualidade da informação, em especial à reconciliação e coaptação de dados e à detecção e identificação de erros grosseiros. Inicialmente é apresentado o aplicativo ***Reconciliare***, mostrando seu desenvolvimento através dos requisitos definidos para sua construção e das ferramentas nela utilizadas, para então seguir com a descrição de sua modelagem através dos diagramas da linguagem de modelagem UML, culminando com a apresentação comentada de sua interface gráfica. Em seguida, é apresentado o aplicativo **Servidor de Dados OPC**, usado para gerar dados simulados e publicá-los por meio de um servidor OPC para que sejam captados no aplicativo ***Reconciliare***. Na seqüência são mostradas algumas aplicações dos dois *softwares* trabalhando em conjunto.

8.1 O aplicativo *Reconciliare*

Um dos principais objetivos deste projeto é definir uma série de requisitos para a criação e desenvolvimento de um *software* aplicado à análise da qualidade de informação de plantas químicas. Este *software*, batizado como ***Reconciliare***, foca nos amplos tópicos da reconciliação e coaptação de dados e da detecção e identificação de erros grosseiros e teve o início do seu desenvolvimento relatado em Barbosa (2003).

O aplicativo ***Reconciliare*** é um desenvolvimento com propósitos acadêmicos, mas que se pautou em várias características do desenvolvimento de *softwares* comerciais voltados para **CAE** (*Computer Aided Engineering* – Engenharia Assistida por Computadores) e **CAPE** (*Computer Aided Process Engineering* – Engenharia de Processos Assistida por Computadores). A seguir são listados os requisitos do seu desenvolvimento, mostrada a sua modelagem e as técnicas e ferramen-

tas de programação subjacentes à sua construção.

8.1.1 Requisitos e ferramentas de desenvolvimento

Para indicar as técnicas e ferramentas de programação a serem utilizadas e determinar o escopo de aplicação do *software*, foram escolhidos alguns requisitos de desenvolvimento elencados e comentados a seguir.

- Modularidade;
- Extensibilidade;
- Facilidade de uso;
- Interação com sistemas de informação.

A *modularidade* diz respeito à maneira como o aplicativo é desenvolvido e é resultado do uso de paradigmas de programação que conferem um certo grau de independência entre as suas partes. O aplicativo ***Reconciliare*** foi construído usando duas linguagens de programação: **FORTRAN 90/95** e **Delphi**. Na parte desenvolvida em Delphi, foi usado o paradigma da programação orientada a objetos (POO), pois este paradigma apresenta os elementos necessários para fazer a representação das entidades de um problema e descrever suas interrelações de modo natural, conferindo a característica da modularidade.

No aplicativo ***Reconciliare***, as partes de descrição dos modelos e preparação das tarefas, do controle das análises e da interação com sistemas externos são desenvolvidas sob o paradigma da programação orientada a objetos, usando o compilador de Delphi **Borland Developer Studio 2006**. As subrotinas matemáticas (inversão de matrizes, fatoração QR, SQP, etc.) são todas estruturadas[1], programadas em FORTRAN 90/95, sendo que algumas utilizam a biblioteca de métodos numéricos **IMSL** que acompanha a distribuição **Compaq FORTRAN 6.6C** usada neste trabalho.

[1] Segundo Boratti (2002), a programação estruturada consiste em uma forma de resolução de problemas que procura dividir o problema maior em problemas menores. Cada problema menor, dependendo do caso, pode também ser dividido em outros problemas e assim sucessivamente. A solução do problema maior consiste então na solução, de modo ordenado, de cada um dos problemas menores. A resolução de cada problema menor passa pela identificação dos dados necessários – denominados *entradas*, pela identificação dos resultados que se deseja obter – denominados *saídas*, e a definição de qual processamento deve ser aplicado às entradas para se obter as saídas.

Segundo Boratti (2002) e Sonnino (2003) uma linguagem de programação orientada a objetos tem que implementar os seguintes conceitos fundamentais: suporte a classes, encapsulamento, hereditariedade e polimorfismo. A classe é uma estrutura que abstrai um conjunto de objetos com características similares. Uma classe define o comportamento de seus objetos através de métodos e os estados possíveis destes objetos através de atributos. Em outros termos, uma classe descreve os serviços providos por seus objetos e quais informações eles podem armazenar. Uma classe define o estado e o comportamento de um objeto geralmente implementando métodos e atributos, nomenclatura que é usada na maioria das linguagens de programação modernas. Os atributos, também chamados de campos, indicam as possíveis informações armazenadas por um dado objeto de uma classe, representando seu estado. Os métodos são procedimentos (funções e subrotinas) que formam os comportamentos e serviços oferecidos pelos objetos de uma classe.

O encapsulamento, ou ocultação de dados, é a criação de níveis de visibilidade, também chamados de escopos, para os dados dentro de um programa. O conceito subjacente é limitar o acesso às informações de modo a proteger a integridade da entidade sendo representada. Por exemplo, na representação de um reator tanque de mistura contínua, a variável da temperatura de saída pode ser encapsulada e alterada somente pelos cálculos que definem essa temperatura, fazendo assim com que esse comportamento emule o comportamento real do equipamento, pois de fato a temperatura do reator é alterada apenas pela manipulação direta de outras variáveis, como vazão de reagentes ou do fluido refrigerante. De um modo geral são criados métodos (funções ou subrotinas) para acessar essas variáveis ocultas e assim há uma oportunidade de avaliar se um determinado valor pode ser de fato atribuído àquela variável.

A hereditariedade é um mecanismo que faz com que uma classe possa ser derivada de outra, herdando todas as características (campos e métodos) da outra classe. Isto faz com que não seja necessário copiar o código fonte de uma classe quando se deseja fazer alterações e também que eventuais correções se propaguem por todas as classes herdeiras. Assim, a hereditariedade possibilita que um código já escrito seja reutilizado de maneira muito mais eficiente do que na programação estruturada. Um programa orientado a objeto costuma implementar verdadeiras árvores genealógicas de classes, com vários níveis de herança.

O polimorfismo é a capacidade dos objetos de assumirem várias formas, através de um mecanismo que permite que referências de tipos de classes mais abstratas representem o comportamento das classes concretas que a referenciam. Assim, um mesmo método pode apresentar várias formas, de acordo com seu contexto.

O aplicativo ***Reconciliare*** é modelado e documentado com a linguagem **UML** (*Unified Modeling Language* – Linguagem de Modelagem Unificada) que é uma linguagem visual utilizada para modelar sistemas computacionais e nos últimos anos se consolidou como uma ferramenta padrão na modelagem de negócios[2], fundamentalmente sob o paradigma da programação orientada a objetos (MEDEIROS, 2004; GUEDES, 2004). A vantagem do uso da UML está na capacidade de lidar com a complexidade natural do projeto, fornecendo para cada aspecto do *software* uma visualização do que tem que ser feito, além de, no decurso do planejamento e execução, gerar a própria documentação do desenvolvimento, algo que seria inviável em um trabalho desta natureza com os clássicos fluxogramas da programação estruturada, cuja semântica, por sua simplicidade, não tem meios de representar esta complexidade natural.

É importante esclarecer que a UML não é uma metodologia de desenvolvimento, ou seja, ela não define como os modelos serão implementados, bem como ela não é uma linguagem de programação, mas sim uma linguagem de modelagem, cujo objetivo é auxiliar os engenheiros de *software* a definir as características do *software* a ser desenvolvido, tais como seus requisitos, seu comportamento, sua estrutura lógica, a dinâmica de seus processos e até mesmo suas necessidades físicas em relação aos equipamentos sobre o quais o sistema deverá ser implantado. Todas estas características são definidas por meio da UML antes do *software* começar a ser realmente desenvolvido. Além disso, ela pode ser também usada em uma via de mão dupla, na qual os modelos UML se convertem em código fonte e alterações no código fonte retornam à modelagem UML. Isto é feito de forma automatizada por algumas ferramentas **CASE** (*Computer Aided Software Engineering* – Engenharia de *Software* Assistida por Computadores).

A necessidade de uma ferramenta tão exigente em seus formalismos, como é a UML, se deve ao fato de que a criação, desenvolvimento e evolução de *software* tornou-se nas últimas décadas uma tarefa de grande complexidade. Da mesma forma que uma pequena edificação pode prescindir de uma planta para sua execução, um aplicativo com propósitos limitados também não exige um planejamento maior do que o puramente mental. A situação se torna muito diferente quando a escala se amplia. No caso da edificação, um prédio de vários andares é virtualmente impossível de ser erguido apenas baseado em planos mentais, com toda a integração de projetos elétrico, hidráulico e estrutural e ainda sob as restrições impostas pelas autoridades do governo. Do mesmo modo, a gerência de um projeto de *software* complexo demanda algum tipo de metodologia no planejamento e na documentação.

[2]O conceito de "negócios" deve ser entendido da forma mais ampla possível neste contexto e significa qualquer atividade ordenada que envolva a execução de diversas tarefas sujeitas a restrições e que apresente variações de estado.

A linguagem UML está atualmente em seu padrão 2.0, constituído de 13 diagramas (Atividades, Casos de Uso, Classes, Objetos, Seqüência, Comunicação, Estado, Pacotes, Componentes, Implantação, Interação, *Timing* e *Composite Structure Diagram*), sendo que estes são usados conforme a necessidade/complexidade do projeto a ser modelado. Neste trabalho foram usados essencialmente dois diagramas: o de casos de uso e o de classes.

A literatura disponível sobre UML é bastante numerosa e diversificada. Além de livros inteiros devotados ao assunto como Guedes (2004), Medeiros (2004) e Booch *et al.* (2000), também aparece como ferramenta em livros dedicados a outros conceitos específicos no âmbito do desenvolvimento de *software*, como em Boratti (2002) e Lee (2001).

O requisito da *extensibilidade* é definido pela possibilidade de serem acrescentadas novas características ao *software* com o menor impacto possível de retrabalho sobre o código fonte. A extensibilidade, que se apóia na modularidade discutida anteriormente, foi desenvolvida no aplicativo ***Reconciliare*** através de um conjunto de interfaces que permitem a criação de novos módulos que conferem novas funcionalidades ou modificam as existentes, inclusive por terceiros que não precisam ter acesso ao código fonte do programa principal, nem usar as mesmas linguagens de programação e ferramentas usadas neste trabalho. Os conceitos e técnicas associadas à extensibilidade serão aprofundados mais a frente neste capítulo.

A *facilidade de uso* foi considerada fundamental no desenvolvimento deste trabalho. Esse requisito relaciona uma série de características como clareza e rapidez na interação com o usuário (usabilidade[3]), ordenamento das ações do usuário e integração com outros sistemas. O requisito da facilidade de uso está longe de ser supérfluo, pois é essencial para mitigar a possibilidade de eventuais enganos cometidos pelo usuário que comprometeriam os resultados da aplicação do *software*.

Uma característica importante relacionada à facilidade de uso é a interação do usuário com o aplicativo de modo geral e, em particular, como o problema de interesse é modelado. A descrição de plantas químicas via *software* tem, nos dias de hoje, um paradigma bem estabelecido que é o

[3] A ISO (*Internacional Standard Organization*) dispõe de duas definições de usabilidade: **ISO/IEC 9126** "A usabilidade refere-se à capacidade de um *software* de ser compreendido, aprendido, utilizado e ser atrativo para o usuário, em condições específicas de uso" Esta definição enfatiza os atributos internos e externos do produto, os quais contribuem à sua usabilidade, funcionalidade e eficiência. A usabilidade depende não só do produto mas também do usuário. É por esta razão que um produto nunca é intrinsecamente usável, ele só terá a capacidade de ser utilizado num contexto particular e por usuários particulares. A usabilidade não pode ser avaliada estudando um produto de forma isolada (BEVAN; MACLEOD, 1994). A **ISO/IEC 9241** tem sua definição centrada no conceito de qualidade de utilização, isto é, refere-se a como o usuário realiza tarefas específicas em cenários específicos com efetividade. "Usabilidade é a efetividade, eficiência e satisfação com que um produto permite atingir objetivos específicos a usuários específicos num contexto de utilização específico".

dos simuladores de processo, nos quais as unidades de processo são colecionadas em paletas e dispostas na área de diagramação (*flowsheeting*) através do recurso de arrastar e soltar (*drag and drop*), para então serem conectadas por correntes de massa e energia. Este paradigma, seguido por todos os simuladores seqüenciais modulares[4], é intuitivo para qualquer profissional com um mínimo de experiência em simulação de processos. Essa argumentação justifica sua adoção neste trabalho.

O descritivo visual do processo, para as necessidades específicas da reconciliação de dados e detecção de erros grosseiros, precisa de níveis diferentes de abstração dos elementos físicos reais. Por exemplo, para o caso mais simples, de subrotinas para o tratamento de restrições lineares e bilineares, que na prática dizem respeito a balanços mássicos globais e de componentes (Capítulo 3) e balanços de energia sob determinadas circunstâncias (Seção 4.3.4), podem ser usados na descrição blocos sem especificação quanto ao tipo de unidade (nó) de processo que representam, mas para problemas não lineares, é necessária uma adaptação para especializar o bloco com as relações não lineares de processo e isto é feito através da associação de código ao bloco, que pode ser compilado em bibliotecas ou interpretado, e que é depois chamado pelos procedimentos matemáticos.

A partir de uma descrição visual do processo, são construídas as matrizes pertinentes e feitas as associações de dados que serão tratados pelos algoritmos de reconciliação e detecção de erros grosseiros. Outro item importante é a capacidade de salvar este descritivo do processo em arquivo para que possa ser usado em outras ocasiões ou que possam ser feitas alterações com segurança, preservando informações.

Finalmente, a interface com o usuário deve fornecer meios para a visualização e alguma forma de interação com os resultados das análises matemáticas. A visualização pode ser através de gráficos combinados, onde se faz uma comparação direta entre o que é medido e o que é estimado, ou pode ser na forma tabular, acessando e exibindo diretamente os valores numéricos. Em relação à interação com os resultados, a interface tem que dar meios para que estes sejam projetados em outros sistemas de informação ou para que sejam usados na forma de elementos de decisão, dentro do próprio aplicativo, para análises subseqüentes.

[4]Simulador seqüencial modular é aquele no qual as operações unitárias, reatores ou conjuntos de unidades são agrupados em módulos (subrotinas) e executados por um programa mestre, seguindo um fluxo de dados relacionado com o fluxo de massa e energia, de modo seqüencial. Exemplos de simuladores deste tipo são o **Hysys**, o **ASPEN** e o **Design II**. Esta abordagem é antagônica à simulação baseada em equações, na qual a modelagem é feita diretamente sobre as equações e a resolução é sobre o sistema completo resultante. O **gProms** é um exemplo de simulador deste tipo.

Para que as subrotinas desenvolvidas possam ser simuladas com flexibilidade e verossimilhança e/ou analisadas dentro de um sistema real, é necessária a comunicação com *softwares* de uso corrente na indústria. Por causa desta necessidade, foi colocado como requisito a *integração com sistemas externos*, a exemplo da comunicação com banco de dados relacionais[5] e com servidores **OPC**[6], descritos a seguir.

As variáveis tratadas dentro do aplicativo ***Reconciliare*** fazem acesso de leitura e escrita nos referidos sistemas externos para receber valores medidos, escrever valores estimados, indicar variáveis com erros grosseiros, ler parâmetros de projeto e/ou operação, levantar alarmes, etc. Este tipo de comunicação de dados padronizada com sistemas externos faz com que o *software* possa ser usado indistintamente em modo de operação ou em modo de simulação, pois a origem dos dados é indiferente, bastando que estes dados sejam disponibilizados sob o padrão adotado.

O conjunto de especificações da *OPC Foundation*[7] tem o propósito de resolver o problema da padronização da transferência de dados entre *softwares*, equipamentos e redes de diferentes fornecedores da área de automação e controle industrial. Entre as empresas que lançaram o projeto estão a Honeywell, Siemens, Emerson, Rockwell, GE Fanuc, National Instruments, Toshiba, Microsoft, Aspen e Matrikon. Originalmente baseado na tecnologia **OLE** (*Object Linked and Embedded* – Objeto Ligado e Embarcado), da Microsoft, teve a primeira versão do seu padrão lançada em 1996 e, desde então, tem se tornado cada vez mais abrangente e aceito. A tecnologia OLE passou a se chamar **COM** (*Component Object Model*) e sua transferência via rede de **DCOM** (*Distributed* COM). Atualmente, o padrão OPC começa a se afastar da tecnologia proprietária Microsoft e são feitos os primeiros ensaios de um padrão independente, baseado em tecnologias livres como o protocolo **http** e a linguagem de marcação **XML** (*Extensible Markup Language*), o que vai tornando o OPC universal tanto no nível dos equipamentos e redes industriais quanto no nível dos sistemas operacionais.

A idéia central do padrão OPC é que os fabricantes de dispositivos industriais, como **PLCs**, escrevam *drivers* OPC para seus equipamentos e assim a camada de *software* (concentradores de

[5] O banco de dados relacional escolhido é o **Firebird**, por ser um projeto *opensource*. Não se trata de um gerenciador de banco de dados de uso corrente na indústria (como **Sybase**, **Oracle**, **MS Acces**), mas observando-se certos padrões, nomeadamente a linguagem SQL, a transição para os aplicativos citados requer um mínimo de recodificação.

[6] O padrão OPC significava em sua origem *OLE for Process Control* (OLE para Controle de Processos), mas esse acrônimo perdeu o sentido pois a tecnologia OLE da Microsoft mudou de nome e os dispositivos e *softwares* que aderem ao padrão não são usados apenas na indústria de processos, mas também na indústria de manufatura.

[7] www.opcfoundation.org

dados, *data historians*, **HMIs**, **SCADAs**[8], etc) em conformidade com o padrão pode se comunicar com qualquer equipamento de modo indistinto, ao invés de ter que escrever um *driver* para cada equipamento individualmente, como era necessário antes do padrão OPC. Isso libera o consumidor final para escolher o seu *software* baseado nas características que considera desejáveis e não na disponibilidade de *drivers* para o seu equipamento específico[9].

A vantagem de aderir a um padrão deste tipo é que um aplicativo que troque dados em conformidade com ele tem a garantia de poder se comunicar com uma vasta gama de sistemas de *hardware* e *software* e necessitar o mínimo de customização quando instalado em um ambiente industrial. O aplicativo ***Reconciliare*** é um *cliente* OPC e troca dados com servidores do tipo *Data Access*[10].

8.1.2 Modelagem e desenvolvimento do aplicativo *Reconciliare*

As abordagens para resolução dos problemas de reconciliação de dados e detecção de erros grosseiros alcançaram um nível importante de refinamento teórico, como é extensamente reportado na literatura mais recente. Contudo, existem ainda algumas lacunas entre esta teoria e a prática industrial. A proposta deste trabalho é diminuir algumas destas lacunas, investigando, propondo e implementando soluções que sejam atraentes para a indústria de processos químicos e petroquímicos.

Um dos diferenciais deste trabalho é o desenvolvimento de uma aplicação genérica. Um grande número de *softwares* desenvolvidos em trabalhos acadêmicos nesta área (PLÁCIDO, 1995; MENDES, 1995; BARBOSA JÚNIOR, 1996; TEIXEIRA, 1997) é voltado para um processo específico ou se apóia em meios pouco intuitivos como a entrada dos dados da matriz de incidência, da matriz de covariância e os dados a serem reconciliados na forma tabular, por entrada manual ou arquivos de texto. Um aplicativo genérico deve dar meios de nele próprio modelar o problema de interesse e de

[8] HMI – Human Machine Interface (Interface Homem-Máquina) e SCADA – *Supervisory Control and Data Acquisition* (Controle Supervisório e Aquisição de Dados)

[9] Uma analogia a esta situação é feita com o suporte a impressoras. Primeiramente, o suporte à impressão era feito pelo desenvolvedor do aplicativo e ele deveria criar um *driver* de impressão para cada modelo de cada marca de impressora para a qual se pretendesse dar suporte. Assim, um aplicativo de CAD ou um editor de textos tinham seus próprios *drivers* para vários modelos de impressora. A situação mudou quando a Microsoft colocou o suporte à impressão no próprio sistema operacional. Desta forma, todo aplicativo que necessitasse suporte à impressão se comunicaria com o sistema operacional e este é que faria a impressão de fato. Neste modelo de desenvolvimento mais racional, quem fornece o *driver* é o fabricante da impressora.

[10] o padrão OPC comporta uma série de especificações: *Data Access, Alarms & Events, Batch, Data eXchange, Historical Data Access, Security, XML-DA, Complex Data* e *Commands*

analisar os resultados, de preferência simultaneamente com os próprios cálculos.

Nas seções seguintes são apresentados os detalhamentos da modelagem do aplicativo ***Reconciliare*** através dos diagramas de casos de uso da UML, seguido de um descritivo das principais partes que o compõem, para serem então detalhadas as classes que representam as diversas entidades dentro do aplicativo, finalizando com uma descrição da sua interface gráfica com o usuário.

8.1.2.1 Diagrama de Casos de Uso

Segundo Medeiros (2004), o **Diagrama de Casos de Uso** é um dos principais diagramas na modelagem e construção de um *software* orientado a objetos utilizando a UML. Este diagrama é o mais geral e informal da UML, sendo utilizado normalmente nas fases de levantamento e análise de requisitos do sistema, embora venha a ser consultado durante todo o processo de modelagem e possa servir de base para outros diagramas. Apresenta uma linguagem simples e de fácil compreensão para que os desenvolvedores possam ter uma idéia geral de como o sistema irá se comportar. Nele são descritas todas as possíveis interações do aplicativo com os atores envolvidos. O ator pode ser um usuário humano em uma determinada tarefa, um sistema ou uma entidade externa, como um dispositivo físico de medição colocado em campo. Um caso de uso é interpretado como uma macroatividade que encerra diversas tarefas ou atividades menores. Essas tarefas visam à consecução da macroatividade. Os elementos notacionais do Diagrama de Casos de Uso são os *stickmen* para os atores e elipses para os casos. Esta semântica básica está ilustrada na Figura 8.1.

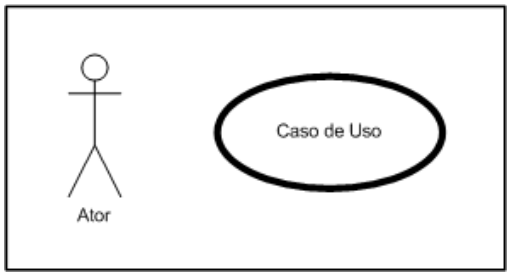


Figura 8.1: Semântica básica do diagrama de casos de uso

O diagrama da Figura 8.2 mostra os casos de uso do aplicativo ***Reconciliare***: o cadastro de fontes de dados, o cadastro de dados locais, a descrição do processo (modelagem), a programação de tarefas, a realização de análises e a comunicação com o usuário cliente e o sistema de informações. Além disso são mostrados também os três atores envolvidos: usuário programador, usuário cliente e o sistema de informações. Os atores usuário programador e usuário cliente do diagrama

podem representar a mesma pessoa atuando de formas diferentes.

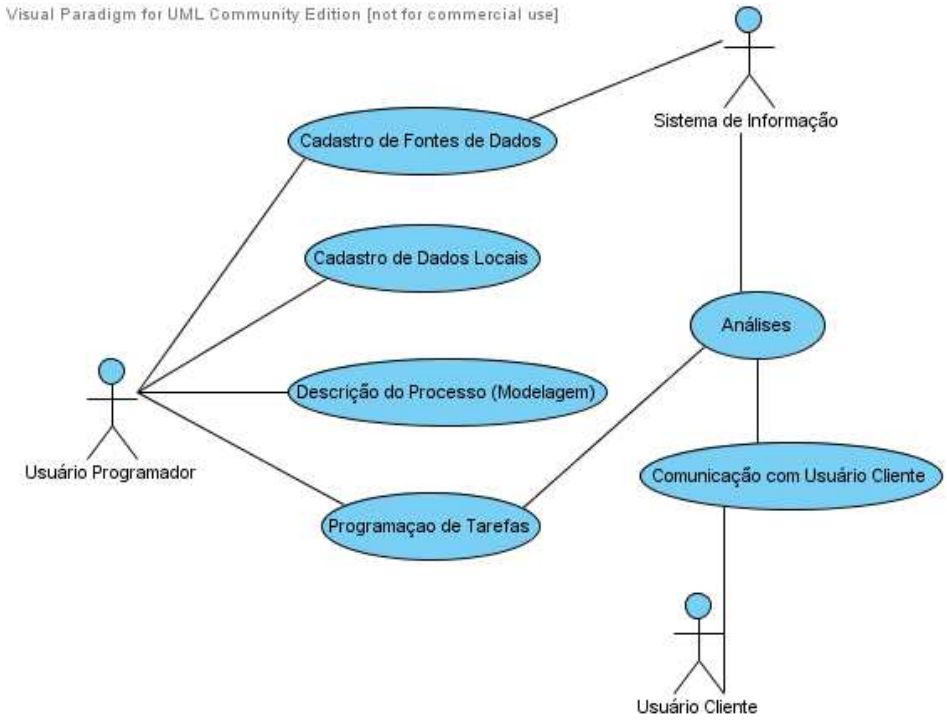


Figura 8.2: Diagrama de casos de uso do aplicativo ***Reconciliare***

O centro do aplicativo é o caso de uso da realização de análises. Cada tarefa analítica (média móvel, reconciliação de dados, detecção de erros grosseiros, etc.) se está ativa, é passível de ser executada continuamente em *loop* infinito, a depender de algum condicional como um agendamento para o um determinado horário ou algum evento específico disparado por outra análise. O resultado de uma tarefa analítica pode ser algum tipo de crítica sobre o resultado, o desvio do fluxo de tarefas, a geração de um alarme ou relatório, etc.

A programação da lista de tarefas analíticas é feita pelo usuário programador, que faz ainda a modelagem dos processos, associa os modelos às análises e faz as conexões das variáveis locais (vazões, composições, temperaturas, etc.) às fontes externas de dados, representadas pelo ator sistema de informações no diagrama.

O usuário comum pode ser desde um gerente de manutenção a um operador da área, recebendo informações pertinentes que sugiram algum tipo de ação humana sobre a planta ou que dêem subsídios para decisões gerenciais.

O último caso de uso é a ação do aplicativo sobre o sistema de informações, por exemplo escrevendo as estimativas da reconciliação ou coaptação de dados diretamente no servidor de dados,

tornando estas estimativas disponíveis para os sistemas de controle e otimização da planta.

8.1.2.2 Estrutura do aplicativo

O aplicativo ***Reconciliare*** é baseado em uma estrutura simples de representação das quatro principais entidades com as quais o programa lida e o relacionamento entre elas. Quais sejam:

- Dados externos;
- Dados locais;
- Modelos;
- Tarefas.

Os dados externos, como já foi colocado anteriormente, podem vir de bancos de dados ou principalmente de fontes de dados OPC. Os dados locais são variáveis de visibilidade e acesso total em todo o programa e podem ser criados sem limitação de número. A cada dado externo de interesse é criado pelo menos um dado local para que este seja efetivamente manipulado pelas análises. Além da cópia local dos dados externos, podem ser criadas outras variáveis para armazenar resultados intermediários e/ou finais das análises. O modelo, que é criado visualmente, descreve o relacionamento entre as entidades físicas do processo como correntes de massa e energia, unidades de separação, divisores de corrente, *mixers*, etc. Essa descrição visual gera as matrizes que são usadas pelas análises matemáticas. No modelo também são feitas as associações entre os dados locais e as entidades físicas. Finalmente, as tarefas representam os métodos matemáticos que são chamados para operar sobre o conjunto do modelo com os dados.

Quando o cliente OPC do aplicativo se conecta a um servidor de dados (Seção 8.2, à página 232) a cada renovação dos valores é gerado um evento que é usado para renovar os dados também internamente. Essa renovação dos dados externos pode se dar duas formas: na freqüência de atualização do servidor ou numa taxa predeterminada no cliente. A escolha entre estas possibilidades está relacionada com a dimensão do problema, de modo que o peso computacional pode ser aliviado escolhendo-se uma taxa mais lenta que a da atualização dos dados no servidor. Por outro lado, os dados locais que sejam espelho dos dados externos têm sua atualização atrelada à atualização dos dados externos.

A modelagem do aplicativo foi feita de modo a permitir a sua extensibilidade pela criação de novas análises que podem ser facilmente incorporadas, inclusive sem a necessidade de recompilação. Quando uma nova tarefa é criada, na sua configuração o usuário é chamado a escolher a partir de uma lista qual a análise que aquela tarefa vai executar. Para cada subrotina de análise é prevista também uma janela gráfica de interface com o usuário para que a análise seja configurada. A análise e as janelas de configuração estão compiladas em subrotinas armazenadas em arquivos do tipo `*.DLL`[11]. Para que um procedimento seja chamado e executado com sucesso, basta que estas subrotinas atendam aos padrões de passagem de parâmetros predeterminados no aplicativo ***Reconciliare***. Desta forma, o aplicativo se abre para o desenvolvimento de terceiros, bastando que sejam fornecidos os arquivos com as subrotinas e janelas de configuração.

8.1.2.3 Diagrama de Classes

O **Diagrama de Classes** é o diagrama mais utilizado e importante da UML, servindo de apoio para a maioria dos outros diagramas. Ele define a estrutura das classes que descrevem o sistema, determinando os campos e métodos de cada classe, além de estabelecer como as classes se relacionam e trocam informações entre si. Uma classe é representada por um retângulo com até três divisões, descritas a seguir (com referências ao exemplo ilustrado na Figura 8.3):

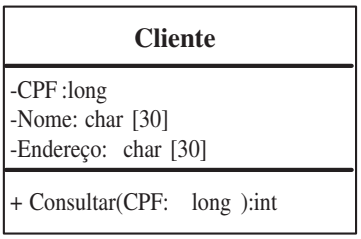


Figura 8.3: Semântica básica do diagrama de classes

- A primeira contém a descrição ou o nome da classe (*Cliente* no exemplo);
- A segunda armazena os campos e os seus tipos de dados (atributos *CPF*, do tipo `long`, *Nome* e *Endereço* do tipo `char`).

[11] *Dynamic Link Library* – Biblioteca de Vínculo Dinâmico. É a implementação feita pela Microsoft para o conceito de bibliotecas compartilhadas. O formato do arquivo DLL é o mesmo dos arquivos executáveis para Windows. Assim como os EXEs, as DLLs podem conter códigos, dados e recursos (ícones, fontes, cursores, entre outros) em qualquer combinação. As DLLs provêem os benefícios comuns de bibliotecas compartilhadas, como a modularidade. Esta modularidade permite que alterações sejam feitas no código ou dados em uma DLL auto-contida, compartilhada por vários aplicativos, sem que qualquer modificação seja feita nos aplicativos em si. Essa forma básica de modularidade permite a criação de *patches* e *service packs* relativamente pequenos.

- A terceira divisão lista os métodos da classe (método *Consultar* recebendo o *CPF*:`long` e retornando um `int`).

O Diagrama de Classes ilustrado na Figura 8.4 mostra as principais classes do aplicativo ***Reconciliare***. A seguir são descritas brevemente algumas das classes[12] que aparecem neste diagrama.

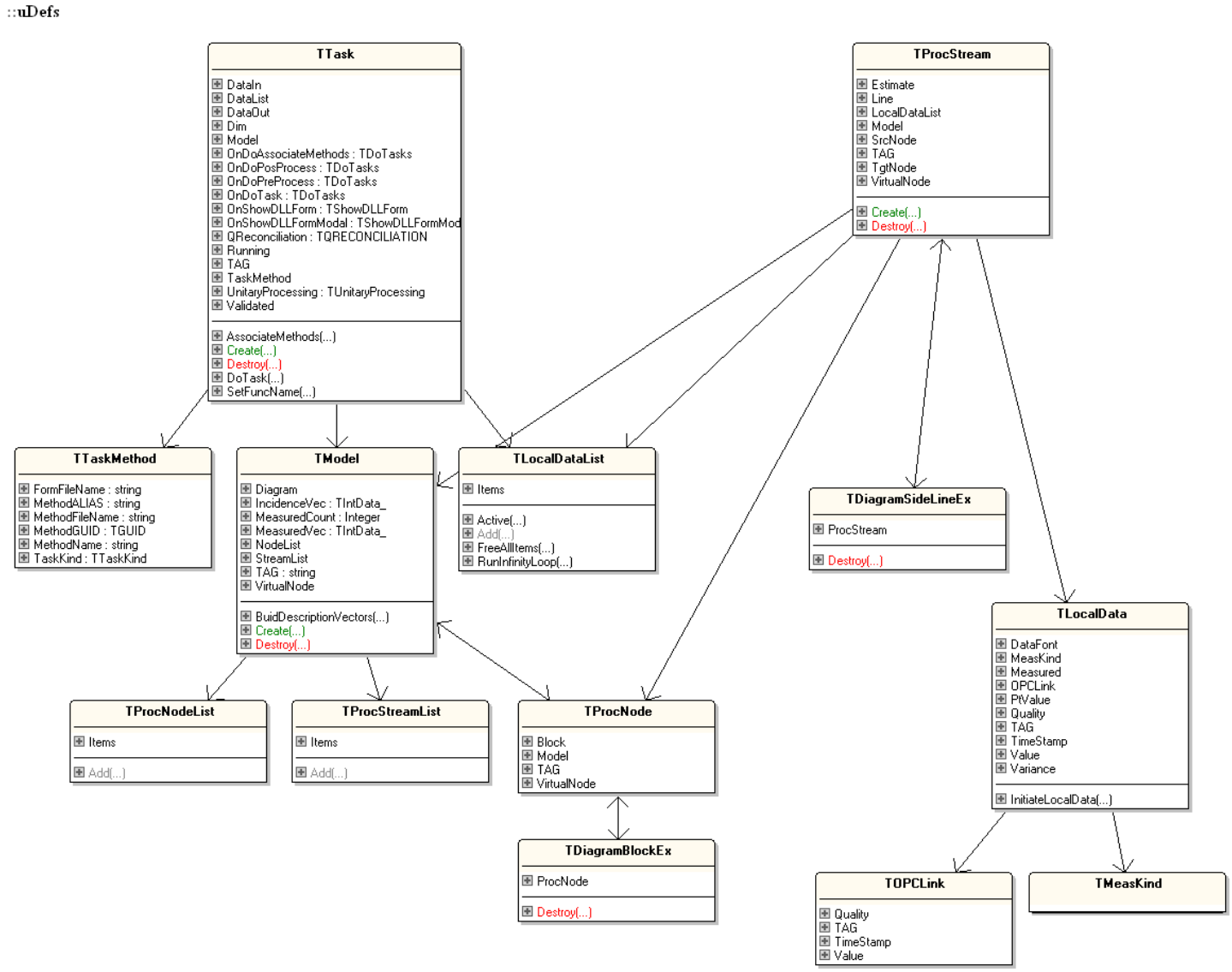


Figura 8.4: Diagrama de Classes (UML) parcial do aplicativo ***Reconciliare***

TOPCLink Esta classe descreve as ligações individuais aos dados publicados em um servidor OPC. Ela tem quatro campos principais: `TAG` para o nome da variável, `Value` para o valor

[12]Por tradição na programação em Delphi, costuma-se começar os nomes das classes com a a letra ***T***, de *tipo*

da variável, `TimeStamp` para data e hora da leitura e `Quality` para informações sobre a qualidade da leitura. A informação em `Quality` pode ser usada em tomadas de decisão quanto a considerar a variável como medida ou não medida e seus possíveis estados são definidos pelo padrão OPC. Os objetos definidos pela classe TOPCLink são colecionados em um objeto global da classe **TOPCLinkList**.

TLocalData Esta classe descreve os dados locais que são efetivamente usados nas análises feitas pelo programa. Ela possui os campos mostrados na classe **TOPCLink** e também o campo `Variance` para guardar a variância do dado, `DataFont` que descreve se o dado vem de um OPCLink, de um banco de dados ou se é um dado local, `MeasKind` que é o tipo da medida associada à variável: temperatura, pressão, vazão, concentração, etc., `Measured` indicando se a variável é medida ou não, `OPCLink` que aponta para o objeto da classe **TOPCLink** quando se trata de um espelho do dado externo. Neste contexto, apontar significa guardar o endereço de memória do objeto para recuperar seus campos e métodos quando for necessário. Isto é diferente de manter uma cópia do dado em outro objeto, pois assim qualquer variação posterior de estado no objeto original subseqüentemente à cópia será perdida.

TatDiagram Esta é a classe fundamental do componente visual **DiagramStudio**. Este componente é usado no aplicativo ***Reconciliare*** para criar e manipular gráficos que representam o processo sendo analisado (ver (1) na Figura 8.11). Ele é produzido pela **Automa Software**[13] e distribuído pela **TMS Software**[14] como *freeware* para uso não comercial e fornece todos os elementos necessários para o desenho e relacionamento gráfico entre as setas (correntes de processo) e os blocos (unidades de processo), a partir dos quais é feita a descrição do problema físico na forma de matrizes através da interpretação dos elementos do *flowsheet*.

TProcStream Esta classe representa as correntes do processo. Os seus campos principais são: `TAG` que é o nome da corrente, `Line` que aponta para uma seta no *flowsheet* do processo (desta forma a descrição lógica é atrelada à descrição visual), `SrcNode` que aponta para o nó do processo do qual a corrente parte e `TgtNode` que aponta para o nó do processo para o qual a corrente se dirige, `LocalDataList` que é uma lista de referências que apontam para os dados do tipo **TLocalData** da lista de dados globais, mas somente com elementos pertinentes à corrente. `Model`, do tipo **TModel** que será discutido mais a frente, que aponta para o modelo ao qual a corrente pertence.

[13] http://www.automa.com.br/
[14] http://www.tmssoftware.com/site/diagram.asp

TProcNode Esta classe representa os nós do processo como divisores de corrente, *mixers*, colunas de separação, reatores, etc. Entre seu campos principais estão a `TAG` que é o nome que identifica o nó, `Model` (do tipo **TModel**) que aponta para o modelo ao qual o nó pertence e `Block` que aponta para o bloco visual no *flowsheet* do processo.

TModel Esta classe contém os elementos que fazem a descrição de um modelo de processo. Seus campos principais são a `TAG` que é nome de identificação do modelo, `Diagram` que é um objeto da classe **TatDiagram** que controla os gráficos do *flowsheet*, além de duas listas para os elementos do *flowsheet*: uma para as correntes e a outra para os nós (`NodeList` e `StreamList`).

TTask Esta classe é a principal classe do aplicativo ***Reconciliare***. As informações das demais classes se associam a esta que trata da chamada de algumas das subrotinas matemáticas descritas nos Capítulos 3 a 7. A **TTask** tem um conjunto de métodos para pré-processar e pós-processar as análises. Cada análise é disparada em uma *thread* separada com um *loop* condicional ao campo booleano `FRunning`. Há um pré-processamento anterior ao *loop* e um interno ao *loop*, antes da chamada da rotina principal que é codificada e disponibilizada em arquivos `*.DLL`. Há também um pós-processamento posterior à chamada da rotina principal e ainda dentro do *loop* e, finalmente, um pós-processamento depois do *loop*. Essas chamadas a pré e pós-processamentos são opcionais e para não serem executadas, basta que não seja associado um método.

As quatro principais listas do aplicativo (lista de items OPC, lista de dados locais, lista de modelos e lista de tarefas) são objetos globais e únicos, o que significa que os objetos mantidos nestas listas são também únicos e alcançáveis a partir de qualquer parte do *software*. Isto implica que os elementos da lista de variáveis locais mantida em cada uma das correntes do modelo simplesmente aponta para um objeto que já foi previamente instanciado na memória, fazendo com que alterações nessas variáveis sejam refletidas em todas as partes do *software* que se utilizarem delas.

8.1.2.4 Apresentação do aplicativo *Reconciliare*

A Figura 8.5 mostra a interface gráfica principal do aplicativo ***Reconciliare***. Quando este é iniciado, um arquivo de configurações gerais (`Reconciliare.xml`) é lido. Esse arquivo é mostrado parcialmente na Listagem 8.1. Nele estão, por exemplo, os nomes das subrotinas das análises

matemáticas e o caminho para os arquivos `*.DLL` onde estão compiladas estas subrotinas. Em um dos detalhes da Figura 8.5 está o menu `File` (Arquivo). Neste menu, que segue o padrão dos menus de manipulação de arquivo, pode-se escolher entre criar um arquivo novo, abrir um existente, salvar o atual, salvar o atual com um novo nome, fechar o arquivo e sair do aplicativo. Na mesma figura, em outro detalhe, é mostrada a caixa de diálogo para abrir arquivos de *setup* de uma seção de trabalho do aplicativo ***Reconciliare***. Esse arquivo tem a extensão `*.REC` (*reconciliare case file*) e é estruturado como um arquivo XML. Estes arquivos `*.REC` contêm toda a descrição de conexões OPC, items locais, *flowsheet* de processo e análises de interesse de uma seção de trabalho do aplicativo.

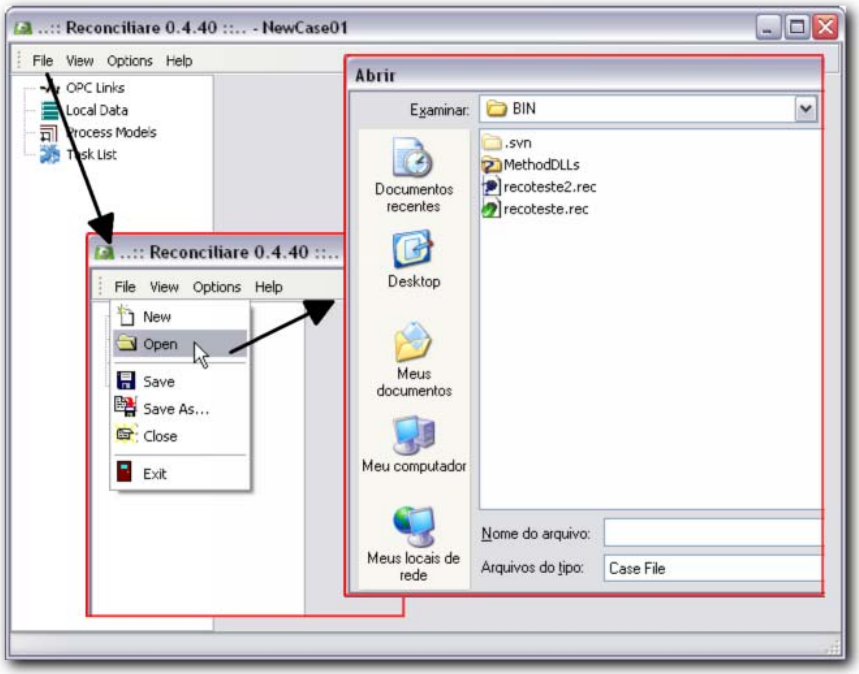


Figura 8.5: Interface gráfica do aplicativo ***Reconciliare***

Listagem 8.1: Reconciliare.xml

```
<ReconciliareConfig>
  <MethodList>
    <Method MethodName="Moving Average" MethodGUID="{1BC1FB24-20C6-4DC3-9B1F-9C0CC695624A}" MethodFileName="..\MethodDLLs\
        TesteMedia.dll" ConfigFormFileName="..\MethodDLLs\TesteMediaFORM.dll" TaskKind="tkUnitaryProcessing" ALIAS="TESTEMEDIA"/>
    <Method MethodName="Simple Reconciliation" MethodGUI="{48BDDAEB-B809-40B9-A1A0-FC9DD6052916}" MethodFileName="..\MethodDLLs\
        QReconciliation.dll" ConfigFormFileName="..\MethodDLLs\QReconciliationFORM.dll" TaskKind="tkQReconciliation" ALIAS="
        QRECONCILIATION" />

    (...)

  </MethodList>
  <Config>
  <LocalDataGraphic Color="$FF0000" PtPg="250" />

  (...)
```

A Figura 8.6 mostra a lista de items OPC carregados a partir do servidor de dados. No painel à esquerda (1) está a lista dos items. Os botões `Connect` e `Disconnect` (2) controlam a conexão e desconexão ao servidor de dados e as TAGs, valores, qualidades e *time stamps* das variáveis externas podem ser observados em (3) .

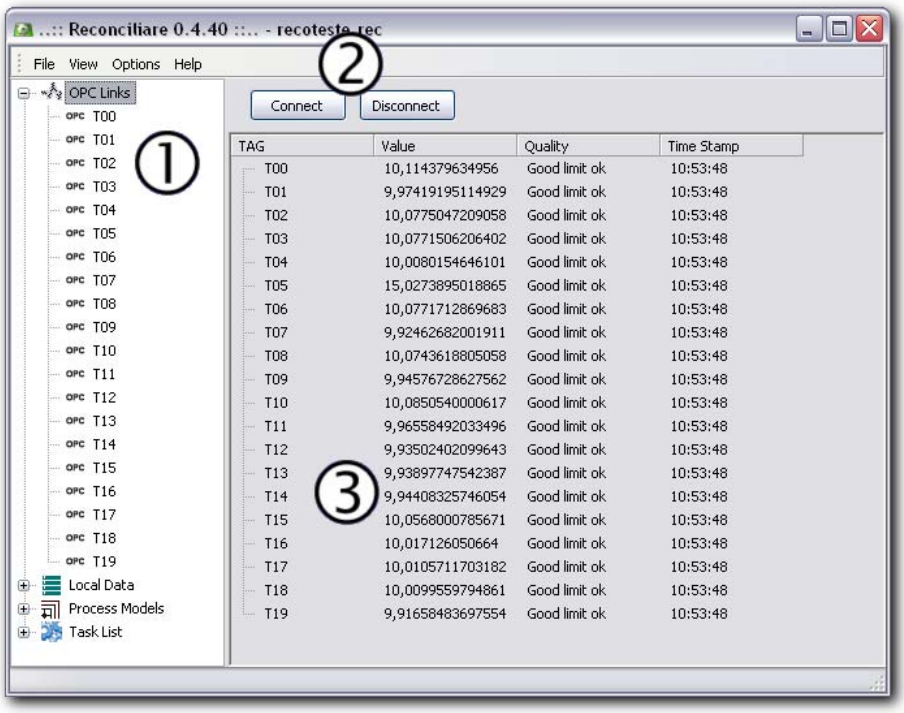


Figura 8.6: Lista de items OPC conectados

A lista de items observada em (3) (Figura 8.6) não pode ser editada. As TAGs, valores, qualidades e *time stamps* vêm do servidor de dados que pode ser um simulador ou um servidor de dados reais medidos em campo.

A Figura 8.7 mostra no painel à direita os dados de um item OPC selecionado na lista à esquerda. Podem ser observadas a TAG da variável, seu valor numérico e no painel abaixo o gráfico dos valores da variável evoluindo no tempo.

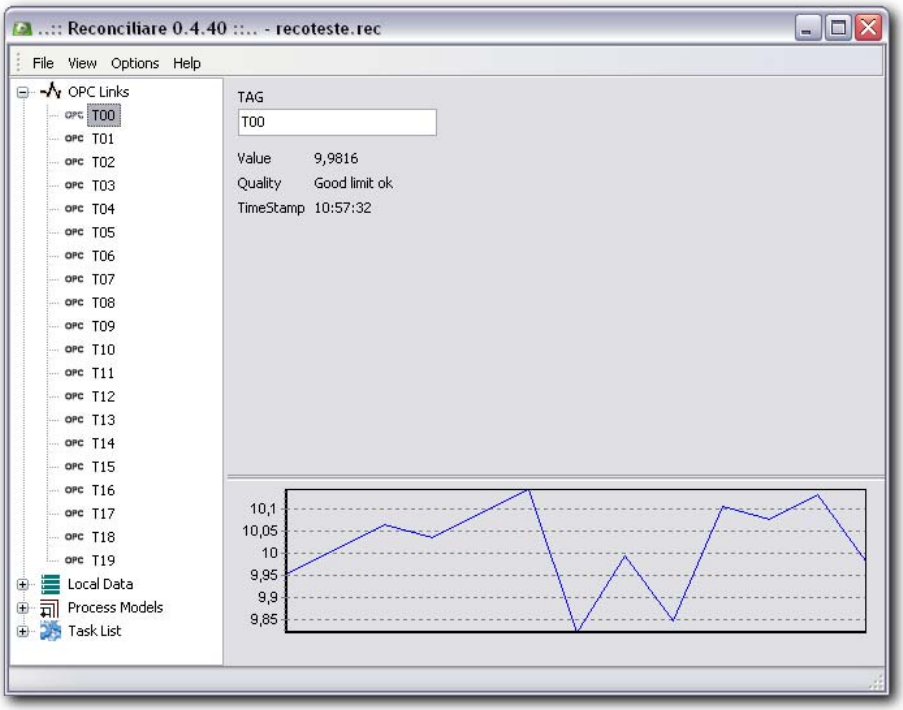


Figura 8.7: Características de um item OPC

A Figura 8.8 mostra a lista de dados locais ao aplicativo. No painel a esquerda (1) está a lista dos items, os botões `Add New` e `Delete` (2) gerenciam a inclusão e exclusão de items. As TAGs, valores, qualidades e *time stamps* das variáveis globais do aplicativo podem ser observados em (3) .

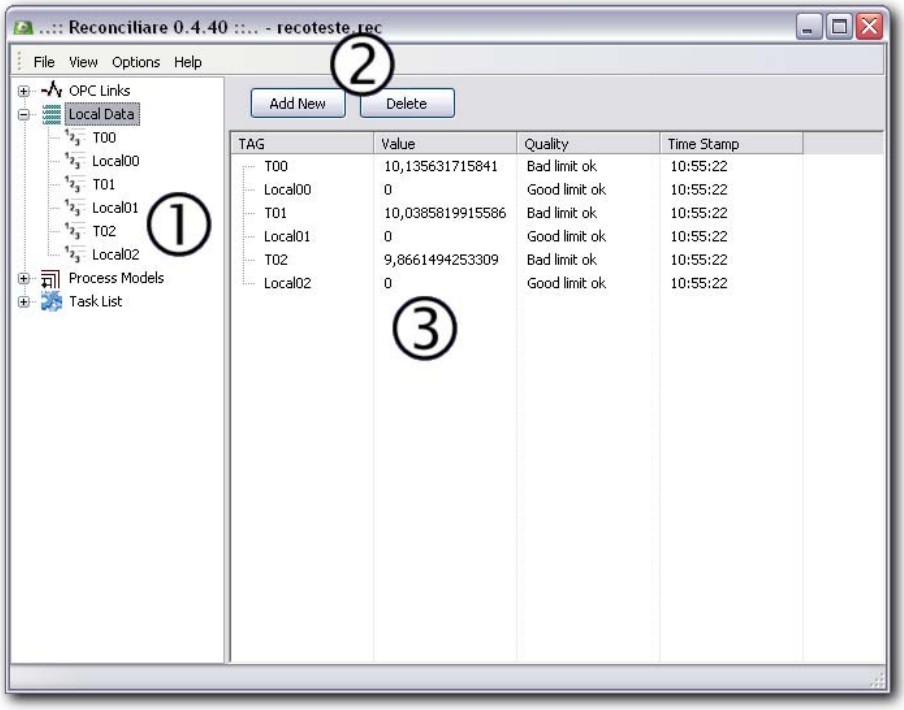


Figura 8.8: Lista de dados locais

A Figura 8.9, do mesmo modo que para o item OPC, mostra no painel à direita os dados de um item de dado local selecionado na lista à esquerda. Podem ser observadas a TAG da variável, seu valor numérico, a fonte de dados e a variância. No painel abaixo é mostrada a evolução gráfica dos valores da variável. Esse painel é também uma ferramenta de edição do item. Se a fonte de dados for um item OPC, apenas a TAG não pode ser editada e mantém o mesmo nome da variável externa que espelha.

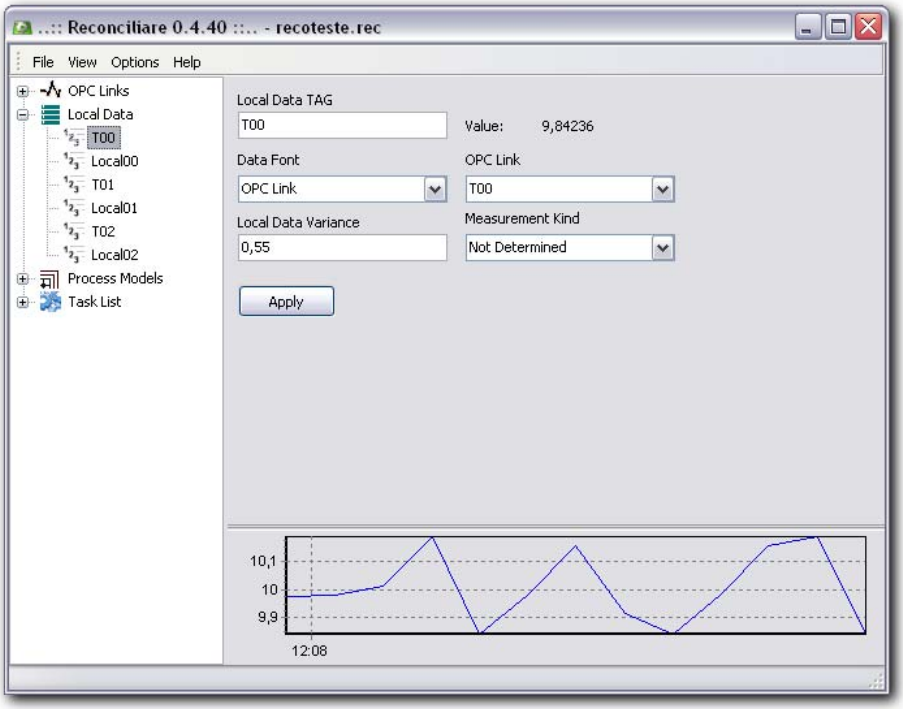


Figura 8.9: Características de um dado local

A Figura 8.10 mostra a lista de modelos de processo no aplicativo ***Reconciliare***. No painel à direita estão os botões de inclusão de um novo modelo e exclusão dos modelos selecionados.

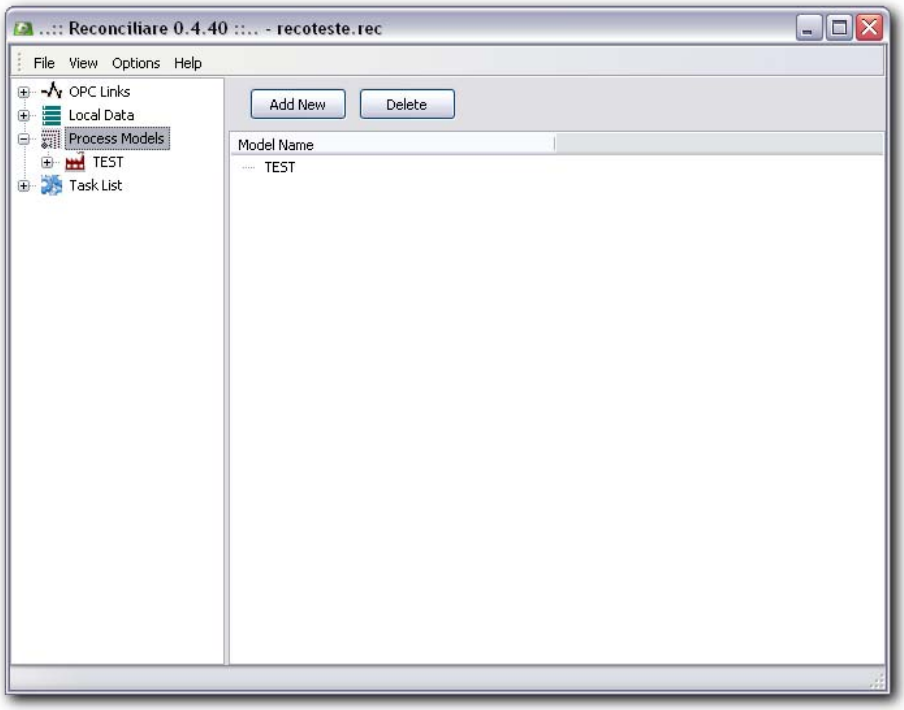


Figura 8.10: Lista de modelos

A Figura 8.11 mostra na lista à esquerda (1) a estrutura lógica de um `Model`, com seus nós e correntes de processo. Em (2) está a barra de botões para criação de novos nós e correntes de processo e sua validação. Em (3) está a área de desenho (*flowsheet*) do processo sendo analisado. As novas correntes que são criadas podem ser arrastadas com o mouse até ancorarem em pontos de ligação nos blocos. Quando isto ocorre é feita também uma conexão lógica que gera a matriz de incidência, como se pode observar em (4) .

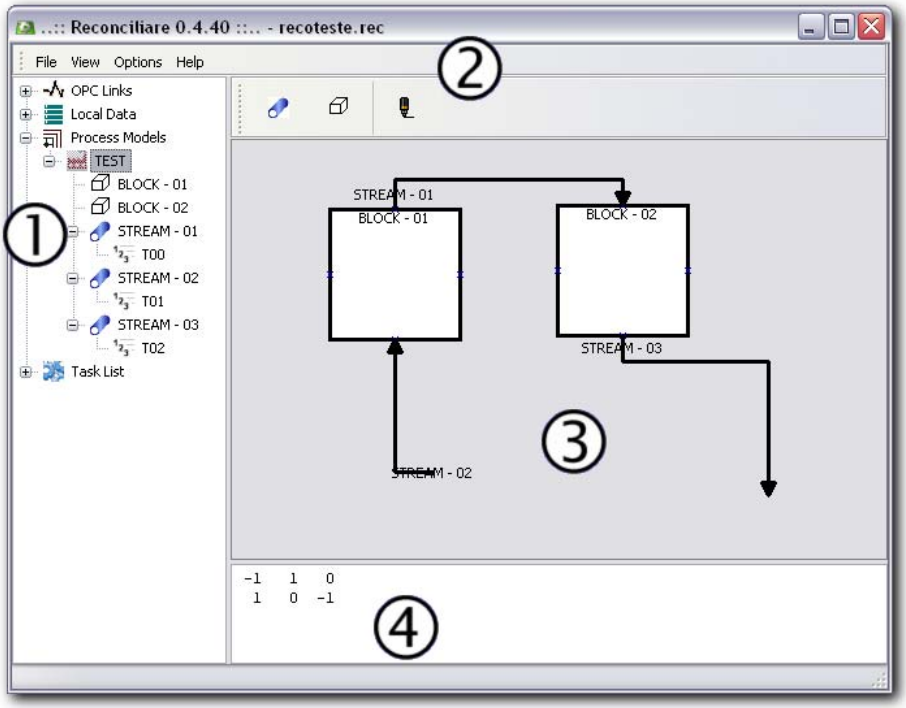


Figura 8.11: *Flowsheet* de processo associado ao modelo

A Figura 8.12 mostra a relação de associação entre uma corrente e as variáveis da lista de variáveis locais. Na lista à esquerda é mostrada, pendendo da corrente, a variável associada a ela e na lista do painel à direita essa variável se mostra com uma marca de checagem (*check mark*). A lista de variáveis associadas a uma corrente de processo pode ser editada selecionado e desselecionando os items. Cada corrente pode ter uma série de variáveis associadas como valores de temperatura, pressão, etc. Esta versão do aplicativo ainda não faz a verificação de associação de uma variável local a mais de uma corrente e isto deve ser, então, verificado pelo usuário.

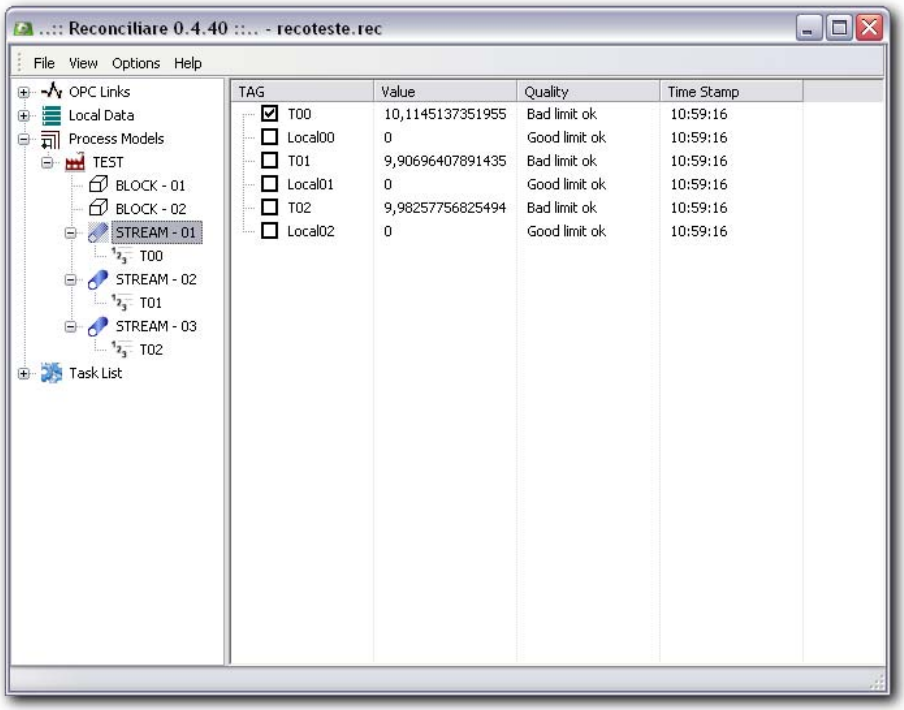


Figura 8.12: Lista de dados locais associados à corrente do modelo

A Figura 8.13 mostra a lista de tarefas disponíveis para serem usadas sobre um determinado modelo. Na lista à direita são mostrados o nome da tarefa e se está ativa ou não. Um duplo clique sobre o item, chama a janela de configuração da tarefa (Figura 8.14). Uma vez que a tarefa seja ativada, uma série de eventos pode ocorrer. Primeiro é feita uma chamada para um pré-processamento antes do *loop* da tarefa. Esse pré-processamento é executado somente uma vez a cada mudança de status para ativo. Se a tarefa sendo configurada não necessitar desse pré-processamento ou se este for opcional, não é associado nenhuma função à esta chamada e nada ocorre. Uma vez dentro do *loop*, há também um pré-processamento que antecede à chamada da função principal e um pós-processamento que a sucede. Do mesmo modo, se não houver necessidade, essas chamadas podem não disparar evento algum. Finalmente, depois do *loop* (que é interrompido mudando o valor do status), pode ser executado um último pós-processamento. Todas essas chamadas são codificadas para marcar lugar e assegurar flexibilidade de aplicação.

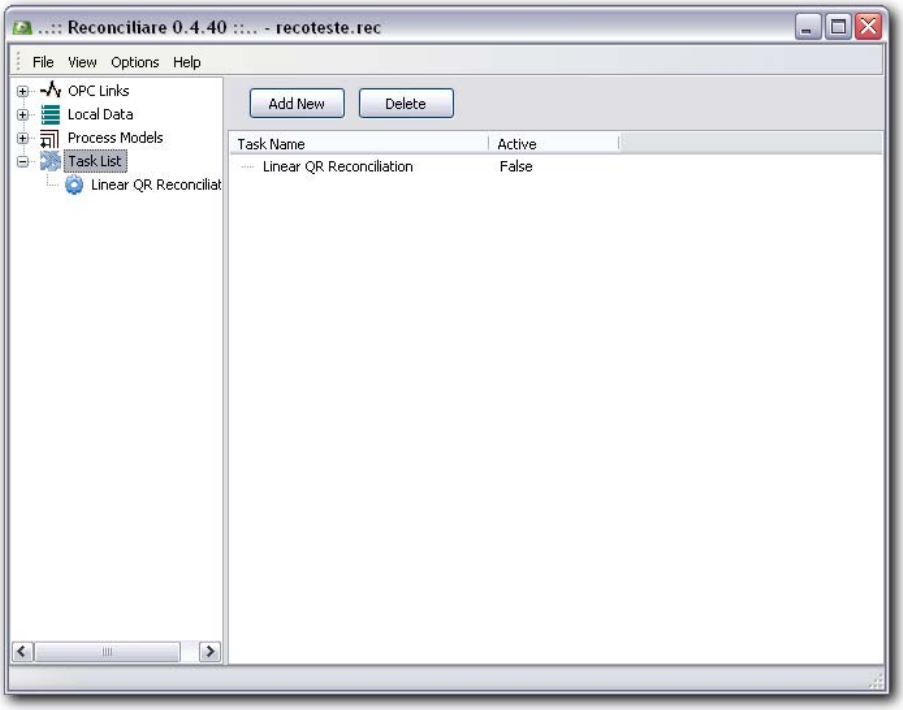


Figura 8.13: Lista de tarefas

A Figura 8.14 mostra a janela de adição e edição de uma nova tarefa. Nesta janela é escolhido o nome, o modelo sobre o qual a `Task` vai operar, a análise, da lista de análises disponíveis (em `Method`) e o período entre cada disparo da análise. Uma vez que uma análise tenha sido escolhida, é necessário clicar no botão `Config Task`, para que a janela de configuração da análise seja chamada. O período entre disparos da análise funciona como um temporizador e pode ser ajustado para uma freqüência semelhante à da atualização dos dados.

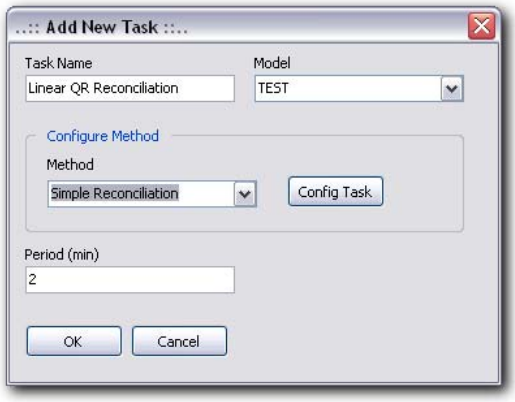


Figura 8.14: Janela de configuração de uma tarefa

Seguindo os casos de uso associados ao usuário programador (Seção 8.1.2.1 e Figura 8.2), pode-se sugerir o seguinte roteiro de trabalho, considerando-se a definição de um problema desde o início.

Passo 1 Abrir uma seção do aplicativo ***Reconciliare*** e iniciar uma conexão com um servidor OPC (Figura 8.6). Uma vez que a conexão esteja aberta, observar se os items estão todos carregados e se de fato há alguma alteração nos valores que evidencie que o servidor está ativo (Figura 8.7);

Passo 2 Selecionar o nó `Local Data` (Figura 8.8) e começar a criar os dados locais. Criar pelo menos um espelhamento local para cada item OPC com o qual se deseje trabalhar nesta seção. Após a inserção de um novo dado, ele pode ser editado para refletir as necessidades do usuário (Figura 8.9). É importante fazer a atribuição da variância manualmente, uma vez que nesta versão do aplicativo não há ferramentas para depreender ou construir esta informação a partir dos próprios dados de entrada;

Passo 3 Selecionar o nó `Process Models` e adicionar um novo modelo (Figura 8.10). Iniciar a reprodução do *flowsheet* criando correntes e nós do processo e os arrastando e conectando para reproduzir as conexões existentes;

Passo 4 Configurar em cada corrente quais são os dados pertinentes. A quantidade de dados associados é limitada pela memória da máquina utilizada e é importante fazer as associações corretamente para reproduzir as relações existentes no processo real (Figura 8.12);

Passo 5 Criar uma `Task` (Figura 8.13) e associar a ela um modelo da lista de modelos e uma análise da lista de análises disponíveis. Com tudo definido, pode-se então ativar as tarefas (`Tasks`) de interesse.

Passo 6 Salvar a seção em arquivo (`*.REC`) para posterior retorno e/ou edição das características.

8.2 O aplicativo Servidor de Dados OPC

Uma ferramenta indispensável para o projeto de sistemas de monitoramento da qualidade dos dados de processo, como a reconciliação de dados e detecção de erros grosseiros, deve ser o sistema de testes, ou seja, meios de simular dos dados da forma como estes podem surgir nas leituras de um processo real.

Grande parte dos avanços nas áreas da reconciliação de dados, detecção de erros grosseiros, diagnóstico de falhas e filtros encontrados na literatura se baseia em simulações, como por exemplo em Vachhani *et al.* (2001). O que leva à conclusão de que, antes de ser uma parte acessória no desenvolvimento das técnicas citadas, o projeto de um simulador adequado e re-utilizável para o maior número possível de situações é fundamental na análise do desempenho destes procedimentos. Narasimhan e Jordache (2000) listam algumas informações necessárias para simulações voltadas à reconciliação de dados.

i. O *flowsheet* do processo indicando o número de unidades do processo, as correntes e sua conectividade. O tipo da unidade de processo não precisa ser especificado quando se faz somente a reconciliação de fluxos globais.

ii. Os valores "verdadeiros" ou "nominais" das variáveis para todos os fluxos das correntes. Estes valores devem ser consistentes com o balanço mássico e são úteis para o julgamento da variação de exatidão alcançada através de reconciliação de dados.

iii. Os conjuntos dos fluxos medidos do processo e os desvios padrão do erro em cada medida. O desvio padrão pode ser expresso como uma fração dos valores verdadeiros ou especificado como um valor absoluto.

É importante também considerar a melhor forma de aproveitar o código fonte legado que esteja disponível em um grupo de pesquisa, evitando ao máximo o retrabalho de código já estável e confiável.

Partindo dessas premissas, foi criado um aplicativo auxiliar chamado **Servidor de Dados OPC** para fornecer dados com ruído ajustável em uma freqüência controlada. A Figura 8.15 mostra a janela principal do aplicativo.

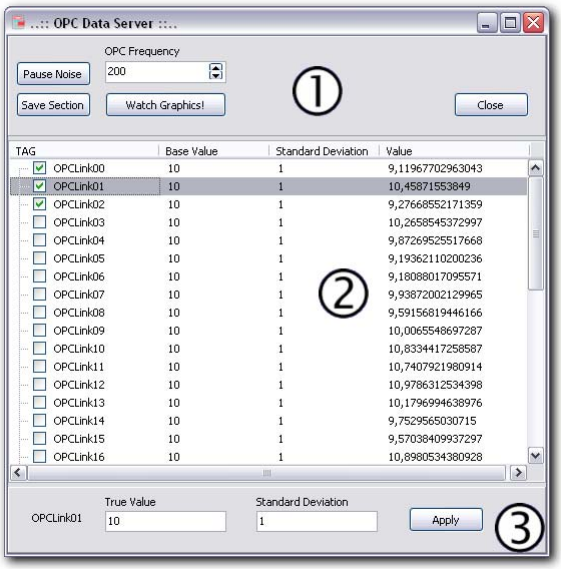


Figura 8.15: Servidor de Dados OPC

Este aplicativo é um servidor de dados do tipo *Data Access* (ver nota de rodapé à página 214) e foi baseado no componente **prOPCKit** criado e distribuído pela **PREL**[15]. A versão usada neste trabalho é livre para uso não comercial, suporta até 40 pontos e cada ponto pode ser definido como somente leitura ou de leitura e escrita.

A adição de ruído ao sinal base, ou verdadeiro, se dá por meio de chamada a uma subrotina em FORTRAN 90/95 (compilada como DLL no Compaq Fortran 6.6C), usando a biblioteca IMLS. A Listagem 8.2 mostra o código desta subrotina.

Listagem 8.2: NoiseGen.F90

```
SUBROUTINE NoiseGen(ExChange, std)

!DEC$ ATTRIBUTES DLLEXPORT, ALIAS:'NOISEGEN' :: NoiseGen

USE box_muller

IMPLICIT NONE

REAL*8  :: ExChange(40), std(40), RND(40)
INTEGER :: I

CALL random_seed

CALL random_number(RND)

DO I=1, 40

      ExChange(I) = ExChange(I) + (1.0d0 - 2.0d0*RND(I))*std(I)

ENDDO
```

[15] http://www.production.robots.btinternet.co.uk/

```
RETURN

ENDSUBROUTINE NoiseGen
```

A janela principal é dividida em três partes. Em (1) está o botão para ativar/pausar a adição de ruídos, o botão para salvar a seção atual, o controle do intervalo de tempo (em milisegundos) no qual são gerados e publicados os dados, o botão de fechar o aplicativo e o botão de chamar as janelas de gráficos (`Show Graphics`). As janelas de gráficos são criadas com os gráficos das variáveis marcadas (*check mark*) na lista em (2). Esta janela pode ser vista na Figura 8.16.

Figura 8.16: Gráficos selecionados no Servidor de Dados OPC

A lista das variáveis publicadas como pontos OPC ((2) na Figura 8.15) mostra a TAG do ponto, seu valor verdadeiro, o desvio padrão em valores absolutos e na última coluna, o valor final que é publicado como sendo o valor verdadeiro acrescido de uma fração aleatória do desvio padrão (Listagem 8.2).

No painel inferior (3) , a linha selecionada na lista de pontos pode ser editada, alterando o valor verdadeiro e o desvio padrão. É neste painel que são preparados os valores para uma determinada simulação.

8.3 Aplicação

É apresentada a seguir uma aplicação do uso conjugado dos aplicativos **Servidor de Dados OPC** e ***Reconciliare***, mostrando a configuração nos dois programas e os resultados das simulações. O exemplo retorna ao problema mostrado no Exemplo 4.4, o processo de flotação de minério, ilustrado na Figura 4.7. Será considerado aqui somente o balanço global de massa. Na Figura 8.17, o processo está representado na tela do aplicativo *Reconciliare*, usando os blocos genéricos (sem associação de código) por se tratar de um problema de reconciliação de dados linear.

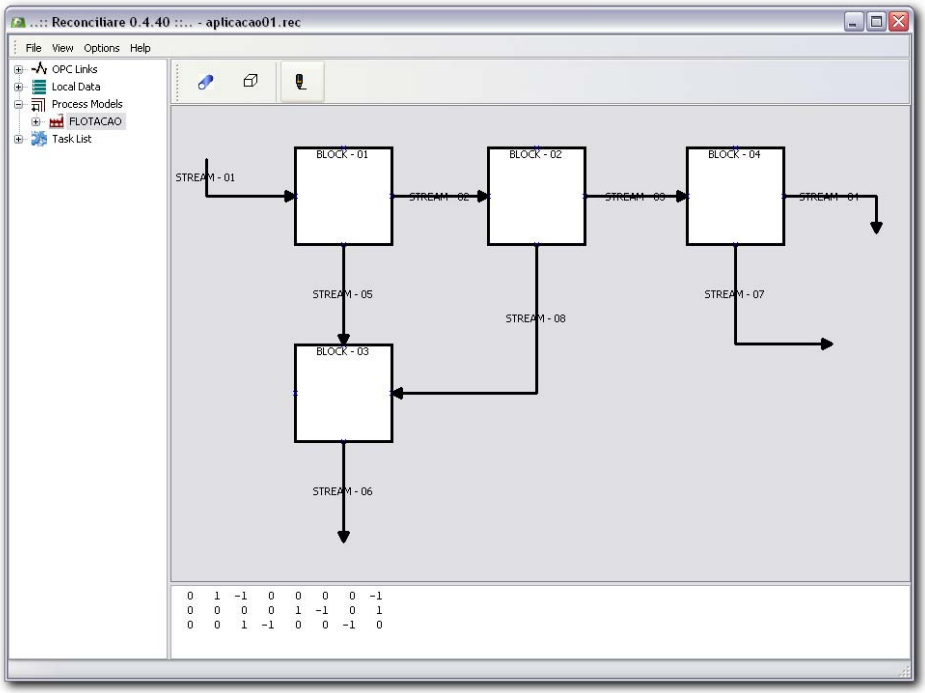


Figura 8.17: Caso 1 modelado no aplicativo ***Reconciliare***

Foram gerados dados aleatórios no aplicativo **Servidor de Dados OPC** mostrado na Seção 8.2. Na Tabela 8.1 estão os valores usados para gerar estes dados (valor verdadeiro base e variância). O valor da corrente 1 foi considerado não medido dentro do aplicativo ***Reconciliare***, e seu valor foi coaptado.

Tabela 8.1: Vazões mássicas - valor verdadeiro e variância - considerados no estudo de caso 1

Corrente	**Valor verdadeiro**	**Variância**
1	1,0	NÃO MEDIDO
2	0,5	0,002512
3	0,25	0,002289
4	0,125	0,002534
5	0,5	0,003289
6	0,75	0,002556
7	0,125	0,004289
8	0,25	0,004289

As Figuras 8.18 a 8.21 mostram os resultados da aplicação da rotina de reconciliação de dados, comparando os valores “verdadeiros”, lidos (gerados com adição de ruído ao valor “verdadeiro”) e os valores reconciliados.

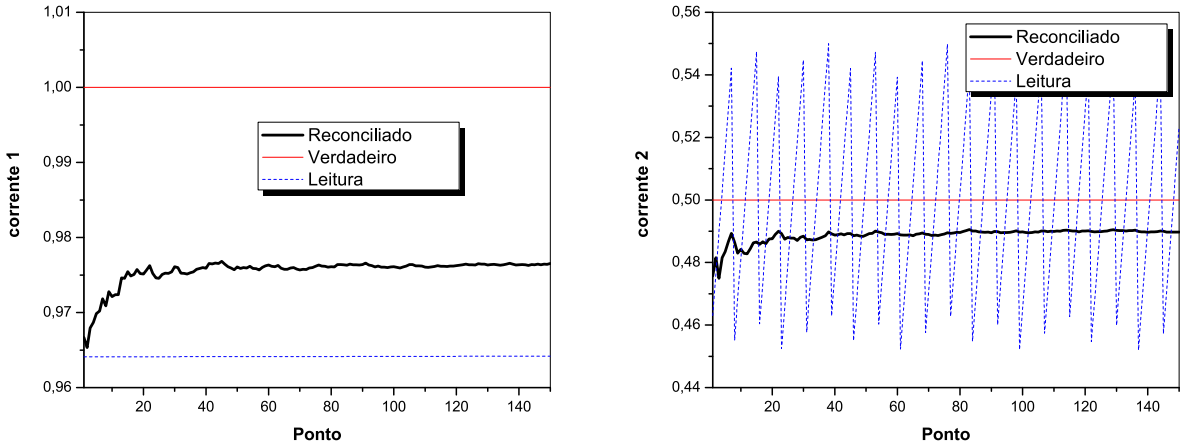


Figura 8.18: Valores verdadeiros, reconciliados e corrompidos – correntes 1 e 2

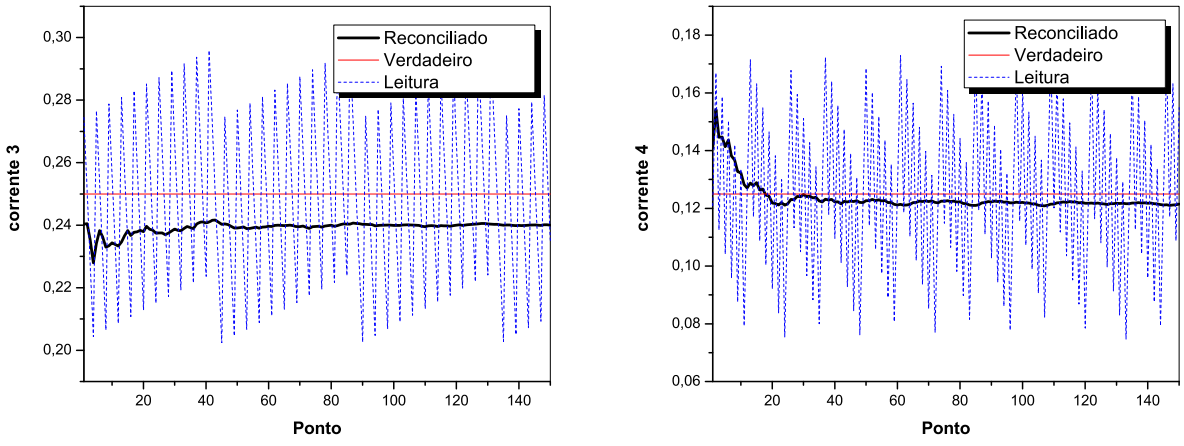


Figura 8.19: Valores verdadeiros, reconciliados e corrompidos – correntes 3 e 4

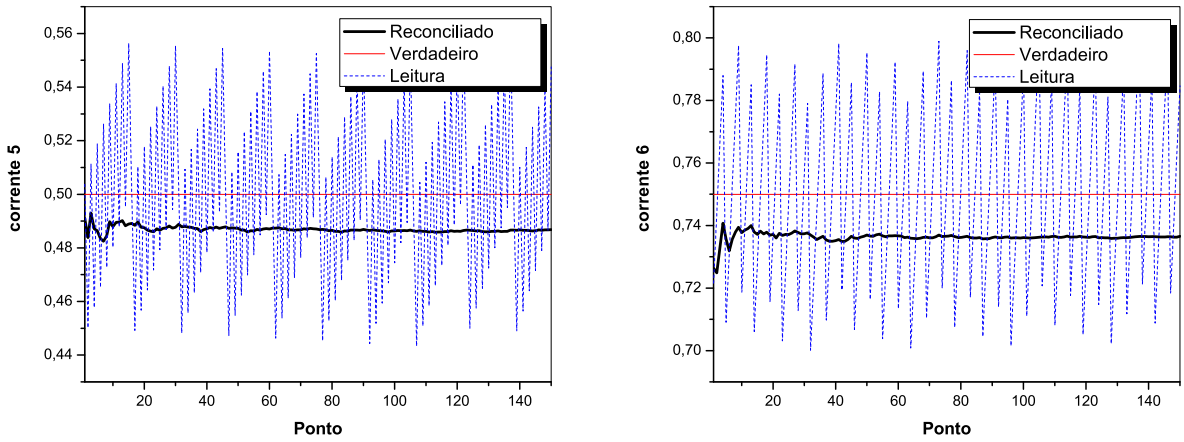


Figura 8.20: Valores verdadeiros, reconciliados e corrompidos – correntes 5 e 6

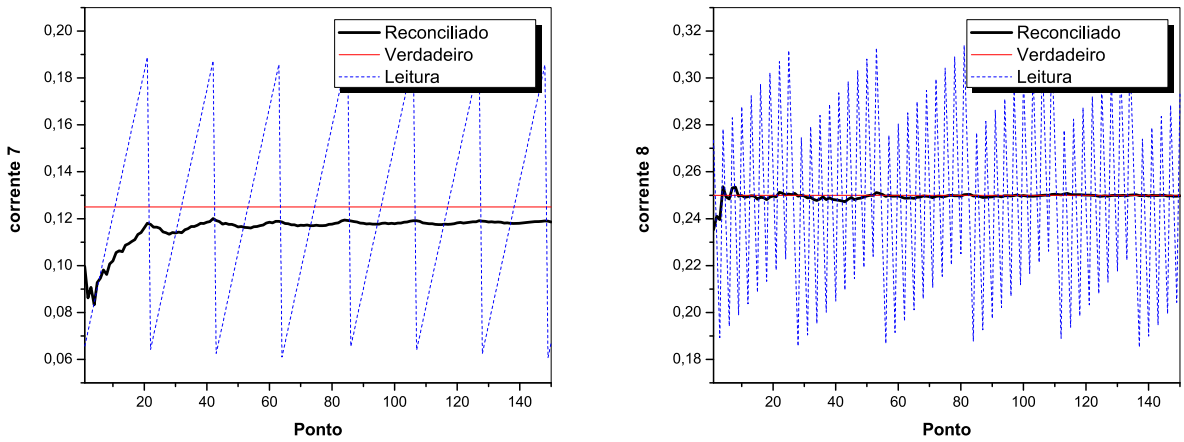


Figura 8.21: Valores verdadeiros, reconciliados e corrompidos – correntes 7 e 8

8.4 Conclusões

Neste capítulo foram apresentados detalhes da concepção e desenvolvimento dos aplicativos que foram construídos neste trabalho e que são o principal objetivo desta tese. Foram mostradas as ferramentas e conceitos subjacentes a cada um destes detalhes e para isso houve a necessidade da introdução de vários aspectos teóricos da engenharia de *software*, como por exemplo os principais conceitos da programação orientada a objetos (classes, objetos, encapsulamento, hereditariedade e polimorfismo).

Uma discussão importante foi feita em relação aos requisitos da modularidade e extensibilidade como um suporte à evolução do trabalho apresentado aqui, no sentido que estes requisitos facilitam a continuidade do trabalho por fragmentarem o problema e permitirem a contribuição bastante localizada de outros pesquisadores que se envolvam com o seu desenvolvimento.

Foram discutidas também as ferramentas de modelagem de *software* onde se contrapuseram as diferenças entre ferramentas de modelagem e documentação como os fluxogramas estruturados, voltados para linguagens de programação estruturadas como o FORTRAN e linguagens de modelagem como a UML, voltadas para o paradigma da programação orientada a objetos. A conclusão é que a capacidade de descrição de uma ferramenta deve ser do mesmo nível de complexidade do problema modelado. A proposta deste trabalho de desenvolver um aplicativo com interface gráfica, comunicação com banco de dados e servidores OPC, com suporte a arquivos e com intensivo gerenciamento de objetos criou a necessidade do uso da UML como ferramenta de modelagem e documentação.

Outra característica importante foi a facilidade de uso. Mesmo sendo um desenvolvimento estritamente acadêmico, houve em cada detalhe a intenção de agregar características de *softwares* comerciais usados na prática industrial. Entre outros motivos, isto se julgou desejável porque a facilidade de uso significa também a condução do usuário no sentido de diminuir eventuais enganos na descrição e execução de suas investigações. O pré-condicionamento das informações referentes à modelagem do problema físico, como por exemplo a matriz de ocorrência, e mesmo a associação com uma fonte de dados de operação/simulação, que são geralmente realizados de forma pouco intuitiva em outros desenvolvimentos acadêmicos, no aplicativo ***Reconciliare*** são rápidos e intuitivos, o quê diminui a possibilidade de se cometer erros na fase de preparação.

Atendendo ao objetivo de criar uma ferramenta com características semelhantes às encontra-

das no ambiente industrial, foi colocado como requisito também a comunicação com sistemas de informação para acesso a dados. Estas conexões externas podem ser via banco de dados ou principalmente via OPC, padrão extensamente discutido neste capítulo. Essa característica faz com que do ponto de vista do aplicativo ***Reconciliare*** seja indiferente operar com dados simulados ou reais.

Neste capítulo foram apresentadas as principais classes do aplicativo ***Reconciliare*** e o relacionamento entre elas. Algumas das funcionalidades foram fornecidas por classes pertencentes a componentes desenvolvidos por terceiros. Estes componentes foram todos indicados, mostrando o nome do desenvolvedor/distribuidor e suas condições de uso. É importante ressaltar que todas as condições de uso foram respeitadas neste trabalho.

Na seqüência, a interface gráfica do aplicativo ***Reconciliare*** foi apresentada, mostrando suas principais características. Foi apresentado também um roteiro de utilização passo a passo.

Outro desenvolvimento importante foi o aplicativo **Servidor de Dados OPC**, desenvolvido para gerar dados via OPC para serem captados e utilizados no aplicativo ***Reconciliare***. Este aplicativo dispõe 40 pontos OPC de leitura e escrita e pode ser usado também para fazer comparações gráficas entre os seus items.

9 *Conclusões e Sugestões para Trabalhos Futuros*

Neste capítulo são discutidas conclusões a respeito dos aspectos teóricos envolvidos na reconciliação e coaptação de dados e na detecção e identificação de erros grosseiros vistos nos capítulos iniciais desta tese. Em seguida são discutidas também conclusões sobre os desenvolvimentos apresentados no capítulo anterior. Finalmente, são feitas sugestões para trabalhos futuros.

9.1 Reconciliação e coaptação de dados

Depois de apresentar os principais elementos da terminologia e alguns resultados relevantes da literatura no Capítulo 2 e conceitos teóricos mais básicos na primeira parte do Capítulo 3, foram introduzidas as técnicas de reconciliação de dados voltada para problemas lineares, que são essencialmente fruto de balanços globais de massa, primeiro para problemas com todas as variáveis medidas e em seguida para problemas com as variáveis parcialmente disponíveis. A introdução de um problema de reconciliação de dados com variáveis parcialmente medidas leva à necessidade da decomposição, primeiro por uma questão de abordagem metodológica de solução, mas também por uma questão de alívio computacional, pois a decomposição do problema de reconciliação promove uma redução de dimensionalidade.

A solução do problema de reconciliação linear com todas as variáveis medidas é encontrada pelo método dos multiplicadores de lagrange (Seção 3.7.1) e envolve a avaliação direta de uma única expressão, na qual o único esforço computacional é relacionado a inversões de matrizes. Por outro lado, a solução do problema parcialmente medido, como foi colocado anteriormente, envolve a decomposição do problema em duas partes: reconciliação e coaptação de dados. A reconciliação vai operar sobre o conjunto redundante de dados e a coaptação vai dividir os dados entre apenas determinado (a variável é medida, mas não há redundância disponível para reduzir a variância),

observável (o sistema completo dá meios para estimar a variável não medida) e não observável (onde não é possível prover estimativa de espécie alguma).

A decomposição do problema geral de estimativa é feita por duas famílias de fatoração: a técnica da projeção de matrizes de Crowe (CROWE *et al.*, 1983; CROWE, 1986, 1989a) e a fatoração QR (SÁNCHEZ; ROMAGNOLI, 1996). Neste trabalho, foi usada esta última abordagem que envolve uma série de operações matriciais para condicionar e resolver o problema.

No Capítulo 4 são apresentadas soluções de reconciliação bilinear como alternativa aos métodos não lineares, pois quando o seu emprego é possível, pode-se efetivamente alcançar resultados com exatidão semelhante a dos métodos não lineares, mas com maior rapidez. Um conjunto de restrições independentes tem que ser imposto para cada unidade de processo na formulação do problema de reconciliação de dados bilinear. Diferentes conjuntos podem ser impostos, sendo que alguns são mais convenientes que outros. É necessário incluir as restrições de normalização sobre as composições para garantir que as estimativas reconciliadas as satisfaçam.

É importante frisar também que os métodos especiais desenvolvidos para resolver os problemas de reconciliação de dados bilinear são eficientes, mas não tratam todos os tipos de unidades nem lidam com restrições de factibilidade como limites sobre as variáveis. Além disso, as técnicas de reconciliação de dados não linear podem ser usadas para resolver problemas bilineares. Estas técnicas são menos eficientes mas não têm as mesmas limitações da reconciliação bilinear.

No Capítulo 5 foi visto que as restrições de um problema de reconciliação de dados não linear podem abarcar restrições de igualdade (balanços materiais, balanços de energia, restrições de equilíbrio e correlações variadas) e de desigualdade (limites nas variáveis e restrições de factibilidade termodinâmica). Os problemas de reconciliação de dados não linear que contenham somente restrições de igualdade podem ser resolvidos usando técnicas iterativas baseadas em linearizações sucessivas e a solução analítica do problema de reconciliação de dados linear. Os problemas de reconciliação de dados não linear contendo restrições de desigualdade somente podem ser resolvidos usando técnicas de otimização não linear sujeita a restrições. Se são impostos limites sobre as variáveis não medidas, então as variáveis não podem ser eliminadas por nenhuma técnica de fatoração ou projeção para obter o problema reduzido e às vezes é necessário que sejam impostos estes limites sobre as variáveis para se obter estimativas factíveis.

Foram vistas duas técnicas de otimização não linear (métodos GRG e SQP) usadas para resolver problemas de reconciliação de dados não linear. Dentre os métodos apresentados, destaca-se

pela abrangência e aplicabilidade o SQP. Contudo, devem ser tomadas algumas precauções, apontadas na Seção 5.3.1, no sentido de adaptar o *solver* SQP às necessidades específicas dos problemas de reconciliação de dados. É desejável que o SQP seja associado a algum outro método, de modo a reduzir o espaço de busca e garantir o ótimo global.

No Capítulo 6 foi vista a importância da reconciliação de dados para aplicações de controle de processos, pois o uso de estados estimados no lugar das medidas pode levar a um controle mais eficiente se houver uma redução sensível de variância. Foi visto também que para explorar a redundância temporal dos dados, foram usados modelos dinâmicos que descrevem o comportamento das variáveis de estado em conjunto com as medições.

A técnica mais abrangente apresentada foi a do filtro de Kalman, que pode ser usado para estimar variáveis de estado em sistemas dinâmicos não lineares. Se perturbações nas variáveis de estado forem ignoradas, então o filtro de Kalman é equivalente à reconciliação de dados. Além disso, a estimativa de estados em sistemas dinâmicos não lineares pode ser realizada usando um filtro de Kalman estendido ou suas variantes, mas esses métodos não tratam restrições de factibilidade sobre as variáveis. Por outro lado, métodos de otimização não linear podem ser usados para a reconciliação dinâmica de dados em processos não lineares e constituem uma importante alternativa aos filtros de Kalman. Estes métodos podem contabilizar restrições de desigualdade, mas são menos eficientes que os filtros de Kalman estendidos.

9.2 Detecção e identificação de erros grosseiros

No Capítulo 7 foram apresentados vários conceitos básicos e resultados fundamentais como a definição dos dois tipos de erros associados com qualquer teste estatístico: O Erro do Tipo I (quando o teste detecta um erro que não existe de fato), o Erro do Tipo II (quando o teste falha na detecção de um erro que de fato está presente) e o fato de que qualquer estratégia de abordagem aos erros grosseiros precisa detectar e também identificar a sua localização.

Foram introduzidos os quatro testes básicos (global, nodal, da medida e o GLR). Dentre eles, somente o teste da medida e o teste GLR podem diretamente identificar a localização de um erro grosseiro (por uma simples regra de identificação). O teste GLR é o único teste que pode identificar tanto viéses nas medidas quanto vazamentos pelo mesmo tipo de teste. A estratégia de detecção de erros grosseiros pelo teste GLR envolve também a estimativa de suas magnitudes.

Os testes baseados em componentes principais não podem identificar diretamente a localização do erro grosseiro pois requerem uma análise adicional para encontrar a restrição ou medida que contribui majoritariamente com o componente principal que falhou no teste.

Já em relação às técnicas de identificação, foi visto que a redução na estatística do teste global depois da eliminação de uma medida é igual à estatística do teste GLR. A eliminação serial pode ser usada para identificar os erros grosseiros detectados pelo teste global e abordagem combinatória MT-NT se mostra excelente do ponto de vista computacional, agregando vantagens e cancelando fraquezas dos testes MT e NT.

A detectabilidade de um erro grosseiro depende principalmente de sua magnitude e localização. Alguns erros grosseiros podem ser detectados, mas nem sempre identificado apropriadamente.

9.3 Desenvolvimento dos *softwares*

Foram mostradas as ferramentas e conceitos subjacentes ao processo de modelagem e construção dos aplicativos desenvolvidos neste trabalho e para isso houve a necessidade da introdução de vários aspectos teóricos da engenharia de *software*, a exemplo dos principais conceitos da programação orientada a objetos vistos no Capítulo 8.

O desenvolvimento do aplicativo ***Reconciliare***, objetivo principal deste trabalho, foi guiado por uma série de requisitos. Dentre eles, foram discutidos os requisitos da modularidade e extensibilidade como um suporte à evolução do trabalho, no sentido que estes requisitos facilitam a sua continuidade por fragmentarem o problema e permitirem a contribuição bastante localizada de outros pesquisadores que se envolvam com o seu desenvolvimento.

Outro requisito importante foi a facilidade de uso. O desenvolvimento dos aplicativos apresentados nesta tese foi pautado por características de *softwares* comerciais usados na prática industrial. Considerou-se que a facilidade de uso significa que o usuário é conduzido no sentido de diminuir eventuais enganos na descrição e execução de suas investigações. O pré-condicionamento das informações referentes à modelagem do problema físico, como por exemplo a matriz de ocorrência e a associação com uma fonte de dados, que são geralmente realizados de forma pouco intuitiva em outros desenvolvimentos acadêmicos, no aplicativo ***Reconciliare*** são rápidos e intuitivos, o quê diminui a possibilidade de se cometer erros na fase de preparação.

Um outro requisito foi a comunicação com sistemas de informação para captação de dados.

Estas conexões externas são feitas principalmente via OPC. Essa característica faz com que do ponto de vista do aplicativo ***Reconciliare*** seja indiferente operar com dados simulados ou reais.

Outro desenvolvimento importante relatado no Capítulo 8, foi o aplicativo **Servidor de Dados OPC**, desenvolvido para gerar dados via OPC para serem captados e utilizados no aplicativo ***Reconciliare***. Este aplicativo dispõe 40 pontos OPC de leitura e escrita e pode ser usado também para fazer comparações gráficas entre os seus items.

Por tudo quanto foi mostrado, conclui-se que os desenvolvimentos relatados constituem importantes ferramentas na investigação de subrotinas de reconciliação e coaptação de dados e de detecção e identificação de erros grosseiros. O objetivo central da tese, o de criar um aplicativo genérico, voltado para monitoramento e análise da qualidade da informação em plantas de processos químicos e petroquímicos, foi alcançado com a criação de uma ferramenta que pode ser usada para agregar inclusive trabalhos futuros.

9.4 Sugestões para trabalhos futuros

No decurso da pesquisa e do desenvolvimento realizados neste trabalho, uma série de novas questões e oportunidades de aprofundamento foram surgindo. Por uma questão de escopo, algumas destas questões e oportunidades não foram abordadas, mas são elencadas a seguir, sendo algumas delas discutidas como sugestões para trabalhos futuros.

- Aprimorar as atuais e adicionar novas subrotinas de tratamento de dados;
- Estudar o impacto de distribuições não normais de erros;
- Reformar o gerenciamento da lista de tarefas;
- Aprimorar a interface gráfica e criar uma versão para web;
- Avaliar técnicas de otimização;
- Aprimorar as ferramentas de simulação usadas em conjunto com o aplicativo ***Reconciliare***;

Durante este trabalho foram levantadas e testadas uma série de técnicas de reconciliação de dados e detecção de erros grosseiros, mas nem todas puderam ser integradas no aplicativo ***Reconciliare***. Sugere-se prosseguir com a pesquisa destas técnicas, principalmente aquelas devotadas à

reconciliação de dados de sistemas dinâmicos, com sua implementação e integração no aplicativo tomando partido justamente de seu modelo de desenvolvimento que facilita a inclusão de novas análises. Outra sugestão é investigar a distorção provocada por modelos de medidas não-lineares sobre distribuições normais de erros.

O gerenciamento das tarefas desenvolvido no aplicativo ***Reconciliare*** é executado como um aplicativo normal. Sugere-se transformar esse módulo em um serviço do sistema operacional (**MS Windows Service**), conferindo assim maior estabilidade, robustez e rastreabilidade. Esta separação do aplicativo em duas partes (o gerenciamento das análises em um serviço do Windows e o cadastro e modelagem em uma aplicação normal) seria mais facilmente "embarcável" em outras soluções existentes. Para promover isto, seria feita uma revisão de toda a modelagem do aplicativo e de seu código fonte para que o cerne do programa, que é o gerenciamento das tarefas, seja colocado de forma independente do resto do aplicativo.

Sugere-se aperfeiçoamentos na interface que facilitem o uso do aplicativo, como, por exemplo, opcionalmente automatizar a criação de variáveis locais espelhando os items OPC conectados e implementar verificações de integridade dos modelos, fazendo com que o modelo só esteja disponível se passar por essas verificações. Sugere-se também iniciar uma versão do aplicativo para ser executado via inter/intranet por meio de um *browser*. Esse tipo de abordagem vem se tornando cada vez mais comum no meio industrial e apresenta uma série de vantagens do ponto de vista dos usuários.

É importante também que sejam avaliadas técnicas de otimização a serem usadas principalmente nas subrotinas de reconciliação de dados.

Referências

ABADIE, J. Design and implementation of optimization software. In: ____. Holland: Sijthoff and Noordhoff, 1978. cap. The GRG Method for Nonlinear Programming.

AGUIRRE, L. A. *Introdução à Identificação de Sistemas*. [S.l.]: Editora UFMG, 2000.

ALBUQUERQUE, J. S.; BIEGLER, L. T. Data reconciliation and gross error detection for dynamic systems. *AIChE Journal*, v. 42, n. 10, p. 2841–56, 1996.

ALMASY, G. A.; SZTANO, T. Checking and correction of measurements on the basis of linear system model. *Problems of Control and Information Theory*, n. 4, p. 57–69, 1975.

ANDERSON, B. D. O.; MOORE, J. B. *Optimal Control: Linear Quadratic Methods*. Englewood Cliffs, NJ: Prentice Hall, 1989.

AO, I. J. G. C. *Implantação de Sistema de Reconciliação de Dados em uma Refinaria de Petróleo*. Dissertação de Mestrado — Escola de Administração/UFBA, 2005.

BADELL, M.; ESPUNA, A.; PUIGJANER, L. Using erp systems with budgeting optimization tools for investiment decision making. In: *Annual AIChE Meeting*. Miami: [s.n.], 1998b. p. 239b.

BADELL, M.; PUIGJANER, L. Short-term planning from the business level in erp systems with vertical integration. In: *Annual AIChE Meeting*. Miami: [s.n.], 1998a. p. 240j.

BAGAJEWICZ, M. J. On the probability distribution and reconciliation of process plant data. *Computers Chem. Engng.*, v. 20, n. 6/7, p. 813–19, 1996.

BAGAJEWICZ, M. J. *Process Plant Instrumentation: Design and Upgrade*. Lancaster, Pennsylvania: Technomic Publishing Co. Inc., 2001.

BAGAJEWICZ, M. J.; JIANG, Q. An integral approach to dynamic data reconciliation. *AIChE Journal*, v. 43, p. 2546–58, 1997.

BAGAJEWICZ, M. J.; JIANG, Q. Gross error modeling and detection in plant linear dynamic reconciliation. *Computers Chem. Engng.*, v. 22, n. 12, 1998.

BAGAJEWICZ, M. J.; JIANG, Q. Comparison of steady state and integral dynamic data reconciliation. *Computers Chem. Engng.*, v. 24, p. 2367–83, 2000.

BAGCHI, A. *Optimal Control of Stochastic Systems*. Hertfordshire: Pretice-Hall, 1993.

BARBOSA, A. G. *Desenvolvimento de um Software para Reconciliação de Dados de Processos Químicos e Petroquímicos*. Dissertação de Mestrado — Faculdade de Engenharia Química/UNICAMP, 2003.

BARBOSA JÚNIOR, V. P. *Reconciliação de Dados por Programação Quadrática de Reatores de Polimerização*. Tese de Doutorado — FEQ/UNICAMP, 1996.

BENQLILOU, C. *Data Reconciliation as a Framework for Chemical Processes Optimization and Control*. Thesis — Department of Chemical Engineering – Universitat Politècnica de Catalunya, 2004.

BENSON, R. S. Computer integrated management: An industrial perspective on the future. *Comp. Chem. Eng.*, v. 19 Suppl., p. S543–51, 1995.

BEVAN, N.; MACLEOD, M. Usability measurement in context. *Behaviour and Information Technology*, v. 13, n. 1, p. 132–145, 1994.

BIEGLER, L. T.; DAMIANO, J. J.; BLAU, G. E. Non-linear parameter estimation: A case study comparison. *AIChE Journal*, n. 32, p. 29–43, 1986.

BOOCH, G.; RUMBAUGH, J.; JACOBSON, I. *UML - Guia do Usuário*. Rio de Janeiro: Editora Campus, 2000.

BORATTI, I. C. *Programação Orientada a Objetos usando Delphi*. 2a. ed. [S.l.]: Visual Books, 2002.

BORRIE, J. A. *Stochastic Systems for Engineers - Modeling Estimation and Control*. Hertfordshire: Pretice-Hall, 1992.

BRITT, H. I.; LUECKE, R. H. The estimation of parameters in nonlinear implicit models. *Technometrics*, v. 2, 1973.

BROWN, J.; GLAZIER, E. V. D. *Signal Analysis*. [S.l.]: Reinhold, 1964.

BROYDEN, C. G. A class of methods for solving nonlinear simultaneous equations. *Math Comp*, p. 577, 1965.

BRYDGES, J. A.; HRYMAK, A.; MARLIN, T. Real time optimization of a fcc recovery section. In: *FOCAPO*. [S.l.: s.n.], 1998.

BUNCH, P. R. Integration of planning and scheduling systems with manufacturing processes. In: *Annual AIChE Meeting*. Miami: [s.n.], 1998. p. 235g.

CHARPENTIER, V.; CHANG, L. J.; SCHWENZER, G. M.; BARDIN, M. C. An on-line data reconciliation system for crude and vacuum units. In: *NPRA Computer Conference*. Huston, TX: [s.n.], 1991.

CHEN, H. S.; STADTHERR, M. A. Enhancements of han-powell method for successive quadratic programming. *Computers Chem. Engng.*, v. 8, n. 3-4, 1984.

CONSUL, C. M. D. *Técnicas Estatísticas Multivariadas para o Monitoramento de Processos Industriais Contínuos*. Tese de Doutorado — Faculdade de Engenharia Química/UNICAMP, 2002.

CROWE, C. M. Reconciliation of process flow rates by matrix projection. part ii: The non-linear case. *AIChE Journal*, n. 32, p. 616–23, 1986.

CROWE, C. M. Recursive identification of gross errors in linear data reconciliation. *AIChE Journal*, v. 34, p. 541–50, 1988.

CROWE, C. M. Observability and redundancy of process data for steady state reconciliation. *Chem. Eng. Sci.*, v. 44, n. 12, 1989a.

CROWE, C. M. Test of maximum power for detection of gross errors in process constraints. *AIChE Journal*, n. 35, p. 869–72, 1989b.

CROWE, C. M. Data reconciliation. progress and challenges. In: *Process Systems Engineering Interational Symposium*. Kyongju, Korea: [s.n.], 1994.

CROWE, C. M. Data reconciliation - progress and challenges. *J. Proc. Cont.*, v. 6, n. 2/3, p. 89–98, 1996.

CROWE, C. M.; CAMPOS, Y. A. G.; HRYMAK, A. Reconciliation of process flow rates by matrix projection. part i: Linear case. *AIChE Journal*, v. 29, n. 6, 1983.

DAHLQUIST, G.; BJORK, A. *Numerical Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

DAROUACH, M.; ZASADZINSKI, M. Data reconciliation in generalized linear dynamic systems. *AIChE Journal*, v. 37, p. 193–201, 1991.

DEMPF, D.; LIST, T. On-line data reconciliation in chemical plants - industrial application of known methods. *Computers Chem. Engng.*, v. 22, p. S1023–25, 1998. Suppl.

DENNIS, J. J. E.; SCHNABEL, R. B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

DUNIA, R.; QIN, J.; EDGAR, T. F.; MCAVOY, T. J. Identification of faulty sensors using principal component analysis. *AIChE Journal*, v. 42, p. 2797–812, 1996.

EDGAR, T. F.; HIMMELBLAU, D. M. *Optimization of Chemical Processes*. New York: McGraw-Hill, 1988.

FISHER, D. G.; SEBORG, D. E. *Multivariable Computer Control - A Case Study*. Amsterdam: North Holland, 1976.

FORBES, J. F.; MARLIN, T. E. Design cost: A systematic approach to technology selection for model-based real-time optimization systems. *Comp. Chem. Eng.*, v. 20, n. 6/7, p. 717–34, 1996.

FRANKLIN, G. F.; POWELL, M. J. D.; WORKMAN, M. L. *Digital Control of Dynamic Systems*. [S.l.]: Addison-Wesley, 1980.

GARCIA, C. *Modelagem e simulação de processos industriais e de sistemas eletromecânicos*. 2a. ed. [S.l.]: Edusp, 2005.

GELB, A. *Applied Optimal Estimation*. Cambridge, MS: MIT Press, 1974.

GONZAGA, J. C. B.; MELEIRO, L. A. C.; KIANG, C.; MACIEL FILHO, R. Ann-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. *Computers and Chemical Engineering*, in press.

GROSDIDIER, P. Understand operation information systems. *Hydrocarbon processing*, v. 77, n. 9, p. 67–78, 1998.

GUEDES, G. T. A. *UML – Uma Abordagem Prática*. [S.l.]: NOVATEC Editora, 2004.

GUPTA, G.; NARASIMHAN, S. Application of neural networks for gross error detection. *Ind. Eng. Chem. Res.*, v. 32, n. 8, p. 1651–7, 1993.

HAN, S. P. A globally convergent method for nonlinear programming. *J. Optimization Theory and Applications*, v. 22, p. 297, 1977a.

HARKINS, B. Turning knowledge into profit. *Chemical Engineering*, p. 92, 1999.

HEENAN, W. A.; SERTH, R. W. Detecting errors in process data. *Chem. Engng*, p. 99–103, 1986.

HIMMELBLAU, D. M. *Process Analysis by Statistical Methods*. [S.l.]: Wiley, 1970.

HIMMELBLAU, D. M. Rectification of data in a dynamic process using artificial neural networks. In: *Proceedings of the Process Systems Engineering International Symposium*. Kyongju, Korea: [s.n.], 1994.

HONG, S. J.; JUNG, J. H.; HAN, C. A design methodology of a soft sensor based on local models. *Comp. and Chem. Eng. Suppl.*, p. S351–4, 1999.

JIANG, Q.; BAGAJEWICZ, M. On a strategy of serial identification with collective compensation for multiple gross error estimation in linear data reconciliation. *Ind. and Eng. Chem. Res.*, v. 38, n. 5, p. 2119–28, 1999.

JOHNSTON, L. P. M.; KRAMER, M. A. Maximum likelihood data rectification. steady state systems. *AIChE Journal*, v. 41, n. 11, p. 2415–26, 1995.

JORDACHE, C. I.; MAH, R. S. H.; TAMHANE, A. C. Performance studies of the measurements test for detection of gross erros in process data. *AIChE Journal*, v. 31, n. 7, p. 1187–201, 1985.

JORDACHE, C. I.; TILTON, B. Gross error detection by serial elimination: Principal component measurement test versus univariate measurement test. In: *AIChE Spring National Meeting*. Huston, TX: [s.n.], 1999.

JORIS, P.; KALITVENTZEFF, B. Process measurements analysis and validation. *Proc. CEF'87: Use Comput. Chem. Eng.*, p. 41–6, 1987.

KALMAN, R. E. Contributions to the theory of optimal control. *Boletin de la Sociedad Matematica Mexicana*, n. 5, p. 102–19, 1960a.

KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, n. 82, p. 35–45, 1960b.

KALMAN, R. E. New results in linear filtering and prediction problems. *Journal of Basic Engineering*, n. 83, p. 95–108, 1961.

KAO, C. S.; TAMHANE, A. C.; MAH, R. S. H. Gross error detection in serially correlated process data. *Ind. Eng. Chem. Res.*, v. 29, n. 6, p. 1004–12, 1990.

KARJALA, T.; HIMMELBLAU, D. Dynamic data rectification by recurrent neural networks vs. traditional methods. *AIChE Journal*, v. 40, p. 1865–75, 1994.

KARJALA, T.; HIMMELBLAU, D. Dynamic rectification of data via recurrent neural nets and the extended kalman filter. *AIChE Journal*, v. 42, p. 22–5, 1996.

KELLER, J. Y. Analytical estimator of measurement error variances in data reconciliation. *Computers and Chem. Engng.*, v. 16, n. 3, p. 185–8, mar. 1992.

KELLER, J. Y.; DAROUACH, M.; KRZAKALA, G. Fault detection of multiple biases or process leaks in linear steady state systems. *Computers and Chem. Engng.*, v. 18, p. 1001, 1994.

KIM, I. W.; KANG, M. S.; PARK, S.; EDGAR, T. F. Robust data reconciliation and gross error detection: The modified mimt using nlp. *Computers and Chem. Engng.*, v. 21, n. 7, p. 775–82, 1997.

KIM, I. W.; LIEBMAN, M. J.; EDGAR, T. F. Robust error in variables estimation using nonlinear programming techniques. *AIChE Journal*, v. 36, p. 985–93, 1990.

KIM, I. W.; LIEBMAN, M. J.; EDGAR, T. F. A sequential error in variables estimation method for nonlinear dynamic systems. *Computers and Chem. Engng.*, v. 15, p. 663–70, 1991.

KNEILE, R. Wring more information out of plant data. *Chem. Engng.*, p. 110–6, mar. 1995.

KNEPPER, J. C.; GORMAN, J. M. Statistical analysis of constrained data sets. *AIChE Journal*, v. 26, n. 2, p. 260–4, mar. 1980.

KRAMER, M. A.; MAH, R. S. H. Model-based monitoring. In: *FOCAPO Proceedings*. Crested Butte: [s.n.], 1993.

KRETSOVALIS, A.; MAH, R. S. H. Observability and redundancy classification in multicomponent process networks. *AIChE Journal*, n. 33, p. 70–82, 1987.

KRETSOVALIS, A.; MAH, R. S. H. Observability and redundancy classification in generalized process networks. i: Theorems. *Comput. Chem. Eng.*, n. 12, p. 671–87, 1988a.

KRETSOVALIS, A.; MAH, R. S. H. Observability and redundancy classification in generalized process networks. ii: Algorithms. *Comput. Chem. Eng.*, n. 12, p. 689–703, 1988b.

LASDON, L. S.; WAREN, A. D. Design and implementation of optimization software. In: ____. Holland: Sijthoff and Noordhoff, 1978. cap. Generalized Reduced Gradient Software for Linearly and Nonlinearly Constrained Problems.

LAUKS, U. E.; VANBINDER, R. J.; VALKENBURG, P. J.; LEEUWEN, C. van. On-line optimization of an ethylene plant. *Comp. and Chem. Eng. (ESCAPE 1) Suppl.*, v. 16, p. S213–20, 1992.

LEE, R. C. (Ed.). *UML e C++ – Guia Prático de Desenvolvimento Orientado a Objetos*. São Paulo: Makron Books, 2001.

LIEBMAN, M. J.; EDGAR, T. F. Data reconciliation for nonlinear processes. In: *AIChE Annual Meeting*. Washington, DC: [s.n.], 1988.

LIEBMAN, M. J.; EDGAR, T. F.; LADSON, L. S. Efficient data reconciliation and estimation for dynamic process using nonlinear programming techniques. *Comput. Chem. Eng.*, v. 16, n. 10/11, p. 963–86, 1992.

LJUNG, L. *System Identification - Theory for the User*. 2nd. ed. [S.l.]: Prentice Hall, 1999.

MACDONALD, R. J.; HOWAT, C. S. Data reconciliation and parameter estimation in plant performance analysis. *AIChE Journal*, v. 34, n. 1, p. 1–8, 1980.

MADRON, F. A new approach to the identification of gross errors in chemical engineering measurements. *Chem. Eng. Sci.*, v. 40, n. 10, p. 1855–60, 1985.

MADRON, F. *Process Plant Performance: Measurement and Data Processing for Optimization and Retrofits*. [S.l.]: Ellis Horwood Ltd., 1992. (Ellis Horwood Series in Chemical Engineering).

MAH, R. S. Design and analysis of process performance monitoring systems. *Engineering Foundation Conf. on Chemical Process Control II*, p. 525–40, 1981.

MAH, R. S. H. *Chemical Process Structures and Information Flows*. Stoneham: Butterworths, 1990.

MAH, R. S. H.; STANLEY, G.; DOWNING, D. Reconciliation and rectification of process flow and inventory data. *Ind. Eng. Chem. Process Des. Dev.*, v. 15, 1976.

MAH, R. S. H.; TAMHANE, A. C. Detection of gross erros in process data. *AIChE Journal*, v. 28, n. 5, p. 828–30, 1982.

MAH, R. S. H.; TAMHANE, A. C. Generalized likelihood ratio method for gross error identification. *AIChE Journal*, v. 33, n. 9, p. 1514–21, 1987.

MARQUES, J. A. *Reconciliação de dados na identificação e caracterização de balanços hídricos em plantas industriais*. Dissertação de Mestrado — COPPE/UFRJ, 2006.

MARTIN, G. Consider soft sensors. *Chem. Eng. Prog.*, p. 66–70, July 1997.

MEDEIROS, E. *Desenvolvendo software com UML 2.0 - definitivo*. São Paulo: Pearson - Makron-Books, 2004.

MELEIRO, L. A. C.; MACIEL FILHO, R. A self-tuning adaptive control applied to an industrial large scale ethanol production. *Computers and Chemical Engineering*, v. 24, p. 925–30, 2000.

MENDES, T. F. *Reconciliação e Retificação de Dados e Classificação de Variáveis de Processo*. Tese de Doutorado — Faculdade de Engenharia Química/UNICAMP, 1995.

MEYER, M.; ENJALBERT, M.; KOEHRET, B. Computer applications in chemical engineering. In: ____. Amsterdam: Elsevier, 1990. cap. Data Reconciliation on Multicomponent networks Using Obserbility and Redundancy Classification.

MEYER, M.; KOEHRET, B.; ENJALBERT, M. Data reconciliation on multicomponent network process. *Comput. Chem. Eng.*, v. 17, n. 8, p. 807–17, 1993.

MILES, J.; JELFFS, P. A. M. Computer aided loss investigation and monitoring. In: JELFFS, P. A. M. (Ed.). *The Second Oil Loss Conference*. [S.l.]: John Wiley and Sons, Ltd., 1988.

MILLER, R. W. *Flow Measurement Engineering Handbook*. [S.l.]: McGraw Hill, 1996.

MULLICK, S. Rigorous on-line model (rom) for crude oil planning, scheduling, engineering and optimization. In: *AIChE Spring Meeting*. [S.l.: s.n.], 1993.

MURTAGH, B. A.; SAUNDERS, M. A. *MINOS 5.0 User's Guide*. Report sol. California, 1983.

MUSKE, K. R.; EDGAR, T. F. Nonlinear process control. In: ____. New Jersey: Prentice-Hall, 1997. cap. Nonlinear State Estimation, p. 311–70.

NAIR, P.; JORDACHE, C. On-line reconciliation of steady-state process plants applying rigorous model-based reconciliation. In: *AIChE Spring National Meeting*. Orlando, FL: [s.n.], 1990.

NAIR, P.; JORDACHE, C. Rigorous data reconciliation is key to optimal operations. *Control for the Process Industries*, IV, n. 10, p. 118–23, 1991. Chicago - Putnam Publ.

NARASIMHAN, S. Maximum power tests for gross error detection using likelihood ratios. *AIChE Journal*, v. 36, n. 7, p. 1589–91, 1990.

NARASIMHAN, S.; JORDACHE, C. *Data Reconciliation and Gross Error Detection - an intelligent use of process data*. [S.l.]: Gulf Publishing Company, 2000.

NARASIMHAN, S.; MAH, R. S. H. Generalized likelihood ratio method for gross error identification. *AIChE Journal*, v. 33, n. 8, p. 1514–21, 1987.

NARASIMHAN, S.; MAH, R. S. H. Generalized likelihood ratios for gross error identification in dynamic processes. *AIChE Journal*, v. 34, n. 8, 1988.

NARASIMHAN, S.; MAH, R. S. H. Treatment of general steady state process models in gross error detection. *Comput. Chem. Engng.*, v. 13, n. 7, p. 851–3, 1989.

NATORI, Y.; TJOA, I. B. To innovate chemical plant operation by applying advanced technology and management. In: *Proceedings of FOCAPO*. [S.l.: s.n.], 1998.

OGATA, K. (Ed.). *Engenharia de Controle Moderno*. 4a. ed. São Paulo: Pearson/Prentice Hall, 2003.

OLIVEIRA JÚNIOR, A. M. d. *Estimação de parâmetros em modelos de processo usando dados industriais e técnicas de reconciliação de dados*. Tese de Doutorado — COPPE/UFRJ, 2006.

ÖZYURT, D. B.; PIKE, R. W. Theory and practice of simultaneous data reconciliation and gross error detection for chemical process. *Comput. Chem. Engng.*, v. 28, p. 381–402, 2004.

PAI, C. C. D.; FISHER, G. Application of broyden's method to reconciliation of nonlinearly constrained data. *AIChE Journal*, v. 34, n. 5, p. 873–76, 1988.

PELHAM, R.; PHARRIS, C. Refinery operations and control: A future vision. *Hydrocarbon Processing*, July 1996.

PERRY, R. H.; GREEN, D. (Ed.). *Perry's Chemical Engineers' Handbook*. 6th. ed. New York: McGraw-Hill, 1984.

PLÁCIDO, J. *Desenvolvimento e Aplicação Industrial da Técnica de Reconciliação de Dados*. Tese de Mestrado — POLI/USP, 1995.

POWELL, M. J. D. A fast algorithm for nonlinearly constrained optimization calculations. In: WATSON, G. A. (Ed.). *Numerical Analysis Conference*. Dundee: [s.n.], 1977.

PRATA, D. M. *Reconciliação de Dados em um Reator de Polimerização*. Tese de Mestrado — COPPE/UFRJ, 2005.

QIN, S. J.; YUE, H.; DUNIA, R. Sef-validating inferential sensors with application to air emission monitoring. *Ind. Eng. Chem. Res.*, v. 36, n. 5, p. 1675–85, 1997.

RAMAMURTHI, Y.; BEQUETTE, B. W. Data reconciliation of systems with unmeasured variables using nonlinear programming techniques. In: *AIChE Spring National Meeting*. Orlando, FL: [s.n.], 1990.

RAMAMURTHI, Y.; SISTU, P. B.; BEQUETTE, B. W. Control relevant dynamic data reconciliation and parameter estimation. *Comput. Chem. Engng.*, v. 17, p. 41–59, 1993.

RAO, C. R. *Linear Statistical Inference and its Applications*. [S.l.]: Wiley, 1973.

RAVIKUMAR, V.; SINGH, S. R.; GARG, M. O.; NARASIMHAN, S. Rage - a software tool for data reconciliation and gross error detection. In: RIPPING, D. W. T.; HALE, J. C.; DAVIS, J. F. (Ed.). *Foundations of Computer-Aided Process Operations*. Amsterdam: CACHE/Elsevier, 1994. p. 429–36.

REDDY, V. N.; MAVROVOUNIOTIS, M. L. An input-training neural network approach for gross error detection and sensor replacement. *Inst. of Chem. Eng. Trans IchemE*, v. 76 Part A, May 1998.

REILLY, P. M.; CARPANI, R. E. Application of statistical theory of adjustments to material balances. In: *13th Canadian Chemical Engineering Conference*. Ottawa, Canada: [s.n.], 1963.

REKLAITIS, G. V. *Introduction to Material and Energy Balances*. [S.l.]: John Wiley & Sons, 1983.

RIPPS, D. L. Adjustment of experimental data. *Chem. Eng. Progress - Symp. Series*, v. 55, n. 61, p. 8–13, 1965.

ROLLINS, D. K.; CHENG, K. Y.; DEVANATHAN, S. Intelligent selection of hypothesis tests to enhance gross error identification. *Comput. Chem. Engng.*, v. 20, p. 517–30, 1996.

ROLLINS, D. K.; DAVIS, J. F. Unbiased estimation of gross errors in process measurements. *AIChE Journal*, v. 38, n. 4, p. 563–72, 1992.

ROLLINS, D. K.; DEVANATHAN, S. Unbiased estimation in dynamic data reconciliation. *AIChE Journal*, v. 39, p. 1330–4, 1992.

ROMAGNOLI, J. A. On data reconciliation: Constraints processing and treatment of bias. *Chem. Eng. Sci.*, v. 38, n. 7, p. 1107–17, 1983.

ROMAGNOLI, J. A.; SÁNCHEZ, M. C. *Data Processing and Reconciliation for Chemical Process Operations*. [S.l.]: Academic Press, 2000.

ROMAGNOLI, J. A.; STEPHANOPOULOS, G. On the rectification of measurements errors for complex chemical plants. *Chem. Eng. Sci.*, v. 35, n. 5, p. 1067–81, 1980a.

ROMAGNOLI, J. A.; STEPHANOPOULOS, G. A general approach to classify operational parameters and rectify measurement errors for complex chemical processes. *Comp. Appl. to Chem. Engng.*, p. 153–74, 1980b.

ROMAGNOLI, J. A.; STEPHANOPOULOS, G. Rectification of process measurements data in the presence of gross errors. *Chem. Eng. Sci.*, v. 36, n. 11, p. 1849–63, 1981.

ROSENBERG, J.; MAH, R. S. H.; JORDACHE, C. Evaluation of schemes for detecting and identification of gross errors in process data. *Ind. & Eng. Chem. Proc. Des. Dev.*, n. 26, p. 555–64, 1987.

SAGE, A. P.; MELSA, J. L. *Estimation Theory with Applications to Communications and Control*. New York: McGraw-Hill, 1971.

SÁNCHEZ, M. C.; BANDONI, A.; ROMAGNOLI, J. A. Pladat - a package for process variable classification and plant data reconciliation. *Comput. Chem. Engn.*, n. S16, p. S499–506, 1992.

SÁNCHEZ, M. C.; ROMAGNOLI, J. Monitoreo de processos continuos: Análisis comparativo de técnicas de identificación y cálculo de bias en los sensores. In: *AADECA 94 - XIV Simposio Nacional de Control Automático*. Argentina: [s.n.], 1994.

SÁNCHEZ, M. C.; ROMAGNOLI, J. A. Use of orthogonal transformations in data classification-reconciliation. *Comput. Chem. Engng.*, v. 20, n. 5, 1996.

SÁNCHEZ, M. C.; ROMAGNOLI, J. A.; JIANG, Q.; BAGAJEWICZ, M. Simultaneous estimation of biases and leaks in process plants. *Comput. Chem. Engng.*, v. 23, n. 7, 1999.

SERTH, R. W.; HEENAN, W. A. Gross errors detection and data reconciliation in steam-metering systems. *AIChE Journal*, v. 32, 1986.

SERTH, R. W.; VALERO, C. M.; HEENAN, W. A. Detection of gross errors in nonlinearly constrained data: A case study. *Chem. Eng. Comm.*, v. 51, p. 89–104, 1987.

SMITH, H. W.; ICHIYEN, N. Computer adjustment of metallurgical balances. *CIM Bull*, p. 97–100, 1973.

SMITH, O. Development and benefits of on-line process modeling applications. In: *AIChE Spring Meeting*. [S.l.: s.n.], 1996.

SÖDERSTRÖM, T.; STOICA, P. *System Identification*. Englewood Cliffs: Prentice-Hall, 1989.

SÖDERSTRÖM, T. A.; HIMMELBLAU, D. M.; EDGAR, T. F. A mixed integer optimization approach for simultaneous data reconciliation and identification of measurements bias. *Chem. Eng. Practice*, v. 9, p. 869–76, 2001.

SONNINO, B. *Delphi e Kylix – Dicas para turbinar seus programas*. São Paulo: Makron Books/-Pearson Education, 2003.

SORENSON, H. W. *Kalman Filtering: Theory and Applications*. New York: IEEE Press, 1985.

STANLEY, G.; MAH, R. S. H. Estimation of flows and temperatures in process networks. *AIChE Journal*, v. 23, p. 642–50, 1977.

STANLEY, G.; MAH, R. S. H. Observability and redundancy in process data estimation. *Chem. Eng. Sci.*, v. 36, p. 259–83, 1981a.

STANLEY, G.; MAH, R. S. H. Observability and redundancy classification in process networks. *Chem. Eng. Sci.*, v. 36, p. 1941–54, 1981b.

STEPHENSON, G. R.; SHEWCHUCK, C. F. Reconciliation of process data with process simulation. *AIChE Journal*, v. 32, n. 2, p. 247–54, 1986.

SWANSON, I.; STEWART, R. Towards full plant-wide management and control. In: *PetrochemAsia 94 International Conference*. Singapore: [s.n.], 1994.

SWARTZ, C. L. Data reconciliation for generalized flowsheet applications. In: *197th American Chem. Society National Meeting*. Dallas, TX: [s.n.], 1989.

TAMHANE, A. C. A note on use of residuals for detecting an outlier in linear regression. *Biometrika*, v. 69, n. 2, p. 448–9, 1982.

TAY, M. E. Keeping tabs on plant energy and mass flows. *Chemical Engineering*, Sep. 1996.

TEIXEIRA, A. C. *Reconciliação de Dados de Processos e Detecção de Erros Grosseiros em Sistemas com Restrições Não-Lineares*. Dissertação de Mestrado — Faculdade de Engenharia Química/UNICAMP, 1997.

THAM, M. T.; PARR, A. Succeed at on-line validation and reconstruction of data. *Chemical engineering progress*, v. 90, n. 5, p. 46–56, 1994.

TJOA, H.; BIEGLER, L. T. Simultaneous strategies for data reconciliation and grosse error detection of nonlinear systems. *Comput. Chem. Engng.*, v. 15, p. 679–90, 1991a.

TJOA, H.; BIEGLER, L. T. Simultaneous solution and optimization strategies for parameter estimation of differential-algebraic equation systems. *Ind. & Eng. Chem. Research*, v. 30, n. 2, p. 376–85, 1991b.

TONG, H.; CROWE, C. M. Detection of gross errors in data reconciliation by principal component analysis. *AIChE Journal*, v. 41, n. 1712-22, 1995.

VACHHANI, P.; RENGASWAMY, R.; VENKATASUBRAMANIAN, V. A framework for integrating diagnostic knowledge with nonlinear optimization for data reconciliation and parameter estimation in dynamic systems. *Chem. Eng. Sci.*, v. 56, p. 2133–48, 2001.

VÁCLAVEK, V. Studies on system engineering. iii. optimal choice of the balance measurements in complicated chemical engineering systems. *Chem. Eng. Sci.*, v. 24, p. 947–55, 1969.

VÁCLAVEK, V.; LOUCKA, M. Selection of measurements necessary to achieve multicomponent mass balances in chemical plants. *Chem. Engn. Sci.*, n. 31, p. 1199–205, 1976.

VASANTHARAJAN, S.; BIEGLER, L. T. Large-scale decomposition for successive quadratic programming. *Computers Chem. Engng.*, v. 12, n. 11, p. 1087–101, 1988.

VEVERKA, V. V.; MADRON, F. *Material and Energy Balances in the Process Industries - From Microscopic Balances to Large Plants*. [S.l.]: Elsevier, 1997. (Computer Aided Chemical Engineering).

YANG, Y.; TEN, R.; JAO, L. A study of gross erros detection and data reconciliation in process industries. *Computers Chem. Engng.*, v. 19, n. suppl., p. S217–22, 1995. European Symposium on Computer Aided Process Engineering.

ZHANG, P.; RONG, G.; WANG, Y. A new method of redundancy analysis in data reconciliation and its application. *Comp. and Chem. Engng.*, v. 25, p. 941–9, 2001.

ZHANG, Z.; PIKE, R. W.; HERTWIG, T. A. An approach to online optimization of chemical plants. *Comp. and Chem. Engng.*, Suppl. 19, p. S305–10, 1995.