DESENVOLVIMENTO DE UMA METODOLOGIA PARA AVALIAÇÃO E MONITORAMENTO DE SISTEMAS DE MEDIÇÃO DE ÁGUAS INDUSTRIAIS, SUJEITOS A BAIXA QUALIDADE DE INFORMAÇÃO

Edson Cordeiro do Valle

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

Orientadores: Ricardo de Araújo Kalid
Argimiro Resende Secchi
Asher Kiperstok

Salvador

Junho de 2013

DESENVOLVIMENTO DE UMA METODOLOGIA PARA AVALIAÇÃO E MONITORAMENTO DE SISTEMAS DE MEDIÇÃO DE ÁGUAS INDUSTRIAIS, SUJEITOS A BAIXA QUALIDADE DE INFORMAÇÃO

Edson Cordeiro do Valle

TESE SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA INDUSTRIAL DA UNIVERSIDADE FEDERAL DA BAHIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM ENGENHARIA INDUSTRIAL.

Aprovada por:

Prof. Ricardo de Araújo Kalid, D.Sc.

Prof. Asher Kiperstok, D.Sc.

Prof. Argimiro Resende Secchi, D.Sc.

Prof. Rafael de Pelegrini Soares, D.Sc.

Prof. Leizer Schnitman, D.Sc.

Pleycienne Trajano Ribeiro, D.Sc.

Prof. Marcelo Escobar Aragão, D.Sc.

SALVADOR, BA – BRASIL

JUNHO DE 2013

Valle, Edson Cordeiro do

Desenvolvimento de uma metodologia para avaliação e monitoramento de sistemas de medição de águas industriais, sujeitos a baixa qualidade de informação/Edson Cordeiro do Valle. – Salvador: UFBA, 2013.

XXXIV, 256 p.: il.; 29,7cm.

Orientadores: Ricardo de Araújo Kalid
Argimiro Resende Secchi
Asher Kiperstok

Tese (doutorado) – UFBA/Programa de Engenharia Industrial, 2013.

Referências Bibliográficas: p. 246 – 256.

1. Águas Industriais. 2. Sistemas de Medição. 3. Monitoramento e Avaliação de Instrumentos de Medição. 4. Detecção de Erros Grosseiros. I. Kalid, Ricardo de Araújo *et al.*. II. Universidade Federal da Bahia, Programa de Engenharia Industrial. III. Título.

Ao meu pai, Antonio

(em memória).

À minha esposa Silvana,

à minha filha, Clarissa

ao meu filho, Guilherme

Agradecimentos

Gostaria de agradecer inicialmente a Deus, por ter me guiado, protegido e colocado as pessoas certas em meu caminho ao longo dessa jornada. Entre essas pessoas, minha mãe, Isamara e meu pai, Antonio, que nos deixou ano passado.
Gostaria de agradecer meus queridos parentes por tudo que fizeram por minha família, e o auxílio que nos ofereceram durante minhas ausências devido ao doutorado e trabalho. Entre os parentes que mais nos apoiaram está minha sogra, 'Nilda', meu 'compadre' Geraldo e minhas 'comadres' Adriana e Lariana, obrigado por toda atenção e carinho que deram aos meus filhos nesse período. Deixo um agradecimento especial ao amor da minha vida, minha esposa Silvana, que deu muito de si para me ajudar na conclusão desse trabalho, cuidando de nossos filhos e dos afazeres domésticos, sem dúvida uma heroína.
Entre os meus amigos, gostaria de agradecer Fernando Majolo e sua esposa Michele, por terem me acolhido com muito carinho em sua moradia durante minhas diversas estadias em Salvador. Das diversas pessoas que conheci em Salvador, duas merecem destaque por sua amizade e companheirismo, Fábio Menezes e Marcos Narciso. Agradeço também a secretária do PEI, Tatiane, pelas ajuda nos procedimentos de matricula, Tiago Finkler pela ajuda com os PLS's. Agradeço também, Victor Aranovich, por ter me ajudado a superar uma fase bastante estressante de minha vida. Meus sinceros agradecimentos aos meus orientadores pelo tempo e dedicação investido, além dos ensinamentos que contribuíram para minha formação e meu amadurecimento: Ricardo de Araújo Kalid por ter me recebido de portas abertas na UFBA e acreditado em meu potencial, por todas as conversas e cafés. À Argimiro Resende Secchi, que, desde a graduação, vem sendo uma pessoa muito especial em minha formação de pesquisador. Por último, aos meus sócios e amigos da VRTech, Rafael e Paula, por terem compreendido a aceitado minhas ausências na empresa durante a elaboração desse trabalho.

Se For Embora

Rafael - Chimarruts

Se for embora, leve um sorriso meu,
Guardado no teu peito,
Teu coração agora é meu
Mas se ficar, ganha um beijo meu
Com sabor de alegria
É um presente que Deus me deu
Sim, já passei por tudo nesta vida
Mas foi a fé que deu forças pra lutar
Batalhas ganhei, outras perdidas
Ai! Que bom, se pudesse ficar
Se a vida veio lhe chamar
Se tem que ir e não pode mais ficar
Se o destino não se encarregar
Da gente se ver, eu preciso te falar...

A Vitória Demora Mas Vem

Juninho Thybau - Baiaco e Luis Caffe

...Guerreiro não foge a luta
Eu tô sempre na disputa
Em prol de um amanhã melhor (bem melhor)
Pego firme na labuta
Malandro é quem escuta bom conselho
Pra não ficar na pior (na pior)
Eu vou compondo minha história
Guardo em minha memória
Quem sempre me fez o bem (fez o bem)
Aprendi pra ensinar o ensinamento
Que tudo na vida tem seu tempo, tem
A vitória demora mas vem

Aquarela

Toquinho - Vinicius de Moraes - M. Fabrizio - G. Morra

...Um menino caminha e caminhando chega no muro
E ali logo em frente, a esperar pela gente, o futuro está.
E o futuro é uma astronave que tentamos pilotar,
Não tem tempo nem piedade, nem tem hora de chegar.
Sem pedir licença muda nossa vida, depois convida a rir ou chorar.
Nessa estrada não nos cabe conhecer ou ver o que virá.
O fim dela ninguém sabe bem ao certo onde vai dar.
Vamos todos numa linda passarela
De uma aquarela que um dia, enfim, descolorirá.

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

DESENVOLVIMENTO DE UMA METODOLOGIA PARA AVALIAÇÃO E MONITORAMENTO DE SISTEMAS DE MEDIÇÃO DE ÁGUAS INDUSTRIAIS, SUJEITOS A BAIXA QUALIDADE DE INFORMAÇÃO

Edson Cordeiro do Valle

Junho/2013

Orientadores: Ricardo de Araújo Kalid
Argimiro Resende Secchi
Asher Kiperstok

Programa: Engenharia Industrial

O setor industrial consome grande quantidade de água, sendo esta um dos itens fundamentais em muitos processos industriais, além disso, é um dos principais responsáveis pela poluição dos corpos hídricos. A solução desses problemas passa pelo gerenciamento adequado dos recursos hídricos industriais e pela aplicação de ferramentas para reúso de águas (integração mássica), porém, nem sempre o uso de tais técnicas são possíveis devido à falta ou baixa qualidade de informação do circuito de uso de água. Nesse contexto, o objetivo deste trabalho é a seleção de índices para acompanhamento do sistema de medição de águas em plantas industriais em estado estacionário, assim como, o desenvolvimento de uma metodologia que, baseado nos índices selecionados, e, atuando em conjunto com modelos empíricos multivariáveis, seja capaz de indicar possíveis falhas dos sensores e detectar vazamentos em equipamentos. Para cumprir tais objetivos, inicialmente foram selecionados 16 casos teste

para uma validação rigorosa do método proposto. Em seguida, foram escolhidos índices para caracterizar a qualidade dos balanços de massa. Os índices de qualidade selecionados foram as estatísticas de testes para detecção de erros grosseiros: Teste Global, Teste Nodal e Teste de Medida. O modelo empírico, que tem como objetivo final determinar a posição e magnitude dos erros grosseiros a partir dos índices de qualidade, foi gerado a partir do método de regressão em espaço reduzido de mínimos quadrados parciais de ordem quadrática (PLS quadrático ou QPLS). Em seguida, procedeu-se com o ajuste dos parâmetros e validação dos modelos para cada problema teste, tanto para falhas em medidores de vazão quanto para vazamentos, em suas formas individuais ou combinadas. Os resultados do método proposto para detecção de erros grosseiros únicos foram comparados com os métodos de Razão de Verosimilhança Generalizada (GLR para erros em medições e vazamento) e Teste Nodal (NT) usando a abordagem clássica de Monte Carlo para avaliação de métodos de detecção de erros grosseiros nos 16 problemas testes. Para avaliação dos resultados do método proposto para erros múltiplos, o método proposto foi aplicado a um exemplo de literatura e comparado com 3 métodos: estimação simultânea de erros grosseiros, SEGE, com o método de estimação não-tendenciosa, UBET, e com o método GLR. Para erros grosseiros únicos, no conjunto de testes onde foi aplicado, o método proposto apresentou resultados superiores em relação ao método GLR para erros grosseiros em medições e apresentou resultados equivalentes para vazamentos (tanto para o GLR quanto para o NT). Para detecção de erros múltiplos, no respectivo conjunto de dados, o método proposto também apresentou resultados superiores para detecção de erros grosseiros múltiplos envolvendo erros em medições e combinação de erros em medições e vazamentos.

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

DEVELOPMENT OF A NEW METHOD FOR EVALUATION AND MONITORING OF INDUSTRIAL WATER SYSTEMS SUBJECT TO LOW QUALITY OF INFORMATION

Edson Cordeiro do Valle

June/2013

Advisors: Ricardo de Araújo Kalid
Argimiro Resende Secchi
Asher Kiperstok

Department: Industrial Engineering

The industry uses high amount of water in the production processes since this raw material is an essential item in several industrial processes. Beyond, they are also responsible for the pollution of water. These problems can handled by the appropriate use of water management tools and by the application of water reclamation tools. However, the use of these tools cannot be applied in any industrial case, due to the lack of information regarding the water material balance. In this context, the goal of this work is to select index to evaluate the monitoring system of water use circuit in process industry in steady-state. Beyond, this work aims to develop methods that, based on the selected index and acting together with multivariate empirical models, can detect the possibility of measurement sensor failures and equipment leaking. To reach these goals, initially, it was selected 16 benchmark test cases for rigorous validation of the method. Then, some index were selected that could represent the quality of the flowsheet mass balance. The selected index

were the statistics of the Global Test, Nodal Test and Measurement test, then, the empirical model selection were carried out. The goal of the empirical model is to find the gross errors position and magnitude based on these quality index which was built using the reduced space regression model named Quadratic Partial Least Squares (Q-PLS) technique. In the next step, the estimation of the PLS models parameters were carried out, both for measurement bias and equipment leaking, either for single and multiple gross errors. The results of the proposed method were compared with the Generalized Likelihood Ratio Test (GLR for measurement error and leaking) and the Nodal Test (NT), for the 16 test problems, using the classical Monte Carlo approach for evaluation of the gross error detection methods. For multiple gross errors, the proposed method was applied to one literature example and compared with 3 methods: unbiased estimation technique, UBET, simultaneous estimation of gross errors, SEGE and GLR method. For single gross errors detection, among the test problems where the proposed method was applied, it presented superior results when comparing with the GLR for measurement error, and equivalent results for leaking detection (with GLR and NT). For multiple gross errors detection where the method was compared, it also presented a superior result for measurement bias and simultaneous bias and leaking identification.

Sumário

Lista de Figuras **xvii**

Lista de Tabelas **xix**

Lista de Abreviaturas **xxviii**

Lista de Símbolos **xxx**

1 Introdução **1**

1.1 Água industrial e seu uso racional 1

1.2 Integração de processos, uma ferramenta para a redução do consumo de água na indústria 3

1.3 Objetivos do trabalho 9

1.4 Estrutura da tese 10

2 Revisão Bibliográfica **11**

2.1 Coleta de informações de balanços hídricos 11

2.2 Métodos para correção de incoerências em balanços de massa 14

2.3 Reconciliação de Dados e Metrologia 19

2.4 Classificação de variáveis 20

2.5 Erros grosseiros e métodos para detecção e identificação 24

2.5.1 Teste global 28

2.5.2 Teste nodal 30

2.5.3 Teste de medida 31

2.5.4 Teste de razão de verossimilhança generalizada 33

2.5.5 Teste dos componente principais para equipamentos, medidas e global 34

2.5.6 Identificação e localização de erros grosseiros múltiplos usando estratégias simultâneas 35

2.5.6.1 Estratégias usando erros grosseiros individuais 35

2.5.6.2 Estratégias usando hipóteses combinatoriais 36

2.5.6.3 Identificação usando estimação simultânea das magnitudes dos erros grosseiros 36

2.5.7 Identificação e localização de erros de grosseiros múltiplos usando estratégias em série 37

2.5.7.1 Teste de medida iterativo 37

2.5.7.2 Estratégia da compensação serial simples 38

2.5.7.3 Estratégias de combinação de nodos 38

2.5.8 Detectabilidade e identificabilidade de erros grosseiros 39

2.5.9 Medidas de performance das técnicas de detecção de erros grosseiros 40

2.5.9.1 Ferramentas de mineração de dados 41

2.5.10 Métricas para detecção de erros de medição 42

2.6 Contexto do trabalho 43

3 Materiais e Métodos **44**

3.1 Visão geral do método proposto 44

3.2 Escolha do método de otimização 47

3.3 Seleção e formulação de estudos de caso 48

3.4 Seleção e cálculos de índices de qualidade do balanço para ajuste do modelo de PLS 51

3.5 Seleção da técnica de ajuste do modelo de DEG 54

3.5.1 Regressão linear múltipla 54

3.5.2 Técnicas de redução de dimensionalidade 55

3.6 Ajuste do modelo de DEG 58

3.6.1 Escolha da função objetivo 58

3.7 Escolha das incertezas 60

3.8 Geração e seleção de dados para ajuste dos modelos PLS 66

3.8.1 Geração e seleção de dados com erros puramente aleatórios . . 66

3.8.2 Geração e seleção de dados com erros grosseiros nos medidores - erros únicos 69

3.8.3 Geração e seleção de dados com erros grosseiros de vazamentos - erros únicos 71

3.8.4 Geração e seleção de dados com erros grosseiros nos medidores e vazamentos - erros múltiplos 74

3.9 Seleção de dados para ajuste dos parâmetros do modelo e testes . . . 75

3.10 Avaliação do modelo de DEG 75

4 Resultados e Discussões **81**

4.1 Obtenção dos resultados 81

4.2 Erros únicos 84

4.2.1 Problema 1 87

4.2.2 Problema 2 89

4.2.3 Problema 3 90

4.2.4 Problema 7 93

4.2.5 Problema 8 96

4.2.6 Problema 9 98

4.2.7 Problema 14 99

4.3 Erros múltiplos 105

4.3.1 Comparação com outros métodos 109

4.4 Considerações finais 113

5 Conclusões e Sugestões para Trabalhos Futuros 117

5.1 Conclusões 117

5.1.1 Principais contribuições 117

5.1.2 Publicações geradas ao longo do trabalho 118

5.2 Sugestões para trabalhos futuros 119

A Descrição dos Problemas em Forma de Artigo 121

B Estimadores Robustos 193

C Influência dos Desvios Padrões no Poder Total do Teste 194

C.1 Problema 1 194

C.2 Problema 8 199

C.3 Problema 9 206

D Resultados Completos 214

D.1 Erros Únicos 214

D.1.1 Problema 1 214

D.1.2 Problema 2 215

D.1.3 Problema 3 216

D.1.4 Problema 3 - Incerteza-Padrão = 5 t/h 217

D.1.5 Problema 4 218

D.1.6 Problema 5 219

D.1.7 Problema 6 220

D.1.8 Problema 7 221

D.1.9 Problema 8 222

D.1.10 Problema 9 223

D.1.11 Problema 10 224

D.1.12 Problema 11 225

D.1.13 Problema 12 226

D.1.14 Problema 13 228

D.1.15 Problema 14 230

D.1.16 Problema 14 - Incerteza- Padrão = 2 t/h 232

D.1.17 Problema 15 234

D.1.18 Problema 16 236

D.2 Erros Múltiplos 239

D.2.1 Problema 1 239

D.2.2 Problema 2 240

D.2.3 Problema 3 240

D.2.4 Problema 4 240

D.2.5 Problema 5 241

D.2.6 Problema 6 241

D.2.7 Problema 7 241

D.2.8 Problema 8 242

D.2.9 Problema 9 242

D.2.10 Problema 10 242

D.2.11 Problema 11 243

D.2.12 Problema 12 243

D.2.13 Problema 13 243

D.2.14 Problema 14 244

D.2.15 Problema 15 244

D.2.16 Problema 16 244

Referências Bibliográficas **246**

Lista de Figuras

1.1 Reúso, regeneração-reciclo, regeneração-reúso e segregação com tratamento distribuído. 4

2.1 Comparação entre estimador de Cauchy (C_C igual a 2,3849, determinado por Özyurt e Pike (2004)) e Mínimos Quadrados Ponderados. . 17

2.2 Comparação entre estimador de Lorenz (C_L igual a 2,6, determinado por Özyurt e Pike (2004)) e Mínimos Quadrados Ponderados. 18

2.3 Exemplo de análise de observabilidade e redundância. 21

2.4 Exemplo de graduação da QI 23

2.5 Distribuição Normal com média μ e desvio padrão σ. 28

2.6 Esquema para DEG usando o Teste Global. 30

2.7 Esquema para DEG usando o Teste Nodal. 31

2.8 Esquema para DEG usando o Teste de Medida. 32

3.1 Esquema simplificado da metodologia proposta visando gerar modelos para predizer os erros grosseiros. 45

3.2 Esquema simplificado da aplicação da metodologia proposta com os modelos gerados. 46

3.3 Exemplo de processo para elaboração da tabela de mapeamento de índices de qualidade. 52

3.4 Gráfico do OP do MT utilizando o estimador robusto de Cauchy para o Problema 1. 60

3.5 Fluxograma do Problema 9. 61

3.6 Algoritmo para avaliação de técnicas de DEG para erros em medições. 76

3.7 Algoritmo para avaliação das técnicas de DEG para vazamentos em equipamentos. 78

4.1 Fluxograma de processo do Problema 1. 88

4.2 Fluxograma de processo do Problema 2. 89

4.3 Fluxograma de processo do Problema 3. 91

4.4 Fluxograma de processo do Problema 7. 93

4.5 Fluxograma de processo do Problema 8. 97

4.6 Fluxograma de processo do Problema 9. 98

4.7 Fluxograma de processo do Problema 14. 100

4.8 Fluxograma de processo do Problema 15. 107

4.9 Fluxograma de processo do Problema 10. 108

Lista de Tabelas

1.1 Maiores geradores de efluentes do Rio Grande do Sul em volume. . . 2

1.2 Ganho com a aplicação de ferramentas de integração mássica. 5

1.3 Redução de consumo de água e descarte de efluentes obtida com técnicas de reúso de águas industriais nas indústrias do Estado de São Paulo. 6

1.4 Redução de consumo de água e descarte de efluentes obtida com técnicas de reúso de águas industriais. Continuação 7

1.5 Redução de consumo de água e descarte de efluentes obtida com técnicas de reúso de águas industriais. Continuação 8

2.1 Técnicas de identificação de erros grosseiros e o número de exemplos usados para a testar as mesmas. 26

2.2 Técnicas de identificação de erros grosseiros e o número de exemplos usados para a testar as mesmas. Continuação 27

3.1 Exemplo de tabela de dados usada para estimar os parâmetros de modelos multivariáveis usado pelo método proposto. 45

3.2 Características dos problemas selecionados. 50

3.3 Tabela para mapeamento de índices de qualidade. 52

3.4 Tabela para mapeamento de índices de qualidade do exemplo apresentado. 53

3.5 Estatística de Teste de Medida para o Problema 1. 59

3.6 Poder Total do MT para o Problema 9 usando o estimador de mínimos quadrados ponderados. 61

3.7 Poder Total do MT para o estimador robusto de Cauchy, usando as incertezas originais. 62

3.8 Poder Total do MT para o estimador robusto de Cauchy, usando incertezas unitárias em todas as correntes. 63

3.9 Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicado por 0,1. 65

3.10 Poder Total do teste de medida utilizando a incerteza-padrão unitária multipicado por 5. 65

3.11 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicado por 0,1. 65

3.12 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicado por 5. 66

3.13 Tabela para mapeamento de índices de qualidade para erros puramente aleatórios para um processo com um equipamento e três correntes. 68

3.14 Tabela para mapeamento de índices de qualidade após a geração de dados com erros aleatórios e erros grosseiros nas medições para processo com um equipamento e três correntes. 71

3.15 Tabela para mapeamento de índices de qualidade após a geração de dados com erros aleatórios e erros grosseiros nas medições e vazamentos para processo com um equipamento e três correntes. 74

4.1 Resultados do AVT1, tempo computacional e R^2. 86

4.2 Exemplo de tabela de resultados para os problemas com erros únicos de medição. 87

4.3 Exemplo de tabela de resultados para vazamentos 87

4.4 Resultados do Problema 1 para erros grosseiros em medições. 88

4.5 Resultados do Problema 1 para vazamentos. 89

4.6 Resultados do Problema 2 para erros em medições. 90

4.7 Resultados do Problema 2 para vazamentos. 90

4.8 Resultados do Problema 3 para erros em medições. 91

4.9 Resultados do Problema 3 para vazamentos. 92

4.10 Resultados do Problema 3 para erros de medição com incerteza-padrão igual a 0,1 para todas as correntes. 92

4.11 Resultados do Problema 3 para vazamentos com incerteza-padrão igual a 0,1. 93

4.12 Resultados do Problema 7 para erros de medição. 94

4.13 Resultados do Problema 7 para vazamentos. 95

4.14 Resultados dos erros grosseiros de medição para o Problema 7 usando incerteza-padrão de 0,1 $t.h^{-1}$ para todas as correntes. 95

4.15 Resultados de vazamentos para o Problema 7 usando incerteza-padrão de 0,1 $t.h^{-1}$ para todas as correntes. 96

4.16 Resultados do Problema 8 para erros de medição. 97

4.17 Resultados do Problema 8 para vazamentos. 98

4.18 Resultados do Problema 9 para erros de medição. 99

4.19 Resultados do Problema 9 para vazamentos. 99

4.20 Conversão de dados de QI para incertezas-padrão para uso da função objetivo de Cauchy. 101

4.21 Resultados do Problema 14 para erros de medição. 102

4.22 Resultados do Problema 14 para erros de medição com incerteza-padrão igual a 5 $t.h^{-1}$. 103

4.23 Resultados do Problema 14 para vazamentos utilizando incertezas-padrão unitários. 104

4.24 Resultados do Problema 14 para vazamentos e incertezas-padrão de 5 t.h^{-1}. 105

4.25 Cenário de distúrbios com erros múltiplos e os resultados dos ajustes dos modelos. 106

4.26 Exemplo de tabela de resultados dos erros grosseiros múltiplos. 106

4.27 Resultados dos erros grosseiros múltiplos do Problema 15. 108

4.28 Resultados dos erros grosseiros múltiplos do Problema 10. 109

4.29 Comparação entre OP de diferentes métodos para DEG múltiplos para o Problema 6. 110

4.30 Comparação entre AEE de diferentes métodos para DEG múltiplos para o Problema 6. 111

4.31 Comparação entre OP de diferentes métodos para DEG múltiplos com vazamento para o problema 6. 112

4.32 Comparação entre AEE de diferentes métodos para DEG múltiplos com vazamento para o problema 6. 113

4.33 Técnicas de identificação de erros grosseiros e o número de exemplos usadas para teste das mesmas. 116

C.1 Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 0,1. 194

C.2 Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 0,1. 194

C.3 Poder Total do teste de medição utilizando a incerteza-padrão original. 195

C.4 Poder Total do teste nodal utilizando a incerteza-padrão original. . . 195

C.5 Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 5. 196

C.6 Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 5. 196

C.7 Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 0,1. 197

C.8 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 0,1. 197

C.9 Poder Total do teste de medição utilizando a incerteza-padrão unitária. 197

C.10 Poder Total do teste nodal utilizando a incerteza-padrão unitária. . . 198

C.11 Poder Total do teste de medida utilizando a incerteza-padrão unitária multiplicada por 5. 198

C.12 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 5. 198

C.13 Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 0,1. 199

C.14 Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 0,1. 199

C.15 Poder Total do teste de medição utilizando a incerteza-padrão original. 200

C.16 Poder Total do teste nodal utilizando a incerteza-padrão original. . . 201

C.17 Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 5. 201

C.18 Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 5. 202

C.20 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 0,1. 202

C.19 Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 0,1. 203

C.21 Poder Total do teste de medição utilizando a incerteza-padrão unitária. 204

C.22 Poder Total do teste nodal utilizando a incerteza-padrão unitária. . . 204

C.23 Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 5. 205

C.24 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 5. 206

C.25 Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 0,1. 206

C.26 Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 0,1. 207

C.27 Poder Total do teste de medição utilizando a incerteza-padrão original. 208

C.28 Poder Total do teste nodal utilizando a incerteza-padrão original. . . 208

C.29 Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 5. 209

C.30 Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 5. 210

C.31 Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 0,1. 210

C.32 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 0,1. 211

C.33 Poder Total do teste de medição utilizando a incerteza-padrão unitária. 212

C.34 Poder Total do teste nodal utilizando a incerteza-padrão unitária. . . 212

C.35 Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 5. 213

C.36 Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 5. 213

D.1 Resultados do Problema 1 para erros em medições. 214

D.2 Resultados do Problema 1 para erros em medições. 214

D.3 Resultados do Problema 2 para erros em medições. 215

D.4 Resultados do Problema 2 para vazamentos. 215

D.5 Resultados do Problema 3 para erros em medições. 216

D.6 Resultados do Problema 3 para vazamentos. 216

D.7 Resultados do Problema 3 para erros de medição com incerteza-padrão padrão igual a 5. 217

D.8 Resultados do Problema 3 para vazamentos com incerteza-padrão padrão igual a 5. 217

D.9 Resultados do Problema 4 para erros de medição. 218

D.10 Resultados do Problema 4 para vazamentos. 218

D.11 Resultados do Problema 5 para erros de medição. 219

D.12 Resultados do Problema 5 para vazamentos. 219

D.13 Resultados do Problema 6 para erros de medição. 220

D.14 Resultados do Problema 6 para vazamentos. 220

D.15 Resultados do Problema 7 para erros de medição. 221

D.16 Resultados do Problema 7 para vazamentos. 221

D.17 Resultados do Problema 8 para erros de medição. 222

D.18 Resultados do Problema 8 para vazamentos. 222

D.19 Resultados do Problema 9 para erros de medição. 223

D.20 Resultados do Problema 9 para vazamentos. 223

D.21 Resultados do Problema 10 para erros de medição. 224

D.22 Resultados do Problema 10 para vazamentos. 225

D.23 Resultados do Problema 11 para erros de medição. 225

D.24 Resultados do Problema 11 para vazamentos. 226

D.25 Resultados do Problema 12 para erros de medição. 226

D.26 Resultados do Problema 12 para vazamentos. 227

D.27 Resultados do Problema 13 para erros de medição. 228

D.28 Resultados do Problema 13 para vazamentos. 229

D.29 Resultados do Problema 14 para erros de medição. 230

D.30 Resultados do Problema 14 para vazamentos. 231

D.31 Resultados do Problema 14 para erros de medição com incerteza-padrão padrão igual a 5 t/h. 232

D.32 Resultados do Problema 14 para vazamentos e incerteza-padrão padrão de 5 t/h. 233

D.33 Resultados do Problema 15 para erros de medição. 234

D.34 Resultados do Problema 15 para erros de medição. Continuação . . . 235

D.35 Resultados do Problema 15 para vazamentos. 235

D.36 Resultados de erros grosseiros de medição para o Problema 16. 236

D.37 Resultados de erros grosseiros de medição para o Problema 16. Continuação 237

D.38 Resultados de vazamentos para o Problema 16. 238

D.39 Cenário de distúrbios com erros múltiplos e os resultados dos ajustes dos modelos. 239

D.40 Resultados dos erros grosseiros múltiplos do Problema 1. 239

D.41 Resultados dos erros grosseiros múltiplos do Problema 2. 240

D.42 Resultados dos erros grosseiros múltiplos do Problema 3. 240

D.43 Resultados dos erros grosseiros múltiplos do Problema 4. 240

D.44 Resultados dos erros grosseiros múltiplos do Problema 5. 241

D.45 Resultados dos erros grosseiros múltiplos do Problema 6. 241

D.46 Resultados dos erros grosseiros múltiplos do Problema 7. 241

D.47 Resultados dos erros grosseiros múltiplos do Problema 8. 242

D.48 Resultados dos erros grosseiros múltiplos do Problema 9. 242

D.49 Resultados dos erros grosseiros múltiplos do Problema 10. 242

D.50 Resultados dos erros grosseiros múltiplos do Problema 11. 243

D.51 Resultados dos erros grosseiros múltiplos do Problema 12. 243

D.52 Resultados dos erros grosseiros múltiplos do Problema 13. 243

D.53 Resultados dos erros grosseiros múltiplos do Problema 14. 244

D.54 Resultados dos erros grosseiros múltiplos do Problema 15. 244

D.55 Resultados dos erros grosseiros múltiplos do Problema 16. 245

Lista de Abreviaturas

AEE	Erro Médio de Estimação de Erro Grosseiro
AIC	Critério de Informação de Akaike
AVT1	Média dos Erros Estatísticos Tipo I
DEG	Detecção de Erro Grosseiro
DMT	Teste de Medida Dinâmico
DP	Desvio Padrão
EMT	Teste de Medida Estendido
GLR	Teste da Razão de Verossimilhança Generalizada
GT	Teste Global
IMT	Teste de Medida Iterativo
MGLR	Teste da Razão de Verossimilhança Generalizada Modificado
MI	Mista-Inteira
MQP	Mínimos Quadrados Ponderados
MSEGE	Técnica de Estimação de Erros Grosseiros Simultâneos Modificada
MT	Teste de Medida
MUBET	Técnica de Estimação Não-Tendenciosa Modificada
NIPALS	PLS Não-Linear Iterativo
NT	Teste Nodal

OP	Poder Total
PCA	Análise dos Componentes Principais
PCR	Regressão por Componentes Principais
PLS	Mínimos Quadrados Parciais
QI	Qualidade da Informação
RD	Reconciliação de Dados
SEGE	Estimação de Erros Grosseiros Simultâneos
UBET	Técnica de Estimação Não-Tendenciosa

Lista de Símbolos

C_C	Constante do estimador robusto de Cauchy
C_L	Constante do estimador robusto de Lorenz
H_0,H_1	Hipóteses para testes estatísticos
NA	Número de simulações no conjunto de geração de dados com erros puramente aleatórios
NB	Número de simulações no conjunto de geração de dados com erros grosseiros em medidas
NS	Número de simulações
NST	Número de simulações no conjunto de testes
NV	Número de simulações no conjunto de geração de dados com vazamentos
QI_i	Coeficiente de qualidade da informação relativa à medição i
R^2	Coeficiente de determinação de ajuste de um modelo empírico
VM_i	Vazão mapeada da corrente i
VR_i	Vazão reconciliada da corrente i
$V_{j,j}$	Elemento j,j, ou seja, da diagonal, da matriz $\mathbf{V}$
Δ_{bias}	Intervalo entre os valores superiores e inferiores dos erros de medição
Δ_{leak}	Intervalo entre os limites inferiores e superiores para os vazamentos

Σ	Matriz variância-covariância das medições
α	Significância estatística para o teste global, de medição ou teste nodal
α_i	Constante da expressão do QI em função da incerteza
χ^2	Teste Chi-Quadrado
$\chi^2_{(1-\alpha,\nu)}$	Valor crítico do teste χ^2 com significância α e ν graus de liberdade
δ	Erro sistemáticos de medição
$\delta_{i,ns}$	Erro de medição a ser adicionado na medição i na simulação ns
$\delta^r_{j,ns}$	Erro grosseiro no equipamento j na simulação ns
$\dot{m}_{i,entra}$	Vazão mássica que entra em um equipamento de processo
$\dot{m}_{j,sai}$	Vazão mássica que sai de um equipamento de processo
ϵ_r	Erro aleatório da medição
$\epsilon_{i,ns}$	Erro aleatório para a medida i na simulação ns
γ	Valor da estatística de teste do Teste Global
γ_k	Estatística de teste usada no teste de componentes principais baseada no balanço global
$\hat{\mathbf{b}}$	Vetor de parâmetros do modelo multivariável
$\hat{\mathbf{X}}_a$	Decomposição da matriz $\hat{\mathbf{X}}$ em suas sub matrizes não colineares
$\hat{\mathbf{X}}$	Vetor das variáveis independentes
$\hat{\mathbf{Y}}$	Vetor das variáveis dependentes
$\hat{b}_i$	Parâmetro do modelo multivariável
$\hat{x}_k$	Valor da variável independente k em uma regressão múltipla

$\hat{x}_{i,k}$	Valor da variável independente medida k no experimento i em uma regressão múltipla
$\hat{y}_k$	Valor da variável dependente k em uma regressão múltipla
$\mathbf{A}$	Matriz de Incidência (matriz Jacobiana das restrições)
$\mathbf{V}$	Matriz de variância-covariância dos resíduos, $\mathbf{R}$
$\mathbf{W}$	Inverso da matriz Σ
ϵ_r	Vetor de erro aleatório da medição
$\overline{\mathbf{X}}_{NOM}$	Vetor das vazões nominais do processo
$\mathbf{m}_j$	Vetor com valor 1 na posição j do vazamento e zero nas demais posições do teste GLR
μ	Média da distribuição Normal
ν	Graus de liberdade do Teste Chi-Quadrado
$\overline{\mathbf{W}}$	Matriz Variância-Covariância dos ajustes
$\overline{w_{i,i}}$	Elemento i,i da matriz $\overline{\mathbf{W}}$
$\overline{x}_{NOM,i}$	Valor da vazão nominal i na simulação ns
$\overline{x}_{NOM}$	Valor real da variável tomada pelo instrumento
ρ	Função objetivo genérica da reconciliação de dados
σ_i	Desvio-padrão dos dados de medição i
$\mathbf{Ad}$	Vetor dos ajustes
$\mathbf{F}(\mathbf{X})$	Vetor das restrições do processo relacionado com balanços de massa energia e quantidade de movimento
$\mathbf{R}$	Vetor dos resíduos do Balanço de Massa
$\mathbf{W}_a$	Autovetores da matriz $\overline{\mathbf{W}}$ do teste de componentes principais das medições

$\mathbf{W}_r$	Autovetores da matriz $\mathbf{V}$ do teste de componentes principais dos resíduos
$\mathbf{X}$	Vetor da medições reconciliadas
$\mathbf{Y}$	Vetor das medições
$\mathbf{e}_i$	Vetor com valor 1 na posição i do erro de medição e zero nas demais posições do teste GLR
$\mathbf{f}$	Vetor de assinatura do erro do teste GLR
$\mathbf{g}$	Vetor de parâmetros da regressão de componentes principais (PCR)
$\mathbf{p}_{ai}$	Vetor de componentes principais das medidas
$\mathbf{p}_n$	*loading vectors* da decomposição em componentes principais
$\mathbf{p}_r$	Vetor de componentes principais dos resíduos das restrições
$\mathbf{t}_n$	*score vectors* da decomposição em componentes principais
u_{x_i}	Incerteza-padrão da variável reconciliada i
ε_i	Diferença entre a variável medida e reconciliada, ponderada pelo inverso do desvio-padrão.
ad_i	Ajuste entre a medida y_i e a variável reconciliada x_i
b	Magnitude do erro grosseiro (medição ou vazamento) do teste GLR (escalar)
$bias_{lower}$	Limite inferior para o erro grosseiro de medição
$bias_{upper}$	Limite superior para o erro grosseiro de medição
l,up	Fator multiplicativo para o limite inferior e superior do erro grosseiro de medição das simulações, respectivamente
l_j^r,up_j^r	Fator multiplicativo para o limite inferior e superior do vazamento das simulações, respectivamente

$leak_{lower}$	Limite inferior para os vazamentos
$leak_{upper}$	Limite superior para os vazamentos
nc	Número de correntes
ne	Número de equipamentos
$r_{j,ns}$	Resíduo do equipamento j na simulação ns
r_j	Resíduo do balanço j (linha j do vetor $\mathbf{R}$
u	Incerteza-padrão
u_c	Incerteza-padrão combinada
u_i	Incerteza-padrão da variável i
u_{c_i}	Incerteza-padrão combinada da variável i
x_i	Valor reconciliado da medida i
$x_{i_{max}}$	Limite superior da variável reconciliada x_i
$x_{i_{min}}$	Limite inferior da variável reconciliada x_i
y_i	Medição i
y_m	Valor mais provável da medição
$y_{i,ns}$	Valor da medição i na simulação ns
$z_{a,c}$	Valor crítico do teste de medição da corrente i
$z_{a,i}$	Valor da estatística de teste do Teste de Medida da corrente i
$z_{am,i}$	Valor da estatística de teste do Teste de Medida modificado da corrente i
$z_{r,c}$	Valor crítico do teste de vazamento do equipamento j
$z_{r,j}$	Valor da estatística de teste do Teste Nodal do equipamento j
$z_{rm,j}$	Valor da estatística de teste do Teste Nodal modificado do equipamento j

Capítulo 1

Introdução

1.1 Água industrial e seu uso racional

Economicamente, a água é uma matéria-prima estratégica na agricultura e indústria (VALLE, 2005; CANADA, 2009; Van Der Leeden et al., 1990; AUSTRALIA, 2006; KENNY et al., 2009; EUROPEAN COMMISION, 2010; UNITED KINGDON, 2011). Na indústria, empresas de transformação dos setores de química, petroquímica e de alimentos, não só utilizam grandes volumes deste recurso, como também são os principais agentes poluidores, como pode ser comprovado através de um estudo da Fundação Estadual de Proteção Ambiental do Rio Grande do Sul(FEPAM, 2001), apresentado na Tabela 1.1 .

Além da grande extensão, a solução do problema de poluição dos corpos hídricos por parte da indústria é de grande complexidade, porém existem ferramentas para a minimização do uso da água e despejo de efluentes líquidos industriais(MIRRE et al., 2009; MANN; LIU, 1999; EL-HALWAGI, 1997; SMITH, 2005; ALMATO et al., 1997; BAGAJEWICZ et al., 2002; BAGAJEWICZ, 2000; JODICKE et al., 2001). Para resolver tal problema com qualquer metodologia de minimização do uso de água, se faz necessário o conhecimento prévio do balanço hídrico industrial, como sua topologia (rede de conexões), vazões e cargas de poluentes. Além disso, muitas oportunidades para a redução do consumo de água e redução da geração de efluentes líquidos podem ser avaliadas sem a aplicação desse tipo de metodologia de minimização, porém, é necessário informações confiáveis sobre o balanço hídrico. De posse de informações apropriadas do balanço hídrico, diversas respostas podem

Tabela 1.1: Maiores geradores de efluentes do Rio Grande do Sul em volume.

Indústria	Setor	Vazão ($m^3.dia^{-1}$)
Riocell S/A (Atual Aracruz)	Papel e Celulose	31200
Gerdau S/A	Metalúrgica	10200
REFAP S/A	Petroquímica	9431
Bunge Alimentos - Proteína Atual Solae	Alimentos	8800
Celulose Cambará	Papel e Celulose	7325
Cervejaria Brahma	Alimentos	7280
Fabrica de Papel Santa Therezinha	Papel e Celulose	5760
COPESUL (atual Braskem UNIB Triunfo)	Petroquímica	5540
	Total	85536

ser fornecidas à equipe de processos, entre elas:

- Existem pontos de perdas ou vazamentos no sistema?
- Existem falhas na instrumentação de medição?
- Em qual faixa de eficiência o equipamento está operando?
- Qual o equipamento *gargalo* do sistema?
- Como distúrbios de vazão e composição se propagam no circuito?
- Qual o período de maior e menor demanda de água?
- Qual o melhor momento de troca de um equipamento para manutenção?
- Qual o melhor ponto para adicionar novos instrumentos de medição?
- Qual a margem que o processo deve operar para que as emissões de poluentes não ultrapassem os exigidos pela legislação ambiental?

- Qual a melhor forma de minimizar ou reaproveitar o uso da água no sistema?

No que diz respeito a metodologias disponíveis para a atenuação do problema de consumo e descarte excessivo de água na indústria, as mais recomendadas são as que utilizam os princípios da produção limpa[1], citadas na próxima seção, embora atualmente as mais utilizadas são as soluções de fim-de-tubo[2].

1.2 Integração de processos, uma ferramenta para a redução do consumo de água na indústria

O uso racional da água na indústria vem sendo estudado na área de engenharia química desde a década de 80. Diversas ferramentas para produção limpa estão disponíveis na literatura, entre as quais, a integração de processos, que merece destaque. A integração mássica de processos consiste no reúso de efluentes das unidades geradoras com o objetivo de reaproveitar as correntes com algum potencial para transferir massa, ou seja, receber carga poluidora, antes do descarte. Utilizando os princípios de integração mássica, diversas técnicas pode ser utilizadas com o intuito de minimizar o uso de águas industriais, entre elas:

- Reúso: Efluentes de processos podem ser reutilizados diretamente em outras unidades em que o nível de poluente não interfere nas operações dos equipamentos da unidade, apresentada na Figura 1.1.a.

- Regeneração-reúso: Efluentes de processos podem ser regenerados através do tratamento parcial ou total para remoção de poluentes, para então serem reutilizados em outras unidades, apresentada na Figura 1.1.b.

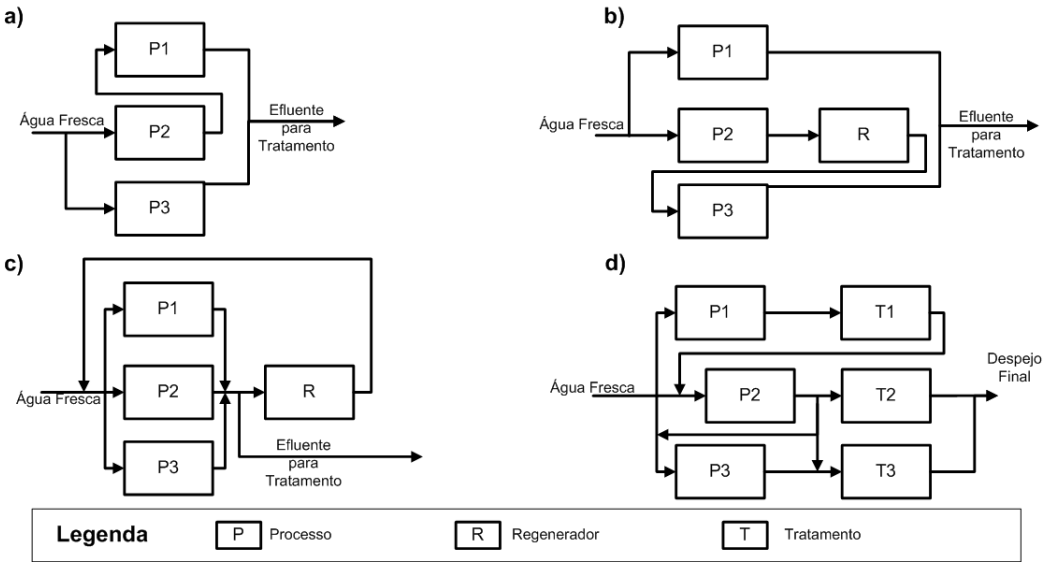
- Regeneração-reciclo: Efluentes de processos podem ser regenerados através do tratamento parcial ou total para remoção de poluentes para serem reciclados na mesma unidade, apresentada na Figura 1.1.c.

[1] Ações preventivas que buscam evitar a geração de resíduos por meio do aproveitamento máximo das matérias-primas utilizadas durante o processo produtivo(CNTL, 2003).

[2] Ações que diminuem o impacto ambiental de resíduos ou dão-lhes disposição final(CNTL, 2003). Exemplo: tratamento de efluentes ou disposição de resíduos sólidos em aterros sanitários.

- Segregação: Consiste, simplesmente, em evitar a mistura de correntes. Com a segregação é possível alocar as unidades de tratamento de efluentes mais próximo a sua fonte geradora, denominado tratamento distribuído, conforme a característica do efluente em relação ao tipo de contaminante, concentração, etc. A segregação quando aplicada em conjunto com o tratamento distribuído pode permitir o reaproveitamento de uma corrente de processo como nas técnicas de regeneração-reúso e regeneração-reciclo. A segregação de correntes com diferentes composições, apresentada na Figura 1.1.d, também evita diluições desnecessárias, reduzindo o custo de remoção de poluente a partir de uma corrente mais concentrada.

Figura 1.1: Reúso, regeneração-reciclo, regeneração-reúso e segregação com tratamento distribuído.



Fonte: Próprio autor

Existem ganhos econômicos significativos quando se realiza a implementação prática de técnicas de reúso de águas e efluentes industriais, conforme apresentado na Tabela 1.2. A Federação das Indústrias do Estado de São Paulo, FIESP, promove anualmente um concurso para estimular a conservação de recursos hídricos industriais nesse Estado, sendo os principais casos de sucesso apresentados em seu sítio na internet (FIESP, 2013), e compilados na Tabela 1.3, Tabela 1.4 e Tabela 1.5. A redução de uso de água e descarte de efluentes apresentados nas tabelas previamente

mencionadas poderiam ser resultantes da redução da produção das indústrias desses segmentos, porém, de acordo com dados de ABIQUIM (2013) a indústria química brasileira registrou um crescimento de 29,6 % entre os anos de 2010 e 2012, logo, os valores apresentados nas tabelas mencionadas são reduções reais devido a aplicações das técnicas de reaproveitamento de águas ou efluentes.

Tabela 1.2: Ganho com a aplicação de ferramentas de integração mássica.

Fonte	Área	Insumo	Economia(%)
Fontana (2002)	Teórico	Água	89
Farenzena et al. (2004)	Curtumes	Água	69
Raghava Rao et al. (2003)	Curtumes	Água e Insumos	60% de água e 24% de insumos
Aquim (2009)	Curtumes	Água	34
Mann e Liu (1999)	Petroquímica	Água	17
El-Halwagi e Manousiouthakis (1989)	Petroquímica	Água	58
El-Halwagi (1997)	Petroquímica	Água	60
Wang e Smith (1994)	Teórico	Água	58
Grande e Sales (2008)	Celulose e Papel	Água e Efluentes	10% de água e 14% de efluentes
Cunha et al. (2008)	Petroquímica	Efluentes	48
Hoinkis e Panten (2008)	Têxtil	Água	90
Shukla et al. (2011)	Papel	Água	42
		Efluentes	71
Tokos et al. (2012)	Alimentos	Água	37
Khor et al. (2012)	Petroquímica	Efluentes	32
Kiperstok et al. (2013b)	Petroquímica	Água	20
Kiperstok et al. (2013b)	Petroquímica	Efluentes	40
Kiperstok et al. (2013b)	Petroquímica	Água	42
Thevendiraraj et al. (2003)	Alimentos	Água	30
Gomes (2002)	Teórico	Água	37

Tabela 1.3: Redução de consumo de água e descarte de efluentes obtida com técnicas de reúso de águas industriais nas indústrias do Estado de São Paulo.

Empresa	Ramo	Insumo	Economia %
3M do Brasil	Químicos	Água	25
		Insumos	44
		Efluentes	33
AGCO Corporation	Mecânico	Efluentes	36
All Washed	Lavanderia	Água	50
Lavanderia Industrial		Efluentes	80
AmBev	Alimentos	Água	14
ArcelorMittal Brasil	Siderurgia	Água	20
Baxter hospitalar	Médico-Hospitalar	Água	70
Brasilit Ind. e Comp.	Construção Civil	Água	34
Braskem	Químicos	Água	85
		Efluentes	50
BSH Continental	Eletrodomésticos	Efluentes	30
Eletrodomésticos		Água	30
Cabot Brasil Ind. E Com.	Químicos	Água	30
Cargill Agrícola	Alimentos	Efluentes	20
Cargill Agrícola	Alimentos	Efluentes	29
Caterpillar	Mecânico	Água	60
Cebrace Cristal Plano	Transformação	Água	10
		Efluentes	47
Copebrás	Químicos	Água	81
Cosipa	Siderurgia	Água	96
Covolan Indústria Têxtil	Têxtil	Água	70
Dana Indústrias	Mecânico	Água	26
Dana Indústrias		Efluentes	28
Dedini Indústrias de Base	Fundição	Água	30
Elekeiroz	Químicos	Água	2
Electrolux do Brasil	Eletrodomésticos	Água	9

Continua

Tabela 1.4: Redução de consumo de água e descarte de efluentes obtida com técnicas de reúso de águas industriais. Continuação

Empresa	Ramo	Insumo	Economia %
Eluma Ind. e Com.	Metalurgia	Efluentes	98
Emicol Eletrônica	Eletroeletrônico	Água	70
Destilaria Alcídia	Eletroeletrônico	Água	80
Geni Porcelana	Cerâmicos	Água	50
Gnatus Equip.	Médico-Hospitalar	Água	70
Médicos Odontológicos		Efluentes	85
Ind. Mineradora Pagliato	Químicos	Água	15
Iochpe-Maxion	Mecânico	Água	59
		Efluentes	60
Kimberly Clark Brasil	Papel	Água	50
Logictel	Telecomunicações	Água	93
Lwarcel Celulose e Papel	Papel	Água	43
Mahle Metal Leve	Mecânico	Água	40
Metalúrgica Inca	Metalurgia	Água	89
Metalúrgica Inca	Metalurgia	Água	90
Metalúrgica Inca	Metalurgia	Água	84
MWM International Ind. de Motores	Mecânico	Água	19
Nestlé Brasil	Alimentos	Água	23
Opersan Resíduos Ind.	Tratamento	Água	22
	de efluentes	Efluentes	6
Orsa Celulose Papel e Embalagens	Papel	Água	75
Panozon Ambiental	Têxtil	Água	90
Papirus Ind. de Papel	Papel	Água	17
Refinaria de Paulínia	Petroquímica	Água	38
Pirelli Pneus	Transformação	Água	35
Saint-Gobain-Brasilit	Transformação	Água	34
Sanofi-Aventis Farmacêutica	Farmacêutica	Água	20

Continua

Tabela 1.5: Redução de consumo de água e descarte de efluentes obtida com técnicas de reúso de águas industriais. Continuação

Empresa	Ramo	Insumo	Economia %
Santista Têxtil Brasil	Têxtil	Água	25
		Efluentes	21
São Carlos S/A Ind.	Papel	Água	70
de Papel e Embalagens		Efluentes	70
Spal Ind. Brasileira de Bebidas	Alimentos	Água	75
Termogal Tratamento de Superfícies	Metalurgia	Água	95
Uniforme Lavanderia e Locação	Lavanderia industrial	Água	25
Unilever Brasil	Domissanitários	Efluentes	60
Usina Colombo Açúcar e Álcool	Sucroalcooleiro	Água	15
Valeo Sistemas Automotivos	Mecânico	Água	52
Voltra Brasil	Mecânico	Água	22
Votorantim Celulose e Papel	Celulose e papel	Água	86
White Martins Gases Industriais	Química	Água	22

Para aplicar ferramentas de reaproveitamento de água (KIPERSTOK et al., 2013b; KIPERSTOK et al., 2013a; MIRRE et al., 2009), é necessário que informações a respeito de seu uso sejam conhecidas, entre elas:

- Conexão entre os equipamentos.
- Vazão mássica ou volumétrica das correntes e dos equipamentos.
- Carga de poluentes gerada por cada equipamento.
- Limites máximos e mínimos de vazão das correntes.
- Limites máximos e mínimos de vazão na entrada ou saída dos equipamentos.
- Limites máximos e mínimos de concentração de poluentes na entrada ou saída dos equipamentos.

- Perdas para o meio ambiente, se existirem.
- Custos de novas conexões.
- Custo de instalação de novos equipamentos.

A falta de algumas destas informações, como, por exemplo, a vazão das correntes ou equipamentos, dificulta, ou pode inviabilizar, a aplicação direta de metodologias de reaproveitamento de águas industriais, sendo necessário o uso de outros métodos para avaliar o balanço mássico da unidade industrial. Além disso, é necessário que as informações mencionadas anteriormente tenham a melhor qualidade possível, sendo que a confiabilidade e a robustez da rede de reaproveitamento resultante está diretamente relacionada à qualidade das informações disponíveis. Kiperstok et al. (2013b) apresenta diversos estudos de casos industriais visando a minimização de uso e descarte de águas, onde salienta que medições de vazão e composição do circuito de uso da água em tais casos eram escassas ou inexistentes. Tal problema tem sua origem na conjuntura histórica da concepção dos processos químicos: no Brasil, as plantas industriais atuais foram concebidas na década de 60 e 70, quando a água era considerada um recurso abundante e o investimento em redes de monitoramento do uso da água era visto como um gasto desnecessário. A falta de manutenção apropriada das redes de medição de água também levam à degradação da informação do circuito de uso de água prejudicando a aplicação de metodologias de minimização de consumo de água.

1.3 Objetivos do trabalho

O presente trabalho tem como objetivo propor uma metodologia para realizar a avaliação e monitoramento de sistemas de medição de vazão total de água em indústrias de processos em estado estacionário. Os objetivos específicos do presente trabalho são:

1 Seleção de índices para acompanhamento do sistema de medição de vazão total de águas em plantas industriais.

2 Desenvolvimento de uma metodologia que, baseado nos índices selecionados, e, atuando em conjunto com modelos preditivos, seja capaz de indicar os sensores com falhas e detectar vazamentos em equipamentos de processo.

1.4 Estrutura da tese

Esta tese está dividida em cinco capítulos, arranjados da seguinte forma: No presente capítulo, é apresentado o problema e os objetivos a serem abordados neste trabalho.

No Capítulo 2 são apresentadas as bases das teorias para avaliação de balanços mássicos: reconciliação de dados e detecção de erros grosseiros, além das variantes e limitações destas técnicas.

No Capítulo 3 é apresentada a metodologia proposta para avaliação de balanços hídricos e detecção de erros nos mesmos.

No Capitulo 4 são avaliados os resultados do método proposto a partir da aplicação do mesmo em 16 problemas testes. Ao longo dos resultados, também são discutidas algumas particularidades relativas a cada problema.

No Capítulo 5 são apresentadas as principais conclusões deste trabalho, as publicações desenvolvidas a partir do mesmo e sugestões para trabalhos futuros.

Capítulo 2

Revisão Bibliográfica

2.1 Coleta de informações de balanços hídricos

Conforme mencionado no capítulo introdutório, a coleta de informações sobre balanços hídricos para a aplicação de metodologias de reaproveitamento de água passa por diversas dificuldades, desde a falta de informações até a baixa confiabilidade nos dados de processo. Devido à complexidade de sítios petroquímicos, a simples coleta e levantamento dos equipamentos e correntes relacionadas aos balanços hídricos é uma tarefa de difícil execução, pois é necessária uma grande interação com a equipe de engenharia e operação da planta para o levantamento de tais informações, tanto em nível técnico quanto em nível gerencial. Como exemplo, podemos citar as tarefas chave usadas por Mann e Liu (1999) em seus 5 trabalhos de aplicação do método do *Pinch* de água em complexos químicos em Taiwan de 1996 a 1997. Tais tarefas são descritas abaixo:

1 Selecionar a unidade industrial para o projeto;

- 1.a Divulgar os objetivos do projeto para as unidades industriais interessadas;
- 1.b Organizar um comitê diretor para o projeto em cada unidade industrial interessada;
- 1.c Desenvolver um programa de incentivos para promover o sucesso do projeto nas unidades industriais interessadas;

2 Ajudar as unidades industriais na coleta de dado;

2.a Compilar e completar o diagrama de tubulação e instrumentação do circuito de uso de água;

2.b Pesquisar as fontes de água e pontos de uso (demanda e suprimento) de acordo com a categoria do processo ;

3 Identificar os potenciais de reúso de água e sua classificação de prioridade conforme determinado pelo comitê diretor;

3.a Consultas entre os membros do comitê diretor do projeto para identificar os potenciais de reúso de água;

3.b Classificação de importância das opções identificadas;

4 Medição das vazões e padrões de qualidade das fontes e pontos de uso;

5 Desenvolver as propostas de reúso de água baseadas em técnicas sistemáticas em conjunto com a experiência da equipe de engenharia;

5.a Aplicação de métodos como diagrama de fontes de água, *Pinch* para água, programação matemática, entre outros;

5.b Evolução dos resultados obtidos anteriormente com a experiência da equipe de engenharia;

6 Executar o projeto de engenharia preliminar e avaliação econômica das propostas;

6.a Projeto de engenharia preliminar da nova rede de reúso proposta e das mudanças estruturais necessárias em outros equipamentos de processo;

6.b Análise dos benefícios econômicos diretos e indiretos, com o uso de indicadores econômicos, como Valor Presente Líquido, Taxa Interna de Retorno, entre outros;

7 Detalhamento técnico final da nova rede de reúso e demais modificações necessárias;

8 Implementação detalhada da nova rede de reúso e demais modificações necessárias;

8.a Auxiliar os trabalhos da equipe de engenharia e a partida das unidades que farão uso da água de reúso;

8.b Acompanhamento dos benefícios econômicos e operacionais;

9 Divulgação dos resultados;

9.a Publicação de relatórios e documentos com a experiência dos projetos;

9.b Patrocínio de eventos relacionados à área de reúso de águas e apresentação dos resultados nos mesmos.

Também cabe citar a sistemática usada pela Rede de Tecnologias Limpas da Bahia (TECLIM) desenvolvida com base em 13 trabalhos na área ambiental em complexos químicos na região de Salvador. As 9 etapas da técnica denominada 'Auditoria de Fontes de Alimentação de Efluentes' (OLIVEIRA, 2011) são:

1 Caracterização das regiões próximas às fontes geradoras;

2 Avaliação das condições de amostragem quanto a distúrbios que possam mascarar resultados;

3 Levantamento das vazões medidas e estimadas de efluentes;

4 Amostragem das canaletas e bacias do sistema de efluentes orgânicos;

5 Análise das amostras coletadas nas canaletas e bacias de efluentes;

6 Realização do balanço de massa das correntes;

7 Cálculo do acúmulo de orgânicos sólidos na bacia de efluentes;

8 Análise dos dados;

9 Elaboração de programas de ação.

Além disso, Kiperstok et al. (2013b), Kiperstok et al. (2013a) apresentaram as principais etapas e dificuldades para aplicação de técnicas de uso racional de água na indústria petroquímica, destacando a importância da coleta e análise sistemática de dados de processo para esse fim. Nesse mesmo trabalho, são apresentadas as

ferramentas gerenciais e técnicas desenvolvidas ao longo de 11 anos de atuação em projetos com empresas do ramo petroquímico.

Nesse contexto, é importante salientar o papel fundamental da elaboração correta dos balanços hídricos industriais para aplicação de qualquer metodologia de otimização ambiental dos mesmos.

2.2 Métodos para correção de incoerências em balanços de massa

Após o levantamento topológico e das vazões das correntes e equipamentos de processo, é necessária testar as mesmas. O teste se faz necessário, pois, em muitos casos, onde as medições de processo estão presentes, frequentemente são encontradas incoerências nos balanços mássicos quando calculados a partir de medições. A principal fonte de incoerência no balanço mássico é a própria precisão do equipamento de medição, porém é possível que o mesmo apresente alguma falha, ou ainda, exista um vazamento material em algum ponto do sistema. Para melhor compreender o papel da medição na detecção de incoerências em balanços de massa, é apresentado o conceito de medição segundo Romagnoli e Sánchez (1999):

$$y = \overline{x}_{NOM} + \epsilon_r + \delta, \quad y \in \mathbb{R}, \quad \overline{x}_{NOM} \in \mathbb{R}, \tag{2.1}$$

em que y é o valor da medição, $\overline{x}_{NOM}$ é o valor real da variável tomada pelo instrumento, ϵ_r é o erro aleatório da medição e δ é o erro sistemático da medição. Os erros aleatórios implicam que a magnitude e o sinal do erro de cada repetição não podem ser previstos, sendo a mesma caracterizada por uma variável aleatória a qual segue uma função de densidade de probabilidade (FDP). Em condições normais, estes tipos de erros não podem ser removidos e não são desejáveis, porém, é possível aplicar técnicas de filtragem por equipamentos ou por programas de computador. Os erros sistemáticos são causados por eventos não aleatórios como um defeito no instrumento de medição, instrumento mal calibrado, por corrosão, deformação mecânica ou más condições de operação. Este tipo de erro tem, a cada repetição da medida, sinal previsível e também não são desejáveis. Os erros sistemáticos são

chamados, por diversos autores, de erros grosseiros, o qual não está de acordo com os preceitos da Metrologia, porém, como a literatura especializada em estudo de falhas em balanços e instrumento de medição de processos químicos industriais usa ambas como sinônimos, esta equivalência de terminologia será adotada neste trabalho. Em processos industriais, deseja-se detectar os efeitos dos erros grosseiros para acionar procedimentos de manutenção e manter as informações do processo sempre em níveis confiáveis. Na prática, busca-se sempre isolar os efeitos dos erros aleatórios e grosseiros e, para distinguí-los, aplicam-se técnicas matemáticas conhecidas como reconciliação de dados (RD) e detecção de erros grosseiros (DEG) descritas em detalhes a seguir.

Como todo processo químico, balanços hídricos seguem relações de conservação de massa, e podem ser considerados em estado estacionário e sem reações químicas, conforme:

$$\sum_{i=1}^{n} \dot{m}_{i,entra} = \sum_{j=1}^{n'} \dot{m}_{j,sai}, \tag{2.2}$$

em que $\dot{m}_{i,entra}$ são as vazões das correntes que entram no equipamento e $\dot{m}_{j,sai}$ as vazões das correntes j que saem do equipamento. Embora trivial, o modelo de conservação de massa é usado como base para corrigir as vazões de entrada e saída dos equipamentos através da técnica de reconciliação de dados, que pode ser combinada com outras técnicas estatísticas para detecção de incoerências maiores em balanços, como, por exemplo, vazamentos em equipamentos ou falhas em instrumentos de medição.

A reconciliação de dados pode ser formulada como um problema de otimização com restrições, onde essas últimas provém da termodinâmica, balanços de massa, energia ou quantidade de movimento do sistema em estudo. A função objetivo da otimização é a minimização do erro quadrático entre as variáveis medidas e reconciliadas, multiplicado pelo inverso da variância das medidas (NARASIMHAN; JORDACHE, 2000), conhecida, também, como estimador de mínimos quadrados

ponderados (MQP), conforme:

$$\begin{cases} \min\limits_{x_i \in \Re^n} \sum\limits_{i=1}^{n} \left(\dfrac{y_i - x_i}{\sigma_i} \right)^2 \\ \text{sujeito a} \\ \mathbf{F}\left(\mathbf{X}\right) = 0 \\ x_{i_{min}} \leq x_i \leq x_{i_{max}}, \end{cases} \tag{2.3}$$

em que σ_i é o desvio-padrão das vazões medidas, y_i são as vazões medidas, x_i são as vazões reconciliadas, $x_{i_{min}}$ são os limites inferiores das vazões reconciliadas, $x_{i_{max}}$ são os limites superiores das vazões reconciliadas e $\mathbf{F}\left(\mathbf{X}\right)$ são as restrições impostas pela termodinâmica, balanços de massa, energia e quantidade de movimento. Mais especificamente para balanços hídricos industriais, abordado no presente trabalho, as equações de restrições, $\mathbf{F}\left(\mathbf{X}\right)$, são os balanços de vazão total em base mássica ou volumétrica. Como resultado da resolução do problema de RD, as variáveis reconciliadas, x_i, são ajustadas de forma ponderada pelo inverso de seu desvio-padrão, respeitando as equações de restrição. Note que, medições com o módulo de suas magnitudes muito elevadas, que podem ser indicativos de erros grosseiros, fazem com que as variáveis reconciliadas sofram um grande ajuste em relação ao seu valor medido, y_i, para minimizar esses desvios. Considerando que na RD, as variáveis estão conectadas pelos balanços de massa, o ajuste das variáveis reconciliadas, x_i, acaba afetando o valor final das demais variáveis, sendo esse efeito é denominado pela literatura como efeito de espalhamento ou *smearing*. Na presença de erros grosseiros ou quando a distribuição dos erros de medição não seguem a distribuição Normal, outros estimadores podem ser mais adequados. Os estimadores robustos são funções objetivo que reduzem o efeito dos erros grosseiros no problema de reconciliação de dados, amortecendo o valor das variáveis ajustadas à medida que a mesma aumenta, reduzindo o efeito do espalhamento e tornando mais confiável os resultados da RD mesmo na presença de erros grosseiros. Para apresentar os estimadores robustos, seja o regressor genérico:

$$\varepsilon_i = \left(\frac{y_i - x_i}{\sigma_i} \right), \tag{2.4}$$

em que y_i é o valor da medida i e x_i é o valor da medição i reconciliada. A equação regressora genérica, $\rho(\varepsilon_i)$, é definida para cada estimador. O estimador padrão de

mínimos quadrados ponderados é definido da seguinte forma:

$$\rho(\varepsilon_i) = \frac{1}{2}\varepsilon_i^2, \tag{2.5}$$

já o estimador robusto de Cauchy é assim definido:

$$\rho(\varepsilon_i) = C_C^2 ln\left(1 + \frac{\varepsilon_i^2}{C_C^2}\right), \tag{2.6}$$

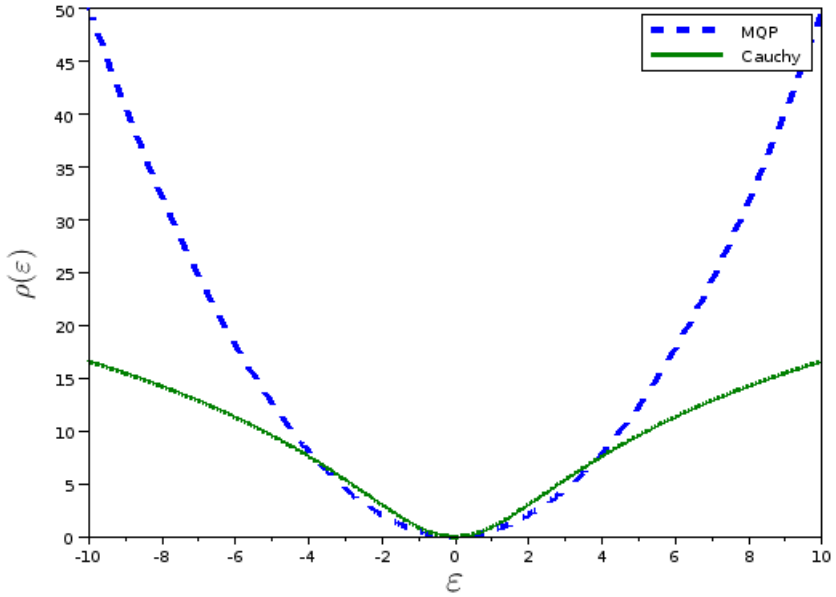
enquanto que o estimador robusto de Lorenz é definido como:

$$\rho(\varepsilon_i) = -\frac{1}{1 + \left(\frac{\varepsilon_i^2}{2C_L^2}\right)}, \tag{2.7}$$

em que C_C e C_L são parâmetros das funções.

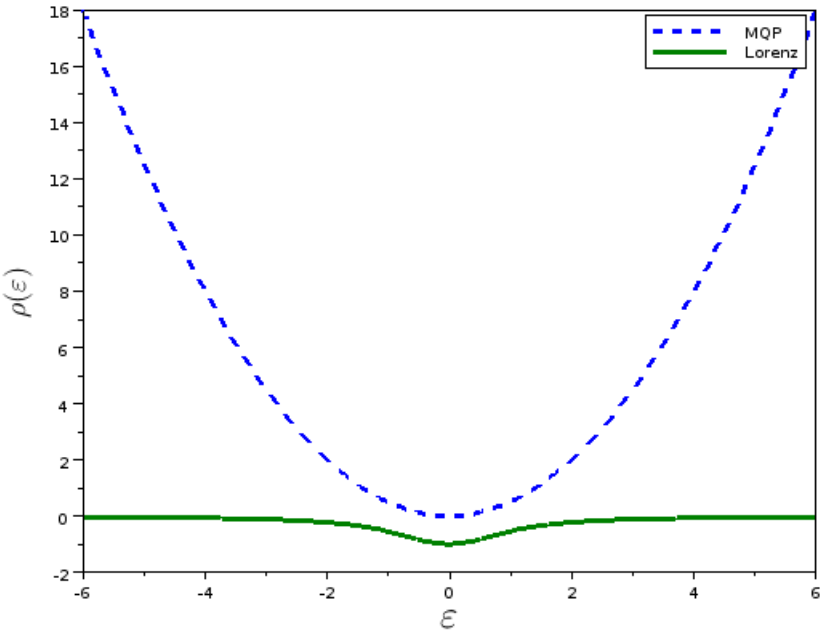
Para entender melhor o efeito dos estimadores robustos sobre o erro, são apresentados um gráfico de ρ em função do erro-padrão, ε, para o estimador de Mínimos Quadrados Ponderados (MQP) e de Cauchy na Figura 2.1 e o estimador MQP e de Lorenz na Figura 2.2.

Figura 2.1: Comparação entre estimador de Cauchy (C_C igual a 2,3849, determinado por Özyurt e Pike (2004)) e Mínimos Quadrados Ponderados.



Fonte: Próprio autor

Figura 2.2: Comparação entre estimador de Lorenz (C_L igual a 2,6, determinado por Özyurt e Pike (2004)) e Mínimos Quadrados Ponderados.



Fonte: Próprio autor

Como pode ser observado, com o estimador de mínimos quadrados ponderados, a função ρ, por ter sua forma quadrática, tem sua magnitude ampliada de forma muito elevada para grandes erros-padrão. Já no caso das funções com estimadores robustos, como Cauchy e Lorenz, é possível observar que, a medida que o erro-padrão aumenta, a magnitude de ρ aumenta de forma não tão elevada. Na prática, os estimadores robustos amortecem o efeito dos erros grosseiros no problema de otimização, resultando em ajustes mais proporcionais às magnitudes dos erros, evitando assim, o espalhamento dos erros grosseiros através de outras variáveis. A avaliação de diversos outros estimadores robustos com aplicação à reconciliação de dados é apresentada em Özyurt e Pike (2004). Nesse extenso trabalho, os estimadores foram aplicados em diversos problemas clássicos de reconciliação e detecção de erros mostrando um desempenho semelhante e, em alguns casos, até melhores que os métodos sequenciais de detecção de erros, com a vantagem de ser necessário somente uma resolução de um problema de otimização. Os autores também apresentam a importância de seleção adequada dos parâmetros dos estimadores, como C_C, para Cauchy e C_L para Lorenz. Arora e Biegler (2001) e Zhang et al. (2010) tam-

bém apresentaram resultados de estimadores aplicados à RD e DEG para processos industriais onde destacam resultados positivos em relação ao uso dos mesmos.

2.3 Reconciliação de Dados e Metrologia

Na função objetivo da reconciliação de dados, o regressor genérico, apresentado na Equação (2.4), representa a diferença entre o valor observado e o valor reconciliado ponderado pelo recíproco do desvio-padrão. O desvio-padrão é uma medida de dispersão de dados em torno da média dos mesmos e pode ser aplicado a qualquer tipo de dados com distribuição Normal, obtidos a partir de sistemas de medição ou não. Em reconciliação de dados é assumido que os dados das medições do processo partem de instrumentos de medição, onde, além das dispersão dos dados, devem ser considerados a variabilidade relativa ao instrumento propriamente dito. O parâmetro que caracteriza a dispersão dos valores atribuídos a uma grandeza medida é denominado incerteza (INMETRO, 2012). A incerteza de medição geralmente engloba muitas componentes, que são geralmente divididas em duas categorias: Tipo A e Tipo B. Avaliação da componente da incerteza de medição do Tipo A é realizada por uma análise estatística dos valores medidos, obtidos sob condições definidas de medição, e pode ser caracterizada por desvio-padrão. Já a incerteza do Tipo B é toda a incerteza avaliada ou obtida por meios diferentes daquele adotado para uma avaliação do Tipo A, por exemplo:

- Associada a valores publicados por autoridade competente;
- Associada ao valor de um material de referência certificado;
- Obtida a partir de um certificado de calibração ou classe de exatidão de um instrumento de medição verificado;
- Obtida a partir de limites deduzidos da experiência pessoal.

Embora a literatura de reconciliação de dados utilize amplamente a incerteza-padrão na definição de seus regressores genéricos, segundo os preceitos da metrologia, o parâmetro mais apropriado a ser usado é a incerteza-padrão, definida como(INMETRO,

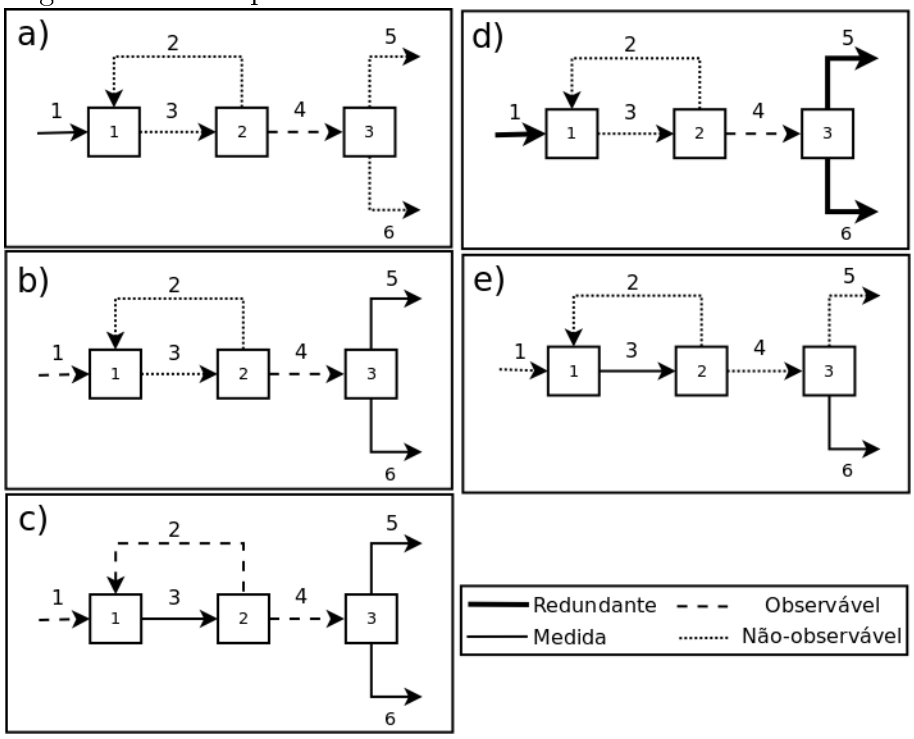
2012): 'Incerteza de medição expressa na forma de um desvio-padrão'. A incerteza-padrão envolve a incerteza da variabilidade dos dados em conjunto com a variabilidade do instrumento de medição e é representada pelo símbolo u. Para caso em que a incerteza-padrão u depende de uma soma de outras incertezas individuais, esta é denominada de incerteza-padrão combinada e é representada pelo símbolo u_c (INMETRO, 2012; BIPM, 2008; BIPM, 2009). Nesse contexto, o presente trabalho, passa a mencionar o desvio-padrão de medições por sua terminologia metrológica apropriada, ou seja, incerteza-padrão.

2.4 Classificação de variáveis

Como pode ser observado na Equação (2.3) as medições das variáveis são incluídas na função objetivo com a respectiva variável a ser reconciliada. No fim da otimização, as variáveis terão seu valor ajustado, satisfazendo as restrições do processo. O número de variáveis que podem ser efetivamente reconciliadas depende do número de medições, da posição no fluxograma de processo e da topologia do processo em si (restrições). Em alguns casos, uma medida, por exemplo, vazão total, não oferece informação suficiente para realizar a RD, neste caso, esta é classificada como 'medida' e 'não redundante' e é apresentada na Figura 2.3, caso 'e' (correntes 3 e 6). No caso que uma variável possui informações para efetuar a RD, esta é dita 'redundante', como no caso 'd' (correntes 1, 5 e 6) da Figura 2.3, consequentemente terá seu valor ajustado. Embora para uma melhor confiabilidade nos resultados da aplicação de RD seja necessário ter um grande número de variáveis medidas, isso não é viável na prática por questões de custo. Quando uma variável, por exemplo, uma vazão total, não pode ser medida, mas pode ser estimada através do balanço de massa, esta é classificada como 'observável' e pode ser exemplificada através do caso 'c' (correntes 1, 2 e 4) e 'b' (correntes 1 e 4) da Figura 2.3. Variáveis observáveis não terão seus valores ajustados pela técnica de RD, mas serão estimadas. Em alguns casos, algumas variáveis não medidas não podem ser observadas, pois, as equações que as relacionam com outras variáveis (balanços de massa) não permitem tal inferência, sendo tais variáveis classificadas como 'não-observáveis', conforme apresentado no caso 'e' (correntes 1, 2, 4 e 5) na Figura 2.3. Correntes não-observáveis

não podem ser estimadas por restrições de processo, e, embora alguns métodos de RD apresentem valores numéricos para as mesmas, esses não tem sentido. Embora sejam necessárias variáveis redundantes para a realização da reconciliação de dados, dependendo da técnica de RD escolhida, a técnica de classificação de variáveis em si não se faz necessária, entretanto, a classificação oferece informações importantes sobre a estrutura do problema de RD. As técnicas de classificação de variáveis mais usadas são baseadas na análise da matriz Jacobiana do modelo do processo (ALBUQUERQUE; BIEGLER, 1996; CROWE, 1989; VEVERKA; MADRON, 1997; SWARTZ, 1989; SÁNCHEZ; ROMAGNOLI, 1996), embora alguns autores utilizem análise de grafos para tal classificação (STANLEY; MAH, 1980; KRETSOVALIS; MAH, 1987).

Figura 2.3: Exemplo de análise de observabilidade e redundância.



Fonte: Próprio autor

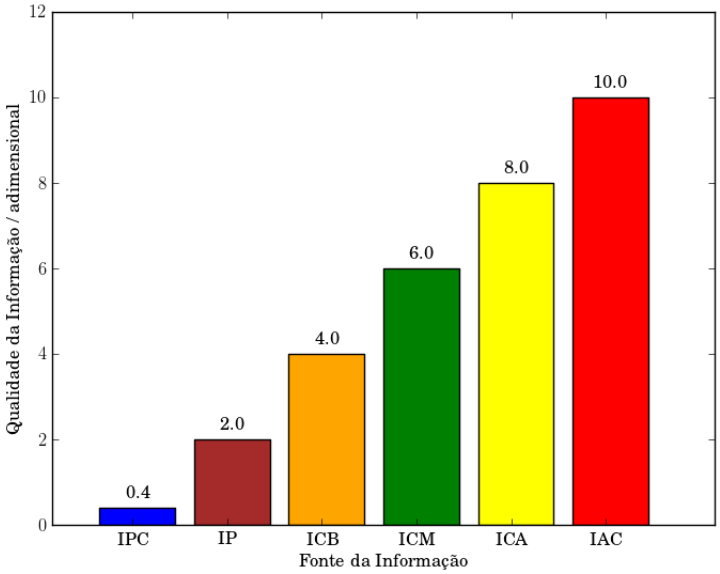
Conforme mencionado anteriormente, um problema frequente encontrado em balanços hídricos industriais é a falta de medições, que restringe ou impossibilita a aplicação de técnicas de RD, visto que é necessário um número mínimo de variáveis redundantes para aplicar tal técnica. Para contornar esse problema, Fontana et al. (2004a), Fontana et al. (2004b) propuseram um método para casos onde diversas va-

riáveis de processos não são medidas. Nesses trabalhos, posteriormente aprimorado por Martins et al. (2010), é definido o conceito de Qualidade de Informação, ou QI, para introduzir uma redundância artificial no sistema de medição, tanto para vazões medidas quanto para não-medidas. A QI é um conceito subjetivo e adimensional, que utiliza diversas fontes de informações disponíveis para inferir a sua incerteza da variável associada. Para que isso ocorra, todo tipo de informação pode ser avaliada, como:

- Estimativas;
- Dados de projeto;
- Documentação existente;
- Variáveis medidas;
- Dados técnicos como, por exemplo, simulações computacionais ou inferências a partir de outras medições;
- Informações da equipe de trabalho da planta, etc.

Cada informação obtida recebe um nível de confiança, que é um conceito definido em acordo entre a equipe de pesquisa e a equipe de engenharia da planta, onde os dados provenientes de medições diretas também são classificadas conforme seu índice de confiança. Um exemplo de como pode ser definida a graduação de QI é apresentada na Figura 2.4.

Figura 2.4: Exemplo de graduação da QI .



Fonte: (NARCISO, 2012).

No caso da Figura 2.4, a QI foi graduada em seis conceitos:

- **IPC**, ou informação pouco confiável, a qual possui valor 0,4;
- **IP**, ou informação precária, a qual possui valor 2,0;
- **ICB**, ou informação de confiança baixa, a qual possui valor 4,0;
- **ICM**, ou informação de confiança média, a qual possui valor 6,0;
- **ICA**, ou informação de confiança alta, a qual possui valor 8,0;
- **IAC**, ou informação de altíssima confiança, a qual possui valor 10,0.

Segundo a metodologia, uma vazão medida pode ser redundante ou não, ou ainda, as medidas podem ser redundantes mas sua incerteza associada desconhecida, sendo assim, a vantagem da metodologia é atribuir pesos relativos para cada fonte de informação. Usando a função objetivo com a incorporação do conceito da QI, é possível obter em sistemas sem redundância de medição, um resultado similar a reconciliação de dados clássica(FONTANA et al., 2004b; FONTANA et al., 2004a; SOUZA et al., 2010), conforme:

$$\min_{VR_i \in \Re^n} \sum_{\mathbf{i=1}}^{n} \frac{QI_i^2}{VM_i}(VR_i - VM_i)^2, \tag{2.8}$$

em que VR_i é a vazão reconciliada, VM_i é a vazão mapeada obtida através de alguma fonte de informação associada ao QI_i. A relação entre QI e a incerteza é dada por (FONTANA et al., 2004b; FONTANA et al., 2004a; MARTINS et al., 2010; SOUZA et al., 2010):

$$QI_i = \alpha_i . \frac{y_i}{k_i . u_{c_i}}, \tag{2.9}$$

em que u_{c_i} é a incerteza-padrão combinada e α_i/k_i é uma constante de proporcionalidade. Martins et al. (2010) provaram que para balanços de massa total, a razão α_i/k_i é constante e igual a 0,1, logo, é possível utilizar outros estimadores na função objetivo com a conversão apropriada entre QI e a incerteza, ou ainda, usar diretamente a função objetivo da Equação (2.8). Martins et al. (2010) citam, ainda, a aplicação desta metodologia com sucesso em casos industriais reais.

2.5 Erros grosseiros e métodos para detecção e identificação

Em muitos processos onde se dispõe de um sistema de medição, observa-se que, ao calcular o balanço de massa ou energia com informações das respectivas medições, muitas vezes esses balanços não são respeitados exatamente. Tal fato pode ocorrer por diversos motivos, entre eles:

1 Oscilação aleatória (ruído branco) do circuito que realiza a medição;

2 Variações nas condições climáticas (temperatura, umidade, etc) que interferem na precisão de determinados equipamentos de medição;

3 Falta de calibração e/ou manutenção periódica no medidor;

4 Medidor danificado;

5 Vazamento em algum equipamento ou tubulação.

Os itens acima mencionados são indesejáveis do ponto de vista de operação de processos químicos, porém, somente os itens 3, 4 e 5 podem ser eliminados através da

manutenção apropriada dos equipamentos de medição ou de processo e são denominados erros grosseiros. Segundo Narasimhan e Jordache (2000), qualquer técnica de detecção de erros grosseiros deve, preferencialmente, possuir as seguintes características:

1 Detectar a presença de um ou mais erros nos dados.

2 Identificar o tipo (vazamento ou falha na medição) e a localização do erro.

3 Localizar e identificar múltiplos erros de medição que poderão estar presentes nos dados.

4 Estimar a magnitude do erro.

Para detectar e identificar anormalidades em sensores de medição ou em balanços de massa (vazamentos), diversas técnicas estão disponíveis, muitas possuem a capacidade somente de detectá-las no balanço de massa do processo como um todo, enquanto outras de identificá-las e localizá-las em pontos únicos. Outras técnicas podem identificar ou localizar erros grosseiros em múltiplos locais. Os métodos de DEG também se distinguem pela forma de aplicação, sendo que alguns aplicam o procedimento uma única vez enquanto outros são baseados em testes iterativos, realizando diversos testes intermediários em versões modificadas do balanço de massa. Narasimhan e Jordache (2000) apresentam em um livro texto bem amplo, o detalhamento e características de diversos métodos para DEG. Algumas das técnicas mais consolidadas para DEG são apresentadas na Tabela 2.1 e Tabela 2.2, onde também é possível observar que muitos autores testam seus métodos de DEG em um universo muito restrito de problemas.

Tabela 2.1: Técnicas de identificação de erros grosseiros e o número de exemplos usados para a testar as mesmas.

Autor	Teste para Detecção/Identificação de Erros Grosseiros	Número de Exemplos Testados
Romagnoli e Stephanopoulos (1981)	Método Algorítmico Proposto	3
Rosenberg et al. (1987)	Teste de Medida dinâmico e estendido	7
Narasimhan e Mah (1987)	Razão de Verosimilhança Generalizada (GLR)	2
Reddy e Mavrovouniotis (1998)	Abordagem de Redes Neuronais	1
Rollins e Davis (1992)	Método da Estimação Não-Tendenciosa (UBET)	2
Sánchez et al. (1999)	Estimação Simultânea de Erros Grosseiros Modificada UBET e GLR modificados	2
Bagajewicz et al. (1999)	Avaliação de Análise de Componentes Principais (PCA) para DEG	1
Kongsjahju et al. (2000)	Modificação do Método UBET para sistemas em série	1
Soderstrom et al. (2001)	Abordagem Inteira-Mista	4
Arora e Biegler (2001)	Estimadores Robustos para RD e DEG	3
Özyurt e Pike (2004)	Reconciliação Robusta	7
Wang et al. (2004)	Aprimoramento do MT e Teste Nodal (NT)	1
Devanathan et al. (2005)	Métodos Bayesianos e de Verossimilhança	1

Continua

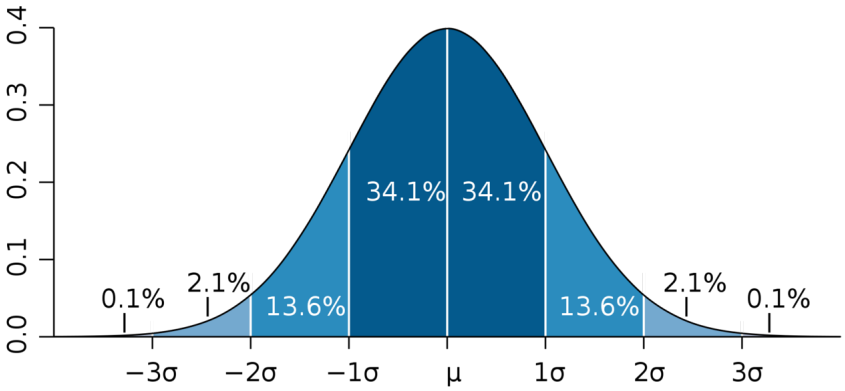
Tabela 2.2: Técnicas de identificação de erros grosseiros e o número de exemplos usados para a testar as mesmas. Continuação

Autor	Teste para Detecção/Identificação de Erros Grosseiros	Número de Exemplos Testados
Mei et al. (2006)	NT-MT para DEG	2
Alhaj-Dibo et al. (2008)	RD usando distribuições robustas	1
Maronna e Arcas (2009)	RD e DEG baseado em Modelos de Regressão	2
Zhang et al. (2010)	Estimadores Robustos	4
Farias (2009)	Teste de Medida	2
	Iterativo (IMT) Robusto	2

Embora as técnicas de RD com o estimador de MQP sejam importantes para ajustar balanços de massa, esta, por si só, não é capaz de realizar a detecção de erro grosseiro (DEG) e deve ser usada em conjunto com outros métodos. As técnicas de detecção existentes são baseadas em testes estatísticos aplicados aos resultados do problema de reconciliação de dados. Tais testes são necessários porque existe uma componente aleatória no valor medido, relativo à aleatoriedade do sistema (instrumentação e processo), a qual se deseja, dentro de um intervalo de confiança, separar do erro grosseiro. O teste mais amplamente usado para DEG é o teste de hipóteses. Para a detecção de erros grosseiros, são formuladas duas hipóteses: H_0 e H_1. A hipótese H_0 considera que não há presença de erros grosseiros, ou, existe a presença somente de erros aleatórios. Como o teste de hipótese usa uma significância estatística, definida pelo indivíduo que aplicará o teste, existe a possibilidade que, em um caso real onde existam somente erros aleatórios, ocorra a falsa detecção de um ou mais erros grosseiros. Tal comportamento pode ser melhor entendido analisando a Figura 2.5 onde é apresentado a Função de Distribuição de Probabilidade Normal com média μ e desvio padrão σ. Supondo uma medição de processo que siga essa distribuição, embora 99,8% das medições estejam entre o intervalo de -3σ e +3σ

existe a probabilidade de 0,2% das medições estarem fora desses intervalos. Se for usado, por exemplo, a faixa de intervalo -3σ e +3σ para classificar uma medição como puramente aleatória, em um período muito longo de amostragem de medidas, eventualmente ocorrerão dados que estarão fora dessa faixa, acusando um falso alarme. Tal adversidade é também conhecida como falha de teste de hipótese do 'Tipo I', quando a hipótese H_0 é verdadeira, ou seja, não existem erros grosseiros, mas os mesmos são detectados, acionando um falso alarme.

Figura 2.5: Distribuição Normal com média μ e desvio padrão σ.



Fonte: Wikipedia. Autor: Petter Strandmark.

A segunda hipótese do teste é denominada H_1 e parte do princípio que existe um ou mais erros grosseiros. Também existe a possibilidade de declarar que não existe a presença de erro, quando na verdade há, denominado de falha do teste estatístico do *Tipo II*, quando a hipótese H_1 é falsa. As características de cada teste para detecção de erros grosseiros associadas aos testes de hipóteses estatísticos são apresentadas nas próximas seções e é importante ressaltar que todas as técnicas exigem que os erros de medição sigam a distribuição de probabilidade Normal, com desvio padrão conhecido.

2.5.1 Teste global

O Teste Global (MADRON, 1985) apud (NARASIMHAN; JORDACHE, 2000), ou *Global Test* (GT), usa o resultado dos resíduos dos balanços, ponderados pela

matriz de variância-covariância dos balanços e é realizado por uma sequência de cálculos. Inicialmente, o vetor de resíduos dos balanços, $\mathbf{R}$ é calculado conforme:

$$\mathbf{R} = \mathbf{A} \cdot \mathbf{Y}, \tag{2.10}$$

em que $\mathbf{A}$ é a matriz Jacobiana do balanço de massa do processo (constante quando é considerada a vazão total) e $\mathbf{Y}$ é o vetor das vazões medidas. Em seguida, a matriz variância-covariância dos resíduos do balanço de massa, $\mathbf{V}$, é assim calculada:

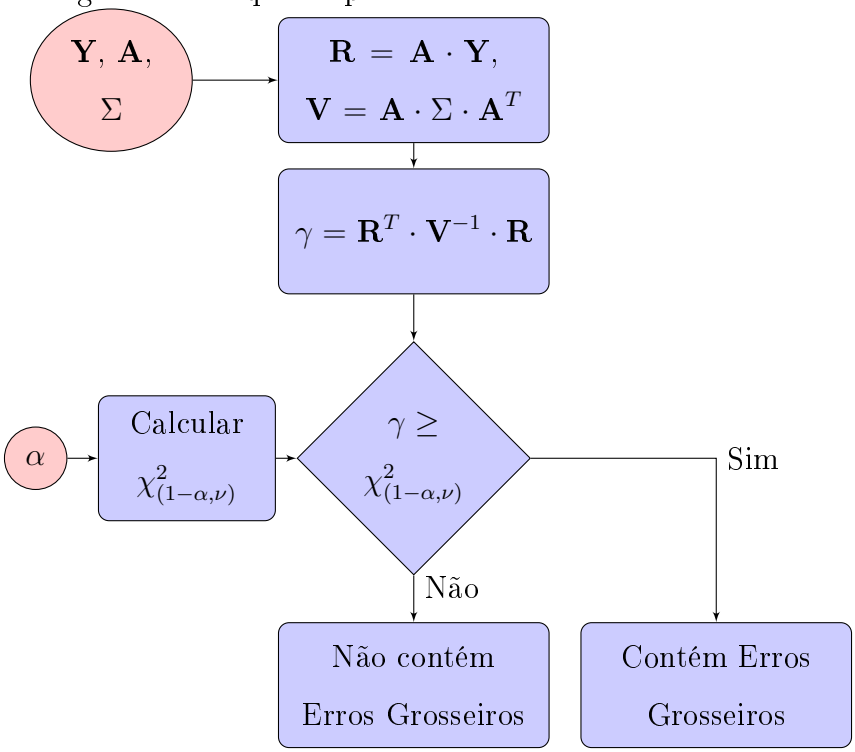
$$\mathbf{V} = cov(\mathbf{R}) = \mathbf{A} \cdot \Sigma \cdot \mathbf{A}^T, \tag{2.11}$$

em que Σ é a matriz variância-covariância das medidas. O valor da estatística de teste, γ, é calculado da seguinte maneira:

$$\gamma = \mathbf{R}^T \cdot \mathbf{V}^{-1} \cdot \mathbf{R}. \tag{2.12}$$

Como γ considera simultaneamente as interações de todo o balanço, trata-se de uma estatística multivariável que segue uma distribuição χ^2 com ν graus de liberdade, onde ν é o *rank* da matriz Jacobiana do balanço, $\mathbf{A}$. Para aceitar ou rejeitar a hipótese H_0, seleciona-se uma significância estatística, α, e, utilizando a informações da distribuição χ^2 com ν graus de liberdade através de dados tabelados ou da avaliação da própria função, encontra-se o valor crítico do teste, ou critério do teste, cuja magnitude deve ser comparado com γ. Se γ é maior que o valor crítico do teste, $\chi^2_{(1-\alpha,\nu)}$, o balanço como um todo, dentro da significância estatística escolhida, é classificado como contendo um ou mais erros grosseiros, conforme apresentado na Figura 2.6. Se γ é menor que o valor crítico do teste, o balanço como um todo, é classificado como contendo erros de medição puramente aleatórios. Também é possível provar que γ é o valor final da função objetivo de mínimos quadrados ponderados da reconciliação de dados (NARASIMHAN; JORDACHE, 2000). O teste global é capaz de detectar a presença de erro no fluxograma de processo como um todo, porém, somente com seu resultado não é possível apontar a localização do erro, o tipo (vazamento ou falha de medição) nem mesmo se estão presentes um ou mais erros de medição. Geralmente, o resultado do Teste Global é usado em conjunto com outras técnicas para ajudar a localização e identificação dos erros grosseiros.

Figura 2.6: Esquema para DEG usando o Teste Global.



Fonte: Próprio Autor

2.5.2 Teste nodal

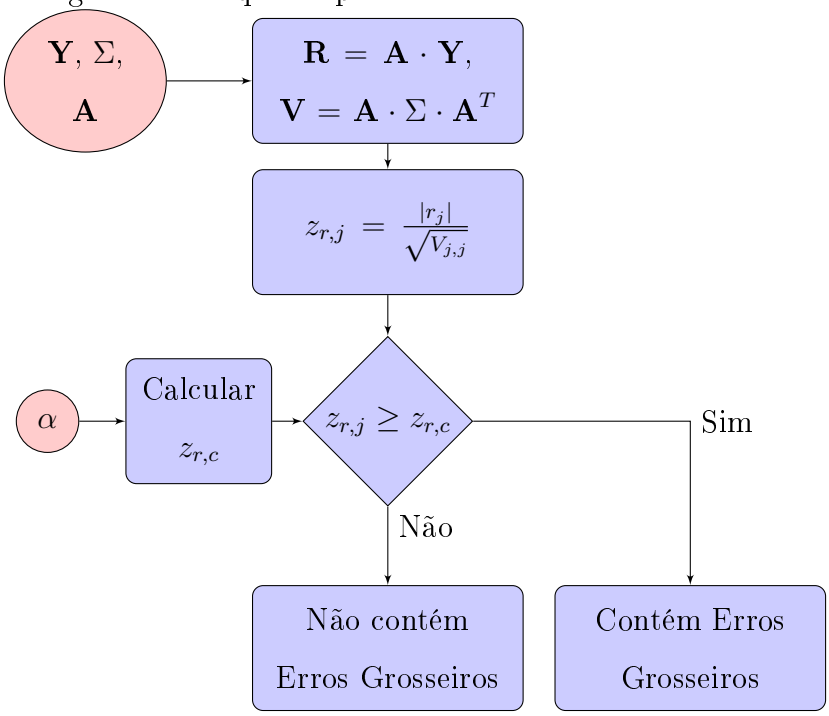
O Teste Nodal (MAH et al., 1976), ou *Nodal Test* (NT), é um teste realizado nos resíduos dos balanços individualmente. Para cada nó, ou equipamento, a variável $z_{r,j}$ é calculada conforme:

$$z_{r,j} = \frac{\mid r_j \mid}{\sqrt{V_{j,j}}}, \tag{2.13}$$

em que r_j é o resíduo do balanço j e $V_{j,j}$ é o elemento j,j, ou seja, da diagonal, da matriz $\mathbf{V}$. O teste de hipótese monovariável z é aplicado considerando normalidade de medidas, baseado no valor da variável $z_{r,j}$. Para aceitar ou rejeitar a hipótese H_0, seleciona-se uma significância estatística, α, e, utilizando a informações da distribuição z, através de dados tabelados ou da avaliação da própria função, encontra-se o valor crítico do teste, ou critério do teste, cuja magnitude deve ser comparado com $z_{r,j}$. Se $z_{r,j}$ é maior que o valor crítico do teste, $z_{r,c}$, o equipamento, dentro da significância estatística escolhida, é classificado como contendo vazamento. Se $z_{r,j}$ é menor que o valor crítico do teste, o equipamento é classificado como não contendo

vazamento, conforme Figura 2.5.2 . É possível detectar mais de um vazamento aplicando o mesmo teste para outros equipamentos simultaneamente.

Figura 2.7: Esquema para DEG usando o Teste Nodal.



Fonte: Próprio Autor

2.5.3 Teste de medida

O teste da medida proposto por Mah e Tamhane (1982) é baseado no ajuste da medida após a reconciliação, sendo o ajuste definido conforme:

$$ad_i = y_i - x_i \tag{2.14}$$

ou na forma vetorial por:

$$\mathbf{Ad} = \Sigma \cdot \mathbf{A}^T \cdot \mathbf{V}^{-1} \cdot \mathbf{R}, \tag{2.15}$$

em que x_i é o valor reconciliado da medida e y_i o valor medido. Inicialmente é construída a matriz de covariância dos ajustes $\overline{\mathbf{W}}$, conforme:

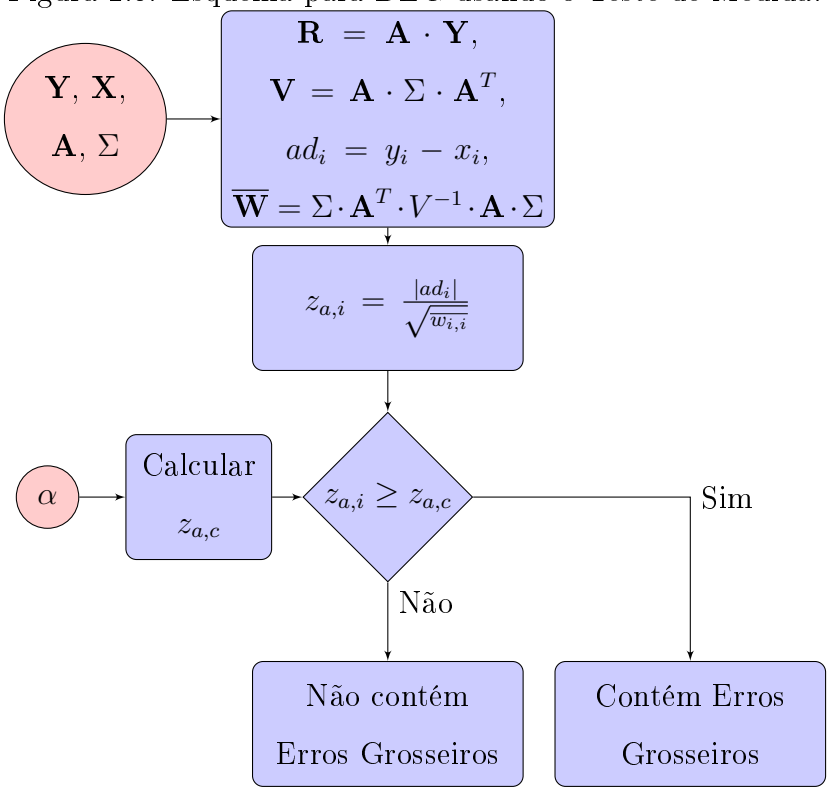
$$\overline{\mathbf{W}} = \Sigma \cdot \mathbf{A}^T \cdot V^{-1} \cdot \mathbf{A} \cdot \Sigma, \tag{2.16}$$

em seguida, é calculado o termo $z_{a,i}$, para a estatística de teste, conforme :

$$z_{a,i} = \frac{\mid ad_i \mid}{\sqrt{\overline{w_{i,i}}}}, \tag{2.17}$$

em que $\overline{w_{i,i}}$ é o elemento i,i, ou seja da diagonal, da matriz $\overline{\mathbf{W}}$. O teste de hipótese monovariável z é aplicado considerando a normalidade das medidas, baseado no valor da variável $z_{a,i}$ e é capaz de detectar falha em uma medição específica. Para aceitar ou rejeitar a hipótese H_0, seleciona-se uma significância estatística, α, e, utilizando a informações da distribuição z, através de dados tabelados ou da avaliação da própria função, encontra-se o valor crítico do teste, ou critério do teste, cuja magnitude deve ser comparado com $z_{a,i}$. Se $z_{a,i}$ é maior que o valor crítico do teste, $z_{a,c}$, a medição, dentro da significância estatística escolhida, é classificada como contendo erro grosseiro. Se $z_{a,i}$ é menor que o valor crítico do teste, a medição é classificada como contendo erros puramente aleatórios, conforme Figura 2.5.3. Caso existam falhas em outras medições, é possível detectá-las aplicando o teste em outras medidas do processo simultaneamente.

Figura 2.8: Esquema para DEG usando o Teste de Medida.



Fonte: Próprio Autor

2.5.4 Teste de razão de verossimilhança generalizada

O Teste da Razão de Verossimilhança Generalizada(NARASIMHAN; MAH, 1987), ou *Generalized Likelihood Ratio Test* (GLR), é formulado baseado no modelo de processo com erros grosseiros. O modelo de erro de medição de magnitude b é definido por:

$$\mathbf{Y} = \overline{\mathbf{X}}_{NOM} + \epsilon_r + b\mathbf{e}_i, \tag{2.18}$$

em que $\overline{\mathbf{X}}_{NOM}$ é o vetor das vazões nominais do processo, ϵ_r é o vetor de erro aleatório e $\mathbf{e}_i$ é um vetor com valor 1 na posição do erro i, e zero nas demais posições e b é a magnitude do erro de medição. Já um vazamento no equipamento j de magnitude desconhecida b é modelado conforme:

$$\mathbf{A} \cdot \mathbf{Y} - b\mathbf{m}_j = 0, \tag{2.19}$$

em que $\mathbf{m}_j$ é um vetor com valor 1 na posição do vazamento e zero nas demais posições e b é a magnitude do vazamento. Pode ser provado (NARASIMHAN; MAH, 1987) que quando um erro de medição ou vazamento está presente, tem-se que:

$$E\left[\mathbf{R}\right] = b.\mathbf{f} \tag{2.20}$$

em que $E\left[\mathbf{R}\right]$ é a função de valor esperado da variável $\mathbf{R}$, $\mathbf{f} = A.\mathbf{e}_i$ para erros grosseiros de medição no instrumento i e $\mathbf{f} = \mathbf{m}_j$ para vazamento no equipamento j. Os vetores $\mathbf{f}$ também são conhecidos como assinatura do erro. Baseado no teste da razão de verossimilhança entre as hipóteses:

H_0 não possui erros grosseiros e

H_1 possui erros grosseiros com magnitude $b.\mathbf{f}$, é formulado um teste estatístico do tipo χ^2 para cada erro (medição ou balanço). O teste da Razão de Verossimilhança Generalizada é um dos mais completos pois, além de localizar e identificar o tipo de erro, pode estimar sua magnitude (cálculo apresentado em Narasimhan e Mah (1987)). O GLR, através de procedimentos iterativos, também pode identificar erros múltiplos, técnica denominada de Estratégia de Compensação Serial (NARASIMHAN; MAH, 1987).

2.5.5 Teste dos componente principais para equipamentos, medidas e global

Ao contrário dos testes anteriormente apresentados que usam informações somente das diagonais das matrizes $\mathbf{V}$ e $\overline{\mathbf{W}}$, os testes dos componentes principais usam informações destas matrizes como um todo, realizando transformações lineares dos resíduos dos balanços ou das medidas ou na estatística do teste global (γ) para a realização dos testes estatísticos (TONG; CROWE, 1995). O teste de componentes principais dos resíduos dos balanços é feito, inicialmente, realizando a transformação de variáveis nos resíduos dos balanços, conforme:

$$\mathbf{p}_r = \mathbf{W}_r^T \cdot \mathbf{R}, \tag{2.21}$$

em que $\mathbf{p}_r$ é o vetor de componentes principais dos resíduos das restrições, $\mathbf{W}_r$ são os autovetores de $\mathbf{V}$. Após o cálculo de $\mathbf{p}_r$, o teste de hipótese monovariável z é aplicado baseado no valor da variável $\mathbf{p}_r$, considerando a normalidade das medidas. O teste de componentes principais dos resíduos é capaz de detectar vazamentos em equipamentos. Caso existam falhas em outros equipamentos, é possível detectá-las aplicando o teste em outros equipamentos.

O teste de componente principais das medidas é realizado através da estatística de teste $\mathbf{p}_{ai}$, calculado conforme:

$$\mathbf{p}_{ai} = \left(W_a^T \cdot \mathbf{Ad}\right)_i, \tag{2.22}$$

em que $\mathbf{W}_a$ são os autovetores de $\overline{\mathbf{W}}$ e $\mathbf{Ad}$ é definido pela Equação (2.15). Após o cálculo de $\mathbf{p}_{ai}$, o teste de hipótese monovariável z é aplicado baseado no valor da variável $\mathbf{p}_{ai}$, considerando a normalidade das medidas. Tal teste é capaz de detectar falhas em uma medição específica. Caso existam falhas em outras medições, é possível detectá-las aplicando o teste em outras medidas do processo.

O teste global baseado em componentes principais, denominado de teste global coletivo (baseado em componentes principais) é realizado, inicialmente, calculando a variável γ_k, conforme:

$$\gamma_k = \mathbf{p}_r^T \cdot \mathbf{p}_r. \tag{2.23}$$

O teste de hipótese monovariável z é aplicado baseado no valor da variável γ_k, considerando a normalidade das medidas. Assim como o teste global tradicional, o

teste global dos componentes principais é capaz de detectar a presença de erro no fluxograma de processo como um todo, porém, somente com seu resultado, não é possível apontar a localização do erro, o tipo, nem mesmo se estão presentes um ou mais erros de medição.

2.5.6 Identificação e localização de erros grosseiros múltiplos usando estratégias simultâneas

Em uma planta industrial existe a possibilidade de falhas simultâneas dos sensores de medição. A maioria das técnicas de detecção de múltiplos erros de medição faz uso das técnicas de detecção de erros de medição individuais e serão apresentadas nas próximas seções.

2.5.6.1 Estratégias usando erros grosseiros individuais

Estratégias simultâneas buscam encontrar erros em medições de uma forma não-iterativa, ou seja, em uma única etapa. Tais técnicas se baseiam no princípio de que os erros grosseiros são detectados simultaneamente em todas as medições ou balanços cuja estatística de teste ultrapassam o valor limite. Algumas das estratégias previamente apresentadas podem ser consideradas como estratégias simultâneas, como, por exemplo, o Teste das Medidas e da Razão de Verossimilhança Generalizada. Serth e Heenan (1986) e Iordache et al. (1985) mostraram resultados em que o Teste das Medidas e da Razão de Verossimilhança Generalizada tendem a apresentar resultados errôneos, classificando equivocadamente correntes sem erros grosseiros como tal. Tal fato ocorre principalmente porque as medições estão relacionadas pelas equações dos balanços, sendo assim um erro grosseiro em uma medição pode se *espalhar* por uma medição vizinha e ser detectado erroneamente nesta. A forma desse *espalhamento* depende de vários fatores, como o nível de redundância, a diferença entre as incertezas de medição e a magnitude do erro grosseiro.

2.5.6.2 Estratégias usando hipóteses combinatoriais

De forma semelhante ao Teste da Razão de Verossimilhança Generalizada, a estratégia de Hipóteses Combinatoriais é formulada de forma a detectar todas as possibilidades de erros grosseiros levando em consideração a presença de dois ou mais erros grosseiros(NARASIMHAN; JORDACHE, 2000; ROSENBERG et al., 1987). Estendendo o Teste da Razão de Verossimilhança Generalizada, a estratégia de Hipóteses Combinatoriais monta uma matriz com as possíveis combinações de erros aplicando uma estatística de teste do tipo χ^2, porém, com k graus de liberdade, em que k é o número de medições presentes, considerando distribuição Normal.
A estratégia de Hipóteses Combinatoriais tem como desvantagem o tempo computacional que tende a ser elevado para balanços com muitas correntes e equipamentos, além disso, não é possível detectar e identificar combinações de erros para k menor que o número de equipamentos.

2.5.6.3 Identificação usando estimação simultânea das magnitudes dos erros grosseiros

Com o objetivo de contornar a limitação da estratégia de Hipóteses Combinatoriais de não detectar e identificar combinações de erros para k menor que o número de equipamentos, Rollins e Davis (1992) formularam um procedimento simultâneo denominado técnica de estimação não tendenciosa (*unbiased estimated technique*, UBET). Nesse procedimento, assume-se que existe uma matriz $F_{n,m}$, onde as colunas são as denominadas *assinatura de erros* (semelhante ao método da Razão da Verossimilhança Generalizada) e m é o número de balanços, sendo a matriz com dimensões n,m com todas as linhas e colunas linearmente independentes. Na UBET, é realizado um teste de hipóteses para avaliar se o erro previamente assumido na *assinatura de erros* possui magnitude zero ou não, aplicando um teste z. Os candidatos iniciais aos erros podem ser obtidos através do Teste da Razão de Verossimilhança Generalizada, Teste Nodal ou testes dos componentes principais.

Sánchez et al. (1999) revisaram o procedimento de estimação simultânea de erros grosseiros (SEGE) e propuseram uma nova técnica para contornar as falhas da metodologia de estimação simultânea, denominando-a de procedimento de estimação

simultânea de erros grosseiros modificada (MSEGE), aplicando-a a um exemplo de literatura onde é observada uma superior melhora em relação aos demais métodos. A nova técnica proposta se aplica ao caso onde existem correntes não medidas e também trata de casos onde um conjunto de erros grosseiros é equivalente a outro conjunto (quando conjuntos de erros grosseiros apresentam o mesmo valor da função objetivo, tornando os dois conjuntos indistinguíveis entre si).

2.5.7 Identificação e localização de erros de grosseiros múltiplos usando estratégias em série

As estratégias em série, em contraste com as simultâneas, identificam os erros grosseiros serialmente, sendo que existem diversas combinações de testes estatísticos e de estratégias sequenciais.

2.5.7.1 Teste de medida iterativo

No algoritmo de Teste de Medida Iterativo, IMT, é aplicado uma série de Testes de Medidas, sendo que, as que apresentam as maiores estatísticas de testes são eliminadas do problema subsequente de reconciliação. A eliminação busca as medições com maior estatística de teste: possíveis candidatas a erro. O procedimento de eliminação de correntes é equivalente à fusão de nós ou equipamentos das correntes que os conectam. O teste em questão pode ser aplicado com redundância de todas as medidas, ou somente com algumas medidas, além disso, também é possível aplicar o Teste Global ao invés do Teste de Medidas. Considerando ainda a eliminação em série de medições, Serth e Heenan (1986) propuseram um modificação do método anterior onde informações sobre os valores inferiores e superiores das vazões são alimentadas no problema, denominado Testes Iterativo das Medidas Modificados. Rosenberg et al. (1987) estenderam o trabalho de Serth e Heenan (1986) acrescentando limites inferiores e superiores não só às variáveis medidas, mas também às variáveis não-medidas, ou seja, estimadas como resultado do problema. Cabe ressaltar que os testes anteriormente descritos são exclusivos para detectar erros em medições. Farias (2009) apresentou a versão robusta para o IMT, denominada r-

IMT, baseado no uso de estimadores robustos na função objetivo, sendo que tal teste apresentou resultados superiores ao IMT.

2.5.7.2 Estratégia da compensação serial simples

O algoritmo da Estratégia da Compensação Serial Simples, inicialmente proposto por Narasimhan e Mah (1987), é usado tanto para detectar erros de medida quanto vazamentos em equipamentos. O procedimento usa iterativamente o Teste de Razão de Verossimilhança Generalizada e baseia-se no princípio de compensar o vazamento ou desvio da medição com a respectiva magnitude do erro grosseiro estimado, seguido da reaplicação das estatísticas de teste. A repetição do teste é realizada até que a estatística de teste máximo não exceda o critério de teste. Rollins e Davis (1992) demonstraram através de estudos de simulação que a técnica proposta pode apresentar diversos erros de predição, tanto quanto ao tipo e localização quanto na estimativa da magnitude do erro, problemas que tentaram ser resolvidos pela técnica de Compensação Serial Modificada (KELLER et al., 1994). De modo semelhante aos Testes de Medida Iterativo, onde se evoluiu para a adição de restrições de vazão superior e inferior, Harikumar e Narasimhan (1993) também apresentaram a versão do Teste de Razão de Verossimilhança Generalizada com informações de restrições de limites superiores e inferiores das medições.

2.5.7.3 Estratégias de combinação de nodos

As técnicas que usam Estratégias de Combinação de Nodos são usadas, principalmente, para detectar desvios em medições. Inicialmente proposto por Mah et al. (1976), tal técnica se baseia no fato de que, se uma corrente apresenta erro grosseiro, esta afetará os nodos, ou equipamentos nos quais a mesma está conectada. Para eliminar o efeito dessa medição com erro grosseiro, a adição dos dois nodos que conectam a corrente com erro de medição, tende a eliminar o efeito deste erro no balanço global. Para explorar esse princípio, as Estratégias de Combinação de Nodos, aplicam sistematicamente a união de um ou mais equipamentos conectados, indicando correntes candidatas a problemas e possíveis isentas de problemas. Outras técnicas podem ser aplicadas subsequentemente para confirmar as suposições. Rol-

lins et al. (1996) sistematizou tal técnica criando a Técnica da Combinação Linear (*Linear Combination Technique* LCT).

2.5.8 Detectabilidade e identificabilidade de erros grosseiros

Um requisito essencial para a detecção de erros grosseiros é a presença de redundância na medição. Na prática, algumas medidas redundantes podem apresentar comportamento de variáveis não redundantes, denominadas praticamente não-redundantes. Tal caso foi apresentado por Crowe (1988), que reportou dificuldades em reconciliar e detectar erros grosseiros em medidas presentes com características específicas. Madron (1992), apud Narasimhan e Jordache (2000), apresentou o conceito de ajustabilidade e Charpentier et al. (1991), apud Narasimhan e Jordache (2000), o conceito de detectabilidade para classificar esses tipos de medidas redundantes. Iordache et al. (1985), através de simulações, observou que em alguns tipos de balanços, as medidas redundantes também se comportam como praticamente não-redundantes, são eles:

- Variáveis com pequena incerteza-padrão em um mesmo balanço com variáveis com grande incertezas-padrão.
- Variáveis com grandes diferenças em ordem de grandeza envolvidas em um mesmo balanço.
- Correntes em paralelo, como, por exemplo, correntes de saída de divisores de correntes.
- Medidas que aparecem somente em uma equação com uma corrente não medida sem restrição de outro balanço ou equação, como por exemplo, fontes e sumidouros.

A detectabilidade é calculada pela seguinte expressão:

$$\text{detect}_i = \sqrt{\left(1 - \frac{u_{x_i}^2}{u_i^2}\right)} \tag{2.24}$$

em que u_i é a incerteza-padrão da medição e u_{x_i} é a incerteza-padrão das variáveis reconciliadas, este último pode ser calculado por (HEYEN et al., 1996)

$$u_{x_i} = \sqrt{\sum_{j=1}^{mc} \frac{\left(M_{i,j}^{-1}\right)^2}{u_j^2}}, \tag{2.25}$$

em que mc é o número de correntes medidas e M é a matriz:

$$\mathbf{M} = \begin{pmatrix} \mathbf{W} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{pmatrix} \tag{2.26}$$

2.5.9 Medidas de performance das técnicas de detecção de erros grosseiros

Para avaliar a performance dos testes de detecção de erros grosseiros, a literatura reporta o uso de experimento numéricos do tipo simulações de Monte Carlo com a localização e magnitude dos erros grosseiros conhecidos (SÁNCHEZ et al., 1999; ÖZYURT; PIKE, 2004; FARIAS, 2009; ROSENBERG et al., 1987; KONGSJAHJU et al., 2000; NARASIMHAN; JORDACHE, 2000). Nesses experimentos numéricos, erros grosseiros de magnitudes conhecidas são sistematicamente adicionados em posições do balanço conhecidas. São realizadas quantidades elevadas de simulações e os resultados são avaliados com algumas métricas. Diversas métricas estão presentes na literatura (NARASIMHAN; JORDACHE, 2000) para avaliar a performance dos testes de detecção de erros de grosseiros. As mais importantes são as seguintes:

- Poder Total(OP) = Número de Erros Grosseiros Corretamente Identificados / Número de Erros Grosseiros Simulados;
- Média do Valor dos Erros do Tipo I (AVT1) = Número de Erros Erroneamente Identificados / Número de Simulações;
- Seletividade = Número de Erros Grosseiros Corretamente Identificados / Número de Erros Grosseiros Detectados;
- Erro Médio de Estimação (AEE) $= \frac{1}{NS} \sum_{i=1}^{NS} \left| \frac{\text{valor estimado}_i - \text{valor real}_i}{\text{valor real}_i} \right|$;

em que NS é o número de simulações de Monte Carlo realizadas.

O Poder Total é uma das principais métricas para avaliação da performance de métodos para DEG, seu valor máximo é 1, quando todos os erros grosseiros foram corretamente identificados nas simulações. Também é usual elaborar uma curva de OP, onde a magnitude do erro grosseiro é variada dentro de um intervalo e o OP do método de DEG é avaliado.

O AVT1 está associado com Erros do Tipo I do teste de hipóteses, ou falso positivo. É importante ressaltar que todo o método de detecção apresentará uma fração de falso positivo, de acordo com a significância estatística usada, mesmo em simulações numéricas. Cabe ressaltar, que cada método de DEG pode usar uma significância estatística diferente, porém, quando o AVT1 é mantido no mesmo valor, diferentes métodos de DEG podem ser comparados na mesma base, independente da significância estatística configurada. Em simulações do tipo Monte Carlo para avaliação da performance de métodos de DEG, a significância estatística do teste de hipótese é configurada realizada por tentativa e erro para que o AVT1 seja aproximadamente 0,1 ou 10%.

A Seletividade avalia se o erro grosseiro adicionado em uma determinada medição ou equipamento em uma simulação foi apropriadamente identificado no mesmo local adicionado. Em casos de erros grosseiros múltiplos, a seletividade é capaz de indicar se os erros grosseiros foram identificados nos locais adicionados ou não.

O Erro Médio de Estimação não é aplicável a qualquer método de DEG, pois, somente alguns métodos podem estimar a magnitude do erro grosseiro adicionado, entre eles o GLR, a técnica de estimação simultânea de erros grosseiros (SEGE), e suas versões modificadas. Cabe ressaltar que essa métrica só é avaliada em casos onde o erro grosseiro foi corretamente identificado. Também é necessário manter o registro da magnitude do erro grosseiro original adicionado para tal comparação.

2.5.9.1 Ferramentas de mineração de dados

As ferramentas de mineração de dados tem como objetivo classificar e segmentar conjuntos de dados para a obtenção de informações úteis nos mesmos. Junior et

al. (2008) apresentaram uma metodologia de classificadores hierárquicos e teoria dos protótipos para classificar o tipo de erro grosseiro, localização e magnitude do mesmo, em conjunto com ferramentas de reconciliação de dados. Prata et al. (2008) apresentaram o uso de análise de *clusters* para a reconciliação e detecção de dados espúrios de um processo dinâmico e não-linear, detecção de erros grosseiros, aplicando o mesmo para um reator de polimerização industrial.

2.5.10 Métricas para detecção de erros de medição

As técnicas para detecção de erros grosseiros anteriormente mencionadas, usam suas respectivas métricas nas estatísticas de testes, sendo as mais importantes abaixo listadas:

- Teste Global;
- Teste Global Coletivo;
- Teste de Medida;
- Teste Nodal;
- Teste da Razão de Verossimilhança Generalizada;
- Teste de componentes Principais das Restrições ou componentes Principais das Medidas.

Através da revisão bibliográfica, foi possível observar que as metodologias de reconciliação de dados e detecção de erros grosseiros usam uma, ou, no máximo, duas métricas em seus testes. Sendo essas métricas relativamente simples de serem calculadas com o uso de pacotes numéricos de álgebra linear, o presente trabalho tem como um dos objetivos avaliar o uso simultâneo de algumas dessas métricas na avaliação de balanços hídricos industriais. determinadas variáveis.

2.6 Contexto do trabalho

Conforme destacado no início do capítulo, na prática, a avaliação de balanços hídricos é uma tarefa bastante trabalhosa, devido à falta de dados e informações, sendo que a avaliação dos problemas nos balanços deve ser feita com informações das equipes de operações e de engenharia da empresa. Nesse contexto, o trabalho desenvolvido busca fundir as informações das diversas métricas existentes na literatura para desenvolver um algoritmo para avaliação e monitoramento de balanços sujeitos a pouca quantidade ou baixa qualidade de informações.

Capítulo 3

Materiais e Métodos

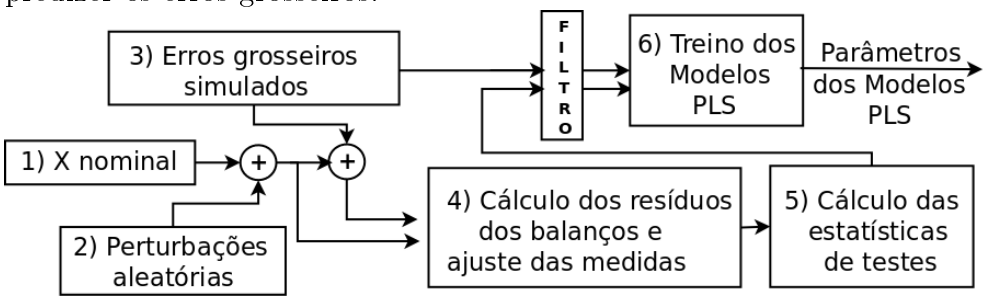
3.1 Visão geral do método proposto

Para o aumento da confiança dos balanços de massa em equipamentos envolvendo balanços hídricos, é necessário o uso de ferramentas de reconciliação de dados em conjunto com técnicas de detecção de erros grosseiros para identificar eventuais falhas em sistemas de medição ou vazamentos.

Além da falta de redundância e da observabilidade nos sistemas industriais, a grande dificuldade para detecção de erros grosseiros é a possibilidade dos mesmos estarem presentes em pontos múltiplos nos sistemas em estudo. Durante a etapa de revisão bibliográfica, foi possível perceber que existe uma lacuna para detectar corretamente erros grosseiros únicos e múltiplos de forma adequada para balanços mássicos de forma genérica, independente da sua topologia, e é sobre esse aspecto que foi proposto um método que funcione em ambientes industriais. Uma das dificuldades que surge dos métodos citados no Capítulo 2 é a característica combinatorial da detecção de erros grosseiros múltiplos, sendo assim, sua resolução fica muito complexa à medida que aumentam o número de correntes e de equipamentos. Também observa-se a ocorrência de falsos positivos para alguns métodos de detecção e, para determinadas topologias, existem falhas na detecção mesmo para erros de medição únicos.

Para contornar o problema, foi desenvolvida a metodologia apresentada na Figura 3.1.

Figura 3.1: Esquema simplificado da metodologia proposta visando gerar modelos para predizer os erros grosseiros.



Fonte: Próprio autor

A metodologia proposta se resume em estimar parâmetros de um modelo empírico multivariável capaz de identificar a existência de um erro grosseiro, sua posição e magnitude, representado pela etapa 6 da Figura 3.1. Inicialmente é necessário gerar dados através de simulações com erros puramente aleatórios e com erros grosseiros em diversas posições, que serão as variáveis dependentes do modelo empírico, as quais se deseja prever, apresentados nas etapas 1 a 3 da Figura 3.1. Em seguida são calculados índices de qualidade dos balanços, variáveis independentes, correspondente ao respectivo erro, aleatório ou grosseiro, apresentadas nas etapas 4 e 5 da Figura 3.1. Em seguida, são aplicados filtros que realizam uma pré-seleção de dados, sendo, em seguida, montada uma tabela, conforme Tabela 3.1. Finalmente, os parâmetros do modelo empírico são estimados, considerando como variáveis independentes os índices de qualidade e como variáveis dependentes a magnitude dos erros.

Tabela 3.1: Exemplo de tabela de dados usada para estimar os parâmetros de modelos multivariáveis usado pelo método proposto.

Entradas					Saídas
Valor do Índice 1	Valor do Índice 2	Valor do Índice 3	Valor do Índice 4	Valor do Índice N	Valor Nominal do Erro i
$I1$	$I2$	$I3$	$I4$	$I5$	$Im1$
$I1_n$	$I2_n$	$I3_n$	$I4_n$	$I5_n$	$Im1_n$

De forma mais detalhada, o método é composto das seguintes etapas:

1 A partir de um fluxograma de processo conhecido, são selecionadas, ou arbitradas estimativas preliminares para as vazões (valores nominais), onde os balanços de massa são todos satisfeitos, denominado $\overline{x}_{NOM,i,ns}$.

2 Erros aleatórios, $\epsilon_{i,ns}$, são gerados e somados à $\overline{x}_{NOM,i,ns}$ e denominados $y_{i,ns}$, ou seja $y_{i,ns} = \overline{x}_{NOM,i,ns} + \epsilon_{i,ns}$.

3 Erros grosseiros, $\delta_{i,ns}$, são adicionados ao termo $y_{i,ns}$.

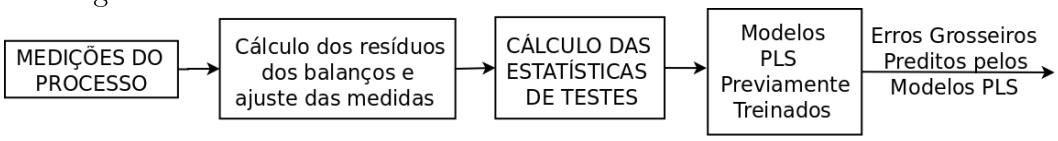
4 Baseado nos valores de $y_{i,ns}$ e $\delta_{i,ns}$, são calculados os resíduos dos balanços de massa e os ajustes das medidas usando um problema de reconciliação de dados modificado.

5 Índices de qualidade dos balanços baseados nas estatísticas de testes dos testes Global (γ), Nodal ($z_{r,j}$) e de Medida ($z_{a,i}$), são calculados a partir dos resíduos dos balanços de massa e dos ajustes das medidas.

6 Os dados calculados, γ, $z_{r,j}$, $z_{a,i}$, $y_{i,ns}$ e $\delta_{i,ns}$, variáveis de dependentes e independentes, são usados para estimar parâmetros de modelos multivariáveis.

Os modelos obtidos pela metodologia tem como objetivo prever a posição e magnitude dos erros grosseiros puros, $\delta_{i,ns}$, a partir de medições de processo, conforme Figura 3.2.

Figura 3.2: Esquema simplificado da aplicação da metodologia proposta com os modelos gerados.



Fonte: Próprio autor

Para consolidar a metodologia de avaliação de balanços de massa e DEG apresentada na Figura 3.1 foram desenvolvidas as seguintes etapas:

1 Escolha do método de otimização.

2 Seleção e formulação de estudos de caso.

3 Seleção de índices de qualidade do balanço para o aprendizado.

4 Seleção de sistemas de aprendizado.

5 Escolha da função objetivo.

6 Escolha das incertezas.

7 Geração e seleção de dados para treino dos modelos PLS.

8 Seleção de dados para ajuste e teste.

9 Ajuste do modelo PLS.

Detalhes de cada uma das etapas estão descritas nas demais seções deste capítulo.

3.2 Escolha do método de otimização

Quando se deseja realizar a reconciliação de dados somente com vazões totais e com a função objetivo de Mínimos Quadrados Ponderados, o problema de otimização resultante possui solução analítica mesmo na presença de variáveis não-medidas (NARASIMHAN; JORDACHE, 2000; ROMAGNOLI; SÁNCHEZ, 1999; CROWE et al., 1983; KELLY, 1999; MITSAS, 2010; SWARTZ, 1989; VEVERKA; MADRON, 1997). Como na metodologia proposta deseja-se explorar outras funções objetivo (estimadores robustos) se faz necessário o uso de um otimizador externo. O otimizador escolhido foi o Ipopt (WÄCHTER; BIEGLER, 2006), capaz de tratar de problemas de Programação Não-Linear (NLP) genéricos com funções objetivos e restrições não-lineares. O Ipopt utiliza um algoritmo de ponto interior primal-dual com filtro para busca em linha (*line-search*) e também é adequado para problemas de grande escala (*large-scale*, ou com dezenas de milhares variáveis e restrições), embora problemas com esse número de variáveis e restrições dificilmente são encontrados na área de RD e DEG e não foram selecionados nesse trabalho. O Ipopt possui interface de comunicação com diversos programas de computador para computação científica, entre eles, o Scilab (INRIA, 2010). O Ipopt usa informações das derivadas primeira (matriz Jacobiana) e segunda (Matriz Hessiana ou sua versão aproximada) para

encontrar o ótimo local do problema de otimização. No trabalho desenvolvido, as informações das matrizes Jacobiana e Hessiana dos estimadores robustos (função objetivo) foram expressas analiticamente, enquanto que a matriz Jacobiana do balanço de massa é constante, porque os balanços de massa total nos equipamentos são funções lineares. Todas essas funções são expressas dentro do Scilab. Cabe ressaltar que outros otimizadores capaz de resolver problemas de programação não-linear também podem ser usados, podendo apresentar resultados superiores em relação ao tempo computacional, mas, devido à facilidade de uso e integração com o *software* Scilab, o Ipopt foi escolhido.

3.3 Seleção e formulação de estudos de caso

A identificação apropriada de erros grosseiros é uma tarefa difícil, pois, os métodos de detecção são influenciados por muitos fatores que podem levar a resultados equivocados. Estudos anteriores elaborados por Rosenberg et al. (1987), Iordache et al. (1985) e Sánchez et al. (1999) indicaram que os algoritmos de detecção e identificação de erros grosseiros únicos e múltiplos podem ser influenciados por fatores como:

1 Magnitude das medições e seus limites;

2 Magnitude da razão entre erro grosseiro e a incerteza da medição;

3 Posição da medição que contém o erro grosseiro;

4 Restrições e estrutura do fluxograma de processo.

Mais especificamente em relação ao fluxograma de processo, Iordache et al. (1985) observou que, em alguns fluxogramas, a redundância necessária para a aplicação de técnicas de DEG pode ser degenerada, comprometendo os resultados das mesmas. As medições para esses fluxogramas são denominadas 'praticamente não-redundantes' e podem estar presentes nos seguintes casos (NARASIMHAN; JORDACHE, 2000; SÁNCHEZ et al., 1999):

- Vazões com incertezas pequenas no mesmo equipamento (ou balanço) que uma vazão com incerteza elevada;
- Correntes paralelas;
- Vazões medidas que aparecem somente em uma equação com uma corrente não-medida;

Embora o tema de RD e DEG datem de mais de 50 anos, ainda existem muitos desafios nessa área (CROWE, 1996; NARASIMHAN; JORDACHE, 2000). A complexidade do tema sugere que os métodos de detecção propostos sejam testados em diversos casos para torná-los mais generalizáveis e robustos. Nesse sentido, o presente trabalho coletou 15 problemas lineares (somente com balanço de vazão total) de literatura, mais um caso industrial, para testar a metodologia aqui desenvolvida. Os problemas foram selecionados para testar as adversidades presentes ao se aplicar técnicas de DEG. As principais características de cada problema selecionado são apresentadas na Tabela 3.2. Os problemas completos são apresentados no Apêndice A em forma de artigo. A Tabela 3.2 é dividida em três colunas, a primeira indica o problema. Na segunda coluna são apresentadas as correntes do processo. Na terceira coluna, são apresentadas as características das respectivas correntes, conforme a seguinte abreviação:

- P: Correntes paralelas;
- R: Reciclo;
- M: Magnitude da incerteza ou medição.

Logo, tomando como exemplo o Problema 5, tem-se as correntes 4 e 5 como paralela e as correntes 2, 6, 7 e 8 como correntes problemáticas por diferença na magnitude de suas incertezas ou de suas medições. As notações com ponto-e-vírgula indicam que mais de um conjunto de correntes possuem a mesma característica, como, por exemplo, o Problema 8 que possui as seguintes correntes em paralelo: 2,3; 6, 7; 8, 9; 10, 11. Correntes entre parênteses indicam um reciclo alternativo, ou seja, com mais de um caminho possível.

Tabela 3.2: Características dos problemas selecionados.

Problema	Correntes	Características
P1	2,3	P
P2	2,3;5,6	P
P3	2,11;9,12;4,5,6,7,8	P
	1,2(11),9(12),3	R
P4		-
P5	4,5	P
	2,6,7,8	M
P6	2,3,4;2,3,5,6	R
P7	5,9;8,10	M
	5,7,8,9;2,3,5,6	R
P8	2,3;6,7;8,9;10,11	P
P9	1,2	P
	8,9,10	R
	3,4,5;5,6,7	M
P10	1,2;8,9,10,13;5,11,12	M
P11	2,3,6,5;2,7,9,8,5;9,12,10	R
	5,6,8;10,11,12	M
P12	5,6,7	P
	1,2,3,4;4,9,10;8,9,11;10,11,12	M
P13	12,13	P
	2,3,4;6,7,10;12(13),14,16	R
	1,2,4;6,7,8;7,9,10;14,15,16	M
P14	1,2,3,4;21,22;23,24	P
	10,11,15;15,16,19	M
P15	10,12,22,18,7,5;26,8,5,1,3,15,26	R
	1,2,3,4;5,6,7,8,9;10,11,12;12,16,22,25	M
P16	319,316,378,312;456,460;592,581;549,550	P
	408,453;441,401,402,404,405,407	R
	503,592,581;537,598,599	M

*P: Correntes paralelas; R: Reciclo

M: Magnitude da incerteza (ou medição)

3.4 Seleção e cálculos de índices de qualidade do balanço para ajuste do modelo de PLS

A detecção de erros grosseiros na metodologia proposta foi desenvolvida através da elaboração de modelos multivariáveis para predição da presença de erros grosseiros, e, caso existentes, sua magnitude e localização. A elaboração desses modelos necessita de dados de medições para predições dos erros grosseiros. Através da revisão bibliográfica, foi observado que algumas estatísticas de testes são um bom indicativo da qualidade do sistema de medição e do balanço de massa sendo possível de serem usados para predição de erros grosseiros. Outro critério para a escolha dos índices de qualidade de cada um dos balanços e medições foi a simplicidade computacional para sua avaliação e a capacidade de capturar erros grosseiros únicos e múltiplos dos sistemas em estudo. As seguintes estatísticas de teste foram escolhidas como índices de qualidade das medições e balanços de massa:

- Estatística de Teste Global.
- Estatística de Teste de Medida.
- Estatística de Teste Nodal

Os índices foram calculados usando o *software* Scilab (INRIA, 2010), que possui pacotes de otimização e de cálculos estatísticos. Cabe ressaltar que a estatística do Teste de Medida e Nodal não foram tomadas com seus valores absolutos, ou seja, os sinais dos ajustes e dos resíduos dos balanços foram considerados no cálculo desses índices, conforme

$$z_{am,i} = \frac{ad_i}{\sqrt{\overline{w_{i,i}}}} \tag{3.1}$$

e

$$z_{rm,j} = \frac{r_j}{\sqrt{V_{j,j}}}, \tag{3.2}$$

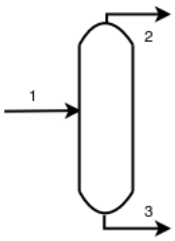
em que $z_{am,i}$ e $z_{rm,j}$ são os índices das estatísticas de teste do MT e NT, respectivamente, sem considerar a operação de módulo em seu numerador. Com os índices calculados e propriamente mapeados, foram montadas tabelas para a coleta de dados e subsequente estimação de parâmetros do modelo preditivo. Nestas tabelas,

os dados de entrada, ou variáveis independentes (apresentadas em colunas), são os índices, enquanto que os dados de saída, ou variáveis dependentes (apresentadas em colunas) são os valores dos erros grosseiros gerados nas simulações, conforme apresentado na Tabela 3.3. Cada linha da tabela correspondente a um experimento numérico, podendo ou não conter erros grosseiros. Para melhor exemplificar a elaboração da Tabela 3.3, seja o equipamento de processo apresentado na Figura 3.3. A tabela para mapeamento de índices de qualidade pode ser representada pela Tabela 3.4. Cada linha da Tabela 3.4 corresponde à um experimento numérico com as primeiras 5 colunas descrevendo as variáveis independentes do modelo (ou entradas), em que $I1$ são os valores de γ; $I2$, $I3$ e $I4$ são os respectivos valores de $z_{am,1}$, $z_{am,2}$ e $z_{am,3}$; $I5$ é o valor de $z_{rm,1}$. Cabe ressaltar que todos os valores de I_n são diferentes para cada linha da tabela. As saídas ou variáveis dependentes são as magnitudes dos erros grosseiros simulados (sem adição de erros aleatórios), tanto para erros de medição ($Bias_1$, $Bias_2$ e $Bias_3$) quanto para vazamentos ($Vazamento$). É possível notar que a magnitude do erro grosseiro de medição pode possuir tanto sinal positivo quanto negativo, enquanto que o vazamento terá sempre o mesmo sinal.

Tabela 3.3: Tabela para mapeamento de índices de qualidade.

Entradas					Saídas
Valor do Índice 1	Valor do Índice 2	Valor do Índice 3	Valor do Índice 4	Valor do Índice N	Valor Nominal do Erro i
$I1$	$I2$	$I3$	$I4$	$I5$	$Im1$
$I1_n$	$I2_n$	$I3_n$	$I4_n$	$I5_n$	$Im1_n$

Figura 3.3: Exemplo de processo para elaboração da tabela de mapeamento de índices de qualidade.



Fonte: Próprio autor

Tabela 3.4: Tabela para mapeamento de índices de qualidade do exemplo apresentado.

Entradas					Saídas			
γ	$z_{am,1}$	$z_{am,2}$	$z_{am,3}$	$z_{rm,1}$	$Bias_1$	$Bias_2$	$Bias_3$	$Vazamento$
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	5	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	-5	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	5	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	-5	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	5	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	-5	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	5
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	5
$I1$	$I2$	$I3$	$I4$	$I5$	6	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	-6	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	6	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	-6	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	6	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	-6	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	6
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	6
$I1$	$I2$	$I3$	$I4$	$I5$	...	...	...	...

A organização dos dados das simulações de erros grosseiros em forma de tabelas, como a Tabela 3.4, permite a alimentação de motores de cálculo para ajuste de modelos de forma simplificada. Diversas técnicas de ajuste de modelos estão disponíveis na literatura cuja seleção e aplicação será abordada na seção seguinte.

3.5 Seleção da técnica de ajuste do modelo de DEG

No caso em estudo, o principal objetivo é identificar erros grosseiros únicos e múltiplos de forma robusta, além de evitar a detecção de erros puramente aleatórios. Para executar tal tarefa, a metodologia em questão busca extrair informações dos dados das simulações (com e sem erros grosseiros) e, em um segundo momento, classificar os erros grosseiros em uma das categorias: medidor, vazamento, únicos ou múltiplos. Uma das alternativas para prever o comportamento de variáveis de processo baseados em variáveis independentes é a regressão multivariável. Em problemas de regressão multivariável, geralmente se deseja prever o comportamento de uma variável de resposta, ou dependente, ou de saída, baseado na magnitude de múltiplas variáveis independentes, ou entradas. Tal problema é consolidado e bem documentado na literatura por diversos autores (AGUIRRE, 2000; WERKEMA; AGUIAR, 2006; RAWLINGS et al., 1998). Nesta abordagem, parâmetros de um modelo multivariável puramente empírico são estimados a partir de variáveis independentes e dependentes. Com a estimação dos parâmetros, consolida-se um modelo matemático capaz de calcular o valor das variáveis dependentes a partir de variáveis independentes.

3.5.1 Regressão linear múltipla

Conforme comentado na seção anterior, o objetivo da regressão múltipla é prever o valor da variável dependente, ou de resposta, denominada $\hat{y}_i$, utilizando valores de variáveis independentes, denominadas $\hat{x}_j$, estimando o valor dos parâmetros do modelo, $\hat{b}_j$. Logo o modelo multivariável para o caso linear assume a forma:

$$\hat{y}_i = \hat{b}_0 + \hat{b}_1\hat{x}_{i,1} + \hat{b}_2\hat{x}_{i,2} + \cdots + \hat{b}_k\hat{x}_{i,k}, \tag{3.3}$$

em que o subscrito i, representa as diferentes observações, que, em nosso caso, serão as simulações numéricas. A Equação (3.3) pode ser escrita pela notação vetorial da seguinte forma:

$$\hat{\mathbf{Y}} = \hat{\mathbf{X}} \cdot \hat{\mathbf{b}}, \tag{3.4}$$

em que $\hat{\mathbf{Y}}$ é o vetor contendo as predições do modelo para o valor da variável de resposta em cada uma das n observações experimentais, conforme:

$$\hat{\mathbf{Y}} = [\hat{y}_1, \hat{y}_2, \hat{y}_3, \cdots, \hat{y}_n]^T, \tag{3.5}$$

em que $\hat{\mathbf{b}}$ é o vetor cujos coeficientes representam as estimativas para cada um dos $k+1$ coeficientes do modelo linear e assume a seguinte forma:

$$\hat{\mathbf{b}} = \left[\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3, \cdots, \hat{b}_k\right]^T, \tag{3.6}$$

e $\hat{\mathbf{X}}$ é a matriz de valores de variáveis independentes cujas linhas são as observações das variáveis independentes (em nosso caso, as simulações numéricas) e as colunas cada uma das variáveis, sendo que esta matriz assume a forma:

$$\hat{\mathbf{X}} = \begin{bmatrix} 1 & \hat{x}_{1,1} & \hat{x}_{1,2} & \cdots & \hat{x}_{1,k} \\ 1 & \hat{x}_{2,1} & \hat{x}_{2,2} & \cdots & \hat{x}_{1,k} \\ 1 & \hat{x}_{3,1} & \hat{x}_{3,2} & \cdots & \hat{x}_{1,k} \\ . & . & . & . & . \\ 1 & \hat{x}_{n,1} & \hat{x}_{n,2} & \cdots & \hat{x}_{n,k} \end{bmatrix}. \tag{3.7}$$

A solução analítica para o problema de mínimos quadrados para a estimação dos coeficientes $\hat{\mathbf{b}}$ pode ser obtida, através da aplicação de álgebra matricial, por:

$$\hat{\mathbf{b}} = \left(\hat{\mathbf{X}}^T\hat{\mathbf{X}}\right)^{-1} \cdot \left(\hat{\mathbf{X}}^T\hat{\mathbf{Y}}\right). \tag{3.8}$$

Além disso, a forma linear apresentada pela Equação (3.3) também pode ser estendida para casos não-lineares nas variáveis independentes através da transformação destas variáveis $\hat{x}_i$ (FINKLER, 2003; WERKEMA; AGUIAR, 2006).

3.5.2 Técnicas de redução de dimensionalidade

Um dos problemas encontrados em técnicas de regressão multivariável é a limitação numérica. Quando as variáveis independentes estão correlacionadas entre si, os métodos envolvidos no problema de estimação de parâmetros podem apresentar problemas (FINKLER, 2003). Na prática, quando as colunas da matriz $\hat{\mathbf{X}}$ são quase colineares, a inversão da matriz $\hat{\mathbf{X}}^T\hat{\mathbf{X}}$ acaba sendo muito sensível aos valores das

variáveis $\hat{x}_{i,k}$. Para resolver tais problemas existem os métodos de redução de dimensionalidade cujo princípio é realizar uma transformação da matriz $\hat{\mathbf{X}}$ em uma nova matriz onde não esteja presente a colinearidade, como, por exemplo, a técnica de Análise de Componentes Principais (PCA), assim apresentada:

$$\hat{\mathbf{X}} = \hat{\mathbf{X}}_1 + \hat{\mathbf{X}}_2 + \cdots + \hat{\mathbf{X}}_a \tag{3.9}$$

ou

$$\hat{\mathbf{X}} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \cdots + \mathbf{t}_n\mathbf{p}_n^T \tag{3.10}$$

ou em notação vetorial

$$\hat{\mathbf{X}} = \mathbf{T} \cdot \mathbf{P}^T. \tag{3.11}$$

Segundo a técnica de PCA, cada um dos vetores $\mathbf{t}_1$,$\mathbf{t}_2, \cdots, \mathbf{t}_n$ (*score vectors*) é composto por n elementos, resultantes das combinações lineares das colunas de $\hat{\mathbf{X}}$ enquanto que os vetores $\mathbf{p}_1^T + \mathbf{p}_2^T + \cdots + \mathbf{p}_n^T$ (*loading vectors*) são compostos de k elementos, resultantes das combinações lineares das linhas de $\hat{\mathbf{X}}$. A partir da Decomposição em Valores Singulares (SVD) da matriz $\hat{\mathbf{X}}$, é possível obter os vetores $\mathbf{T}$ e $\mathbf{P}$ da seguinte forma(VARMUZA; FILZMOSER, 2009):

$$\hat{\mathbf{X}} = \mathbf{T}_0 \cdot \mathbf{S} \cdot \mathbf{P}^T, \tag{3.12}$$

em que $\mathbf{T}_0$ contém os *pca scores* normalizados para o tamanho de 1 e $\mathbf{P}^T$ é a matriz com os *loading vectors*. Os *pca scores*, $\mathbf{T}_0$ são obtidos por:

$$\mathbf{T} = \mathbf{T}_0 \cdot \mathbf{S}, \tag{3.13}$$

enquanto que os *loading vectors* são obtidos por:

$$\mathbf{P} = \hat{\mathbf{X}}^T \cdot \mathbf{T}_0 \cdot \mathbf{S}^{-1}. \tag{3.14}$$

Após a extração da informação útil das variáveis em $\hat{\mathbf{X}}$ é possível realizar uma regressão multivariável nesse novo espaço de variáveis não correlacionadas, denominada de Regressão por Componentes Principais, ou PCR (*Principal Components Regression*). Substituindo $\hat{\mathbf{X}}$ da Equação (3.11) na Equação (3.4), temos:

$$\hat{\mathbf{Y}} = \left(\mathbf{T}\mathbf{P}^T\right) \hat{\mathbf{b}} \tag{3.15}$$

e reagrupando

$$\hat{\mathbf{Y}} = \mathbf{Tg} \tag{3.16}$$

em que $\mathbf{g} = \mathbf{P}^T\hat{\mathbf{b}}$, sendo que o vetor $\mathbf{g}$ pode ser obtido por

$$\mathbf{g} = \left(\mathbf{T}^T\mathbf{T}\right)^{-1} . \left(\mathbf{T}^T\hat{\mathbf{Y}}\right) \tag{3.17}$$

sendo que os parâmetros $\hat{\mathbf{b}}$ originais podem ser obtidos novamente através de

$$\hat{\mathbf{b}} = \mathbf{Pg}. \tag{3.18}$$

O método PCR pode ser aplicado quando a variável de saída é única ou as mesmas não estão correlacionadas. Quando diversas variáveis de saída, $\hat{\mathbf{Y}}$, estão presentes e estas podem estar correlacionadas, é conveniente usar a técnica de mínimos quadrados parciais, ou PLS (*Partial Least Squares*), que consiste em decompor as variáveis $\hat{\mathbf{Y}}$ em seus respectivos componentes principais, conforme:

$$\hat{\mathbf{Y}} = \mathbf{UQ}^T, \tag{3.19}$$

e, em seguida, realizar a regressão nesses novos espaços de variáveis não-correlacionadas, dadas por Equação (3.11).

Detalhes sobre implementações e descrições mais detalhadas da regressão por componentes principais (PCR) e os mínimos quadrados parciais (PLS) são abordados em diversos livros texto (VARMUZA; FILZMOSER, 2009; ADAMS, 2004; MARK; WORKMAN, 2007; JOLLIFFE, 2002).

Como ferramenta computacional para geração de modelos PLS, foi usado o *software* VRAnalyst (FINKLER et al., 2006). Neste programa de computador, os dados são alimentados em forma de tabela e estatísticas básicas em relação à qualidade do modelo podem ser facilmente avaliadas e os resultados com os modelos gerados podem ser simulados e avaliados. Estão presentes diversos tipos de modelos de mínimos quadrados parciais: Linear, Quadrático, Cúbico e Box-Tidel flexível cujos parâmetros são estimados usando o algoritmo NIPALS (*Non-Linear Iterative Partial Least Squares*)(VARMUZA; FILZMOSER, 2009). Os modelos podem ser exportados em formato compatível com o software Scilab. No presente trabalho foram ajustados modelos matemáticos tipo PLS quadráticos pois apresentaram melhor coeficiente de determinação (R^2) que modelos PLS lineares em todos os 16 casos estudados.

Modelos PLS cúbicos também apresentaram um coeficiente de determinação um pouco melhor que modelos PLS quadráticos, porém, como sua avaliação é um pouco mais onerosa computacionalmente, optou-se pelo uso de modelos PLS quadráticos. A vantagem do uso de modelos PLS está no fato de que erros grosseiros múltiplos, dependendo da posição onde os mesmos estão presentes no fluxograma de processo, podem gerar dados das variáveis de entrada (ou saída) correlacionados, dificultando a estimação dos parâmetros em modelos multivariáveis simples.

Os dados de entrada e saída para estimação de parâmetros dos modelos PLS quadráticos foram obtidos através de simulações numéricas conforme metodologia descrita nas próximas seções.

3.6 Ajuste do modelo de DEG

Antes do ajuste dos modelos PLS em si, algumas etapas foram necessárias para que as variáveis de saída, a magnitude dos erros grosseiros, pudessem ser melhor 'explicadas' pelas variáveis de entrada, as estatísticas do GT, NT e MT. Tais etapas são descritas nas próximas seções.

3.6.1 Escolha da função objetivo

Conforme comentado na seção anterior, alguns algoritmos de DEG falham ao encontrar correntes paralelas, que é um caso bastante comum em processos industriais. Um dos principais testes para a detecção de erros em medidores, o Teste de Medida (MT), que é o principal teste usado para detecção de erros únicos ou múltiplos, também apresenta limitações para esse tipo de sistema. A principal limitação do MT em sistemas com correntes paralelas é porque sua estatística de teste apresenta o mesmo valor para todas correntes nessa categoria, não sendo possível distinguir qual a corrente onde o erro grosseiro se encontra. A Tabela 3.5 apresenta o resultado de uma simulação numérica do Problema 1 (apresentado na Figura 3.3), utilizando a função objetivo de mínimos quadrados ponderados, onde um erro grosseiro de medição com a magnitude de nove vezes a incerteza-padrão é adicionado nas correntes

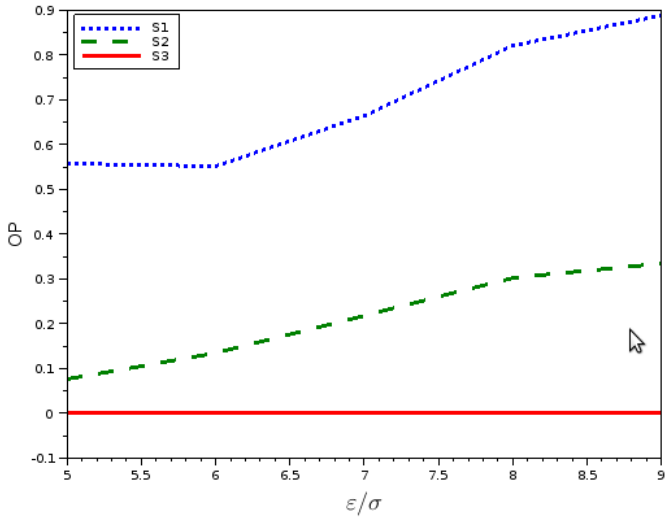
1, 2 e 3 individualmente (apresentado nas linhas). É possível observar que, devido à topologia do sistema, não é possível distinguir a localização do erro grosseiro com este método, pois, todas as correntes possuem a mesma estatística de teste, não importando a localização do erro grosseiro adicionado. Nesse caso, o Poder Total do Teste de Medida, que indica a fração de erros grosseiros corretamente identificada, é zero para todas as correntes, pois não é possível identificar a posição do erro. Por esse motivo, o presente trabalho buscou explorar outras funções objetivo com o intuito de obter uma distinção entre a estatística de teste das correntes em questão, para que o sistema de aprendizado pudesse distinguí-las para determinar a posição correta do erro de medição.

Tabela 3.5: Estatística de Teste de Medida para o Problema 1.

Erro	Estatística de Teste MT		
Grosseiro	$z_{a,1}$	$z_{a,2}$	$z_{a,3}$
Corrente 1	0,1462	0,1462	0,1462
Corrente 2	0,1462	0,1462	0,1462
Corrente 3	0,1462	0,1462	0,1462

Como funções objetivo alternativas, foram testados estimadores robustos, conforme apresentado no Apêndice B. Entre os estimadores utilizados, optou-se pelo estimador de Cauchy, pois o mesmo apresentou o melhor resultado para todos os problemas estudados. No método aqui apresentado foi usada a constante C_c com o valor de 2,3849, determinado por Özyurt e Pike (2004). Como exemplo do uso deste estimador, podemos apresentar a curva de poder de teste, onde é graficado a magnitude da razão δ/u (onde δ é a magnitude do erro grosseiro e u a incerteza-padrão) contra o Poder Total (OP) para o Problema 1, conforme Figura 3.4. Neste gráfico, é possível observar que, ao contrário do que acontece com o estimador MQP, o estimador de Cauchy apresenta um poder total do teste de DEG maior para as correntes 1 e 2, porém, erros de medição para a corrente 3 continuam sendo indistinguíveis. Tal limitação foi solucionada através da modificação das incertezas do problema original, apresentada na próxima seção.

Figura 3.4: Gráfico do OP do MT utilizando o estimador robusto de Cauchy para o Problema 1.

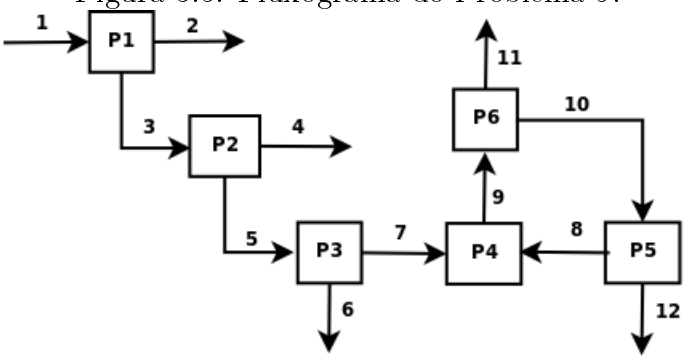


Fonte: Próprio autor

3.7 Escolha das incertezas

Conforme apresentado na seção anterior, somente a modificação da função objetivo não é suficiente para melhorar o OP para o Problema 1. Além disso, para sistemas onde existe reciclo interno de correntes, é possível que, devido à diferença entre as incertezas das correntes presentes no reciclo, a medição com erro seja incorretamente identificada. Tal fenômeno pode ser melhor exemplificado através do Problema 9, apresentado na Figura 3.5. Nesse fluxograma, é possível observar um reciclo formado pelas correntes 8, 9 e 10. Conforme apresentado na Tabela 3.6 a corrente 8 é a de maior precisão entre as correntes que formam o reciclo (menor incerteza-padrão), e, ao ser adicionado um erro grosseiro na mesma, esse erro acaba sendo atribuído às correntes mais imprecisas após a reconciliação de dados, como as correntes 9 e 10, mesmo usando uma função objetivo robusta, conforme apresentado na Tabela 3.7.

Figura 3.5: Fluxograma do Problema 9.



Fonte: Próprio autor

Tabela 3.6: Poder Total do MT para o Problema 9 usando o estimador de mínimos quadrados ponderados.

Corrente	Medida Nominal	Desvio Padrão	OP MT
1	230	38	0
2	21	1.1	0
3	209	5.0	0.507
4	35	1.8	0.104
5	174	2.0	0.384
6	15	0.9	0.036
7	159	7.3	0.839
8	50	1.0	0.149
9	209	5.0	0.755
10	94	2.0	0.310
11	115	18.1	0.853
12	44	2.4	0.403

Tabela 3.7: Poder Total do MT para o estimador robusto de Cauchy, usando as incertezas originais.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0.978	0.992	0.994	0.992	1
2	0	0	0	0	0
3	0.606	0.592	0.67	0.686	0.78
4	0.004	0.032	0.05	0.092	0.124
5	0.278	0.408	0.526	0.6	0.632
6	0.002	0.002	0.006	0.002	0.01
7	0.98	0.992	0.996	0.998	1
8	0	0.004	0.002	0.002	0.004
9	0.964	0.982	0.99	0.992	0.996
10	0.356	0.424	0.456	0.456	0.506
11	0.974	0.988	0.992	0.996	0.998
12	0.456	0.504	0.506	0.522	0.562

Uma alternativa eficiente para remover o problema do espalhamento de erros é a utilização das incertezas das medições iguais em todas as correntes, dando assim pesos iguais a todos os erros e reduzindo o efeito de espalhamento ou distribuição dos erros grosseiros. O uso das incertezas idênticas, em conjunto com o estimador robusto de Cauchy, apresentaram bons resultados para identificação de erros, conforme apresentado na Tabela 3.8.

Tabela 3.8: Poder Total do MT para o estimador robusto de Cauchy, usando incertezas unitárias em todas as correntes.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0.403	0.473	0.487	0.473	0.510
2	0.300	0.453	0.490	0.497	0.503
3	0.763	0.900	0.973	0.977	0.997
4	0.723	0.837	0.947	0.980	1.000
5	0.773	0.927	0.973	1.000	1.000
6	0.737	0.877	0.953	0.993	1.000
7	0.810	0.947	1.000	1.000	1.000
8	0.743	0.893	0.967	0.987	0.993
9	0.760	0.917	0.963	0.990	0.997
10	0.770	0.870	0.953	0.993	1.000
11	0.767	0.920	0.957	1.000	1.000
12	0.777	0.923	0.980	0.983	1.000

Com as técnicas do uso do estimador de Cauchy em conjunto com o uso de incertezas iguais, o poder de detecção de erros para o teste de medição apresentaram bons resultados, pois mimificam o que acontece com erros grosseiros em casos reais, ou seja, em uma planta industrial, onde uma falha em um medidor não afeta os demais medidores. Devido à melhoria no Poder Total do teste estatístico apresentada com o uso do estimador de Cauchy em conjunto com o uso das incertezas unitárias, esta metodologia foi usada para geração de dados para ajuste dos modelos PLS.

Ao longo dos experimentos numéricos desenvolvidos para a escolha das incertezas, outras magnitudes das incertezas originais e unitárias também foram testadas, entretanto, ao se aumentar essas magnitudes, o OP para detecção de vazamentos (Teste Nodal) é comprometido (diminui). É importante observar que, quando a incerteza de uma determinada corrente aumenta, o valor do respectivo elemento da diagonal da matriz de variância das medidas, Σ, aumenta. Consequentemente, a matriz de variância-covariância dos balanços, $\mathbf{V}$ aumenta, conforme a Equação (2.11),

aqui reapresentada:

$\mathbf{V} = cov(\mathbf{R}) = \mathbf{A} \cdot \Sigma \cdot \mathbf{A}^T$,

assim, a estatística de teste $z_{rm,j}$, apresentada na Equação (2.13), diminui, pois o termo $V_{i,i}$ encontra-se no denominador da expressão, aqui reapresentada:

$z_{rm,j} = \frac{|r_j|}{\sqrt{V_{j,j}}}$.

Também é possível observar que o mesmo efeito ocorre com a estatística de teste, $z_{am,i}$, pois, ao aumentar as incertezas, $z_{am,i}$ aumenta, conforme a Equação (2.17), aqui reapresentada:

$z_{am,i} = \frac{|ad_i|}{\sqrt{\overline{w_{i,i}}}}$,

reapresentando a equação que define o respectivo elemento do denominador, $\overline{w_{i,i}}$:

$\overline{\mathbf{W}} = \Sigma \cdot \mathbf{A}^T \cdot V^{-1} \cdot \mathbf{A} \cdot \Sigma$,

logo, tanto $z_{am,i}$ quanto $z_{rm,j}$ sofrem o mesmo efeito com o aumento ou diminuição das incertezas e quanto maior forem suas magnitudes, maior a probabilidade dos erros serem classificados como grosseiros ou vazamentos. Embora o teste estatístico que classifica correntes e balanços com a presença de erros grosseiros dependa da magnitude de $z_{am,i}$ e $z_{rm,j}$ o efeito final sobre o OP, que determina a qualidade do método de DEG, não depende somente da magnitude dessas variáveis. Na prática, o efeito da variação das incertezas sobre as estatísticas de testes $z_{am,i}$ e $z_{rm,j}$ não é difícil de ser obtido analiticamente, porém, o efeito final sobre o OP é complexo. É importante salientar que, ao variar as incertezas, a distribuição de probabilidade Normal, parametrizada pela média e desvio-padrão, sofre mudança e há a necessidade de definir uma nova significância estatística dos testes em questão, por tentativa e erro, para manter o AVT1 em 0,1. Logo, embora o uso de incertezas iguais para todas as correntes distribua os erros de forma mais homogênea, após a reconciliação de dados, entre as correntes, a determinação exata das magnitudes iguais das incertezas para obter uma boa relação de custo benefício entre o MT e NT, devem ser obtidas através de experimentos numéricos. Para exemplificar melhor a influência da magnitude das incertezas idênticas no OP do teste de medida, seja o Problema 1, apresentado na Figura 3.3 cujos dados são apresentadas as Tabelas 3.9 e 3.10. Através deses dados, é possível observar o aumento do OP com o aumento das magnitude das incertezas de 0,1 t.h^{-1} para 5 t.h^{-1}. Em contraponto, o OP do NT é reduzido ao aumentar a magnitude das incertezas idênticas de 0,1t.h^{-1} para 5t.h^{-1},

conforme pode ser observado nas Tabelas 3.11 e 3.12, respectivamente.

Tabela 3.9: Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicado por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,395	0,459	0,495	0,506	0,495
2	0,198	0,348	0,446	0,45	0,484
3	0,339	0,458	0,489	0,449	0,466

Tabela 3.10: Poder Total do teste de medida utilizando a incerteza-padrão unitária multipicado por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,757	0,796	0,82	0,818	0,814
2	0,498	0,531	0,532	0,548	0,512
3	0,538	0,509	0,529	0,471	0,497

Tabela 3.11: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicado por 0,1.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1

Tabela 3.12: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicado por 5.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
0,102	0,099	0,103	0,102	0,105

Segundo testes numéricos conduzidos nos problemas 1, 6 e 9, utilizando as incertezas originais e modificadas, a incerteza unitária para todas as correntes apresenta um bom compromisso para detecção de erros de medição e vazamentos, sendo tais testes apresentados no Apêndice C. Também foi possível observar nesses testes numéricos que, ao usar a incerteza igual para todas as correntes, o OP é influenciado pela magnitude da incerteza escolhida e da vazão da corrente associada. Em processos onde existem um grande número de correntes paralelas ligadas ao mesmo equipamento, também foi observado falhas na detecção, sendo necessário reajustar um novo conjunto de incertezas idênticas para todo o fluxograma, sendo tais casos apresentados e discutidos no próximo capítulo.

3.8 Geração e seleção de dados para ajuste dos modelos PLS

Após a seleção da função objetivo e das incertezas necessárias para distinguir entre os erros grosseiros, inicia-se o processo de seleção dos dados gerados para a estimação de parâmetros dos modelos PLS.

3.8.1 Geração e seleção de dados com erros puramente aleatórios

Nesse tipo de experimento numérico do tipo semelhante à simulação de Monte Carlo, previamente à geração de dados com erros aleatórios, a literatura reporta a

necessidade de determinar uma significância estatística, por tentativa e erro, para os testes GT, MT e NT de forma a manter a Média dos Erros do Tipo I (AVT1) em 0,1 ou 10% (SERTH; HEENAN, 1986; ÖZYURT; PIKE, 2004; ROSENBERG et al., 1987; IORDACHE et al., 1985; ZHANG et al., 2010; NARASIMHAN; JORDACHE, 2000). Em seguida, são gerados e selecionados dados com erros puramente aleatórios através das seguintes etapas:

1 Gerar dados de medições de vazão total com erros puramente aleatórios, utilizando a função Normal, com o desvio-padrão idêntico e unitário, e adicionar a todas as correntes;

2 Reconciliar os dados e calcular as estatísticas de teste para o GT, MT e NT (γ, $z_{a,i}$, $z_{r,j}$).

3 Se as estatísticas dos testes GT, MT (para todas as medições) e NT (para todos os balanços) forem maior que a significância, descartar a simulação.

4 Caso contrário, adicionar ao conjunto de ajuste com erro grosseiro zero.

5 Repetir o procedimento NA vezes.

No procedimento acima descrito, após as etapas 1 e 2 estarão presentes tanto dados com erros puramente aleatórios quanto dados com erros grosseiros, devido a quantidade elevada de experimentos numéricos realizado. Na etapa 3, a eliminação das simulações que ultrapassam a estatística de testes para o GT, MT e NT é necessária, pois, não é desejado que no conjunto de dados com erros puramente aleatórios estejam presentes dados com erros grosseiros para estimação de parâmetros do modelo. Logo, nesse conjunto de dados, somente dados com erros aleatórios são usados nos ajustes dos modelos. O critério de inclusão ou descarte dos dados é baseado na significância estatística dos testes de hipótese previamente determinada de forma a manter o AVT1 em 0,1 ou 10%. Esta significância estatística será a mesma usada para distinguir dados com erros grosseiros e sem erros grosseiros nas seções seguintes. O descarte dos dados que excedem a significância estatística tem um efeito positivo posteriormente na estimação de parâmetros do modelo empírico multivariável, pois qualquer erro aleatório adicionado à variáveis de entrada ou saída

desse tipo de modelo reduzem a qualidade dos resultados dos mesmos(VARMUZA; FILZMOSER, 2009). Tal efeito aumenta em função da magnitude do erro aleatório adicionado.

O número NA de repetições do procedimento está relacionado à obtenção de um conjunto de dados suficientemente grande para evitar o efeito de *overfitting* ou sobre-ajuste. Fixando o parâmetro NA do procedimento em 100, garante a obtenção de, na média, 90 pontos que serão incluídos nesse conjunto de dados. Considerando que para geração de um modelo quadrático multivariável estão presentes $2n\hat{\mathbf{X}} + 1$ parâmetros [1], onde $n\hat{\mathbf{X}}$ é o número de variáveis independentes do modelo, seria necessário $2n\hat{\mathbf{X}} + 2$ pontos para evitar o sobre-ajuste. Para o método proposto, considerando em processo com um equipamento e três correntes, estão presentes 5 variáveis de entradas, sendo que nesse caso, seriam necessários 12 simulações. Considerando que restarão em média 90 simulações após a aplicação do procedimento acima descrito, após o descarte das simulações, é possível observar que o número de simulações é mais que suficiente para evitar o sobre-ajuste. Além disso, é importante salientar que mais dados serão adicionados ao conjunto de estimação após as etapas seguintes.

Após a etapa anteriormente descrita, será construída as primeiras linhas da tabela com dados puramente aleatórios, conforme Tabela 3.13, onde o número de linhas está relacionado ao número de simulações não descartadas.

Tabela 3.13: Tabela para mapeamento de índices de qualidade para erros puramente aleatórios para um processo com um equipamento e três correntes.

Entradas					Saídas			
γ	$z_{am,1}$	$z_{am,2}$	$z_{am,3}$	$z_{rm,1}$	$Bias_1$	$Bias_2$	$Bias_3$	$Vazamento$
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	...	...	...	...

[1] Considerando a ausência de termos quadráticos cruzados

3.8.2 Geração e seleção de dados com erros grosseiros nos medidores - erros únicos

A geração e seleção de dados com erros grosseiros nos medidores devem ser aplicadas sequencialmente em todas as medições, e segue as seguintes etapas descritas abaixo:

1 Usando o conjunto de dados aleatórios gerados na etapa anterior (removidos os falsos positivos) gera-se, para uma determinada corrente, diga-se C, um erro grosseiro

2 Reconciliar os dados e calcular as estatísticas de teste para o GT, MT e NT (γ, $z_{a,i}$, $z_{r,j}$).

3 Usando a significância estatística dos erros puramente aleatórios (seção anterior), se a estatística dos testes GT, MT e NT para a corrente, ou nós que envolvem a corrente, forem menor que a significância, descartar a simulação.

4 Caso contrário, adicionar ao conjunto de ajuste com erro grosseiro diferente de zero.

5 Repetir o procedimento NB vezes com a corrente C.

Os erros grosseiros da etapa 1 foram variados da magnitude de cinco a nove vezes os desvios-padrão. O sinal dos erros foi escolhido utilizando uma distribuição uniforme, com a mesma probabilidade do sinal ser positivo ou negativo. O número de simulações, NB, foi fixado em 100 para obter um conjunto de dados com diversos elementos para evitar o efeito de sobre-ajuste. A significância estatística da etapa de geração de dados com erros puramente aleatório (seção anterior) foi usada como critério para descartar as simulações. Tal procedimento de descarte foi realizado para remover simulações que poderiam entrar na estimação de parâmetros do modelo e, após somar erros aleatórios e grosseiros, os erros de medições e vazamentos serem considerados como erros aleatórios. Cabe ressaltar que esse teste era feito individualmente ou seja, se, ao adicionar um erro na corrente C e a estatística de teste dessa, e somente dessa, corrente fosse inferior ao critério de corte, a simulação

como um todo era eliminada. Testes numéricos foram previamente elaborados e indicaram, conforme esperado, que a inclusão de todos os dados, sem descarte, reduz a qualidade do ajuste do modelo, avaliado através do coeficiente de determinação R^2 e também reduz a capacidade de identificação de erros grosseiros do método proposto. A faixa de erros grosseiros entre cinco a nove vezes os desvios-padrão usada no presente trabalho foi determinada através de experimentos numéricos. Valores dos desvios-padrão menor que cinco reduzem a magnitude de $z_{am,i}$ e $z_{rm,j}$, e, consequentemente a probabilidade de classificação da medição final como erro grosseiro, reduzindo o número de dados no conjunto usado para estimação de parâmetros do modelo. O uso de nove desvios-padrão já incluem uma quantidade suficiente de dados no conjunto de dados para estimação, não sendo necessário usar desvios-padrão de magnitudes maiores.

Além disso, existe um compromisso entre a escolha da magnitude da incerteza idêntica e as estatísticas de teste $z_{am,i}$ e $z_{rm,j}$. Ao se reduzir as magnitudes das incertezas, $z_{am,i}$ aumenta, consequentemente, aumenta a probabilidade de detectar erros grosseiros, incluindo um maior número de dados nesse conjunto. Em contraponto, como $z_{am,i}$ e $z_{rm,j}$ não possuem uma relação diretamente proporcional com as magnitudes das incertezas, com a redução dessas, pode ocorrer a redução de $z_{rm,j}$, reduzindo o número de dados no conjunto para determinação de vazamentos. Cabe ressaltar que, para a estimação de parâmetros de um modelo genérico, quanto maior o número de dados distribuídos de forma homogênea em um intervalo, melhor a qualidade do ajuste(WERKEMA; AGUIAR, 2006; VARMUZA; FILZMOSER, 2009), o que foi confirmado no caso em estudo através de experimentos númericos.

Após essa etapa, o conjunto de dados será concatenado com os dados obtidos na etapa anterior conforme apresentado na Tabela 3.14, onde o número de linhas está relacionado ao número de simulações não descartadas.

Tabela 3.14: Tabela para mapeamento de índices de qualidade após a geração de dados com erros aleatórios e erros grosseiros nas medições para processo com um equipamento e três correntes.

Entradas					Saídas			
γ	$z_{am,1}$	$z_{am,2}$	$z_{am,3}$	$z_{rm,1}$	$Bias_1$	$Bias_2$	$Bias_3$	$Vazamento$
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	5	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	-5	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	5	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	-5	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	5	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	-5	0
$I1$	$I2$	$I3$	$I4$	$I5$	6	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	-6	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	6	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	-6	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	6	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	-6	0
$I1$	$I2$	$I3$	$I4$	$I5$	...	...	...	...

3.8.3 Geração e seleção de dados com erros grosseiros de vazamentos - erros únicos

A geração e seleção de dados com erros grosseiros de vazamentos deve ser aplicada sequencialmente em todos os equipamentos, e segue as seguintes etapas:

1 Usando o conjunto de dados aleatórios gerados na etapa anterior (removidos os falso positivo) gera-se, para um determinado equipamento, diga-se M, um vazamento.

2 Calcular as medições a partir dos resíduos.

3 Reconciliar os dados e calcular as estatísticas de teste para o GT, MT e NT (γ, $z_{am,i}$, $z_{rm,j}$).

4 Usando a significância estatística dos erros puramente aleatórios (seção anterior), se a estatística dos testes GT, MT e NT para as correntes envolvidas no equipamento, ou no próprio equipamento, forem menor que a significância, descartar a simulação.

5 Caso contrário, adicionar ao conjunto de ajuste com erro grosseiro diferente de zero.

6 Repetir o procedimento NV vezes.

Os vazamentos adicionados nos equipamentos eram variados da magnitude de 7% a 15% (com intervalo de 2%) da vazão total que passa pelo equipamento (soma das vazões que entram e que saem, sem considerar o sinal), determinado por experimentos numéricos com o objetivo de obter um conjunto de dados com um número razoável de elementos. O número de simulações, NV, foi fixado em 100 para obter um conjunto de dados com diversos elementos para evitar o efeito de sobre-ajuste. Cabe ressaltar que o vazamento tem sempre sinal negativo em relação ao resíduo aleatório. Como a estatística do Teste Nodal não necessita da reconciliação, é necessário obter as medições a partir dos resíduos dos balanços de massa da seguinte forma:

$$\mathbf{Y} = \Sigma \cdot \mathbf{A}^T \cdot \left(\mathbf{A}\Sigma \cdot \mathbf{A}^T\right)^{-1} \cdot \mathbf{R} + \overline{\mathbf{X}}_{NOM}, \tag{3.20}$$

em que $\overline{\mathbf{X}}_{NOM}$ é o valor nominal da medição. Em seguida, obtida as medições a partir dos vazamentos, as estatísticas de teste γ, $z_{am,i}$ e $z_{rm,j}$ foram calculadas. Na próxima etapa, as simulações onde as correntes ou resíduos dos equipamentos que possuem uma estatística de teste abaixo do critério de corte são descartadas, utilizando a mesma significância estatística do erro puramente aleatório da seção anterior. Tal procedimento foi realizado para remover simulações que poderiam entrar no conjunto de dados e, após somar erros aleatórios e grosseiros, os erros de medições e vazamentos serem considerados como erros aleatórios. Cabe ressaltar que esse teste era feito individualmente ou seja, se, ao adicionar um vazamento no equipamento M e a estatística de teste desse, e somente desse, equipamento

fosse inferior ao critério de corte, a simulação como todo era eliminada. Testes numéricos foram previamente elaborados e indicaram, conforme esperado, que a inclusão de todos os dados, sem descarte, reduz a qualidade do ajuste do modelo, avaliado através do coeficiente de determinação R^2 e também reduz a capacidade de identificação de erros grosseiros. Conforme discutido na seção anterior, existe um compromisso entre a magnitude das incertezas idênticas e as estatísticas de teste $z_{am,i}$ e $z_{rm,j}$. Dependendo da magnitude das incertezas idênticas escolhidas, após o ajuste adequado do AVT1, haverá um maior probabilidade de detecção de erros grosseiros nas correntes ou para os vazamentos, sendo essa relação difícil de deduzir analiticamente. No presente trabalho, conforme mencionado anteriormente, a escolha da magnitude das incertezas foi determinada por experimentos numéricos visando obter um melhor compromisso entre o OP da detecção de erros grosseiros nas correntes e para os vazamentos. É importante ressaltar que, se o OP para o NT é muito baixo, poucos dados serão incluídos no conjunto de dados, reduzindo o OP do método proposto.

Após essa etapa, o conjunto de dados será concatenado com os dados obtidos na etapa anterior conforme apresentado na Tabela 3.15, onde o número de linhas está relacionado ao número de simulações não descartadas.

Tabela 3.15: Tabela para mapeamento de índices de qualidade após a geração de dados com erros aleatórios e erros grosseiros nas medições e vazamentos para processo com um equipamento e três correntes.

Entradas					Saídas			
γ	$z_{am,1}$	$z_{am,2}$	$z_{am,3}$	$z_{rm,1}$	$Bias_1$	$Bias_2$	$Bias_3$	$Vazamento$
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	5	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	-5	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	5	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	-5	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	5	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	-5	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	5
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	5
$I1$	$I2$	$I3$	$I4$	$I5$	6	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	-6	0	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	6	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	-6	0	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	6	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	-6	0
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	6
$I1$	$I2$	$I3$	$I4$	$I5$	0	0	0	6
$I1$	$I2$	$I3$	$I4$	$I5$	...	...	...	...

3.8.4 Geração e seleção de dados com erros grosseiros nos medidores e vazamentos - erros múltiplos

No caso de erros múltiplos, a metodologia de seleção de dados é um pouco diferente da metodologia de erros únicos. É possível ajustar um modelo de DEG com erros múltiplos somente de erros de medição, somente de vazamentos ou ambos

combinados. Também é possível selecionar a quantidade de erros grosseiros para ser adicionada ao fluxograma, gerando assim um problema combinatorial. Como temos 16 problemas testes, a quantidade de simulações necessárias seria bastante grande. Para contornar tal dificuldade o presente trabalho selecionou alguns cenários de erros para cada fluxograma de teste.

O problema de geração de erros grosseiros de medição e vazamentos para erros múltiplos é ligeiramente distinto dos apresentados para erros únicos. A principal diferença nesse caso, é que, ao invés de investigar a estatística de teste somente das correntes, ou equipamentos, onde os erros grosseiros são adicionados, todo o conjunto de estatística de testes de medições e resíduos de balanços de equipamentos é investigado. Somente simulações onde qualquer estatística de teste de medição ou resíduos dos balanços ultrapassar o limite da estatística de teste, essa linha é incluída no conjunto de teste. Esse procedimento foi assim tomado pois, observou-se que, em algumas simulações, erros múltiplos em medidores acabavam se comportando como erros únicos em outros medidores, sendo que esse fenômeno deveria ser capturado e enviado para o modelo que terá seus parâmetros ajustados.

3.9 Seleção de dados para ajuste dos parâmetros do modelo e testes

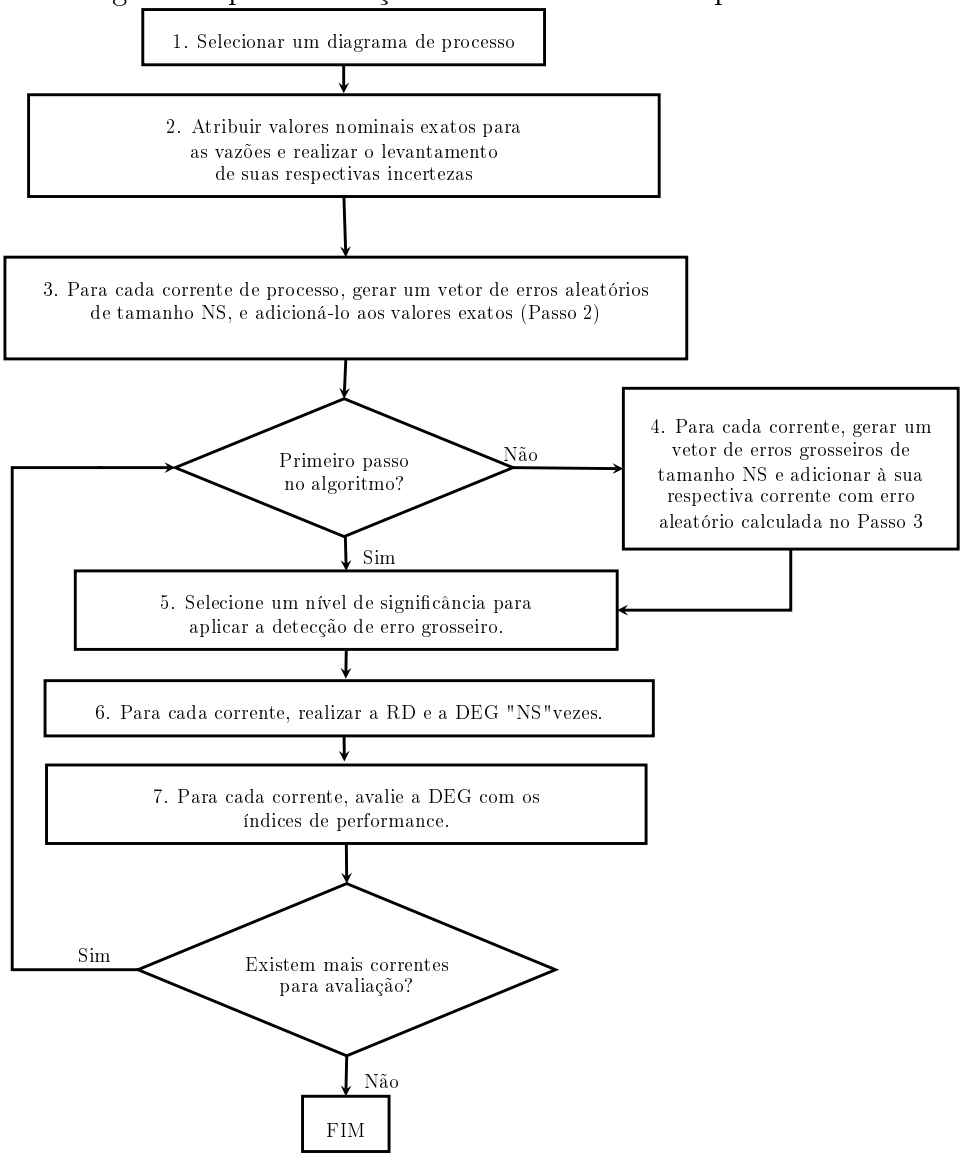
Após montada a tabela de dados, semelhante à Tabela 3.4, uma parte dos dados foi separada para ajuste no modelo de DEG e outra utilizada para teste do mesmo, sendo que essa última não foi inserida no conjunto de dados para ajuste. Para cada problema, 2 modelos foram gerados, um para erros únicos e outro para erros múltiplos.

3.10 Avaliação do modelo de DEG

Após ajustados os modelos, para os 16 problemas teste selecionados, foram utilizadas simulações numéricas para avaliar a capacidade dos métodos de literatura

e proposto de identificar apropriadamente os erros grosseiros. Os testes de literatura utilizados são o Teste Nodal (NT) e Razão de Verossimilhança Generalizada (GLR). As simulações numéricas utilizadas estão de acordo com as metodologias de avaliação de técnicas de identificação de erros grosseiros encontradas na literatura (SERTH; HEENAN, 1986; ÖZYURT; PIKE, 2004; ROSENBERG et al., 1987; IORDACHE et al., 1985; ZHANG et al., 2010; NARASIMHAN; JORDACHE, 2000). O detalhamento das etapas para o teste do método de identificação de erros de medição e vazamentos são apresentadas na Figura 3.6 e Figura 3.7, respectivamente.

Figura 3.6: Algoritmo para avaliação de técnicas de DEG para erros em medições.



Fonte: Próprio autor

As etapas detalhadas do algoritmo apresentado na Figura 3.6 são assim descritas.

1 Selecionar o fluxograma de processo.

2 Selecionar valores nominais para todas as vazões, considerando o sistema em estado estacionário, satisfazendo o balanço de massa de todas as unidades, denominando-as de $\overline{x}_{NOM,i}$. Configurar a incerteza-padrão de cada instrumento de medição.

3 Gerar um vetor de erros aleatórios de tamanho NS (em que NS é o número de testes de simulação), denominando-os de $\epsilon_{i,ns}$, para cada medida, utilizando distribuição Normal com média zero e desvio padrão dado pelo valor da incerteza-padrão configurada na Etapa 2. Adicionar $\epsilon_{i,ns}$ aos valores nominais das medições: $y_{i,ns} = \overline{x}_{NOM,i} + \epsilon_{i,ns}$ (ou $\mathbf{Y}_{ns}$ em notação vetorial).

3.1 Se for a primeira execução do laço, ir para a Etapa 5, caso contrário, ir para a Etapa 4.

4 Para cada corrente, gerar um vetor de erros grosseiros de tamanho NS, denominado $\delta_{i,ns}$. A magnitude e o sinal do mesmo deve seguir uma distribuição uniforme e aleatória escolhida entre os limites superiores e inferiores: $l\,(\overline{x}_{NOM,i} + \epsilon_{i,ns}) \leq \delta_{i,ns} \leq up\,(\overline{x}_{NOM,i} + \epsilon_{i,ns})$. Onde l é o fator multiplicativo para o limite inferior e up é o fator multiplicativo do limite superior do erro grosseiro que serão gerados. Adicionar $\delta_{i,ns}$ ao termo $y_{i,ns}$ (previamente avaliado no passo 3), para obter o valor final da medição. No caso de múltiplos erros grosseiros, a localização dos mesmos deve ser de forma uniforme e aleatória, ou seja algumas medições podem não conter erros grosseiros.

5 Selecionar o nível de significância, α, para o teste estatístico. Essa significância deve ser determinada por tentativa e erro para manter o AVT1 próximo a 0,1 quando não há erros grosseiros presentes.

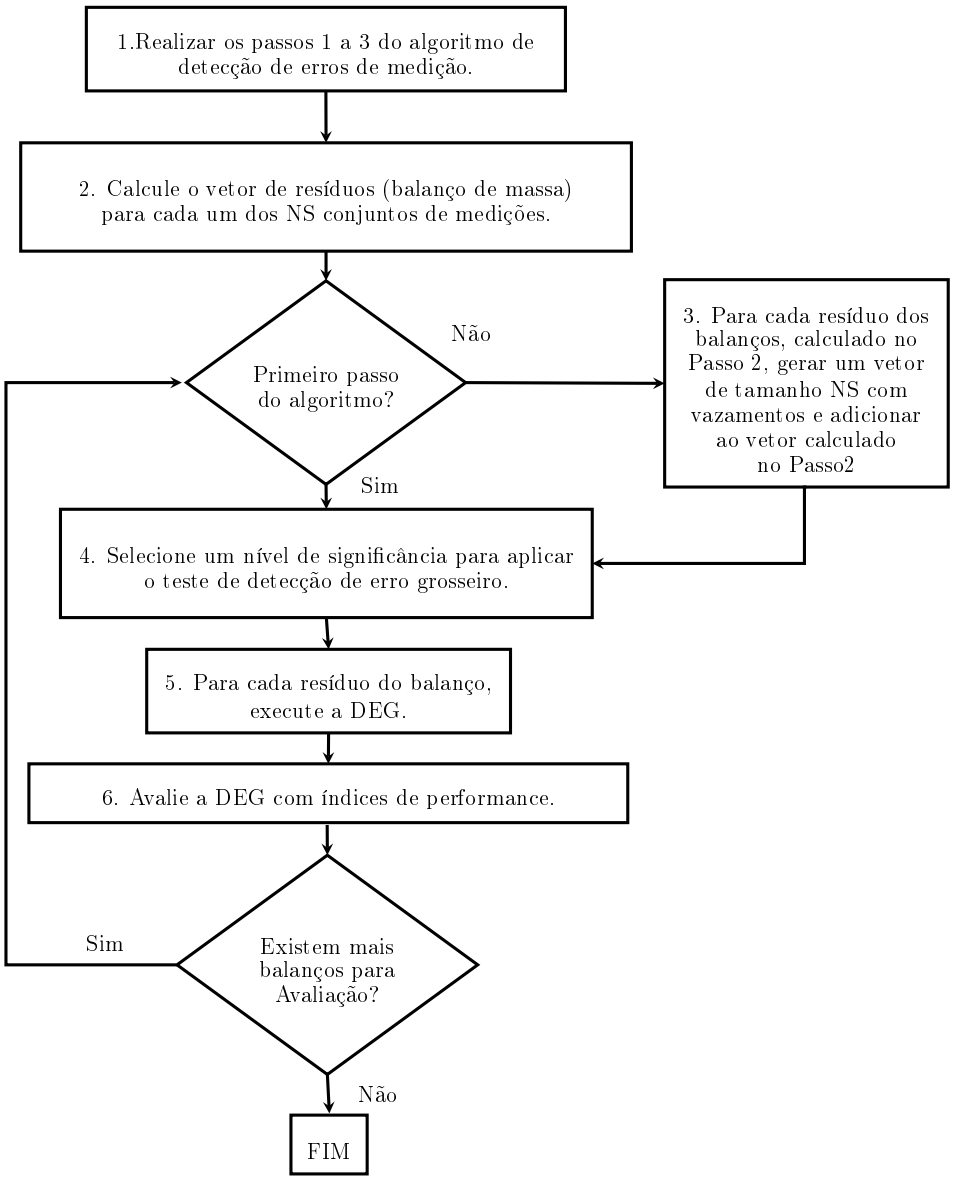
6 Executar o método de RD e DEG a ser estudado. As etapas 3.1 à 6 devem ser repedidas NS vezes.

7 Avaliar o método de DEG em estudo através de índices de performance usando os resultados de todas as NS simulações.

No presente trabalho, os fatores, l e up foram selecionados como $5.u$ e $9.u$, respectivamente.

Também foi realizada simulações para aplicação do Teste Nodal, NT, apresentado na Figura 3.7, que é ligeiramente diferente do procedimento anterior, devido ao fato de que não é necessário aplicar o procedimento de reconciliação de dados à cada simulação, visto que os testes estatísticos são aplicados somente aos resíduos brutos.

Figura 3.7: Algoritmo para avaliação das técnicas de DEG para vazamentos em equipamentos.



Fonte: Próprio autor

As etapas detalhadas do algoritmo apresentado na Figura 3.7 são assim descritas.

1 Executar os passos 1 a 3 do algoritmo apresentado na Figura 3.7.

2 Para cada simulação, NS, use o conjunto de medições ($\mathbf{Y}_{ns}$) e calcule, para cada equipamento o resíduo, denominado $r_{j,ns}$, ou vetor dos resíduos, $\mathbf{R}_{ns}$.

2.1 Se for a primeira execução do laço, ir para a Etapa 4, caso contrário ir para a Etapa 3.

3 Para cada equipamento, gerar um vetor de erros grosseiros de tamanho NS, denominado $\delta^r_{j,ns}$. O sinal do erro grosseiro é sempre negativo, por se tratar de um vazamento. A magnitude do vazamento deve ser escolhida aleatoriamente entre os limites inferiores e superiores: $l^r_j\,(\text{total nf}_j) \leq \delta^r_{j,ns} \leq up^r_j\,(\text{total nf}_j)$, em que l^r_j é o fator multiplicativo do limite inferior de geração de vazamentos, total nf_j é a soma total das vazões que entram e saem do equipamento e up^r_j é o fator multiplicativo do limite superior de geração de vazamentos. Em seguida, somar $r_{j,ns}$ e $\delta^r_{j,ns}$ para obter os resíduos finais (aleatório mais vazamento).

4 Escolher o nível de significância, α, para o teste estatístico. Essa significância deve ser determinada por tentativa e erro para manter o AVT1 próximo a 0,1 quando não há erros grosseiros presentes.

5 Executar a técnica de DEG em estudo. Repetir NS vezes as etapas 1 a 5.

6 Avaliar o método de DEG em estudo através de índices de performance usando os resultados de todas as NS simulações.

Para os problemas em estudo, l^r_j e up^r_j foram mantidos em 0,07 e 0,15, respectivamente.

No que diz respeito ao trabalho em questão, é possível afirmar que, um sistema de detecção de erros grosseiros satisfatório deve atender aos seguintes requisitos:

- Detectar a existência e posição do sensor com falha e seu sinal (positivo ou negativo).
- Detectar a existência e posição do vazamento.

- Não detectar erros aleatórios como falhas (falso positivo).
- Não detectar falso positivos em outras correntes.
- Fornecer uma estimativa da magnitude do erro.

Nesse sentido, conforme apresentado no Capítulo 2, serão utilizados para avaliar a performance dos testes de detecção de erros de grosseiros (NARASIMHAN; JORDACHE, 2000) os seguintes índices: Poder Total(OP), Média do Valor dos Erros do Tipo 1 (AVT1), Seletividade e Erro Médio de Estimação (AEE).

Como, na metodologia proposta, os resultados das medições e resíduos dos balanços dos equipamentos partem de dados preditos por um modelo, algumas definições são propostas para tornar equivalentes os índices de performance de modelos preditivos com os utilizados pela literatura, acima descritos. No conjunto de dados de teste com erros grosseiros (medição e vazamentos) são separados os dados de todas as correntes. O menor valor do erro grosseiro predito e do vazamento passa a ser o critério de corte de falso positivo, ou seja, no conjunto de erros puramente aleatórios, o valor de erro predito que ultrapassar o menor valor do erro grosseiro predito passa a ser considerado um falso positivo e contabilizará a Média do Valor dos Erros do Tipo 1, ATV1. Já a contabilização de um erro grosseiro corretamente identificado (usado na avaliação do Poder Total e Seletividade) é dado quando o valor do erro grosseiro ultrapassa o maior e o menor valor do erro puramente aleatório (pois deve ser também considerado o sinal do mesmo). O Erro Médio de Estimação é aplicado somente em casos onde os erros grosseiros são corretamente identificados.

Os resultados da metodologia são apresentados no próximo capítulo.

Capítulo 4

Resultados e Discussões

4.1 Obtenção dos resultados

Conforme descrito no capítulo anterior, os resultados da metodologia de DEG propostos foram obtidos, resumidamente descrito, em três etapas:

- Geração de dados;
- Ajuste dos modelos PLS para DEG;
- Teste dos modelos PLS.

Os procedimentos descritos no capítulo anterior foram implementados utilizando o programa de computador Scilab (INRIA, 2010). O Scilab possui uma série de comandos e bibliotecas estatísticas e permite a execução dos mesmos em série através de *scripts*, facilitando a criação de algoritmos e execução de tarefas iterativas. Além disso, o Scilab é capaz de se comunicar com o otimizador Ipopt (WÄCHTER; BIEGLER, 2006). Entre as diversas vantagens do aplicativo, é importante citar a sua gratuidade além de possuir código-fonte aberto. O software utilizado para estimação de parâmetros dos modelos PLS foi o VRAnalyst 1.0.23.

Para execução destas etapas, foram utilizados 2 computadores diferentes. O primeiro, onde a maioria dos testes foi executado, um computador PC Core 2 Quad Q8400 2.7 Mhz com 4 GB de memória RAM, utilizando como sistema operacional Ubuntu Linux 10.04. O software utilizado para gerar os dados foi o Scilab versão 5.4.0. O segundo computador, um laptop PC Core 2 Duo 1.7Mhz com 3 GB de memória RAM utilizando como sistema operacional Ubuntu Linux 10.04. O software

utilizado para gerar os dados foi o Scilab versão 5.3.3. Cabe ressaltar que nenhum recurso de paralelização foi utilizado para redução de tempo computacional.

Inicialmente, foram gerados para teste do método proposto, cenários com erros únicos e múltiplos, sendo primeiro apresentado o potencial de predição de erros únicos seguido dos erros múltiplos. Para os erros múltiplos, devido à possibilidade muito grande de combinações (de número e posições), optou-se por selecionar apenas um cenário para a avaliação do método.

Para comparação dos resultados do método proposto, foi escolhido o GLR para teste de erros de medição e vazamentos e Teste Nodal para vazamentos. O Teste de Medida não foi utilizado devido a sua equivalência ao GLR para erros únicos (NARASIMHAN; MAH, 1987; NARASIMHAN; JORDACHE, 2000). Outro motivo par uso do GLR, é a capacidade de estimar a magnitude do erro grosseiro, tanto para medições quanto para vazamentos. O GLR também tem sido usado por alguns autores como teste de referência para comparação de novas técnicas de DEG, devido ao seu potencial (KELLER et al., 1994; SINGTH et al., 2001; SÁNCHEZ et al., 1996; SÁNCHEZ et al., 1999). Para geração de dados usados como comparação para a metodologia proposta (GLR e NT), foi utilizado o procedimento conforme apresentado na Figura 3.6, para erros de medição, e Figura 3.7 para vazamentos, utilizando as incertezas-padrão originais de cada problema. Para avaliação do OP e AEE foram utilizadas 2500 simulações, NS, para cada erro grosseiro de medição ou vazamento. Tal número se mostrou adequado pois, de acordo com Rosenberg et al. (1987), o Poder do Teste estatístico é aproximado pelo OP com um erro menor que 0.02 em aproximadamente 95% das simulações. Os erros grosseiros de medição foram gerados aleatoriamente na faixa de cinco a nove vezes a incerteza-padrão (originais) com o sinal positivo ou negativo com a mesma probabilidade de ocorrer (distribuição uniforme). A faixa de vazamentos gerados ficou entre 0,07 e 0,15 vezes a vazão total que passa pelo equipamento.

Para o ajuste do modelo proposto foram utilizados aproximadamente 180 dados para cada faixa de vazão de corrente mais 180 dados para cada faixa de vazamento e mais 180 dados para erros aleatórios. O número de pontos usado para a estimação

de parâmetros do modelo é dado pela seguinte expressão:

$$\frac{bias_{upper} - bias_{lower} + 1}{\Delta_{bias}}.nc.NS + ne\frac{leak_{upper} - leak_{lower} + 1}{\Delta_{leak}}.NS + NS, \quad (4.1)$$

em que $bias_{upper}$ é o limite superior para o erro grosseiro de medição, $bias_{lower}$ é o limite inferior para o erro grosseiro de medição, Δ_{bias} é o intervalo entre os valores superiores e inferiores dos erros de medição, nc é o número de correntes, ne o número de equipamentos, NS o número de simulações, $leak_{upper}$ o limite superior para os vazamentos, $leak_{lower}$ o limite inferior para os vazamentos e Δ_{leak} o intervalo entre os limites inferiores e superiores para os vazamentos. No estudo em questão o erro grosseiro de medição foi variado entre a faixa de cinco e nove vezes as incertezas-padrão ($bias_{lower}$ e $bias_{upper}$, respectivamente) com intervalo de uma unidade de erro grosseiros (Δ_{bias}). O vazamento foi simulado no intervalo entre 7% a 15% da vazão total que passa pelo equipamento ($leak_{lower}$ e $leak_{upper}$, respectivamente) com intervalo de 2% (Δ_{leak}). Utilizando como exemplo o Problema 1, com 3 correntes e 1 equipamento, tem-se:

$$(9 - 5 + 1).3.180 + 1.5.180 + 180 = 3780, \quad (4.2)$$

dessa forma, tem-se quantidade de dados suficiente para evitar o fenômeno de sobre-ajuste ou *overfitting*.

O número de pontos usado no conjunto de teste é dado pela seguinte equação:

$$\frac{bias_{upper} - bias_{lower} + 1}{\Delta_{bias}}.nc.NST + ne\frac{leak_{upper} - leak_{lower} + 1}{\Delta_{leak}}.NST, \quad (4.3)$$

em que NST o número de simulações no conjunto de testes. No estudo em questão, 10 simulações (NST) de cada cenário foram utilizadas para teste, variando o erro grosseiro de medição na faixa entre cinco e nove vezes a respectiva incerteza-padrão ($bias_{lower}$ e $bias_{upper}$, respectivamente) com intervalo de uma unidade de erro grosseiros (Δ_{bias}). O vazamento foi simulado no intervalo entre 7% a 15% da vazão total que passa pelo equipamento ($leak_{lower}$ e $leak_{upper}$, respectivamente) com intervalo de 2% (Δ_{leak}). Tendo como exemplo o Problema 1, com 3 correntes, 1 equipamento, tem-se:

$(9 - 5 + 1).3.10 + (1.5 + 1).10 = 210.$

Para erros múltiplos segue a fórmula:

$$\frac{bias_{upper} - bias_{lower} + 1}{\Delta_{bias}}.nc, \tag{4.4}$$

ou seja, para o problema 1 tem-se:
$(9 - 5 + 1).3 = 18.$

Devido aos resultados positivos do método proposto e, para facilitar a leitura do texto, optou-se por apresentar e discutir os resultados dos problemas somente onde o método proposto apresentou alguma peculiaridade. Nos demais casos, representados pelos problemas 4, 7, 10, 11, 12, 13 e 15, os resultados são apresentados em forma de tabela no Apêndice D.

Os resultados com erros múltiplos foram comparados com dados de literatura com os valores numéricos obtidos diretamente de Sánchez et al. (1999). Nos resultados com erros múltiplos, também são discutidos alguns casos onde o método proposto apresentou baixa seletividade.

4.2 Erros únicos

Inicialmente são apresentados na Tabela 4.1 os dados do AVT1, tempo computacional para geração e seleção de dados e coeficiente de determinação dos modelos gerados. O coeficiente de determinação ou R^2 é uma medida da qualidade do ajuste de um modelo em relação aos valores observados. Este coeficiente pode variar entre 0 e 1 indicando o quanto o modelo consegue explicar os dados observados de saída, em nosso caso, os erros grosseiros. Quanto mais próximo de 1, melhor a explicação das variáveis de saída pelo modelo.

A configuração da significância estatística para todos os testes é realizada para configurar o AVT1 em aproximadamente 0,1, ou seja, 10% de falso-positivos no conjunto de dados com erros puramente aleatórios. Isso garante que diversos métodos possam ser comparados em uma base comum (NARASIMHAN; JORDACHE, 2000; IORDACHE et al., 1985; ROSENBERG et al., 1987). Também foi registrado o tempo computacional para fins de referência, visto que alguns problemas foram executados em computadores com características bem distintas. Como a definição

de AVT1 para o método proposto se baseia nos resultados de modelos preditos (o menor valor do erro grosseiro predito passa a ser o valor de corte para determinar o falso-positivo), o AVT1 do presente método não pode ser comparado com os demais. Como é possível observar na Tabela 4.1, as significâncias estatísticas foram configuradas para manter a AVT1 próximas a 0,1 (falso-positivos), garantindo assim, a mesma base de comparação entre os métodos (exceto a AVT1 do método proposto). Além disso, como é esperado, o tempo computacional tem a tendência de aumentar conforme o número de correntes e equipamentos. O R^2 apresentou um valor médio de 0,8 para todos os problemas, sendo considerado razoável, exceto para os problemas 8 e 16 em que o R^2 ficou próximo a 0,62 para ambos.

Tabela 4.1: Resultados do AVT1, tempo computacional e R^2.

Problema	AVT1 GT	AVT1 GLR-MT	AVT1 GLR-NT	AVT1 NT	AVT1 Proposto	Tempo (s)	R^2
P1	0,087	0,091	0,087	0,091	0	14,8	
P2	0,113	0,111	0,102	0,116	0	18,7	0,86
P3	0,098	0,091	0,093	0,101	0	735,3	0,85
P3_2	0,102	0,093	0,094	0,102	0	584,8	0,85
P4	0,102	0,110	0,113	0,099	0	168,2	0,88
P5	0,103	0,109	0,109	0,093	0	18,7	0,86
P6	0,093	0,089	0,100	0,090	0	34,7	0,88
P7	0,096	0,102	0,101	0,097	0	494,8	0,70
P7_2	0,094	0,099	0,113	0,108	0	1185,2	0,81
P8	0,101	0,108	0,106	0,107	0	2719,5	0,63
P9	0,102	0,099	0,112	0,104	0	1549,9	0,80
P10	0,105	0,105	0,104	0,118	0	762,0	0,83
P11	0,087	0,094	0,093	0,105	0	557,0	0,82
P12	0,101	0,098	0,104	0,091	0	1796,9	0,80
P13	0.113	0.108	0,112	0,117	0	1823,0	0,80
P14	0,103	0,104	0,088	0,102	0	2252,1	0,78
P14_2	0,103	0,104	0,088	0,102	0	4059,1	0,74
P15	0,095	0,101	0,098	0,096	0	4169,8	0,85
P16	0,101	0,149	0,098	0,112	0	20079,5	0,62
Média	0,1	0,1	0,1	0,1	0	2264	0,80

Para os erros únicos, as tabelas de resultados tem o formato conforme apresentado na Tabela 4.2, onde são apresentados os dados das correntes, como: nome, valor nominal, incerteza-padrão. Na próxima coluna é apresentada a detectabilidade, que é uma medida da *facilidade* de detectar erros grosseiros nas medições (Charpentier et al. (1991), apud Narasimhan e Jordache (2000)). Geralmente, o OP de técnicas como o MT e GLR apresentam uma correlação com a detectabilidade (a detectabilidade é calculada pela Equação 2.5.8), ou seja, quando o OP é elevado,

a detectabilidade também é, sendo o contrário também observado. Em seguida são apresentados o OP do método GLR e do proposto, assim como o erro médio de estimação, AEE. Na última coluna é apresentado a Seletividade, que indica se algum erro grosseiro também foi encontrado em outra corrente. Cabe ressaltar que uma Seletividade menor que 1 indica que também foram encontrados erros grosseiros em outras correntes, porém, se o OP é 1, indica que o maior erro grosseiro foi na corrente onde o tal erro foi adicionado. As unidades de medição do valor nominal de vazão para as correntes e as incertezas-padrão são expressas em $t.h^{-1}$ para todos os problemas, suprimidas das tabelas para redução de espaço.

Tabela 4.2: Exemplo de tabela de resultados para os problemas com erros únicos de medição.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	8,50	0,277	0,846	0	-	1	2,202	0,970

Para vazamentos, os dados são apresentados em tabelas semelhantes à Tabela 4.3

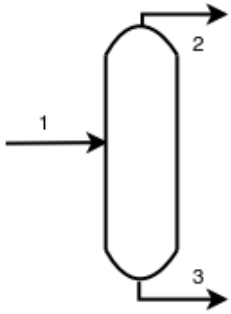
Tabela 4.3: Exemplo de tabela de resultados para vazamentos

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,997	0,997	0	0,84	0,040	1

4.2.1 Problema 1

Recordando que o Problema 1 trata de um divisor simples entre duas correntes, conforme Figura 4.1. Este é um equipamento bastante comum em indústrias de processos e, conforme discutido no Capítulo 3, métodos como o MT e o GLR não conseguem detectar erros em suas correntes. Esse é um dos principais motivos pelo qual grandes diagramas não são divididos em diagramas menores para facilitar a

Figura 4.1: Fluxograma de processo do Problema 1.



Fonte: Próprio autor

detecção ou identificação de erros grosseiros em medições: a divisão de um diagrama grande pode levar à presença de vários divisores, impossibilitando a detecção.

Conforme apresentado na Tabela 4.4, o método proposto apresentou resultados superiores ao método GLR, com a identificação praticamente perfeita e boa seletividade, demonstrando o potencial da técnica proposta.

Tabela 4.4: Resultados do Problema 1 para erros grosseiros em medições.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	8,50	0,277	0,846	0	-	1	2,202	0,970
2	4,50	0,145	0,443	0	-	0,978	2,227	0,813
3	4,00	0,097	0,296	0	-	1	2,315	0,890

Já para o teste de vazamento do Problema 1, o GLR se mostrou um pouco superior ao método proposto, conforme apresentado na Tabela 4.5. Conforme apresentado no Capítulo 3, para o método proposto, existe um compromisso entre o poder de teste do erro de medição e vazamento, e este pode ser ajustado a partir da configuração apropriada da incerteza-padrão da respectiva medida.

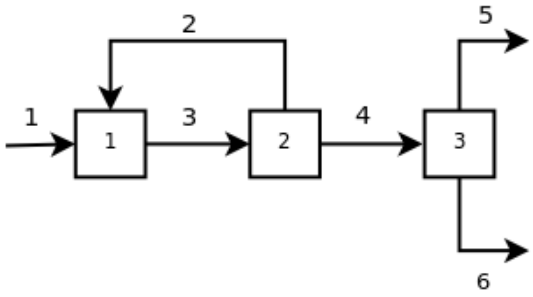
Tabela 4.5: Resultados do Problema 1 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,997	0,997	0	0,84	0,040	1

4.2.2 Problema 2

O Problema 2, conforme apresentado na Figura 4.2, possui dois pares de correntes em paralelo: 2 e 3; 5 e 6. Tal fato leva o GLR a não identificar corretamente erros grosseiros nessas 4 correntes, em contraste com a metodologia proposta que identifica corretamente tais correntes com elevada seletividade, conforme apresentado na Tabela 4.6. Já para a detecção de vazamentos, ambas as metodologias apresentaram resultados equivalentes, conforme apresentado na Tabela 4.7.

Figura 4.2: Fluxograma de processo do Problema 2.



Fonte: Próprio autor

Tabela 4.6: Resultados do Problema 2 para erros em medições.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	11	0,179	0,791	0,981	0,090	1,000	0,343	0,993
2	10	0,161	0,404	0,000	-	0,989	1,503	0,998
3	21	0,346	0,869	0,000	-	0,980	0,980	0,996
4	11	0,182	0,798	0,984	0,087	1,000	0,659	1,000
5	5,5	0,228	0,798	0,000	-	0,940	1,045	0,993
6	5,5	0,122	0,429	0,000	-	1,000	1,048	0,993

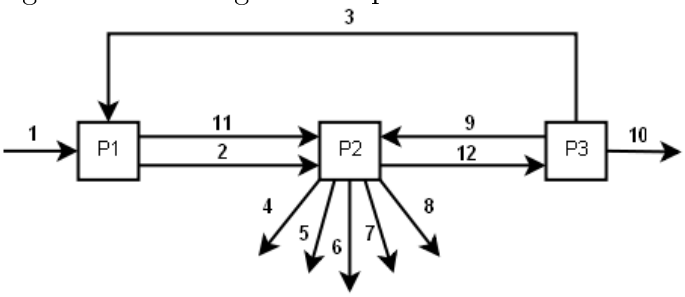
Tabela 4.7: Resultados do Problema 2 para vazamentos.

Equipamento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
	1	1	6,677	1	0,005	1
	1	1	6,810	1	0,0001	1
	1	1	4,848	1	0,001	1

4.2.3 Problema 3

O Problema 3, conforme apresentado na Figura 4.3, possui 3 conjuntos de correntes em paralelo: 2 e 11; 9 e 12; 4, 5, 6, 7, 8. Tal fato leva o GLR a não identificar corretamente erros grosseiros nessas correntes. Já a metodologia proposta apresentou um resultado bem superior, porém, as correntes 4 e 5 apresentaram OP de zero e a corrente 6 um OP de 0,76, considerado baixo, conforme apresentado na Tabela 4.8.

Figura 4.3: Fluxograma de processo do Problema 3.



Fonte: Próprio autor

Tabela 4.8: Resultados do Problema 3 para erros em medições.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
x1	189,980	2,141	0,900	0,994	0,058	1,000	1,436	0,963
x2	174,600	1,948	0,875	0,000	-6,660	1,000	1,218	0,973
x3	3,139	0,034	0,045	0,017	7,821	1,000	0,980	0,981
x4	32,770	0,345	0,183	0,000	-	-	7,526	-
x5	33,470	0,400	0,212	0,000	-	-	7,406	-
x6	7,250	0,087	0,046	0,000	-	0,793	1,854	1,000
x7	0,316	0,004	0,002	0,000	-	1,000	0,448	1,000
x8	92,376	1,059	0,560	0,000	-	1,000	0,131	1,000
x9	28,629	0,361	0,452	0,000	-	1,000	1,030	0,987
x10	23,800	0,300	0,400	0,405	0,246	1,000	0,821	1,000
x11	18,526	0,200	0,090	0,000	-	1,000	0,246	0,987
x12	55,568	0,645	0,810	0,000	-	1,000	1,299	1,000

Já para a detecção de vazamentos, ambas metodologias apresentaram resultados equivalentes, conforme apresentado na Tabela 4.9.

Tabela 4.9: Resultados do Problema 3 para vazamentos.

Equipamento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1,00	1,000	0,033	1	0,007	1
2	1,00	1,000	0,026	1	0,004	1
3	1,00	1,000	0,019	1	0,001	1

Uma alternativa para melhorar a identificação de erros grosseiros nas correntes 4 e 5 foi a de reduzir a incerteza-padrão de 1 $t.h^{-1}$ para 0,1 $t.h^{-1}$ para todas as correntes. Após esse procedimento, foi possível identificar corretamente erros de medição nessas duas correntes, conforme apresentado na Tabela 4.10. Cabe ressaltar que esse procedimento não comprometeu a detecção de vazamentos, conforme apresentado na Tabela 4.11

Tabela 4.10: Resultados do Problema 3 para erros de medição com incerteza-padrão igual a 0,1 para todas as correntes.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
x1	189,980	2,141	0,900	0,996	0,058	1,000	0,789	1,000
x2	174,600	1,948	0,875	0,000	-	1,000	1,387	1,000
x3	3,139	0,034	0,045	0,015	7,648	1,000	0,654	1,000
x4	32,770	0,345	0,183	0,000	-	1,000	0,249	1,000
x5	33,470	0,400	0,212	0,000	-	1,000	0,072	1,000
x6	7,250	0,087	0,046	0,000	-	1,000	1,296	1,000
x7	0,316	0,004	0,002	0,000	-	1,000	0,791	1,000
x8	92,376	1,059	0,560	0,000	-	1,000	1,315	0,995
x9	28,629	0,361	0,452	0,000	-	1,000	0,556	1,000
x10	23,800	0,300	0,400	0,386	0,238	1,000	1,272	1,000
x11	18,526	0,200	0,090	0,000	-	1,000	1,654	0,989
x12	55,568	0,645	0,810	0,000	-	1,000	0,243	0,997

Nota: A incerteza-padrão aqui apresentada refere-se ao problema original.

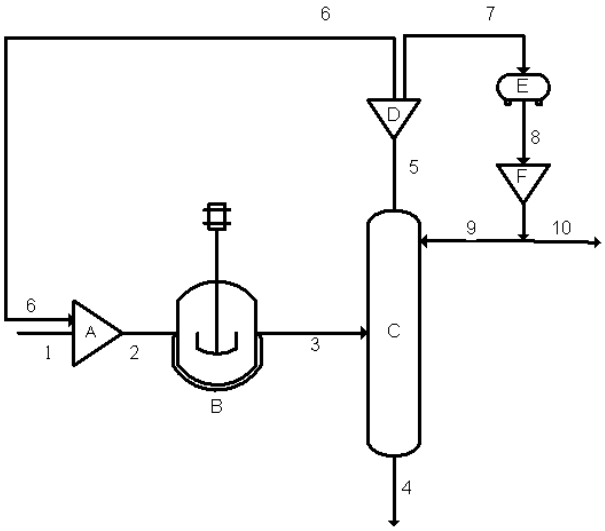
Tabela 4.11: Resultados do Problema 3 para vazamentos com incerteza-padrão igual a 0,1.

Equipmento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1,00	1,000	0,033	1	0,007	1
2	1,00	1,000	0,026	1	0,008	1
3	1,00	1,000	0,019	1	0,001	1

4.2.4 Problema 7

O Problema 7 é apresentado na Figura 4.4.

Figura 4.4: Fluxograma de processo do Problema 7.



Fonte: Próprio autor

A detecção de erros grosseiros de medição do presente problema pela metodologia proposta apresentou bons resultados, conforme apresentado na Tabela 4.12. Já para os vazamentos, o equipamento F não teve seus vazamentos identificados, conforme apresentado na Tabela 4.13. Tal fato ocorreu pois, quando a incerteza-padrão de uma determinada corrente aumenta, o valor do respectivo elemento da diagonal da matriz de variância das medidas, Σ, aumenta. Consequentemente, a matriz de variância-covariância dos balanços, $\mathbf{V}$ aumenta, conforme a Equação (2.11), aqui

reapresentada:

$$\mathbf{V} = cov(\mathbf{R}) = \mathbf{A} \cdot \Sigma \cdot \mathbf{A}^T,$$

assim, a estatística de teste $z_{r,j}$, apresentada na Equação (2.13), diminui pois o termo $V_{j,j}$ encontra-se no denominador da expressão, aqui reapresentada:

$$z_{r,j} = \frac{|r_j|}{\sqrt{V_{j,j}}}.$$

Com a diminuição de $z_{r,j}$, os erros grosseiros adicionados se comportam como erros aleatórios e não são incorporados no conjunto de dados com erros grosseiros, comprometendo o OP do método para detecção de vazamentos. Tal fato não ocorre com o GLR para a detecção de vazamentos, pois em tal teste foram usados as incertezas-padrão originais, com magnitude menor.

Tabela 4.12: Resultados do Problema 7 para erros de medição.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	1,000	0,810	0,985	0,086	1,000	0,648	1,000
2	75	1,000	0,815	0,984	0,084	1,000	1,066	1,000
3	75	1,000	0,815	0,984	0,084	1,000	0,996	0,971
4	48	1,000	0,802	0,988	0,087	1,000	1,166	0,971
5	30	1,000	0,805	0,983	0,088	1,000	0,978	0,971
6	25	1,000	0,810	0,988	0,087	1,000	1,188	1,000
7	5	0,387	0,806	0,981	0,088	1,000	1,460	0,996
8	5	0,387	0,806	0,982	0,086	1,000	1,303	0,971
9	3	0,316	0,617	0,735	0,140	1,000	1,289	1,000
10	2	0,316	0,623	0,753	0,135	1,000	0,921	0,971

Tabela 4.13: Resultados do Problema 7 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
A	1	1,000	0,062	1	0,002	0,971
B	1	1,000	0,052	1	0,001	0,971
C	1	1,000	0,061	1	0,0003	0,971
D	0,9336	0,982	0,141	1,000	0,000	0,971
E	0,37	0,512	0,158	1,000	0,001	0,972
F	0,3252	0,438	0,141	0,000	1,200	0,000

Semelhante ao realizado no Problema 3, tal limitação foi contornada reduzindo a incerteza-padrão de todas as correntes para 0,1 t.h^{-1}, aumentando o sensibilidade para detecção de vazamentos com o aumento da estatística de teste $z_{r,j}$. Os resultados da detecção de erros grosseiros de medição e vazamentos são apresentados na Tabela 4.14 e Tabela 4.15, respectivamente.

Tabela 4.14: Resultados dos erros grosseiros de medição para o Problema 7 usando incerteza-padrão de 0,1 t.h^{-1} para todas as correntes.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	1,000	0,810	0,985	0,087	1,000	1,073	1,000
2	75	1,000	0,815	0,986	0,081	1,000	0,792	1,000
3	75	1,000	0,815	0,985	0,082	1,000	1,498	1,000
4	48	1,000	0,802	0,983	0,088	1,000	1,114	1,000
5	30	1,000	0,805	0,988	0,087	1,000	1,015	1,000
6	25	1,000	0,810	0,988	0,083	1,000	1,079	0,954
7	5	0,387	0,806	0,980	0,088	1,000	1,538	1,000
8	5	0,387	0,806	0,984	0,087	1,000	1,252	1,000
9	3	0,316	0,617	0,749	0,134	1,000	1,534	1,000
10	2	0,316	0,623	0,755	0,138	1,000	1,663	0,995

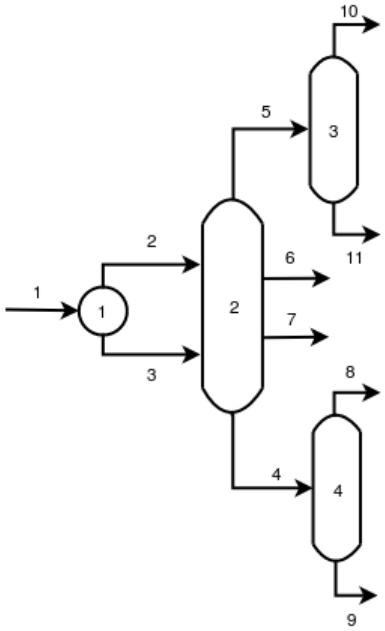
Tabela 4.15: Resultados de vazamentos para o Problema 7 usando incerteza-padrão de 0,1 t.h^{-1} para todas as correntes.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
A	1	1,000	0,061	1	0,003	1
B	1	1,000	0,051	1,000	0,003	1,000
C	1	1,000	0,060	1,000	0,012	1,000
D	0,937	0,982	0,137	1,000	0,012	1,000
E	0,374	0,506	0,155	1,000	0,004	1,000
F	0,325	0,439	0,149	1,000	0,004	1

4.2.5 Problema 8

O problema 8, apresentado na Figura 4.5, possui 4 conjuntos de correntes paralelas, entre elas: 2 e 3; 6 e 7; 8 e 9; 10 e 11. Tal fato leva o GLR à não detectar erros grosseiros nessas correntes. A metodologia proposta apresentou bons resultados, exceto para a corrente 6, que apresentou um baixo OP, conforme apresentado na Tabela 4.16. A melhoria do poder OP para a corrente 6 pode ser obtida através de testes, avaliando o efeito da incerteza-padrão sobre seu OP. Nesse procedimento é necessário avaliar se tal procedimento não afetará o OP da detecção de vazamentos. Já para a detecção de vazamentos, ambas metodologias apresentaram resultados equivalentes, conforme apresentado na Tabela 4.17.

Figura 4.5: Fluxograma de processo do Problema 8.



Fonte: Próprio autor

Tabela 4.16: Resultados do Problema 8 para erros de medição.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	3600	111,210	0,915	0,995	0,052	1,000	1,784	0,984
2	1850	57,000	0,595	0,000	-	1,000	1,393	1,000
3	1750	54,210	0,566	0,000	-	1,000	0,813	1,000
4	2837	87,300	0,859	0,987	0,070	1,000	1,518	1,000
5	730	22,110	0,748	0,802	0,103	1,000	1,886	1,000
6	25	0,780	0,009	0,000	-	0,640	5,508	0,973
7	8	0,228	0,003	0,000	-	0,960	4,510	0,989
8	137	3,150	0,032	0,000	-	1,000	4,938	0,960
9	2700	84,960	0,850	0,000	-	1,000	2,013	0,984
10	58	1,710	0,058	0,000	-	1,000	4,794	0,955
11	672	20,04	0,682	0,000	-	1,000	1,733	0,989

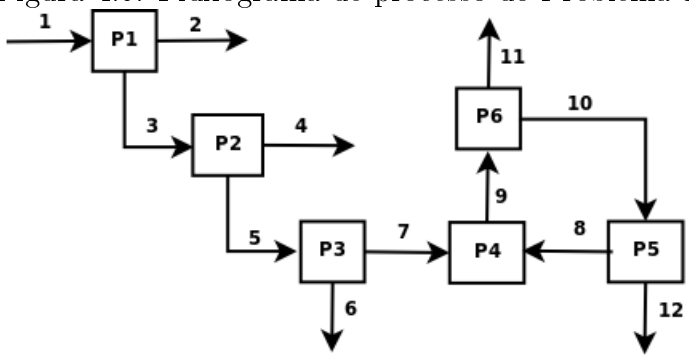
Tabela 4.17: Resultados do Problema 8 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,991	0,995	0,066	1,000	0,040	1,000
2	0,998	0,999	0,086	1,000	0,074	1,000
3	0,988	0,987	0,028	1,000	0,038	1,000
4	0,981	0,993	0,095	1,000	0,027	1,000

4.2.6 Problema 9

O Problema 9 é bastante peculiar no que diz respeito à estrutura de seu diagrama, sendo este apresentado na Figura 4.6. Neste fluxograma, encontram-se correntes paralelas, 1 e 2, e reciclo, formado pelas correntes 8, 9 e 10. No problema original, esse reciclo reduz o OP do GLR, pois a corrente 8 apresenta maior precisão que as demais correntes do reciclo, e, após a reconciliação, os erros simulados nessa corrente acabam se espalhando para as demais, conforme apresentado na Tabela 4.18. Já a metodologia proposta remove esse efeito, pois faz uso das incertezas-padrão unitárias e uso de uma função objetivo robusta, nesse caso, o estimador de Cauchy. Já para a detecção de vazamentos, ambas metodologias apresentaram resultados equivalentes, conforme apresentado na Tabela 4.19.

Figura 4.6: Fluxograma de processo do Problema 9.



Fonte: Próprio autor

Tabela 4.18: Resultados do Problema 9 para erros de medição.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	230	37,575	0,998	0,000	-	1,000	1,103	1,000
2	21	1,08	0,029	0,000	-	1,000	1,491	1,000
3	209	5	0,890	0,665	0,060	1,000	1,511	1,000
4	35	1,825	0,327	0,237	0,367	1,000	1,164	1,000
5	174	2	0,522	0,725	0,174	1,000	1,433	0,999
6	15	0,88	0,193	0,095	1,060	1,000	1,877	0,993
7	159	7,245	0,967	0,997	0,031	1,000	1,677	1,000
8	50	1	0,349	0,370	0,320	1,000	1,288	0,999
9	209	5	0,916	0,974	0,052	1,000	1,501	0,999
10	94	2	0,616	0,548	0,145	1,000	0,928	0,991
11	115	18,1	0,991	0,999	0,016	1,000	1,363	1,000
12	44	2,385	0,733	0,638	0,105	1,000	1,144	0,999

Tabela 4.19: Resultados do Problema 9 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,159	0,161	0,033	1,000	0,014	1,000
2	1,000	1,000	0,020	1,000	0,040	1,000
3	0,973	0,996	0,135	1,000	0,008	1,000
4	0,976	0,993	0,134	1,000	0,004	1,000
5	0,994	0,994	0,014	1,000	0,033	1,000
6	0,546	0,568	0,065	1,000	0,021	1,000

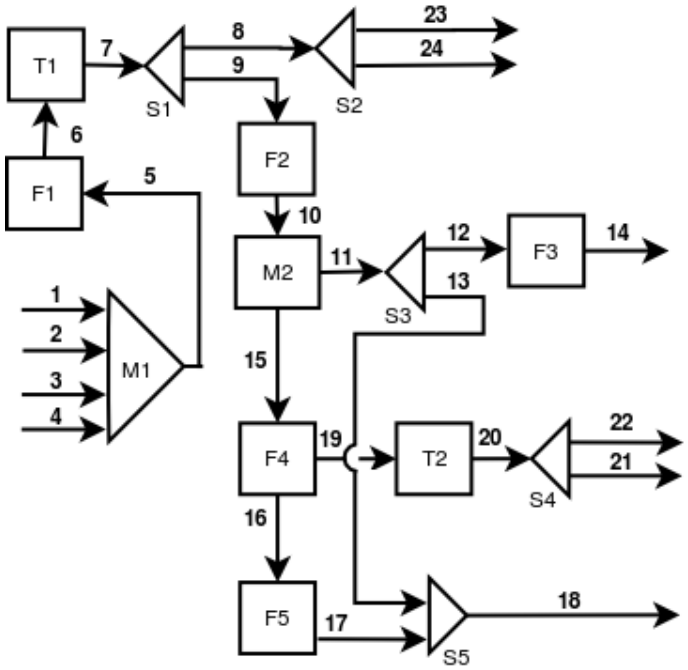
4.2.7 Problema 14

O Problema 14 é baseado em um balanço hídrico real de uma planta de tratamento de água para uma fábrica de fertilizantes e é apresentado na Figura 4.7.

Este é um fluxograma maior que os anteriormente apresentados com 24 correntes e 14 equipamentos. Nesse fluxograma foi aplicada a metodologia do QI para avaliação das incertezas de medição para fins de comparação com o método GLR. Inicialmente, a partir do balanço de massa original e de informações da equipe de responsáveis pela planta, foi organizada as atribuições de QI para cada corrente do processo. Em seguida, a partir do uso da Equação (2.9) foram obtidas as respectivas incertezas-padrão, conforme apresentado na Tabela 4.20.

Em seguida, foi aplicado o método proposto para DEG. Os resultados para os erros grosseiros de medição são apresentados na Tabela 4.21. É possível observar que, como nos casos anteriormente apresentados, as correntes paralelas, 1, 2, 3 e 4, não são corretamente identificadas pelo GLR. Nos resultados do método proposto com incertezas-padrão unitárias, é possível observar que, conforme o ocorrido no Problema 3, quando um número elevado de corrente paralelas estão presentes, o método não identificou corretamente os erros grosseiros nas correntes 3 e 4. Tal dificuldade foi contornada aumentado a incerteza-padrão para 5 t.h^{-1} em todas as correntes, conforme apresentado na Tabela 4.22.

Figura 4.7: Fluxograma de processo do Problema 14.



Fonte: Próprio autor

Tabela 4.20: Conversão de dados de QI para incertezas-padrão para uso da função objetivo de Cauchy.

Corrente	Valor Nominal	QI	Desvio Padrão
1	50	10	0,533
2	150	10	1,460
3	140	10	1,541
4	140	10	1,451
5	480	6	7,611
6	480	6	7,611
7	480	6	8,335
8	220	10	2,136
9	260	10	2,865
10	260	10	2,865
11	85,8	10	0,827
12	15,8	3	0,441
13	70	9	0,772
14	15,8	3	0,441
15	174,2	10	2,038
16	3,484	3	0,095
17	3,484	3	0,095
18	73,484	3	2,852
19	170,716	10	1,858
20	170,716	10	1,999
21	55	10	0,532
22	115,716	10	1,467
23	200	10	1,977
24	20	10	0,160

Tabela 4.21: Resultados do Problema 14 para erros de medição.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	0,533	0,173	0,000	-	1,000	0,057	0,996
2	150	1,460	0,473	0,000	-	1,000	0,416	0,993
3	140	1,541	0,499	0,000	-	0,000	9,000	0,000
4	140	1,451	0,470	0,000	-	0,000	8,000	0,000
5	480	7,611	0,983	0,997	0,022	1,000	2,058	1,000
6	480	7,611	0,983	0,999	0,022	1,000	1,975	0,993
7	480	8,335	0,986	1,000	0,020	1,000	1,918	1,000
8	220	2,136	0,810	0,974	0,085	1,000	2,090	0,994
9	260	2,865	0,949	0,996	0,039	1,000	2,528	0,999
10	260	2,865	0,949	0,998	0,039	1,000	1,976	0,996
11	85,8	0,827	0,745	0,927	0,105	1,000	2,144	0,996
12	15,8	0,441	0,735	0,895	0,103	1,000	3,154	0,988
13	70	0,772	0,701	0,897	0,113	1,000	2,655	0,996
14	15,8	0,441	0,735	0,882	0,104	0,960	3,460	0,987
15	174,2	2,038	0,918	0,996	0,051	1,000	2,711	1,000
16	3,484	0,095	0,708	0,573	0,117	1,000	3,686	0,993
17	3,484	0,095	0,708	0,573	0,113	1,000	3,517	1,000
18	73,484	2,852	0,981	0,997	0,023	1,000	2,508	0,993
19	170,716	1,858	0,900	0,994	0,058	1,000	2,402	0,996
20	170,716	1,999	0,914	0,998	0,053	1,000	2,143	1,000
21	55	0,532	0,292	0,000	-	1,000	2,425	0,996
22	115,716	1,467	0,804	0,000	-	1,000	1,924	0,991
23	200	1,977	0,773	0,000	-	0,960	2,636	0,987
24	20	0,160	0,062	0,000	-	0,860	3,308	1,000

Tabela 4.22: Resultados do Problema 14 para erros de medição com incerteza-padrão igual a 5 t.h^{-1}.

Corrente	Valor Nominal	Desvio Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	0,533	0,173	0,000	-	1,000	1,062	0,997
2	150	1,460	0,473	0,000	-	0,989	0,987	0,988
3	140	1,541	0,499	0,000	-	1,000	1,109	0,990
4	140	1,451	0,470	0,000	-	1,000	0,876	0,991
5	480	7,611	0,983	0,997	0,022	1,000	1,974	0,999
6	480	7,611	0,983	0,999	0,022	1,000	1,716	0,996
7	480	8,335	0,986	1,000	0,020	1,000	1,825	0,988
8	220	2,136	0,810	0,974	0,085	1,000	1,745	0,995
9	260	2,865	0,949	0,996	0,039	1,000	2,291	0,988
10	260	2,865	0,949	0,998	0,039	1,000	1,718	0,992
11	85,8	0,827	0,745	0,927	0,105	1,000	1,343	1,000
12	15,8	0,441	0,735	0,895	0,103	1,000	1,667	0,999
13	70	0,772	0,701	0,897	0,113	1,000	1,703	0,996
14	15,8	0,441	0,735	0,882	0,104	1,000	2,109	0,984
15	174,2	2,038	0,918	0,996	0,051	1,000	2,050	0,988
16	3,484	0,095	0,708	0,573	0,117	1,000	1,619	0,996
17	3,484	0,095	0,708	0,573	0,113	1,000	1,744	1,000
18	73,484	2,852	0,981	0,997	0,023	1,000	1,507	0,990
19	170,716	1,858	0,900	0,994	0,058	1,000	2,001	1,000
20	170,716	1,999	0,914	0,998	0,053	1,000	1,657	0,988
21	55	0,532	0,292	0,000	-	1,000	2,216	0,995
22	115,716	1,467	0,804	0,000	-	1,000	1,735	0,994
23	200	1,977	0,773	0,000	-	1,000	1,747	0,992
24	20	0,160	0,062	0,000	-	0,820	2,227	0,992

No caso de vazamentos, o aumento da incerteza-padrão de 1 t.h^{-1} para 5 t.h^{-1} resultou em falha de identificação de vazamentos nos balanços F3 e F5, conforme apresentado na Tabela 4.23 e Tabela 4.24. Cabe ressaltar que a necessidade de

ajustar as incertezas-padrão de alguns fluxogramas passa a ser uma das limitações do método. Uma outra alternativa possível seria testar outros estimadores robustos para avaliar o efeito dos mesmos para tratar esse tipo de situação.

Tabela 4.23: Resultados do Problema 14 para vazamentos utilizando incertezas-padrão unitários.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
M1	1	1	0,061	1,0	0,013	0,996
F1	1	1	0,086	1,0	0,015	0,996
T1	1	1	0,089	1,0	0,003	0,996
S1	1	1	0,069	1,0	0,000	0,996
F2	1	1	0,049	1,0	0,004	0,996
M2	1	1	0,044	1,0	0,005	0,996
S3	1	1	0,022	1,0	0,002	0,996
F3	0,9764	0,982	0,041	1,0	0,003	0,996
F4	1	1	0,049	1,0	0,030	0,996
F5	0,9804	0,9836	0,007	1,0	0,003	0,996
S5	0,97	0,9756	0,028	1,0	0,003	0,996
T2	1	1	0,050	1,0	0,007	0,996
S4	1	1	0,040	1,0	-0,003	0,996
S2	1	1	0,024	1,0	0,011	0,996

Tabela 4.24: Resultados do Problema 14 para vazamentos e incertezas-padrão de 5 $t.h^{-1}$.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
M1	1	1	0,061	1,0	-0,002	0,992
F1	1	1	0,086	1,0	0,002	0,992
T1	1	1	0,089	1,0	-0,002	0,992
S1	1	1	0,069	1,0	-0,001	0,992
F2	1	1	0,049	1,0	-0,003	0,992
M2	1	1	0,044	1,0	-0,002	0,992
S3	1	1	0,022	1,0	0,000	0,992
F3	0,9764	0,982	0,041	0,0	1,186	0,000
F4	1	1	0,049	1,0	-0,002	0,992
F5	0,9804	0,9836	0,007	0,0	1,100	0,000
S5	0,97	0,9756	0,028	1,0	0,001	0,992
T2	1	1	0,050	1,0	-0,003	0,992
S4	1	1	0,040	1,0	-0,002	0,992
S2	1	1	0,024	1,0	-0,001	0,992

4.3 Erros múltiplos

Os cenários para erros grosseiros de medição e vazamentos, assim como o coeficiente de correlação, são apresentados na Tabela 4.25.

Tabela 4.25: Cenário de distúrbios com erros múltiplos e os resultados dos ajustes dos modelos.

Problema	Corrente com Erro Medição	Equipamento com Vazamento	R^2 Modelo
P1	1	1	0,97
P2	3	3	0,97
P3	3;10	2	0,96
P4	1	4	0,98
P5	2;4	4	0,95
P6	4;5	2	0,95
P7	6;9	5	0,96
P8	4;11	2	0,94
P9	2;3	5	0,94
P10	3;7;12	8	0,97
P11	6;12	3;7	0,97
P12	3;6;11	2	0,95
P13	2;7;13;15	6	0,97
P14	1;8;9;15;19	4;13;14	0,96
P15	1;13;15;19	1;5;10	0,95
P16	1;6;13;15;19;30;43;46;48	1;12;16;18;26	0,96

Os resultados dos problemas são apresentados em tabelas conforme Tabela 4.26, onde, na primeira coluna é indicado onde os erros grosseiros múltiplos foram adicionados: **C** indica corrente e **E** indica equipamento, logo, na Tabela 4.26 foi adicionado um erro grosseiro de medição na Corrente 3 e um vazamento no Equipamento 3.

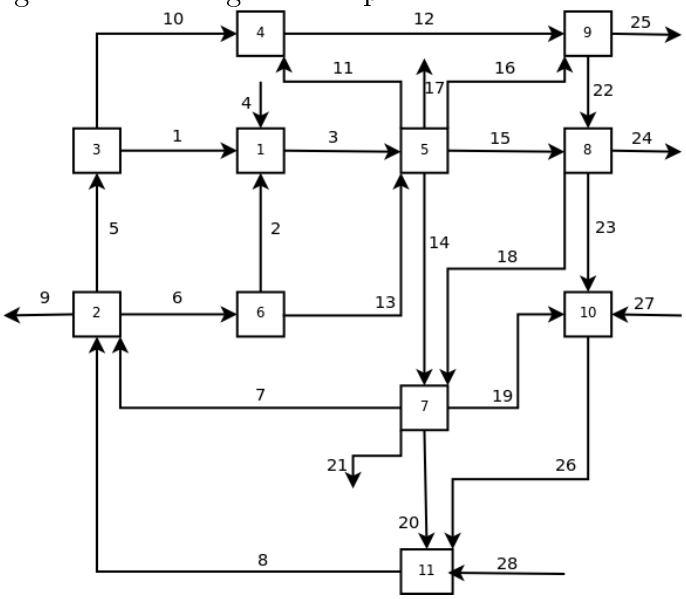
Tabela 4.26: Exemplo de tabela de resultados dos erros grosseiros múltiplos.

Corrente/Equipamento	OP	AEE	Seletividade
C-3	1	0,221	1
E-3	1	0,033	1

Como os resultados dos erros múltiplos foram bastante satisfatórios para os cenários escolhidos, apresentando um elevado OP e baixo AEE, optou-se por comentar somente alguns problemas, principalmente os que apresentaram uma baixa seletividade para erros múltiplos, entre eles, os problemas 1, 4 ,14, 15 e 16. Os resultados dos demais problemas para erros múltiplos são apresentados no Apêndice D.

Os problemas 1, 4 ,14, 15 e 16 apresentaram a seletividade reduzida para detecção dos erros grosseiros múltiplos simulados, indicando que, embora o OP e AEE estejam em uma faixa satisfatória, existem outras correntes ou equipamentos onde os erros grosseiros também excederam os limites de teste, porém, as maiores estimativas foram nas posições dos erros grosseiros adicionados. Tal fato está associado, provavelmente, pela posição de erros grosseiros adicionados, sendo que nesses problemas, os mesmos estão muito próximos. Para evidenciar tal proposição é apresentado o Problema 15, conforme Figura 4.8, onde os erros grosseiros simultâneos foram adicionados às correntes 1, 13, 15 e 19 enquanto que os vazamentos forma adicionados nos equipamentos 1, 5 e 10. A seletividade reduzida do Problema 15 para erros simultâneos pode ser observada na Tabela 4.27.

Figura 4.8: Fluxograma de processo do Problema 15.



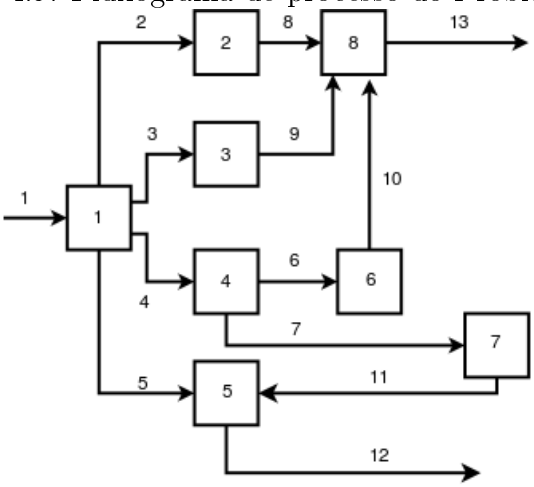
Fonte: Próprio autor

Tabela 4.27: Resultados dos erros grosseiros múltiplos do Problema 15.

Corrente/Equipamento	OP	AEE	Seletividade
C-1	0,944	0,771	0,673
C-13	0,967	0,503	0,673
C-15	1	0,227	0,685
C-19	0,989	0,452	0,649
E-1	1	0,035	0,649
E-5	1	0,035	0,649
E-10	1	0,035	0,649

Em contraponto, problemas onde os erros múltiplos estão posicionados em regiões 'afastadas' do fluxograma, a seletividade apresentou o valor máximo de um, conforme exemplificado através do Problema 10, apresentado na Figura 4.9, onde as correntes com erros grosseiros foram as 3, 7 e 12, enquanto que o vazamento foi adicionado no balanço 8, conforme apresentado na Tabela 4.28.

Figura 4.9: Fluxograma de processo do Problema 10.



Fonte: Próprio autor

Tabela 4.28: Resultados dos erros grosseiros múltiplos do Problema 10.

Corrente/Equipamento	OP	AEE	Seletividade
C-3	1	0,080	1
C-7	1	0,210	1
C-12	1	0,484	1
E-8	1	0,020	1

4.3.1 Comparação com outros métodos

Poucos autores apresentam dados publicados para DEG múltiplos em condições de serem comparados na mesma base do método proposto, visto que muitos apresentam somente o resultado de 1 simulação numérica (WANG et al., 2004; MEI et al., 2006) enquanto outros autores não informam o local exato da adição dos erros grosseiros múltiplos (NARASIMHAN; MAH, 1987). Como base para comparação, foram utilizados dados de Sánchez et al. (1999), que apresentam uma comparação detalhada com erros grosseiros múltiplos comparando três métodos: UBET (ROLLINS; DAVIS, 1992), GLR (NARASIMHAN; MAH, 1987) e SEGE (proposto pelos próprios autores). Embora nem todos os dados obtidos por Sánchez et al. (1999) foram obtidos exatamente nas mesmas condições que o presente trabalho, tais dados são bastante semelhantes e podem ser considerados apropriados para uma comparação.

A Tabela 4.29 apresenta a comparação do OP de diferentes métodos e do método proposto, onde na primeira coluna estão presentes as correntes onde os erros grosseiros simultâneos foram adicionados. A principal diferença entre os dados de Sánchez et al. (1999) e o presente trabalho está na magnitude dos erros grosseiros. Enquanto Sánchez et al. (1999) usou erros grosseiros com magnitudes diferentes para as duas correntes onde foram adicionados erros grosseiros (7 e 4 vezes as incertezas-padrão, respectivamente), o presente trabalho usou 7 vezes a incerteza-padrão para ambas as correntes. É possível observar através da Tabela 4.29 que o método proposto apresentou máximo OP para todos os conjuntos de distúrbios e se mostrou superior

no exemplo estudado em relação aos métodos comparados.

Tabela 4.29: Comparação entre OP de diferentes métodos para DEG múltiplos para o Problema 6.

Corrente	OP SEGE	OP UBET	OP GLR	OP Método
1-7	0,997	0,5	0,504	1
2-3	0,999	0,5	0,997	1
2-4	0,958	0,944	0,5	1
3-4	0,96	0,504	0,5	1
4-5	0,999	0,5	0,999	1
4-6	0,998	0,875	0,577	1
5-6	0,977	0,517	0,5	1
6-7	1	0,499	0,886	1
1-2	0,996	0,999	0,971	1
1-3	1	0,977	0,972	1
1-4	0,974	0,932	0,973	1
1-5	0,704	0,963	0,145	1
2-5	0,999	0,982	0,999	1
2-6	0,987	0,98	0,966	1
2-7	0,999	0,93	0,999	1
3-5	1	0,965	1	1
3-6	0,987	0,972	0,967	1
3-7	0,999	0,976	0,999	1
4-7	0,998	0,6	0,998	1
5-7	0,997	0,946	0,996	1
1-6	0,997	0,966	0,993	1

A Tabela 4.29 compara o AEE para alguns casos de erros grosseiros múltiplos em conjuntos de 2 medições simultâneas, onde é possível observar que o método proposto se apresentou inferior quando comparado com os métodos SEGE e UBET e apresentou equivalente quando comparado com o GLR. Tal fato pode ter ocor-

rido, provavelmente, porque os autores originais não modificaram os sinais dos erros grosseiros, enquanto que no presente trabalho, os sinais dos mesmos são escolhidos aleatoriamente usando uma distribuição uniforme.

Tabela 4.30: Comparação entre AEE de diferentes métodos para DEG múltiplos para o Problema 6.

Correntes com Erros Grosseiros	Corrente	AEE SEGE x 100	AEE UBET x 100	AEE GLR x 100	AEE Método x 100
1-2	1	0,114	0,000	8,000	10,91
	2	0,200	0,067	1,733	7,33
1-3	1	0,114	0,114	8,000	5,42
	3	0,067	0,067	1,667	1,21
1-4	1	0,114	0,114	5,714	2,94
	4	1,000	3,000	2,600	17,18
2-5	2	0,000	0,038	8,419	9,52
	5	0,100	0,000	2,700	3
2-6	2	0,076	0,038	11,238	9,19
	6	0,600	0,200	12,200	0,56
2-7	2	0,000	0,076	5,600	16,43
	7	0,200	0,200	1,800	6,15
3-5	3	0,038	0,038	8,381	4,25
	5	0,200	0,000	2,700	10,16
3-6	3	0,114	0,000	11,200	11,49
	6	0,600	0,200	11,200	5
3-7	3	0,038	0,038	5,600	7,76
	7	0,200	0,000	1,600	14,65
4-7	4	0,000	5,714	2,286	6,7
	7	0,200	4,400	10,800	13,57
5-7	5	0,114	0,229	11,314	7,76
	7	0,200	0,200	6,000	14,46

Na Tabela 4.31 são apresentados resultados de erros grosseiros simultâneos de medição e vazamento, sendo que neste caso, ambos os resultados são equivalentes pois a magnitude do vazamento e do erro de medição são as mesmas apresentadas por Sánchez et al. (1999). É possível observar que o método proposto apresenta desempenho superior, identificando todos os erros grosseiros, resultando em um OP máximo.

Tabela 4.31: Comparação entre OP de diferentes métodos para DEG múltiplos com vazamento para o problema 6.

Erro Grosseiro Múltiplo	OP SEGE	OP UBET	OP GLR	OP Método
E2-C4	0,993	0,961	0,996	1
E2-C5	0,999	0,952	0,837	1
E2-C6	0,996	0,969	0,999	1
E2-C7	0,99	0,951	0,872	1
E3-C2	0,981	0,921	0,999	1
E3-C6	0,999	0,963	0,937	1

A comparação entre o AEE na presença de erros múltiplos é apresentada na Tabela 4.32, onde é possível observar que o método proposto apresenta resultados superiores ao GLR, porém inferiores ao UBET e SEGE.

Tabela 4.32: Comparação entre AEE de diferentes métodos para DEG múltiplos com vazamento para o problema 6.

Erro Grosseiro Múltiplo	Erro Grosseiro	AEE SEGE	AEE UBET	AEE GLR	AEE Método
E2-C4	E2	0,00	0,06	1,17	2,62
	C4	0,16	0,48	0,48	2,78
E2-C5	E2	0,06	0,39	12,33	0,93
	C5	0,00	0,32	29,20	7,58
E2-C6	E2	0,06	0,06	4,39	0,32
	C6	0,00	0,00	4,00	0,82
E2-C7	E2	0,06	0,06	-	0,32
	C7	0,16	0,00	-	0,81
E3-C2	E3	0,00	0,80	12,00	6,19
	C2	0,21	1,44	10,24	1,49
E3-C6	E3	0,08	0,08	1,84	1,27
	C6	0,00	0,00	1,12	0,72

4.4 Considerações finais

Conforme apresentado no Capítulo 3, muitos autores testam suas metodologias para DEG em um universo restrito de problemas, conforme pode ser observado na Tabela 4.33, aqui reapresentada. Entre as razões para esse motivo, estão os recursos computacionais limitados disponíveis no passado, além da ausência de pacotes computacionais (como Matlab, Maple, Scilab, etc.) que facilitam a tarefa de operações matriciais, comuns nesse tipo de problema. Além disso, os problemas encontravam-se distribuídos ao longo da literatura tornando a coleção dos mesmos uma tarefa um pouco difícil. Ao longo de desenvolvimento desse trabalho, buscou-se coletar um conjunto bem heterogêneo de problemas, desde um único equipamento até fluxogramas de processos industriais reais, para teste do método proposto de maneira ampla e robusta. Os problemas testes gerados foram reunidos em um repositório de

problemas teste na internet em Valle (2011) disponível através de *scripts* do *software* Scilab para que outros autores possam testar ou implementar outras técnicas de RD ou DEG.

Além do teste da metodologia proposta em diversos problemas, diversos experimentos numéricos foram conduzidos para escolha do melhor estimador robusto e melhor magnitude da incerteza-padrão. Com diversos experimentos computacionais realizados, comprovou-se que a metodologia apresentada nesse trabalho apresentou um desempenho consideravelmente superior às existentes na literatura em uma ampla faixa de problemas. Tal fato se deve, principalmente, porque o método proposto mimifica o que ocorre nos processos industriais onde um erro em uma medição ou vazamento não se espalha para as demais medições do processo.

Nas metodologias de RD e DEG presentes na literatura, o GT, MT e NT são a base para a identificação de erros únicos e erros múltiplos. A principal limitação, principalmente do MT, é que esse é aplicado aos ajustes das medições, ou seja, após o procedimento de RD. Analisando o procedimento de RD mais detalhadamente, é possível observar que, embora o mesmo possua sólida base estatística, ao aplicar tal técnica, na presença de erros grosseiros, os dados são ajustados de maneira completamente diferente dos fenômenos que ocorrem em processos reais: os erros das medições se espalham para as demais medições; os balanços de massa são forçados a serem satisfeitos o que é inviável de ocorrer na presença de vazamentos reais, também levando ao espalhamento dos erros grosseiros. Tal problema surge da própria base estatística da RD que parte do princípio que, o erro de medição possui distribuição de erro conhecida, geralmente Normal, com desvio padrão conhecido, o que não ocorre na presença de erros grosseiros. É possível notar ainda, que, o NT não sofre essa limitação, pois sua estatística de teste é baseada nos resíduos antes da RD, como pode ser observado nas tabelas de resultados apresentadas. Logo, para erros únicos o NT ou o GLR para detecção de vazamentos podem ser utilizados sem limitações para detecção de erros.

A faixa de confiabilidade dos modelos gerados para mudanças de pontos operacionais não foi explorada nesse trabalho, ficando indicada como sugestões para trabalhos futuros. Como a detecção de medições com erros grosseiros é influenciada

pela magnitude da medição em questão, é de se esperar que, ao se mudar o ponto de operação, seja necessário o ajuste do modelo na nova condição.

Como limitações da técnica proposta, podemos citar a necessidade de avaliação prévia da magnitude da incerteza-padrão, sendo que em alguns casos apresentados, o OP não apresentou bons resultados usando a incerteza-padrão unitária. Uma segunda limitação é a necessidade de geração de um novo conjunto de dados e realização de um novo ajuste dos modelos ao serem modificadas as condições operacionais do processo. Outro problema que pode ocorrer se refere à aplicação da técnica para identificar erros múltiplos, sendo que nesse caso, existe a necessidade de gerar diversas combinações de erros grosseiros múltiplos para estimação de parâmetros do modelo, que pode resultar em um problema combinatorial de grande escala, entretanto, esse número pode ser reduzido se for realizado um procedimento de planejamento ótimo de perturbações para determinar um conjunto mínimo de experimentos numéricos. Uma terceira limitação da técnica é que como a mesma modifica consideravelmente o problema original, tanto em relação à função objetivo quanto à incerteza-padrão, a mesma só pode ser usada especificamente para a DEG no ponto operacional onde os dados foram ajustados, não sendo possível usá-la, por exemplo, para RD. Por esse motivo, para uma futura aplicação industrial, seria necessário ter a técnica de RD sendo executada em paralela com o método proposto.

Tabela 4.33: Técnicas de identificação de erros grosseiros e o número de exemplos usadas para teste das mesmas.

Autor	Teste para Detecção/Identificação de Erros Grosseiros	Número de Exemplos Testados
Romagnoli e Stephanopoulos (1981)	Método Algorítmico Proposto	3
Rosenberg et al. (1987)	DMT and EMT	7
Narasimhan e Mah (1987)	Razão de Verossimilhança Generalizada	2
Reddy e Mavrovouniotis (1998)	Abordagem de Redes Neurais	1
Sánchez et al. (1999)	MSEGE, MUBET and MGLR	2
Bagajewicz et al. (1999)	Avaliação de PCA para DEG	1
Kongsjahju et al. (2000)	Modificação do Método UBET para sistemas em série	1
Soderstrom et al. (2001)	Abordagem Inteira-Mista	4
Arora e Biegler (2001)	Estimadores Robustos para RD e DEG	3
Özyurt e Pike (2004)	Reconciliação Robusta	7
Wang et al. (2004)	Aprimoramento do MT e NT	1
Devanathan et al. (2005)	Métodos Bayesianos e de Verossimilhança	1
Mei et al. (2006)	NT-MT para DEG	2
Alhaj-Dibo et al. (2008)	RD usando distribuições robustas	1
Maronna e Arcas (2009)	RD and DEG baseada em Modelos de Regressão	2
Zhang et al. (2010)	Estimadores Robustos	4
Farias (2009)	IMT Robusto	2

Capítulo 5

Conclusões e Sugestões para Trabalhos Futuros

5.1 Conclusões

5.1.1 Principais contribuições

Conforme apresentado na revisão bibliográfica desse trabalho, as técnicas de DEG podem ser influenciadas por muitos fatores, como conexão entre os equipamentos, incerteza-padrão da medição, posição do erro grosseiro, entre outras. Consequentemente a correta detecção e identificação de erros grosseiros ainda pode ser considerada um desafio (CROWE, 1996; NARASIMHAN; JORDACHE, 2000) mesmo para fluxogramas em estado estacionário com todas as medições redundantes, e, através dos resultados desse trabalho, isso pode ser comprovado ao se usar técnicas de DEG de referências como o GLR.

Nesse trabalho, um cuidado especial foi tomado na seleção de problemas para teste do método proposto de maneira mais geral e robusta. É possível afirmar ainda, que, em relação ao número de problemas e cenários simulados, o presente trabalho é o mais completo apresentado na literatura até o momento.

Conforme os experimentos numéricos conduzidos nos problemas testes, foi possível observar a melhoria no poder do MT ao usar incerteza-padrão unitária e também ao usar estimadores robustos na função objetivo do problema de RD.

Em relação ao método proposto foi observado que, para alguns problemas, como

o Problema 3, Problema 7 e Problema 14, foi necessário a escolha de outros valores para as incertezas-padrão, onde se é possível concluir a importância de uma análise mais detalhada dos valores das incertezas-padrão antes de aplicar a metodologia. A metodologia para geração e seleção de dados também se mostrou apropriada para que a estimação de parâmetros de modelos PLS com um bom coeficiente de determinação fosse realizada para predição da magnitude dos erros grosseiros. Os resultados do método proposto sendo testados com um conjunto de dados não utilizados no conjunto de estimação dos parâmetros do modelo, consolida positivamente a adequação do método proposto para identificação de erros grosseiros e estimação de sua magnitude. Como conclusão final, é possível afirmar que, no conjunto de problemas onde o método foi aplicado, o mesmo se destacou na identificação de erros grosseiros em medidores quando comparado com o método GLR para erros únicos e apresentou um resultado semelhante aos métodos do NT e GLR NT para detecção de vazamentos. Para erros múltiplos foi possível a comparação somente com um exemplo de literatura, onde o método proposto foi comparado com outras técnicas de DEG, sendo elas, GLR, UBET e SEGE, também apresentando resultado superior. Em relação à erros múltiplos, o método proposto apresentou elevado OP para o conjunto de cenários estudados e seletividade reduzida para alguns problemas, sendo necessário, para uma avaliação mais rigorosa, realizar o teste do método em cenários com um conjunto maior de erros múltiplos em cada problema.

5.1.2 Publicações geradas ao longo do trabalho

O presente trabalho gerou uma série de publicações, entre elas:

- Artigo submetido à revista *Industrial and Engineering Chemistry Research* com o título: *Collection of Benchmark Test Problems for Gross Error Detection and Identification.*
- Artigo submetido à revista *Industrial and Engineering Chemistry Research* com o título: *Collection of Benchmark Test Problems for Data Reconciliation.*
- Repositórios de problemas *benchmark* e métodos para RD e DEG (VALLE, 2011).

- Co-autoria no trabalho apresentado no 6o. Congresso Brasileiro de Metrologia (2012), com o titulo: *A Minlp Formulation to Optimize Sensor Allocation Using Reconciled Data in Systems With Fewer Measurements.*

- Co-autoria no trabalho apresentado no *International Conference on Engineering Optimization* - EngOpt2012 , com o titulo: *Optimizing Sensor Allocation Using Reconciled Data in Systems with Scarce Measurements: MINLP and MILP Approaches.*

- Co-autoria no trabalho apresentado no XIX Congresso Brasileiro de Engenharia Química 2012, com o titulo: Validação de Modelo Minlp para Alocação Ótima de Instrumentos Utilizando Dados Reconciliados em Sistemas sem Medições -XIX Congresso Brasileiro de Engenharia Química.

- Pedido de registro de *software* no Instituto Nacional de Propriedade Intelectual (INPI): Módulo de Estudo de Caso no Simulador EMSO.

- Pedido de registro de *software* no (INPI): Módulo de Análise de Sensibilidade no Simulador EMSO.

- Pedido de registro de *software* no (INPI): Método TECLIM de Reconciliação de Dados no Simulador EMSO.

- Pedido de registro de *software* no (INPI): Suplemento de comunicação Excel-EMSO para Otimização e Reconciliação de Dados.

5.2 Sugestões para trabalhos futuros

Embora as contribuições do método proposto sejam significantes, algumas sugestões para trabalhos futuros são apresentadas nessa seção.

A primeira sugestão é a avaliação mais aprofundada dos efeitos da incerteza na detecção de erros grosseiros com objetivo de obtenção de um OP máximo, tanto para erros grosseiros em medições quanto vazamentos, visto que, no presente trabalho, somente experimentos numéricos foram realizados para a determinação das incertezas.

A segunda sugestão é a avaliação da capacidade de predição de modelos para erros únicos a partir de erros grosseiros múltiplos. Também é interessante investigar se modelos ajustados para predizer erros únicos são capazes de prever erros grosseiros múltiplos. Além disso, é possível explorar se ao concatenarmos os dados usados para ajustar modelos de erros grosseiros únicos e múltiplos, a capacidade de predição do modelo tem a tendência de melhorar ou piorar. Outra sugestão, no sentido de explorar a capacidade extrapolativa do método proposto, seria avaliar em que faixa de mudanças em pontos operacionais de cada diagramas o modelo empírico gerado seria capaz de predizer, com alguma confiabilidade, o local e magnitude dos erros grosseiros.

Como terceira sugestão, está um método sistemático para geração de cenários de erros grosseiros múltiplos de forma a gerar os modelos a serem ajustados com um número mínimo de cenários, semelhante à técnica de planejamento de experimentos.

Outra sugestão está o estudo de seleção de variáveis independentes para avaliar a possibilidade de eliminar algumas variáveis de entrada, estudando o comportamento do modelo PLS resultante. Além disso, o presente método usou as estatísticas dos testes GT, MT e NT, mas é possível, ainda usar estatísticas de outros testes, como o PCA para erro de medição, de vazamento, entre outros, para avaliar a possibilidade de gerar modelos eficientes com poucas variáveis de entrada.

Como sugestão final para trabalhos futuros está a avaliação da capacidade de predição de erros grosseiros em cenários com variáveis não medidas e fluxogramas mais complexos envolvendo restrições não-lineares, como balanço de composição, energia e quantidade de movimento.

Apêndice A

Descrição dos Problemas em Forma de Artigo

Collection of Benchmark Test Problems for Gross Error Detection and Identification

Edson Cordeiro do Valle[a,b,*], Ricardo de Araújo Kalid[b,*], Argimiro Resende Secchi[c,*], Asher Kiperstok[b]

[a] *VRTech Tecnologias Industriais Ltda - Rua Guilherme Schmitz, 625, Gravatái - RS - CEP: 94065-200, Brazil*
[b] *Programa de Engenharia Industrial - Escola Politécnica - Universidade Federal da Bahia - Rua Aristides Novis, 2, 6 sexto. andar, Salvador - BA - CEP: 40210-630, Brazil*
[c] *COPPE - Universidade Federal do Rio de Janeiro - Centro de Tecnologia, Bloco G, sala 115 Cidade Universitária, Rio de Janeiro - RJ - CEP 21941-972*

Abstract

Frequently, in a realistic industrial scenario one can find measured data that do not satisfy the mass and energy laws of conservation, which can arise from sensor malfunctioning, process leaking or simply by random noise from measurement sensors. This problem can be approached by applying data reconciliation (DR) and gross error detection and identification (GEDI) techniques, but even with these methods, distinguish between the error sources accurately is not a trivial task. Due to complex nature of GEDI, a general method, that can be applied to any flowsheet, to correctly identify gross error types and location, have not been proposed yet. In this context, the objective of this work is to propose a collection of challenge benchmark problems and data set generation for GEDI in order to help the scientific community to evaluate theirs DR and GEDI methods in a common environment and in

*Corresponding author
Email addresses: edsoncv@gmail.com (Edson Cordeiro do Valle), kalid@ufba.br (Ricardo de Araújo Kalid), arge@peq.coppe.ufrj.br (Argimiro Resende Secchi)

different types of flowsheets. First, challenge problems were collected from literature or proposed and its challenge issues were presented. Then, some GEDI techniques were applied for the benchmark problems. In the results section, the main issues of selected challenge problems are presented and discussed. As a conclusion, the analysis of several challenge test problems together let the reader have an idea of the particularities that appear when applying GEDI techniques. Beyond, applying the GEDI techniques in a common set of problems, leads the validation of the methods in a more general and robust manner.
Keywords: data reconciliation, gross error detection and identification, fault detection, benchmark problems

---

1. Introduction

In the chemical industry, measurement equipments plays a special role in the process monitoring and control. Accurate process data are also important for the application of advanced chemical process tools techniques, such as: model predictive control, real time optimization, state estimators (such as Kalmann Filter), virtual analyzers, steady-state detection, etc. However, some measurement equipments are subjected to climate oscillation and aggressive process conditions which may degenerate their performance. Beyond, it is not always possible to replace a measurement equipment due to the type of the process which it is attached: some chemical processes are large-scale and continuous and cannot be paused or shutdown in order to replace a measurement sensor or equipment. Costs reduction also prevents the appropriate maintenance of measurement sensors. Consequently, process

engineers, occasionally, must accept this performance degeneration or failure along the chemical process operation which can be noticed when mass and energy balances presents discrepancies when calculated based on measured sensor data. However, not all data oscillation are due to failures, but from measurement equipment precision or from leakings that can also leads to mass and energy balances disagreement, when calculated based on sensor measurements. To distinguish between a sensor failure, leaking or data oscillation due to sensor variability, a procedure named data reconciliation (DR) and gross error detection and identification (GEDI) can be applied to process data and has been extensively studied since 60th decade in the chemical and mineral processing areas.

The goal of DR consists in perform a correction in the measurements based on the mass, energy or momentum balance model. To execute this task, optimization methods are applied and the procedure can be expressed in the following way:

$$\begin{cases} \min\limits_{\mathbf{X}\in\Re^n} \left(\mathbf{Y}-\mathbf{X}\right)^T \Sigma^{-1} \left(\mathbf{Y}-\mathbf{X}\right) \\ \text{subject to} \\ F\left(\mathbf{X}\right) = 0 \\ \mathbf{X}_{min} \leq \mathbf{X} \leq \mathbf{X}_{max} \end{cases} \tag{1}$$

where:

Σ: Variance covariance matrix of measurements

$\mathbf{Y}$: Vector of measurements

$\mathbf{X}$: Vector of reconciled measurements

$\mathbf{X}_{min}$: Vector of reconciled measurements lower bounds

$\mathbf{X}_{max}$: Vector of reconciled measurements upper bounds

$F(\mathbf{X})$: Vector of process constraints related with mass, energy balance or momentum balance

In the above DR problem formulation, it is assumed that the measurements are recorded from a steady-state process. The presented objective function is known as weighted least squares (WLS) and is deduced from the maximum likelihood method and assumes that the measurement errors have a normal distribution with known variances. If the measurement errors follow other distribution or the system presents measurement instrument fail, leaking, known as gross errors (GE), other objective functions can be applied, called robust functions[1, 2], or robust estimators, presented in Appendix C. An appropriate robust function must have some characteristics to provide good results in the DR. Considering ε_i as the residual of measured and reconciled data, weighted by the reciprocal of the standard deviation:

$$\varepsilon_i = (y_i - x_i)\,\sigma_i^{-1} \tag{2}$$

If we set $\rho(\varepsilon_i) = 1/2\varepsilon_i^2$ the objective function is the classical WLS estimator which cannot be considered a good estimator because when $|\varepsilon_i| \to \infty$, $\varepsilon_i \to \infty$, which happens in the presence of gross error. According to Özyurt and Pike [2], a good robust estimator must have the following mathematical property:

$$\frac{d\rho(\varepsilon)}{d\varepsilon} \propto \begin{cases} \varepsilon & . \text{ when } |\varepsilon| \text{ is small} \\ c & . \text{ when } |\varepsilon| \to \infty \end{cases} \tag{3}$$

Figure 1: Contaminated Normal and WLS estimators.

The behavior of a good robust estimator can be graphically inspected by plotting a chart of $\rho(\varepsilon)$ versus ε. Figure 1 illustrates the behavior of the Contaminated Normal estimator[2], (see Appendix C) when compared with the WLS, which elucidates the property described in Equation 3.

A robust functions may be used to either remove the effect of the GE in the DR or detect it, if its parameters are appropriately tuned. The derivative presented in the left hand side of Equation 3 is also known as Influence Function (IF) and was used by Özyurt and Pike [2] to tune the robust estimators parameters for GEDI while Arora and Biegler [1] and Zhang et al. [3] used the Akaike information criterion (AIC [4]) for this task. The appropriate selection and tuning of this robust function can be considered an interesting challenge in GEDI area.

Another approach to GEDI is based on the reconciliation adjustment (**Ad**) or on the weighted adjustment (**WAd**), presented in Equation 4 and 5, respectively. With the calculation of Ad and WAd it is possible to have an intuitive idea of the *worst* measurements.

$$\mathbf{Ad} = \mathbf{Y} - \mathbf{X} \tag{4}$$

$$\mathbf{WAd} = \mathbf{W}^{1/2}\left(\mathbf{Y} - \mathbf{X}\right) \tag{5}$$

where $\mathbf{W}$ is the inverse of Σ matrix, or Σ^{-1}.

Analytically, the data reconciliation procedure with WLS estimator cannot

evaluate the presence of gross errors, then to perform such task, statistical hypothesis tests must be applied to the results of DR. These tests are formulated according to the hypotheses: H_0 no gross error is present; H_1 a gross error is present. It is necessary to choose a significance level to accept or reject hypothesis H_0. In fact, the significance level is related to the probability of detecting a gross error when it is not present (also known in statistics as *Type I* error), since the observed measurement can be due to normal instrumentation variability. Some simple GED and GEDI calculations, such as Global Test (GT), Measurement Test (MT) and Nodal Test (NT), are briefly presented in Appendix A. Several statistical tests with its respective technique can be applied to perform GED and GEDI, some of them are presented in Table 1.

In the DR and GEDI literature there are a common terminology to describe the capacity of the methods, and they are described bellow:

- Single or Multiple Gross Error: In real industrial scenarios, the gross errors may be present individually, in more then one location, or combined (leakings or measurement bias).
- Gross Error Detection: Detect the presence of a gross error, without specify its location.
- Gross Error Identification: Identify the type of the error (leaking or measurement bias in a single or multiple equipments) and its position.
- Gross Error Estimation: Estimate the magnitude of the gross error.

The accurate identification of gross errors is a challenge task because many factors affect the detection methods or algorithms and may lead to

Table 1: GEDI technique with the respective statistical tests.

GEDI Test Name	Statistical Test	Gross error type detected
Global (GT)	Chi-squared	Detects a gross error in the mass balance as a whole, do not detect the location
Nodal (NT) or Equipment	z	Detects a gross error in the equipment balance (leaking)
Measurement (MT)	z	Detects a gross error in the measurement
GLR	Chi-squared	Detects equipment leaking or measurement bias and also its magnitude
Principal Components of Constraints	z	Detects an equipment leaking
Principal Components of Measurements	z	Detects a gross error in the measurement

incorrect results. Previous studies performed by Rosenberg et al. [5], Iordache et al. [6], Sánchez et al. [7] indicated that the GEDI and MGEDI (Multiple GEDI) algorithms are influenced by many factors, such as:

1 Magnitude of measurements and bounds:

2 Magnitude of gross errors or the ratio between the gross error and the respective measurement variance:

3 Position of the measurement containing gross error:

4 Constraints and flowsheet structure:

When considering the network structure, the accurate detection and identification of single or multiple gross error can be a complex procedure, because the redundancy needed to perform DR and GEDI (or MGEDI) may be degenerated in some flowsheets which depreciates the results of the techniques. These redundant measures are named *practically nonredundant measurements* (Iordache et al. [6]) and, when considering total flow mass reconciliation, can be present in the following situations (Narasimhan and Jordache [8], Sánchez et al. [7]):

4.1 Flow rates with small absolute standard deviations (or small order of magnitude) in the same balance;

4.2 Parallel streams;

4.3 Measured variables that appear in only one equation with one unmeasured variable;

In real process operation and monitoring, the detection of a single gross error itself does not provide enough information for process analysis. We also want to identify multiple gross errors, its type (measurement bias or leaking) and also estimate the error magnitude. Many strategies to detect and identify multiple gross errors (MGE) are present in the literature [9, 6, 10, 7, 5, 11, 12], some of them perform the identification of the MGE in one step, named simultaneous strategies[8]. Other MGEDI techniques perform iterative GEDI testing, removing *suspicious* measurements or mass balances, repeating the test procedure until identify the appropriate set of gross errors.

When a new or a modified GEDI or MGEDI technique is proposed, it is compared with an existing one in order to validate the method itself. However, when authors propose new techniques for DR and GEDI (or MGEDI), only few cases are used to validate these new methods while the ideal case should be the testing in several cases (flowsheets) in order to verify and to guarantee the generality of the proposed method. On a review of GED methods proposed by literature, one can observe that, in many cases, only a few examples are studied, as presented in Table 2.

Although the DR and GED research topic is not recent, there are many challenges in this field[24, 8], specially regarding MGEDI and flowsheets with both measured and unmeasured streams. The complexity of GEDI or MGEDI shows that the techniques proposed should be tested in a large number of test examples in order to make them more general and robust.

Test problems are well accepted by the scientific community and many research areas have their consolidated benchmarks sets. In the optimization research area the most popular test sets for algorithm testing is Hock and

Table 2: GEDI techniques and the number of examples studied to validate them.

Author	GED Test	Number of Examples Tested
Romagnoli and Stephanopoulos [13]	Algorithmic Method Proposed	3
Rosenberg et al. [5]	DMT and EMT	7
Narasimhan and Mah [11]	Generalized Likelihood Ratio	2
Reddy and Mavrovouniotis [14]	Neural Network Approach to GED	1
Sánchez et al. [7]	MSEGE, MUBET and MGLR	2
Bagajewicz et al. [15]	Evaluation of PCA for GED	1
Kongsjahju et al. [16]	Modification of UBET for serially correlated systems	1
Soderstrom et al. [17]	MI approach to DR and GED	4
Arora and Biegler [1]	Redescending Estimators for DR and GED	3
Özyurt and Pike [2]	Robust Reconciliation	7
Wang et al. [18]	Improvement of Measurement and Nodal Test	1
Devanathan et al. [19]	Likelihood and Bayesian Methods for DR and GED	1
Mei et al. [20]	NT-MT for GED	2
Alhaj-Dibo et al. [21]	Robust DR using contaminated distributions	1
Maronna and Arcas [22]	DR and GED based on Regression Models	2
Zhang et al. [3]	Quasi-weighted least squares estimator	4
Farias [23]	Robust IMT	2

Schittkowski [25], updated later by Schittkowski [26]. Also for numerical package evaluation is available the test set proposed by Mazzia and Magherini [27]. In the control area, some benchmark problems are available: Downs and Vogel [28], Castro [29], and Chen et al. [30]. Other research areas also use benchmark problems to evaluate algorithms, numerical techniques and solving strategies, such as: waste water treatment (the COST waste water treatment plant model, presented by Copp [31]), industrial project and retrofit (in which heat integration benchmark is proposed by Nielsen [32] and mass integration proposed by Fontana [33]), transport phenomena (among them, Leong et al. [34] and Rajeshkanna [35]), molecular simulation (Barth et al. [36]), and others. However, the field of GED lacks of a test collection for algorithm performance evaluation.

In this context, the aim of this work is to collect, the most common, or cited, literature challenge examples for GEDI to help the researchers to test theirs DR and GED techniques. The main challenge related with each type of problem are presented. Beyond, common GEDI techniques are applied to selected problems, their results are presented and discussed.

2. Selection, Development and Implementation of Benchmark Problems

In the present work, 16 linear steady-state challenge test problems were selected. The test problems consist in flowsheets with exact total flow balances and variances which were collected from literature. The problems were selected to test the most common phenomena that influences negatively the performance of GEDI algorithms. The main challenge issue related to each

problem is presented in Table 3 and its influence on the GEDI will be further discussed in the results section.

The flowsheets, flow rates, standard deviation and other data are presented in supplement material due to space limitations. To apply the GEDI algorithm, described in Appendix B, it is necessary to start from a flowsheet with exact flow rates and, when these data are not available in the original paper, it was proposed by the authors.

For 16 linear test problems, a simulation approach was used to evaluate GEDI techniques for measurement bias which is the most accepted and applied systematic methodology to test GEDI performance[9, 2, 5, 6, 3]. The detailed steps for the algorithm is presented in Appendix B and was compiled from Narasimhan and Jordache [8]. It was also performed a simulation approach for Nodal Test (NT) evaluation, which is slightly different from the previous one, since it is not necessary to perform the DR for this test, which is also presented in Appendix B.

One of the steps for the simulation procedure is the application of GED techniques. The selected methods were: Global Test (GT); Measurement Test (MT); Nodal Test (NT); Generalized likelihood ratio test (for measurement bias and leakings, denoted as GLRMT and GLRNT, respectively). GLR can detect, identify and estimate the magnitude of gross errors both for measurement bias and equipment leaking. These tests were chosen because the most common simultaneous or iterative tests are based on them: Kim et al. [37] and Serth and Heenan [9] described tests based on MT; Mei et al. [20] and Wang et al. [18] proposed tests based both on NT and MT while GLR has been used as a reference test when comparing the performance of new

Table 3: Characteristics of linear problems.

Problem	Streams	Characteristics
P1	2,3	S
P2	2,3;5,6	P
P3	2,11;9,12;4,5,6,7,8	P
	1,2(11),9(12),3	R
P4		-
P5	4,5	P
	2,6,7,8	M
P6	2,3,4;2,3,5,6	R
P7	5,9;8,10	M
	5,7,8,9;2,3,5,6	R
P8	2,3;6,7;8,9;10,11	P
P9	1,2	P
	8,9,10	R
	3,4,5;5,6,7	M
P10	1,2;8,9,10,13;5,11,12	M
P11	2,3,6,5;2,7,9,8,5;9,12,10	R
	5,6,8;10,11,12	M
P12	5,6,7	P
	1,2,3,4;4,9,10;8,9,11;10,11,12	M
P13	12,13	P
	2,3,4;6,7,10;12(13),14,16	R
	1,2,4;6,7,8;7,9,10;14,15,16	M
P14	1,2,3,4;21,22;23,24	P
	10,11,15;15,16,19	M
P15	10,12,22,18,7,5;26,8,5,1,3,15,26	R
	1,2,3,4;5,6,7,8,9;10,11,12;12,16,22,25	M
P16	319,316,378,312;456,460;592,581;549,550	P
	408,453;441,401,402,404,405,407	R

GEDI techniques by Keller et al. [38], Singth et al. [39], Sánchez et al. [40] and Sánchez et al. [7]. For all tests, the significance level was chosen as:

- For MT, α (denoted as α_{MT}).
- For NT, α (denoted as α_{NT}).
- For testing measurement bias with GLR, β_{GLRMT}.
- For testing equipment leaking with GLR, β_{GLRNT}.

The significance level, α, were set in order to keep the AVT1, defined below, close to 0.1. After the statistical test execution, indexes to evaluate the performance of the GED detection techniques were calculated. These indexes have been proposed in the literature (Narasimhan and Mah [11], Rosenberg et al. [5], Rollins and Davis [10]):

- Overall Power (OP)= Number of GE correctly identified / Number of GE simulated
- Average number of Type I errors (AVT1) = Number of GE wrongly identified / Number of simulation trials
- Average Error of Estimation (AEE) = $\frac{1}{NS} \cdot \sum_{i=1}^{NS} \mid \frac{\text{estimated value}_i(\text{GLR}) - \text{actual value}_i(\text{Generated})}{\text{actual value}_i(\text{Generated})} \mid$

The OP is computed only for simulation trials in which gross errors are simulated. The AVT1 was tested only on datasets where no gross errors were added (step 4 from the algorithm presented in Appendix B, when null hypothesis is true). AEE is computed only in trials where the gross error

is correctly identified only for GLRMT and GLRNT, since this method can estimate the magnitude of the gross error. The number of simulations trials, NS, was selected as 2500. According to Rosenberg et al. [5], when using this value of NS, the Power is approximated by the OP, with an error $E < 0.02$ in approximately 95% of the estimations.

According to the definition of MT Power by Iordache et al. [6], in a given simulation trial, a gross error is detected only if the MT test statistics ($z_{a,j}$, presented in Appendix A) for a given stream, say j, exceeds the significance level and this MT test statistics is maximum among the MT test statistics of other streams, say k (in this case $j \neq k$). According to this definition, if 2 or more streams have the same MT test statistics (the cases where this occurs will be viewed latter in this paper), the gross error is attributed to the stream where the gross error was simulated. In a realistic industrial scenario, since we do not know previously where the gross error is present, we decided to modify this definition: If 2 or more streams have the same MT test statistics, no gross error is attributed to these streams nor identified among them. This definitions for MT make it similar to GLR where streams with the same GLR test statistics (for measurement bias) are not correctly identified.

In order to measure the improvement in the accuracy of the measurement after the data reconciliation, the adjustability of the measurements were calculated by Equation 6 (Narasimhan and Jordache [8]).

$$\text{adjustability}_i = \left(1 - \frac{\sigma_{x_i}}{\sigma_{y_i}}\right) \tag{6}$$

where σ_{y_i} is the standard deviation of the measurements and σ_{x_i} is standard

deviation of the reconciled variables, computed according to Heyen et al. [41], presented in Equation 7.

$$\sigma_{x_i} = \sqrt{\sum_{j=1}^{m} \frac{\left(M_{i,j}^{-1}\right)^2}{\sigma_{y_j}^2}} \tag{7}$$

where m is the number of measured streams and M is the following matrix:

$$\begin{pmatrix} W & A^T \\ A & 0 \end{pmatrix} \tag{8}$$

The detectabilities of the errors were also calculated and can be interpreted as the capacity of a gross error be detected: the larger the detectability factor, Equation 2, more likely is the gross error detection (Narasimhan and Jordache [8]).

$$\text{detectability}_i = \sqrt{\left(1 - \frac{\sigma_{x_i}^2}{\sigma_{y_i}^2}\right)} \tag{9}$$

Briefly describing, the method for the problem testing consists in 3 steps:

1 Generate the datasets.

2 Perform GED (MT, NT and GLR).

3 Evaluate the GED technique with performance indexes.

The procedure and calculations previously described were implemented in Scilab (INRIA [42]), including the robust estimators described in Appendix C. The data sets generation scripts are available at the Internet at http://code.google.com/p/dr-ged-benchmarks/GED/linear/scilab (Valle

[43]). Scilab includes several statistical functions to describe data properties and its scripts in a pseudo language allows the reader to perform DR and GEDI in a free and open source environment. Other objective functions, rather than WLS, can be used, since the software is linked with the Ipopt solver (Wächter and Biegler [44]). In the implementation, the problems data and GEDI techniques are in separate functions which make the inclusion of new benchmark problems and the creation of new GEDI or MGEDI independent and easily extensible.

3. Results and Discussion

Although the objective of this work is just to present the problems for the scientific community, some of them are solved in order to illustrate the challenges involved. Regarding the robust estimators, although other objective functions were implemented (presented in Appendix C), only the Lorenztian estimator was tested for the sake of space reduction.

3.1. Global Test Results and AVT1

The results of Global Test are presented in Table 4. It is possible to notice that, even keeping the AVT1 very similar to each problem, the OP for each flowsheet is very different. It is important to notice that when the number of streams and equipment increase, the OP for GT also reduces, this may be correlated with the mean detectability (the ratio between detectability and number of streams): there are evidences that when mean detectability reduces, the OP for the GT also reduces (as presented in Table 4), however, further investigations must be carried out to confirm this hypothesis.

Table 4: Global test results.

Problem	Number of Streams	Mean Detectability	OP Global Test	AVT1 Global Test
P1	3	0.528	0.645	0.096
P2	6	0.681	0.709	0.100
P3	12	0.045	0.417	0.107
P4	6	0.816	0.811	0.099
P5	8	0.692	0.721	0.097
P6	7	0.755	0.769	0.101
P7	10	0.771	0.738	0.096
P8	12	0.483	0.521	0.102
P9	12	0.733	0.575	0.098
P10	13	0.772	0.705	0.105
P11	12	0.736	0.670	0.102
P12	12	0.676	0.637	0.104
P13	16	0.699	0.620	0.093
P14	24	0.719	0.591	0.102
P15	28	0.517	0.462	0.103
P16	50	0.653	0.473	0.096

3.2. Results of selected test problems

In this section, due to space limitations, only results of a subset of the selected problems are presented to the reader in order to explore the challenges involved. In this paper, the effect of the adjacent nodes to a particular stream is not analyzed, since, according to Iordache et al. [6], this effect could be masked for complex networks.

3.2.1. Problem 5

This generic process from Yang et al. [45] is presented in Figure 2. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in Table 5 and Table 6, respectively. This example illustrate the challenge related with parallel streams which resulted in a non identifiability of Streams 4 and 5 neither by MT nor by GLRMT. Parallel streams exhibits the same behavior as splitted streams (since splitted streams are also parallel) and is a serious limitation of the MT method, (mentioned in Iordache et al. [6] and Tamhane and Mah [46]) and arise when columns of the incidence matrix, A, are proportional by a positive or negative constant. This behavior also happens in this type of mass balance with the GLR method, since proportional columns results in equal "gross error signature vector", making the identification impossible [11]. This is the main reason why large flowsheets are not divided into small problems when applying these tests: the division would result in small flowsheets with proportional columns in the incidence matrix [46]. Beyond, since many simultaneous or iterative GEDI strategies are based on MT, application of this method for this type of flowsheet structure is very limited. For other streams,

Figure 2: Generic mass balance flowsheet used in GED of Problem 5.

an important result appears when analyzing Table 5: the relation between adjustability, detectability and OP, which follows the same tendency: the streams with highest adjustability/detectability (streams 1 and 3) are the ones with highest OP; while the opposite also occurs (with Stream 6). This behavior is repeated among all other problems tested here. The Nodal Test presented an OP above 0.91 for all tests and nodes, as presented in Table 6, which can be considered a high magnitude for this indicator.

Table 5: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 5

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	98.7	0.994	0.467	0.846	0.774	0.774	10.2
2	41.1	0.641	0.240	0.651	0.596	0.596	18.7
3	78.9	0.888	0.451	0.836	0.784	0.784	11.3
4	30.2	0.550	0.094	0.424	0	0	-
5	109.1	1.045	0.408	0.806	0	0	-
6	19.8	0.445	0.137	0.504	0.374	0.374	31.8
7	57.6	0.759	0.367	0.774	0.71	0.711	13.6
8	37.8	0.615	0.279	0.693	0.626	0.627	17.1

Another test performed in this same flowsheet compares the use of WLS and Lorenztian estimator, keeping the same AVT1 for them, and is presented

Table 6: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 5.

Equipment	OP NT	OP GLR NT	AEE
	α_{NT}	β_{NTGLR}	%
1	0.968	0.968	7.6
2	0.952	0.952	10.1
3	0.983	0.983	6.6
4	0.919	0.918	10.2

in Table 7. It is important to notice the difference in OP for these estimators. Since the role of the robust estimator is to reduce the spread of gross error to other streams, the OP of this estimator is lower than the obtained using the WLS estimator and for this same reason leads the OP of stream 4 and 5 to became different than zero. To make use of this estimator to perform GEDI, it is necessary to tune its parameter and also to determine a criteria for the GEDI detection, as presented by literature [2, 1, 3], which can be considered also a challenge issue in the GEDI area.

Another test was executed to show the influence of the size of the GE on the Overall Power. A chart of the GE versus OP, presented in Figure 3, was built varying the ratio of GE and σ between 2 and 9 for streams 1, 2, 3, 6, 7 and 8 and keeping the AVT1 approximately in 0.1. Streams 4 and 5 were skipped because they OP are zero for all range of GE. It is important to notice that not all OP have the same behavior with the increase of GE and σ ratio. This type of chart is important to determine which streams are

Table 7: Results of MT (testing measurement bias) for measurement bias datasets of Problem 5 using the Lorenztian estimator.

Stream Name	Adjust.	Detect.	OP MT Lorenztian α_{MT}	OP MT WLS α_{MT}
1	0.4670	0.846	0.754	0.774
2	0.2399	0.650	0.451	0.596
3	0.4510	0.836	0.733	0.784
4	0.0943	0.424	0.032	0
5	0.4082	0.806	0.588	0
6	0.1365	0.504	0.124	0.374
7	0.3668	0.774	0.651	0.71
8	0.2794	0.693	0.667	0.626

Figure 3: Overall Power curve for Problem 5.

Figure 4: Generic mass balance flowsheet used in GED of Problem 9.

more leaning for GE detection and determining the factors that influences these curves can be considered a challenge in GEDI area.

3.2.2. Problem 9

The process diagram from Mandel et al. [47] is presented in Figure 4. The streams flow rates and standard deviation units are given in $t.h^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in Table 8 and Table 9, respectively, and the challenges related with flowsheet structure could be well evidenced. Problem 9 have two parallel streams: 1 and 2 (connecting process P1 to the environment node), leading MT and GLR not to identify GE in these streams. Another important issue that can be seen in this problem is the effect of the absolute standard deviation in the streams of the same balance: streams 3, 4 and 5 belongs to the same equipments and since stream 4 have the smallest standard deviation and flow rate, it presents the smallest OP of them. The role of the recycle on the adjustability, detectability and power of the stream with the lower standard deviation can also be observed in this problem: the recycle formed by stream 8, 9 and 10 reduces the adjustability, detectability and power of stream 8. NT for equipments P1 and P6 did not perform as well as P5, because the streams that connects to P1 and P6 have higher standard deviations when compared with other streams and this fact lead the method to attribute the node imbalance to measurement variability.

Table 8: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 9.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	230	38.0	0.933	0.998	0	0	-
2	21	1.1	0	0.029	0	0	-
3	209	5.0	0.544	0.891	0.507	0.507	8.4
4	35	1.8	0.055	0.327	0.104	0.104	78.0
5	174	2.0	0.147	0.522	0.384	0.384	27.5
6	15	0.9	0.019	0.193	0.036	0.036	190.1
7	159	7.3	0.746	0.967	0.839	0.839	4.5
8	50	1.0	0.063	0.349	0.149	0.149	64.0
9	209	5.0	0.599	0.916	0.755	0.755	7.3
10	94	2.0	0.212	0.616	0.310	0.310	20.7
11	115	18.1	0.867	0.991	0.853	0.853	2.4
12	44	2.4	0.322	0.733	0.403	0.403	15.6

Table 9: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 9.

Equipment	OP NT	OP GLR NT	AEE
	α_{NT}	β_{NTGLR}	%
P1	0.044	0.044	3.5
P2	0.722	0.728	4.1
P3	0.429	0.692	23.2
P4	0.406	0.695	23.7
P5	0.565	0.574	2.3
P6	0.106	0.118	7.7

Another test performed to illustrates the challenge involving the flowsheet structure is the influence a measured variable that occurs only in one equation with an unmeasured variable. In this case, the MT was performed considering Stream 7 as unmeasured, evaluating the influence to adjacent balances involving streams 5 and 6. The results are presented in Table 10. It is important to notice that not only Stream 6 had its OP reduced, but also streams 5 and 9. However, other streams presented a significant increase in the OP, such as 4, 8 and 11. It is not trivial to predict this behavior, since some streams with high flow rates (5 and 9) presented a decrease in the OP while stream 6 with a small flow rate also presented this behavior.

3.2.3. Problem 16

This example consists in ethylene plant from Zhang et al. [3], presented in Figure 5. The streams flow rates and standard deviation units are given in

Table 10: Results of MT for measurement bias datasets of Problem 9 with stream 7 unmeasured.

Stream Name	Real Value	Standard Deviation	OP MT Umeas α_{MT}	OP MT Meas α_{MT}
1	230	38.00	0	0
2	21	1.08	0	0
3	209	5.00	0.506	0.507
4	35	1.825	0.122	0.104
5	174	2.00	0.334	0.384
6	15	0.88	0.028	0.036
8	50	1.00	0.165	0.149
9	209	5.00	0.607	0.754
10	94	2.00	0.320	0.310
11	115	18.10	0.867	0.853
12	44	2.39	0.413	0.403

Figure 5: Ethylene plant flowsheet used in GED of Problem 16.

$t.h^{-1}$ (where t denotes *metric ton*). The GED results for measurement test and nodal test are presented in Table 11 and Table 12, respectively. This example illustrates the challenge in perform GEDI in complex flowsheets. Problem 16 has several parallel streams (such as streams S319, S316, S312 and S378; S460, S456 and S452; S592 and S581; S549 and S550), direct recycles (such as formed by streams 453 and 408), and indirect recycles (such as formed by streams 411, 401, 402, 404, 405 and 407). It is possible to notice that some streams have a lower adjustability, detectability and OP (such as streams S599, S536, S525, S384P, S415, S359P2, S358 and S357). The exact cause of this effect is difficult to isolate due to the network complexity but a relationship between the standard deviation, adjustability, detectability and OP could also be noticed here.

3.3. Discussion

As presented in Table 2, many authors test the proposed DR methods with only a few problems. One of the possible reason is the limited computation resources available in the past decade, since it was necessary strong computational effort to perform thousands of simulation runs, which involves random data generation, solving of the DR problem and GED testing. Since GEDI evaluation is a task involving several matrix operations and statistical tests, another reason for the GED test in a few problems could be the lack of good computational packages (such as Matlab, Maple, Scilab, etc.) which, nowadays, made this task simpler. Beyond, the fact that the problems were spread along the literature made the collection assembling a difficult task.

Table 11: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 16.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT $z_{a,j}, \alpha$	OP GLR MT $z_{a,j}, \beta$	AEE %
S319	225.45	2.06	0.007	0.117	0	0	-
S316	167.89	1.31	0.003	0.074	0	0	-
S312	1332.00	11.10	0.223	0.629	0	0	-
S378	1332.00	11.20	0.225	0.632	0	0	-
S336	2276.90	18.60	0.665	0.942	0.751	0.748	5.6
S357	137.50	1.22	0.146	0.520	0.211	0.209	28.9
S346	917.89	7.80	0.556	0.896	0.705	0.702	7.8
S359P1	532.38	4.74	0.285	0.699	0.515	0.512	15.4
S347	385.51	3.32	0.627	0.928	0.745	0.744	6.6
S352	385.51	3.32	0.627	0.928	0.736	0.735	6.5
S356	385.51	3.32	0.627	0.928	0.748	0.747	6.5
S358	100.32	0.85	0.061	0.344	0.092	0.090	75.9
S357P	285.19	2.26	0.485	0.857	0.674	0.670	9.5
S359P2	147.69	1.34	0.183	0.576	0.267	0.266	23
S359	680.07	6.35	0.461	0.842	0.668	0.666	9.9
S338P	683.07	6.35	0.337	0.749	0.541	0.539	13.5
S338	683.07	6.24	0.326	0.738	0.522	0.520	13.4
S341P	1593.80	15.00	0.552	0.894	0.718	0.716	8.2
S341	1593.8	15.00	0.552	0.894	0.708	0.706	7.9
S414	582.60	5.28	0.569	0.902	0.71	0.708	7.7
S502	582.60	5.28	0.569	0.902	0.708	0.706	7.6
S411	1178.20	9.30	0.179	0.571	0.359	0.355	23.2
S401	2872.50	24.00	0.658	0.940	0.748	0.746	5.8
S415	3.84	0.03	0	0.002	0.002	0.002	30000
S402	2876.30	26.00	0.687	0.950	0.736	0.733	5.4
S404	2870.00	26.00	0.687	0.950	0.752	0.75	6.9
S405	2876.30	26.00	0.687	0.950	0.748	0.745	5.5
S407	2876.30	26.00	0.687	0.950	0.746	0.744	5.3
S408	3098.20	26.00	0.547	0.892	0	0	-
S453	1400.00	11.50	0.081	0.395	0	0	-
S460	1337.70	12.20	0.460	0.842	0	0	-
S456	11.36	0.10	0	0.007	0	0	-
S452	349.07	2.98	0.021	0.206	0	0	-
S511	501.44	4.37	0.789	0.978	0.755	0.753	3.5

Table 12: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 16.

Equipment	OP NT	OP GLR NT	AEE
	α_{NT}	β_{NTGLR}	%
DA-301	0.999	1	5.9
Split1	0.990	0.998	5.9
COMP1	0.970	0.997	8.7
EA-32X	0.974	0.998	10.1
FA-309	0.990	1	8.9
Split2	0.992	0.996	6.1
Mixer1	0.975	0.983	4.2
FSPL	0.988	1	9.5
H338	0.957	0.982	6.4
H341	0.954	0.997	11.5
DA-401	0.997	1	8.64
PTC	0.961	0.992	8.7
M1	0.970	1	11.8
448-450	0.951	0.994	13.6
DC-401	0.974	0.997	12.2
452-448	0.963	0.998	12.9
DA-408	0.998	1	8.8
DA-402	0.994	0.999	9.1
DA-404	0.999	1	7.6
DA-405	0.988	0.995	5.3
DC-402	1	1	5.8
DC-40X	0.998	0.999	7.4
DA-407	0.979	0.999	11.4
DA-406	0.988	0.999	8.5
DA-409	0.987	0.988	0.3
Mix2	0.963	0.975	4.7

As afore mentioned, the GED tests were chosen because they are the base for several simultaneous or iterative strategies for GED. If GT, MT or NT does not presents good GEDI performance (low OP), it may indicate that the techniques that use these tests as a base, will also present difficulties in the correct GEDI. For this reason, this benchmark test set could help the scientific community to drive their efforts in the selection of individual or collective statistical test to perform simultaneous or iterative GEDI methods. The evaluation of the OP curve for test problems is an issue that have been poorly explored by literature, since some GEDI method may present a good performance within a range of GE but poorly in another range. This paper shows the importance of this evaluation presenting the OP curve of Problem 5.

Although the implementation itself is not the main focus of this present work, it is important to comment that, in the Scilab scripts available at Valle [43], several customizations are possible, such as:

- Choose multiple GE dataset generation with exact position of them (although no MGEDI technique has been implemented)
- Change the value of standard deviation to test the influence of its magnitude.
- Change the magnitude of gross errors (and consequently the ratio between the gross error and the respective measurement variance) and evaluate all the OP curves.
- Remove a stream from the measured set.

- Use other estimators (all estimators presented in Appendix C are available).

4. Conclusion

As presented in this paper, the DR and GEDI techniques can be influenced by many factors, such as flowsheet connections, measurement variance, position of the gross error, and the unmeasured variables. Consequently, GEDI is still a challenge even for a fully measured linear steady-state diagrams [24, 8], and, through the results of some problems presented, this fact could be confirmed.

In this work, special care was taken in the selection of the benchmark tests. With the appropriate classification according to the its challenge issue, as presented in Table 3, it is possible to evaluate the performance of new tests according to each issue, such as parallel stream, recycles, etc. The benchmark problems also allow the reader to have an idea of the particularities that appear when performing GEDI techniques: for some small problems the causes that leads to a better or worse performance of GEDI techniques could be isolated, while for complex diagrams, the source of increase or decrease in the GEDI performance could not be identified.

The relationship between detectability, adjustability and OP was an important observation among the test problems, since adjustability and detectability calculation does not involve high computational load when compared with the simulation procedure to evaluate the OP.

The literature review shows the need for a common data for validating GED techniques due to the problem complexity, which is presented in this paper.

These data, collected from literature or proposed here, can be considered the most complete benchmark set ever presented in the literature. The public availability (Valle [43]) and the modular implementation of the problems and GEDI techniques in Scilab [42] make the extension of this work (both for new problems and GEDI techniques) a simple task. Beyond, several customizations are allowed to test other particularities of each problem, such as robust estimator selection, multiple gross error data generation, removing of measured streams, etc. With the datasets scheme generation developed, it will be possible to validate, detect and analyze the deficiencies of GEDI techniques in a more general and robust manner.

5. Acknowledgement

The authors would like to thank Zhengjiang Zhang, for the flowsheets diagrams and streams exact values of Problem 16, Prof. Fernando Luiz Pellegrini Pessoa, for suggestions in the paper structure, and Prof. Shankar Narasimhan for explanations about some GED performance indexes.

Appendix A. GEDI Tests

The Global Test (GT) uses the results of the data reconciliation, the variance matrix and the mass balance constraints matrix to perform statistical tests. First, the mass balance residuals matrix, $\mathbf{R}$, is calculated based on measurements and balance matrix, as presented in Equation A.1 (where $\mathbf{A}$ is the process constraints matrix).

$$\mathbf{R} = \mathbf{A}.\mathbf{Y} \tag{A.1}$$

In the next step, as proposed by Madron [48] (apud Narasimhan and Jordache [8]), the variance-covariance matrix is calculated by Equation A.2.

$$\mathbf{V} = \mathbf{cov(R)} = \mathbf{A}.\Sigma.\mathbf{A}^T \tag{A.2}$$

Then, γ, is calculated by Equation A.3.

$$\gamma = \mathbf{R}^T.\mathbf{V}^{-1}.\mathbf{R} \tag{A.3}$$

As γ collects in only one parameter the interactions of the whole mass balance, it is applied a multivariate statistical test χ^2 with ν degrees of freedom, where ν is the *rank* of matrix $\mathbf{A}$. The global test can detect a gross error in the flowsheet as a whole but cannot locate its position or type (leaking or measurement bias).

The nodal test is applied to the balance residuals without consider its iteration[49]. For each equipment, or node, the $z_{r,i}$ is calculated by Equation A.4 and tested with a monovariate statistical hypothesis test z.

$$z_{r,i} = \frac{\mid r_i \mid}{\mid \sqrt{V_{i,i}} \mid} \tag{A.4}$$

where:

r_i: Balance residual, or the ith line of the R matrix

Another test used to detect measurement bias is the Measurement Test (MT) and it is based in the adjustment after the reconciliation (defined by Equation 4). To calculate the MT statistic, $z_{a,j}$ is calculated based on values from **Ad** calculation and $\overline{\mathbf{W}}$ matrix , which is computed by Equation A.5.

$$\overline{\mathbf{W}} = \Sigma\mathbf{A}^T.\mathbf{V}^{-1}.\Sigma \tag{A.5}$$

Figure B.6: Algorithm for evaluation of GED technique for measurement bias (GT, MT and GLRMT).

Then, $z_{a,j}$ is calculated by Equation A.6.

$$z_{a,j} = \frac{\mid a_j \mid}{\mid \sqrt{\overline{w_{j,j}}} \mid} \tag{A.6}$$

where:

a_j: The jth line of the **Ad** matrix

$\overline{w_{j,j}}$: The j, j element of the $\overline{\mathbf{W}}$ matrix

The measurement test is performed using a monovariate statistical hypothesis test z based on the value of the computed $z_{a,j}$ and can detect a gross error in the measurement j.

Appendix B. GED Detection Scheme

When a new GED or GEDI is proposed, the most used method to evaluate it is the Monte Carlo simulation approach and the steps for the algorithm is presented in Figure B.6.

The detailed steps for the algorithm presented in Figure B.6 is described below.

1 Select a process flowsheet.

2 Select exact or nominal steady state values for the flows of all streams, satisfying mass balances in all process units, named $x_{i,ns}$.

3 Set the standard deviation of the errors in each measurement.

4 Generate a random error vector of size NS (where NS is the number of simulation trials), $\varepsilon_{i,ns}$, for each measurements, with normal distribution with zero mean and with the given standard deviation from step 3. Add the $\varepsilon_{i,ns}$ to the true values: $y_{i,ns} = x_{i,ns} + \varepsilon_{i,ns}$ (or $\mathbf{Y}_{ns}$ in the vector notation).

4.1 If it is the first run in the loop, go to step 6, otherwise, go next step.

5 For each stream, generate a vector of gross errors with size of NS, named $\delta_{i,ns}$. The magnitude and the sign of the gross error must be uniformly and randomly chosen between specified upper and lower bound: $l\,(x_{i,ns} + \varepsilon_{i,ns}) \leq \delta_{i,ns} \leq u\,(x_{i,ns} + \varepsilon_{i,ns})$. Where l is the lower fraction and u is the upper fraction of the random gross error value that will be generated. Add $\delta_{i,ns}$, generated in step 5, to $y_{i,ns}$ (previously evaluated in step 4), to obtain the final measurement. In the case of multiple gross error generation, the location must be uniformly and randomly selected, e.g., some measurements will not have such term.

6 Choose the level of the significance, α, for the statistical test. This is generally chosen in order to keep the average Type I error to 0.1 when no gross error are present.

7 Perform DR and GED. The steps from 4.1 to 7 must be repeated NS times.

8 Evaluate the GED using the performance indexes, using the results from NS simulation trials in order to obtain an average error detection value indexes.

Figure B.7: Algorithm for evaluation of GED technique for equipment leaking (NT and GLRNT).

In the present work, the lower and upper bounds multipliers factor, l and u was chosen as $2.\sigma x_{i,ns}$ and $7.\sigma x_{i,ns}$, respectively. The WLS objective function and the reconciliation was carried out using the method from Lagrange Multipliers, were the optimization problem is solved by B.1.

$$\mathbf{X} = \left[\mathbf{I} - \Sigma.\mathbf{A}^T.\left(\mathbf{A}.\Sigma.\mathbf{A}^T\right)^{-1}.\mathbf{A}\right].\mathbf{Y} \tag{B.1}$$

It was also performed a simulation approach for Nodal Test (NT) evaluation, presented in Figure B.7, which is slightly different from the previous one, since it is not necessary to perform the DR for this test.

The algorithm presented in Figure B.7, is detailed below.

1 Perform steps 1 to 4 of the algorithm presented in B.7.

2 For each of the NS simulations, take the measurement data set (Y_{ns}) and calculate, for each equipment, the residual vector, named $r_{j,ns}$ (or R_{ns} in the vector notation).

2.1 If it is the first run in the loop, go to step 5, otherwise, go next step.

3 For each equipment, generate a vector of gross errors with size of NS, named $\delta^r_{j,ns}$. The sign of the gross error in this case is always negative, since we want to simulate a process leak. The magnitude must be randomly chosen between specified lower and upper bounds: $l^r_j\,(\text{total nf}_j) \leq \delta^r_{j,ns} \leq u^r_j\,(\text{total nf}_j)$. Where total nf_j is the sum of total flows of streams entering plus leaving the equipment. Then, perform

the sum of $r_{j,ns}$ and $\delta^r_{j,ns}$ to obtain the final residual (random plus gross error).

4 Choose the level of the significance, α, for the statistical test. This is generally chosen in order to keep the average Type I error to 0.1 when no gross error are present.

5 Perform GED. The steps from 2.1 to 5 must be repeated NS times.

6 Evaluate the GED using the performance indexes using the results from NS simulation trials in order to obtain an average error detection value indexes.

For the problems studied, l^r_j and u^r_j were set as 0.02 and 0.07, respectively.

Appendix C. Robust Estimators

ε Error definition:

$\varepsilon_i = (y_i - x_i)/\sigma_i$

WLS

$^1/_2.\varepsilon_i^2$

Contaminated Normal

$$-ln\left[(1-p_{CN}).exp\left(\frac{-\varepsilon_i^2}{2}\right)+\frac{p_{CN}}{b_{CN}}.exp\left(\frac{-\varepsilon_i^2}{2b_{CN}^2}\right)\right]$$

Cauchy

$c_C^2 \ln\left(1 + \frac{\varepsilon_i^2}{c_C^2}\right)$

Logistic

$2\ln\left(1 + exp\left(\frac{\varepsilon_i}{c_{Lo}}\right)\right) - \frac{\varepsilon_i}{c_{Lo}}$

Lorentzian

$-\frac{1}{1+\left(\varepsilon_i^2/2c_{CL}^2\right)}$

Fair

$2c_F^2\left(\frac{|\varepsilon_i|}{c_F} - \ln\left(1 + \frac{|\varepsilon_i|}{c_F}\right)\right)$

Hampel's redescending M-estimator

$\frac{1}{2}\varepsilon_i^2,\ 0 \leq |\varepsilon_i| \leq a_H$

$a_H.|\varepsilon_i| - \frac{1}{2}.a_H^2,\ a_H < |\varepsilon_i| \leq b_H$

$a_H.b_H - \frac{1}{2}.a_H^2 + (c_H - b_H).\frac{a_H^2}{2}.\left(1 - \left(\frac{c_H - |\varepsilon_i|}{c_H - b_H}\right)^2\right),\ b_H < |\varepsilon_i| \leq c_H$

$a_H.b_H - \frac{1}{2}.a_H^2 + (c_H - b_H).\frac{a_H^2}{2},\ c_H < |\varepsilon_i|$

Quasi-Weighted

$\frac{\varepsilon_i^2}{2+\beta|\varepsilon_i|}$

List of Symbols and Nomenclature

α Significance level for MT or NT

α_{MT} Significance level of MT statistics

α_{NT} Significance level of NT statistics

β_{GLRMT} Significance level of GLR statistics for measurent bias testing

β_{GLRNT} Significance level of GLR statistics for leakings testing

χ^2 Chi-squared test

γ Value of the Global Test statistics

ν Degrees of freedom for Chi-squared test

$\overline{w_{j,j}}$ The j, j element of the $\overline{\mathbf{W}}$ matrix

ρ The generic objective function of DR

Σ Variance covariance matrix of measurements

σ_{x_i} Standard deviation of reconciled variables

σ_{y_i} Standard deviation of the measurements

$\mathbf{R}$ Mass balance residuals matrix

$\mathbf{X}$ Vector of reconciled measurements

$\mathbf{X}_{max}$ Vector of reconciled measurement upper bounds

$\mathbf{X}_{min}$ Vector of reconciled measurement lower bounds

$\mathbf{Y}$ Vector of measurements

ε_i Residual of measured and reconciled data, weighted by the reciprocal of the standard deviation

A Incidence Matrix (Jacobian matrix of the constraints)

a_j: The jth line of the **Ad** matrix

Ad Adjust

$F(\mathbf{X})$ Vector of process constraints related with mass, energy or momentum balance

NS Number of simulation trials

r_i Balance residual, or the ith line of the R matrix

V Covariance matrix of residuals matrix, R

W Inverse of Σ matrix

Wad Weighted Adjust

AEE Average Error of Estimation

AIC Akaike Information Criterion

AVT1 Average Number of Type I Errors

DMT Dynamic Measurement Test

DR Data Reconciliation

EMT Extended Measurement Test

GED Gross Error Detection

GEDI Gross Error Detection and Identification

GLR Generalized Likelihood Ratio

IMT Iterative Measurement Test

MGEDI Multiple Gross Error Detection and Identification

MGLR Modified Generalized Likelihood Ratio

MI Mixed Integer

MSEGE Modified Simultaneous Estimation of Gross Error

MUBET Modified Unbiased Estimation Technique

OP Overall Power

PCA Principal Componets Analysis

SD Standard deviation

SEGE Simultaneous Estimation of Gross Error

UBET Unbiased Estimation Technique

WLS Weighted Least Squares

References

[1] N. Arora, L. T. Biegler, Computers and Chemical Engineering 25 (2001) 1585–1599.

[2] D. B. Özyurt, R. W. Pike, processes. Computers and Chemical Engineering 28 (2004) 381–402.

[3] Z. Zhang, Z. Shao, X. Chen, K. Wang, J. Qian, Computers & Chemical Engineering 34 (2010) 154–162.

[4] H. Akaike, IEEE Transactions on Automatic Control AC-19 (1974) 716.

[5] J. Rosenberg, R. S. H. Mah, C. Iordache, Ind. & Eng. Chem. Proc. Des. Dev, Vol. 26 (1987) 555–564.

[6] C. Iordache, R. S. H. Mah, A. C. Tamhane, AIChE Journal 31 (1985) 1187–1201.

[7] M. Sánchez, J. Romagnoli, Q. Jiang, M. Bagajewicz, M. Sa, Computers & Chemical Engineering 23 (1999) 841 – 857.

[8] S. Narasimhan, C. Jordache, Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data, Gulf Publishing, Houston, 1 edition, 2000.

[9] R. W. Serth, W. A. Heenan, AIChE Journal 32 (1986) 733–747.

[10] D. K. Rollins, J. F. Davis, AIChE Journal 38 (1992) 563–572.

[11] S. Narasimhan, R. S. H. Mah, AIChE Journal 33 (1987) 1514–1521.

[12] P. Harikumar, S. Narasimhan, Computers & Chemical Engineering 17 (1993) 1121–1128.

[13] J. Romagnoli, G. Stephanopoulos, Chemical Engineering Science 36 (1981) 1849–1863.

[14] V. N. Reddy, M. L. Mavrovouniotis, Chemical Engineering Research and Design 76 (1998) 478–489.

[15] M. Bagajewicz, Q. Jiang, M. Sánchez, Computers & Chemical Engineering 23 (1999) S589–S592.

[16] R. Kongsjahju, D. K. Rollins, M. B. Bascuñana, Trans IChemE 78 (2000) 1010–1018.

[17] T. A. Soderstrom, D. M. Himmelblau, T. F. Edgar, Control Engineering Practice 9 (2001) 869–876.

[18] F. Wang, X.-p. Jia, S.-q. Zheng, J.-c. Yue, Computers & Chemical Engineering 28 (2004) 2189–2192.

[19] S. Devanathan, S. B. Vardeman, D. K. Rollins Sr., D. K. Rollins, Chemical Engineering Research and Design 83 (2005) 1391–1398.

[20] C. Mei, H. Su, J. Chu, Chinese Journal of Chemical Engineering 14 (2006) 592–596.

[21] M. Alhaj-Dibo, D. Maquin, J. Ragot, Control Engineering Practice 16 (2008) 159–170.

[22] R. Maronna, J. Arcas, Computers & Chemical Engineering 33 (2009) 65–71.

[23] A. C. Farias, Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros, Ph.D. thesis, 2009.

[24] C. M. Crowe, Journal of Process Control 6 (1996) 89–98.

[25] W. Hock, K. Schittkowski, Lecture Notes in Economics and Mathematical Systems (1981).

[26] K. Schittkowski, Lecture Notes in Economics and Mathematical Systems (1987).

[27] F. Mazzia, C. Magherini, Test set for initial value problem solvers, release 2.4, 2008.

[28] J. J. Downs, E. Vogel, Computers and Chemical Engineering 17 (1993) 245–255.

[29] J. Castro, Journal of Process Control 14 (2004) 329–347.

[30] R. Chen, K. Dave, T. J. McAvoy, M. Luyben, Industrial & Engineering Chemistry Research 2003 (2003) 4478–4487.

[31] J. Copp, The COST Simulation Benchmark Description and Simulator Manual, Technical Report, COST (European Cooperation in the field of Scientific and Technical Research), Brussels, Belgium, 2001.

[32] J. Nielsen, Computers & Chemical Engineering 21 (1997) S469–S474.

[33] D. Fontana, Recuperação de águas de processos, desenvolvimento de um problema padrão, Dissertação (mestrado em pesquisa e desenvolvimento de processos), Universidade Federal do Rio Grande do Sul, 2002.

[34] W. Leong, K. Hollands, a. Brunger, International Journal of Heat and Mass Transfer 41 (1998) 3817–3828.

[35] P. Rajeshkanna, International Journal of Heat and Mass Transfer 49 (2006) 3929–3941.

[36] E. Barth, B. Leimkuhler, S. Reich, Computational methods for macromolecules: challenges and applications. Lect. Notes Comput. Sci. Eng. 24 (2002) 73–103.

[37] I.-w. Kim, M. S. Kang, S. Park, T. F. Edgar, Computers & Chemical Engineering 21 (1997) 775–782.

[38] J. Y. Keller, M. Darouach, G. Krzakala, Comp. & Chem. Eng 18 (1994) 1001–1004.

[39] S. R. Singth, N. Mittal, P. Sen, Minerals Engineering 14 (2001) 809–814.

[40] M. C. Sánchez, G. Sentoni, S. Schbib, S. Toneili, J. Romagnoli, Comp. & Chem. Eng 20 (1996) S1559–S1564.

[41] G. Heyen, E. Marechál, B. Kalitventzeff, Computers & Chemical Engineering 20 (1996) S539–S544.

[42] INRIA, Scilab, 2010.

[43] E. C. Valle, Data Reconciliation and Gross Error Detection Benchmark Problems, 2011.

[44] A. Wächter, L. T. Biegler, Mathematical Programming 106 (2006) 25–57.

[45] Y. Yang, R. Ten, L. Jao, Computers & Chemical Engineering 19 (1995) 217–222.

[46] A. C. Tamhane, R. S. H. Mah, Technometrics 24 (1985) 409–422.

[47] D. Mandel, A. Abdollahzadeh, D. Maquin, J. Ragot, International Journal of Mineral Processing 53 (1998) 157–169.

[48] F. Madron, Chemica Engineering Science 40 (1985) 1855–1860.

[49] R. S. H. Mah, G. M. Stanley, D. M. Downning, Ind. & Eng. Chem. Proc. Des. Dev, Vol. 15 (1976) 175–183.

Collection of Benchmark Test Problems for Gross Error Detection

Edson Cordeiro do Valle[a,b,*], Ricardo de Araújo Kalid[b,*], Argimiro Resende Secchi[c,*], Asher Kiperstok[b]

[a] *VRTech Tecnologias Industriais Ltda - Rua Guilherme Schmitz, 625, Gravatái - RS - CEP: 94065-200, Brazil*
[b] *Programa de Engenharia Industrial - Escola Politécnica - Universidade Federal da Bahia - Rua Aristides Novis, 2, 6 sexto. andar, Salvador - BA - CEP: 40210-630, Brazil*
[c] *COPPE - Universidade Federal do Rio de Janeiro - Centro de Tecnologia, Bloco G, sala 115 Cidade Universitária, Rio de Janeiro - RJ - CEP 21941-972*

Keywords: data reconciliation, gross error detection, benchmark problems

1. Supplement Material

In the present work, 16 test problems from previous paper are presented. The main challenge issue related to each problem is presented in Table 1 and its influence on the GEDI will be further discussed in the results section.

1.1. Problem 1

This example consists in a simple separator, presented in 1 while the GED results for MT and GLRMT is presented in 2. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The OP for the NT (both using α_{NT} and β_{NT}) was 0.7. For the GLR test, considering a possible leaking, the OP was 0.706 with an AEE of 0 (zero).

1.2. Problem 2

Problem 2 is a process with reaction and separation presented in 2. The streams flow rates and standard deviation units are given in $m^3.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 3 and 4, respectively.

*Corresponding author
Email addresses: edsoncv@gmail.com (Edson Cordeiro do Valle), kalid@ufba.br (Ricardo de Araújo Kalid), arge@peq.coppe.ufrj.br (Argimiro Resende Secchi)

Table 1: Characteristics of linear problems.

Problem	Streams	Characteristics
P1	2,3	S
P2	2,3;5,6	P
P3	2,11;9,12;4,5,6,7,8	P
	1,2(11),9(12),3	R
P4		-
P5	4,5	P
	2,6,7,8	M
P6	2,3,4;2,3,5,6	R
P7	5,9;8,10	M
	5,7,8,9;2,3,5,6	R
P8	2,3;6,7;8,9;10,11	P
P9	1,2	P
	8,9,10	R
	3,4,5;5,6,7	M
P10	1,2;8,9,10,13;5,11,12	M
P11	2,3,6,5;2,7,9,8,5;9,12,10	R
	5,6,8;10,11,12	M
P12	5,6,7	P
	1,2,3,4;4,9,10;8,9,11;10,11,12	M
P13	12,13	P
	2,3,4;6,7,10;12(13),14,16	R
	1,2,4;6,7,8;7,9,10;14,15,16	M
P14	1,2,3,4;21,22;23,24	P
	10,11,15;15,16,19	M
P15	10,12,22,18,7,5;26,8,5,1,3,15,26	R
	1,2,3,4;5,6,7,8,9;10,11,12;12,16,22,25,...	M
P16	319,316,378,312;456,460;592,581;549,550,..	P
	408,453;441,401,402,404,405,407;...	R
	503,592,581;537,598,599;...	M

*P: Parallel Streams; R: Recycle
M: Magnitude of SD (or Measurement)

Table 2: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 1.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	8.50	0.277	0.467	0.846	0	0	0	-
2	4.50	0.145	0.104	0.443	0	0	0	-
3	4.00	0.097	0.045	0.296	0	0	0	-

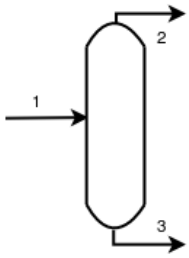


Figure 1: Separator equipment used in GED of Problem 1.

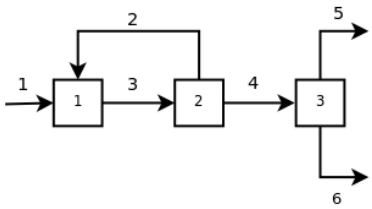


Figure 2: Reaction and separation flowsheet used in GED of Problem 2.

Table 3: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 2.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	11	0.179	0.388	0.791	0.748	0.748	0.748	13.6
2	10	0.161	0.085	0.404	0	0	0	-
3	21	0.346	0.504	0.869	0	0	0	-
4	11	0.182	0.397	0.798	0.767	0.769	0.769	13.1
5	5.5	0.228	0.398	0.798	0	0	0	-
6	5.5	0.122	0.097	0.429	0	0	0	-

Table 4: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 2.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.899	0.899	0.962	16.4
2	0.910	0.910	0.977	17.5
3	0.765	0.766	0.81	9.9

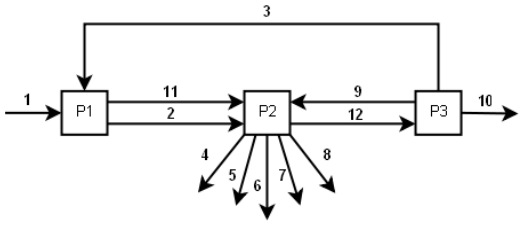


Figure 3: Atmospheric tower flowsheet used in GED of Problem 3.

1.3. Problem 3

This problem consists in an atmospheric tower from Zhang et al. [1] and presented in 3. The streams flow rates and standard deviation units are given in $t.h^{-1}$ (where t denotes *metric ton* or 1000 kg). The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 5 and 6, respectively.

Table 5: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 3

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	189.98	2.14	0.565	0.903	0.831	0.831	0.831	8.4
2	174.61	1.95	0.516	0.875	0	0	0	-
3	3.14	0.034	0.001	0.045	0.016	0.016	0.016	1370
4	32.77	0.345	0.017	0.183	0	0	0	-
5	33.47	0.400	0.023	0.212	0	0	0	-
6	7.25	0.087	0.001	0.046	0	0	0	-
7	0.32	0.004	0	0.002	0	0	0	-
8	92.38	1.059	0.172	0.562	0	0	0	-
9	28.63	0.361	0.108	0.452	0	0	0	-
10	23.81	0.300	0.08	0.401	0.199	0.199	0.199	48.9
11	18.52	0.201	0.004	0.009	0	0	0	-
12	55.57	0.645	0.413	0.810	0	0	0	-

1.4. Problem 4

This example presented by Narasimhan and Jordache [2] is shown in 4 and consists in a process with 4 equipments and 6 streams with exact mass balance. The streams flow rates are given in $kg.s^{-1}$. The standard deviation for all streams is 1 $kg.s^{-1}$ and the adjustability and detectability are all equal for all stream: 0.422 and 0.816, respectively. The GED results for all measurement tests (MT and GLR) and nodal tests (NT and GLR) are presented in 7 and 8, respectively.

Table 6: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 3.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.969	0.969	0.981	8.6
2	0.996	0.996	0.999	7.0
3	0.980	0.980	0.985	5.0

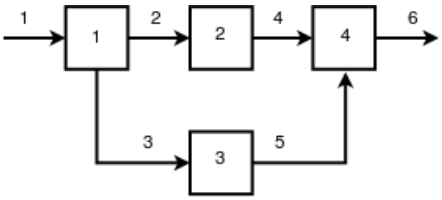


Figure 4: Heat exchanger with by-pass valve flowsheet used in GED of Problem 4.

Table 7: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 4

Stream Name	Real Value	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	100	0.764	0.764	0.764	12.6
2	64	0.768	0.768	0.768	11.8
3	36	0.765	0.765	0.765	12.0
4	64	0.766	0.766	0.766	12.2
5	36	0.778	0.778	0.778	11.8
6	100	0.754	0.754	0.754	12.2

Table 8: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 4.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.937	0.938	0.972	12.4
2	0.864	0.864	0.933	14.6
3	0.535	0.536	0.714	20.4
4	0.943	0.944	0.976	12.4

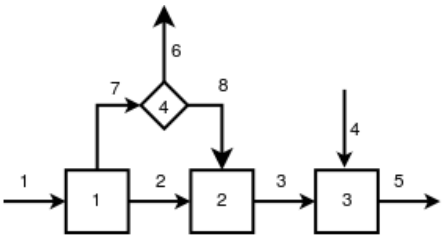


Figure 5: Generic mass balance flowsheet used in GED of Problem 5.

1.5. Problem 5

This generic process from Yang et al. [3] is presented in 5. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 9 and 10, respectively.

Table 9: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 5

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	98.7	0.994	0.467	0.846	0.774	0.774	0.774	10.2
2	41.1	0.641	0.240	0.651	0.596	0.596	0.596	18.7
3	78.9	0.888	0.451	0.836	0.784	0.784	0.784	11.3
4	30.2	0.550	0.094	0.424	0	0	0	-
5	109.1	1.045	0.408	0.806	0	0	0	-
6	19.8	0.445	0.137	0.504	0.374	0.374	0.374	31.8
7	57.6	0.759	0.367	0.774	0.71	0.711	0.711	13.6
8	37.8	0.615	0.279	0.693	0.626	0.627	0.627	17.1

1.6. Problem 6

This flowsheet is a heat exchanger network with recycle from Rosenberg et al. [4] is presented in 6 and was studied by many authors: Narasimhan and Mah [5], Iordache et al. [6], Rosenberg et al. [4], Sánchez et al. [7] and others. The streams flow rates are given in $kg.s^{-1}$. In this problem, the standard deviation is constant and equal to 1 $kg.s^{-1}$ for all streams. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 11 and 12, respectively.

Table 10: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 5.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.968	0.968	0.981	7.6
2	0.952	0.952	0.98	10.1
3	0.983	0.983	0.991	6.6
4	0.919	0.918	0.95	10.2

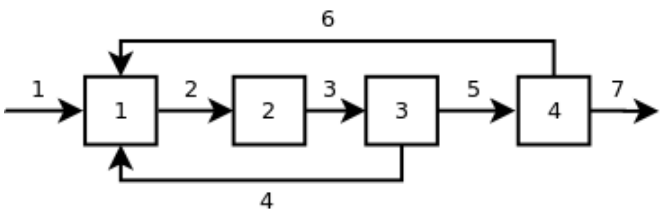


Figure 6: Heat exchanger network with recycle flowsheet used in GED of Problem 6.

Table 11: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 6.

Stream Name	Real Value	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	5	0.383	0.787	0.721	0.721	0.721	12.4
2	15	0.383	0.787	0.725	0.725	0.725	13.2
3	15	0.383	0.787	0.708	0.708	0.708	13.5
4	5	0.276	0.690	0.632	0.631	0.631	16.8
5	10	0.345	0.756	0.676	0.676	0.676	13.9
6	5	0.276	0.690	0.615	0.615	0.615	17.3
7	5	0.383	0.787	0.718	0.718	0.718	13.0

Table 12: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 6.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.062	0.062	0.109	27.8
2	0.116	0.116	0.204	24.2
3	0.074	0.074	0.148	29.4
4	0.044	0.044	0.060	21.1

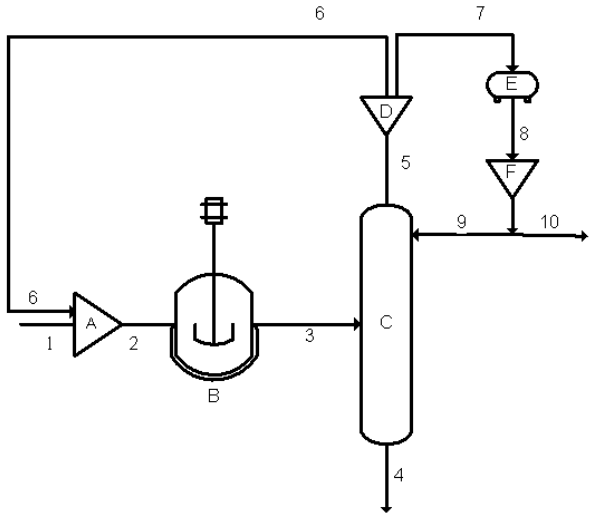


Figure 7: Diagram of a reaction-separation flowsheet used in GED of Problem 7.

1.7. Problem 7

An expanded reaction-separation diagram with recycle is proposed in Problem 7 and its diagram presented in 7. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 13 and 14, respectively.

1.8. Problem 8

This example presented by Meyer et al. [8] (and also studied in Rao and Narasimhan [9]) consists in a juice extraction plant and it is presented in 8. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 15 and 16, respectively.

1.9. Problem 9

The process diagram from Mandel et al. [10] is presented in 9. The streams flow rates and standard deviation units are given in $t.h^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 17 and 18, respectively.

1.10. Problem 10

Problem 10 is a generic mass balance from Martins et al. [11] and presented in 10. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 19 and 20, respectively.

Table 13: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 7.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	50	1	0.413	0.810	0.712	0.712	0.712	12.3
2	75	1	0.421	0.815	0.746	0.747	0.747	11.8
3	75	1	0.421	0.815	0.736	0.738	0.738	11.8
4	48	1	0.403	0.802	0.726	0.727	0.727	12.5
5	30	1	0.406	0.805	0.718	0.719	0.719	12.5
6	25	1	0.414	0.810	0.721	0.722	0.722	11.7
7	5	0.387	0.408	0.806	0.708	0.708	0.708	12.5
8	5	0.387	0.408	0.806	0.736	0.737	0.737	12.1
9	3	0.316	0.213	0.617	0.421	0.421	0.421	19.7
10	2	0.316	0.213	0.623	0.422	0.422	0.422	21.4

Table 14: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 7.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
A	0.824	0.824	0.918	15.0
B	0.904	0.904	0.960	12.6
C	0.825	0.825	0.917	15.2
D	0.330	0.330	0.520	22.8
E	0.063	0.062	0.095	17.8
F	0.058	0.058	0.083	16.3

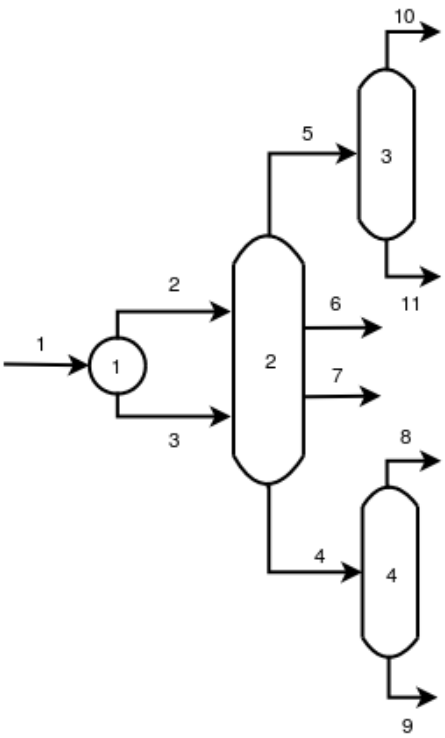


Figure 8: Juice extraction plant flowsheet used in GED of Problem 8.

Table 15: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 8.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	3600	110	0.597	0.915	0.827	0.827	0.827	7.8
2	1850	60	0.196	0.595	0	0	0	-
3	1750	54	0.176	0.566	0	0	0	-
4	2837	90	0.488	0.859	0.801	0.801	0.801	10.4
5	730	22	0.336	0.748	0.548	0.548	0.548	15.2
6	25	0.781	0	0.009	0	0	0	-
7	8	0.233	0	0.003	0	0	0	-
8	137	3.15	0	0.032	0	0	0	-
9	2700	80	0.473	0.851	0	0	0	-
10	58	1.71	0.002	0.058	0	0	0	-
11	672	20	0.268	0.682	0	0	0	-

Table 16: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 8.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.572	0.571	0.623	11.7
2	0.648	0.649	0.786	18.6
3	0.496	0.496	0.502	4.6
4	0.477	0.477	0.576	15.5

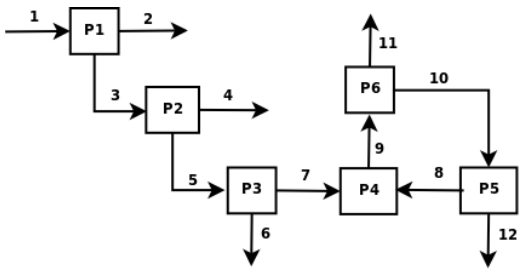


Figure 9: Generic mass balance flowsheet used in GED of Problem 9.

Table 17: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 9.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	230	38	0.933	0.998	0	0	0	-
2	21	1.1	0	0.029	0	0	0	-
3	209	5.0	0.544	0.891	0.507	0.507	0.507	8.4
4	35	1.8	0.055	0.327	0.104	0.104	0.104	78.0
5	174	2.0	0.147	0.522	0.384	0.384	0.384	27.5
6	15	0.9	0.019	0.193	0.036	0.036	0.036	190
7	159	7.3	0.746	0.967	0.839	0.839	0.839	4.5
8	50	1.0	0.063	0.349	0.149	0.149	0.149	64.0
9	209	5.0	0.599	0.916	0.755	0.755	0.755	7.3
10	94	2.0	0.212	0.616	0.310	0.310	0.310	20.7
11	115	18.1	0.867	0.991	0.853	0.853	0.853	2.4
12	44	2.4	0.322	0.733	0.403	0.403	0.403	15.6

Table 18: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 9.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
P1	0.044	0.042	0.044	3.5
P2	0.722	0.719	0.728	4.1
P3	0.429	0.427	0.692	23.2
P4	0.406	0.404	0.695	23.7
P5	0.565	0.560	0.574	2.3
P6	0.106	0.104	0.118	7.7

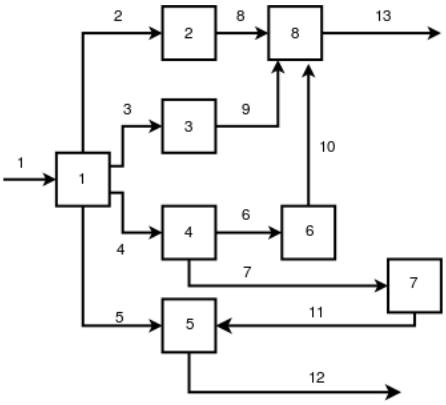


Figure 10: Generic mass balance flowsheet used in GED of Problem 10.

Table 19: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 10.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	28	0.275	0.498	0.865	0.717	0.717	0.717	9.4
2	5	0.050	0.342	0.753	0.554	0.554	0.554	14.2
3	5	0.172	0.587	0.911	0.791	0.791	0.791	7.5
4	7	0.145	0.512	0.873	0.757	0.757	0.757	9.3
5	11	0.372	0.645	0.935	0.792	0.792	0.792	6.5
6	4	0.127	0.572	0.904	0.792	0.792	0.792	8.1
7	3	0.136	0.607	0.919	0.769	0.769	0.769	7.1
8	5	0.045	0.276	0.691	0.486	0.486	0.486	17.4
9	5	0.095	0.251	0.663	0.557	0.557	0.557	17.9
10	4	0.073	0.249	0.661	0.554	0.554	0.554	17.6
11	3	0.064	0.157	0.538	0.403	0.403	0.403	26.0
12	14	0.147	0.159	0.541	0.387	0.386	0.386	28.0
13	14	0.131	0.387	0.790	0.709	0.709	0.709	12.7

Table 20: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 10.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.896	0.896	0.972	14.6
2	0.968	0.968	0.970	3.2
3	0.443	0.443	0.532	13.3
4	0.574	0.575	0.781	22.0
5	0.678	0.679	0.836	17.3
6	0.502	0.504	0.647	16.1
7	0.296	0.296	0.434	17.2
8	0.983	0.983	0.988	5.6

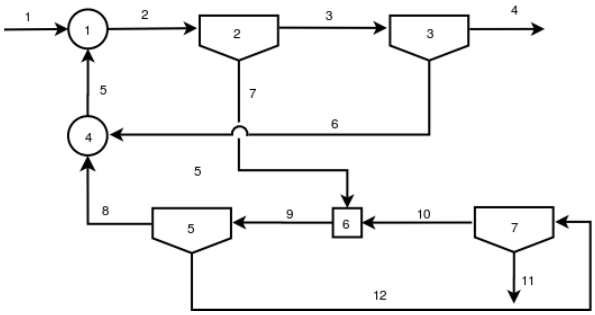


Figure 11: Mineral beneficiation circuit flowsheet used in GED of Problem 11.

1.11. Problem 11

Problem 11 is a typical mineral beneficiation circuit from Rao and Narasimhan [9], reproduced in 11. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 21 and 22, respectively.

Table 21: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 11.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	690	21	0.493	0.862	0.765	0.765	0.763	9.6
2	725	22	0.518	0.876	0.798	0.798	0.796	9.3
3	700	21	0.499	0.865	0.779	0.779	0.778	9.5
4	685	21	0.490	0.860	0.758	0.758	0.757	10.1
5	35	1.076	0.532	0.884	0.441	0.441	0.441	8.8
6	15	0.375	0.049	0.309	0.072	0.072	0.072	90.0
7	25	0.836	0.519	0.877	0.778	0.778	0.777	9.2
8	20	0.701	0.426	0.819	0.723	0.723	0.722	11.6
9	30	0.680	0.403	0.803	0.728	0.728	0.726	11.9
10	5	0.144	0.101	0.438	0.174	0.174	0.174	43.0
11	5	0.136	0.090	0.416	0.147	0.147	0.147	43.0
12	10	0.279	0.431	0.823	0.522	0.522	0.522	11.4

1.12. Problem 12

Mitsas [12] presented a generic process presented in 12. The streams flow rates and standard deviation units are given in $l.min^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 23 and 24, respectively.

Table 22: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 11.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.426	0.427	0.542	15.9
2	0.444	0.444	0.636	21.0
3	0.424	0.425	0.541	15.2
4	0.482	0.482	0.550	10.7
5	0.591	0.592	0.704	16.1
6	0.527	0.527	0.600	12.0
7	0.557	0.557	0.588	7.21

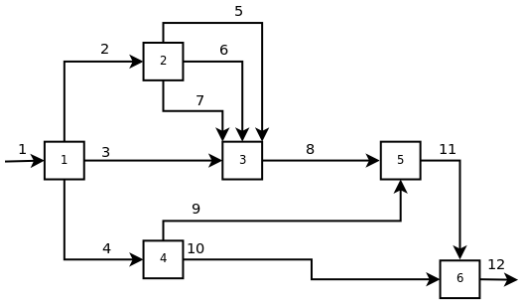


Figure 12: Generic mass balance data flowsheet used in GED of Problem 12.

Table 23: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 12.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	100	1	0.636	0.931	0.813	0.813	0.813	6.7
2	30	0.3	0.274	0.687	0.561	0.561	0.561	17.1
3	40	0.4	0.219	0.620	0.535	0.535	0.535	20.0
4	30	0.3	0.184	0.578	0.437	0.437	0.437	22.0
5	15	0.3	0.242	0.652	0	0	0	-
6	5	0.1	0.024	0.217	0	0	0	-
7	10	0.2	0.100	0.435	0	0	0	-
8	70	0.7	0.543	0.889	0.792	0.792	0.792	8.9
9	10	0.3	0.177	0.569	0.411	0.411	0.411	24.0
10	20	0.4	0.298	0.712	0.538	0.538	0.538	15.5
11	80	0.8	0.538	0.887	0.804	0.804	0.804	8.8
12	100	1	0.636	0.931	0.820	0.820	0.820	6.7

Table 24: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 12.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.991	0.991	0.994	4.4
2	0.939	0.939	0.963	7.5
3	0.982	0.982	0.995	9.3
4	0.889	0.889	0.914	7.4
5	0.972	0.972	0.994	10.6
6	0.978	0.978	0.989	8.0

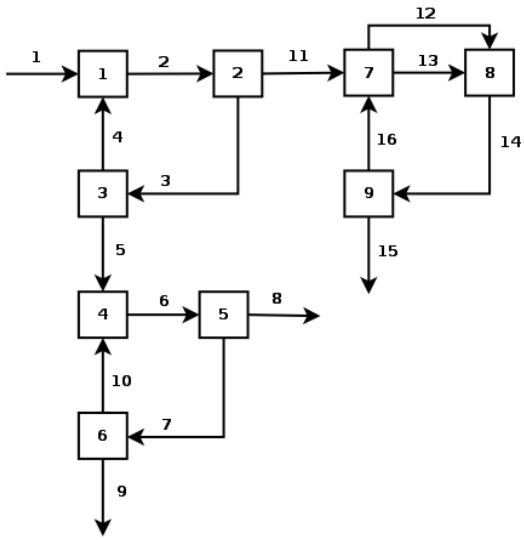


Figure 13: Flotation circuit flowsheet used in GED of Problem 13.

1.13. Problem 13

This diagram is based on a mineral processing plant from Alhaj-Dibo et al. [13] and its diagram is presented in 13. The streams flow rates and standard deviation units are given in $kg.s^{-1}$. The GED results for all measurement tests (MT and GLRMT) and nodal tests (NT and GLRNT) are presented in 25 and 26, respectively.

1.14. Problem 14

A diagram based on a water treatment plant is proposed in this work and presented in 14. The streams flow rates and standard deviation units are given in $m^3.h^{-1}$. The GED results for measurement test and nodal test are presented in 27 and 28, respectively.

Table 25: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 13.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	25	1.000	0.592	0.913	0.774	0.774	0.774	7.3
2	27	1.325	0.668	0.943	0.766	0.766	0.766	5.7
3	22	1.460	0.711	0.957	0.790	0.790	0.790	5.0
4	2	0.200	0.017	0.184	0.032	0.032	0.032	215.7
5	20	0.916	0.576	0.906	0.770	0.770	0.770	8.1
6	24	1.101	0.618	0.924	0.766	0.766	0.766	6.9
7	14	1.040	0.643	0.934	0.759	0.759	0.759	6.5
8	10	0.472	0.201	0.601	0.436	0.436	0.436	22.1
9	10	0.401	0.176	0.567	0.410	0.410	0.410	25.8
10	4	0.207	0.030	0.243	0.060	0.060	0.0604	138.8
11	5	0.300	0.421	0.815	0.685	0.685	0.685	12.0
12	7	0.328	0.345	0.756	0	0	0	-
13	1	0.052	0.007	0.120	0	0	0	-
14	8	0.369	0.420	0.815	0.702	0.702	0.702	11.6
15	5	0.250	0.305	0.719	0.600	0.600	0.600	15.6
16	3	0.385	0.373	0.779	0.674	0.674	0.674	12.8

Table 26: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 13.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.178	0.178	0.360	22.3
2	0.104	0.104	0.392	32.9
3	0.104	0.104	0.332	34.6
4	0.185	0.185	0.468	30.1
5	0.158	0.158	0.464	33.2
6	0.094	0.094	0.237	23.2
7	0.120	0.120	0.286	25.7
8	0.155	0.155	0.387	24.2
9	0.123	0.123	0.307	26.0

Table 27: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 14.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	50	0.533	0.015	0.173	0	0	0	-
2	150	1.460	0.122	0.473	0	0	0	-
3	140	1.541	0.133	0.499	0	0	0	-
4	140	1.451	0.117	0.470	0	0	0	-
5	480	7.600	0.819	0.983	0.815	0.815	0.815	3.1
6	480	7.600	0.819	0.983	0.812	0.812	0.812	3.1
7	480	8.300	0.834	0.986	0.806	0.806	0.806	2.9
8	220	2.136	0.414	0.810	0.694	0.694	0.694	11.9
9	260	2.860	0.685	0.949	0.811	0.811	0.811	5.6
10	260	2.860	0.685	0.949	0.806	0.806	0.806	5.9
11	85.8	0.827	0.333	0.745	0.571	0.571	0.571	14.6
12	15.8	0.441	0.322	0.735	0.549	0.549	0.549	15.2
13	70	0.772	0.287	0.701	0.543	0.543	0.543	15.9
14	15.8	0.441	0.322	0.735	0.545	0.545	0.545	14.7
15	174.2	2.038	0.603	0.918	0.777	0.777	0.777	7.1
16	3.484	0.095	0.294	0.708	0.345	0.345	0.345	16.4
17	3.484	0.095	0.294	0.708	0.369	0.369	0.369	16.4
18	73.484	2.850	0.806	0.981	0.807	0.807	0.807	3.4
19	170.716	1.858	0.565	0.900	0.772	0.772	0.772	8.2
20	170.716	1.999	0.595	0.914	0.773	0.773	0.773	7.4
21	55	0.532	0.043	0.292	0	0	0	-
22	115.716	1.467	0.405	0.804	0	0	0	-
23	200	1.977	0.366	0.773	0	0	0	-
24	20	0.160	0.002	0.062	0	0	0	-

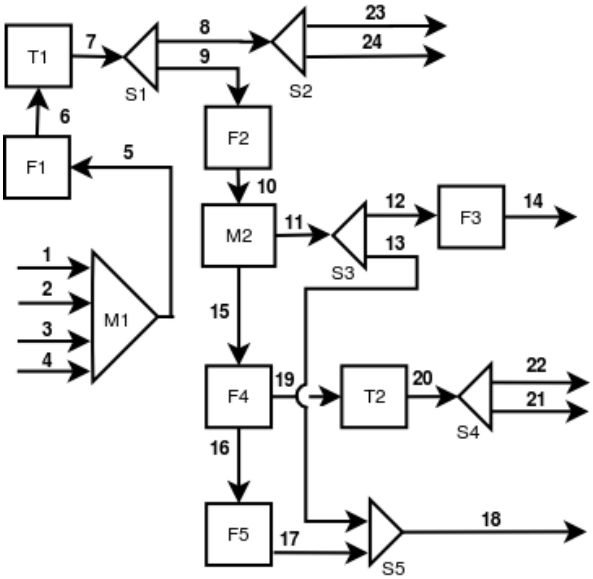


Figure 14: Proposed mass balance of an industrial water treatment unit flowsheet used in GED of Problem 14.

Table 28: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 14.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
M1	0.900	0.900	0.983	15.6
F1	0.774	0.774	0.968	23.0
T1	0.757	0.757	0.962	23.1
S1	0.862	0.862	0.975	17.6
F2	0.922	0.922	0.990	13.9
M2	0.956	0.956	0.995	12.2
S3	0.959	0.959	0.976	5.7
F3	0.382	0.382	0.428	6.6
F4	0.915	0.915	0.989	13.8
F5	0.390	0.390	0.412	1.1
S5	0.366	0.366	0.396	4.3
T2	0.921	0.921	0.983	13.6
S4	0.941	0.941	0.987	10.6
S2	0.972	0.972	0.987	6.4

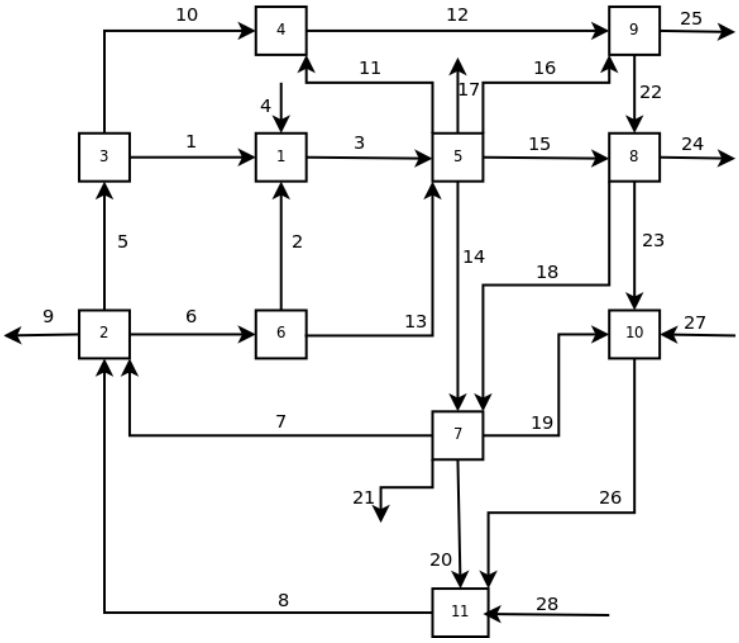


Figure 15: Steam metering system flowsheet used in GED of Problem 15.

1.15. Problem 15

This problem consists in a steam metering system from Serth and Heenan [14], presented in 15. The streams flow rates and standard deviation units are given in $t.s^{-1}$, where t denotes metric tons. The GED results for measurement test and nodal test are presented in 29 and 30, respectively.

1.16. Problem 16

This example consists in ethylene plant from Zhang et al. [1], presented in 16. The streams flow rates and standard deviation units are given in $t.h^{-1}$ (where t denotes *metric ton*). The GED results for measurement test and nodal test are presented in 31 and 32, respectively.

References

[1] Z. Zhang, Z. Shao, X. Chen, K. Wang, J. Qian, Computers & Chemical Engineering 34 (2010) 154–162.

[2] S. Narasimhan, C. Jordache, Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data, Gulf Publishing, Houston, 1 edition, 2000.

Table 29: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 15.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT α_{MT}	OP MT β_{MT}	OP GLR MT β_{MTGLR}	AEE %
1	0.86	0.022	0	0.015	0.004	0.004	0.004	6100
2	1.00	0.025	0	0.010	0.004	0.004	0.004	8100
3	111.82	2.800	0.444	0.831	0.672	0.672	0.672	10.8
4	109.96	2.750	0.435	0.825	0.672	0.672	0.672	11.2
5	53.27	1.332	0.441	0.829	0.660	0.660	0.660	11.0
6	112.27	2.810	0.536	0.886	0.725	0.725	0.725	8.4
7	2.32	0.058		0.027	0.004	0.004	0.004	3200
8	164.05	4.100	0.675	0.946	0.778	0.778	0.778	5.6
9	0.83	0.021	0	0.008	0.003	0.003	0.003	9000
10	52.41	1.310	0.432	0.823	0.672	0.672	0.672	11.4
11	14.86	0.372	0.030	0.242	0.042	0.042	0.042	174.1
12	67.27	1.682	0.539	0.887	0.732	0.732	0.732	8.3
13	111.27	2.780	0.532	0.884	0.702	0.702	0.702	8.6
14	91.86	2.297	0.389	0.792	0.635	0.634	0.634	12.0
15	60.00	1.500	0.388	0.791	0.541	0.541	0.541	12.5
16	23.64	0.591	0.045	0.296	0.062	0.062	0.062	123.2
17	32.73	0.818	0.033	0.255	0.048	0.048	0.048	135.6
18	16.23	0.406	0.003	0.241	0.040	0.040	0.040	156.3
19	7.85	0.196	0.005	0.103	0.006	0.006	0.006	483.0
20	10.50	0.263	0.007	0.121	0.007	0.007	0.007	541.0
21	87.32	2.183	0.359	0.768	0.590	0.590	0.590	12.9
22	5.45	0.136	0.004	0.092	0.007	0.007	0.007	741.0
23	2.59	0.065	0.001	0.044	0.002	0.002	0.002	2700
24	46.63	1.166	0.231	0.639	0.357	0.357	0.357	19.7
25	85.46	2.137	0.575	0.905	0.691	0.691	0.691	7.6
26	81.32	2.033	0.421	0.816	0.669	0.669	0.669	11.4
27	70.79	1.770	0.335	0.747	0.567	0.567	0.567	13.8
28	72.23	1.806	0.237	0.647	0.439	0.439	0.439	18.9

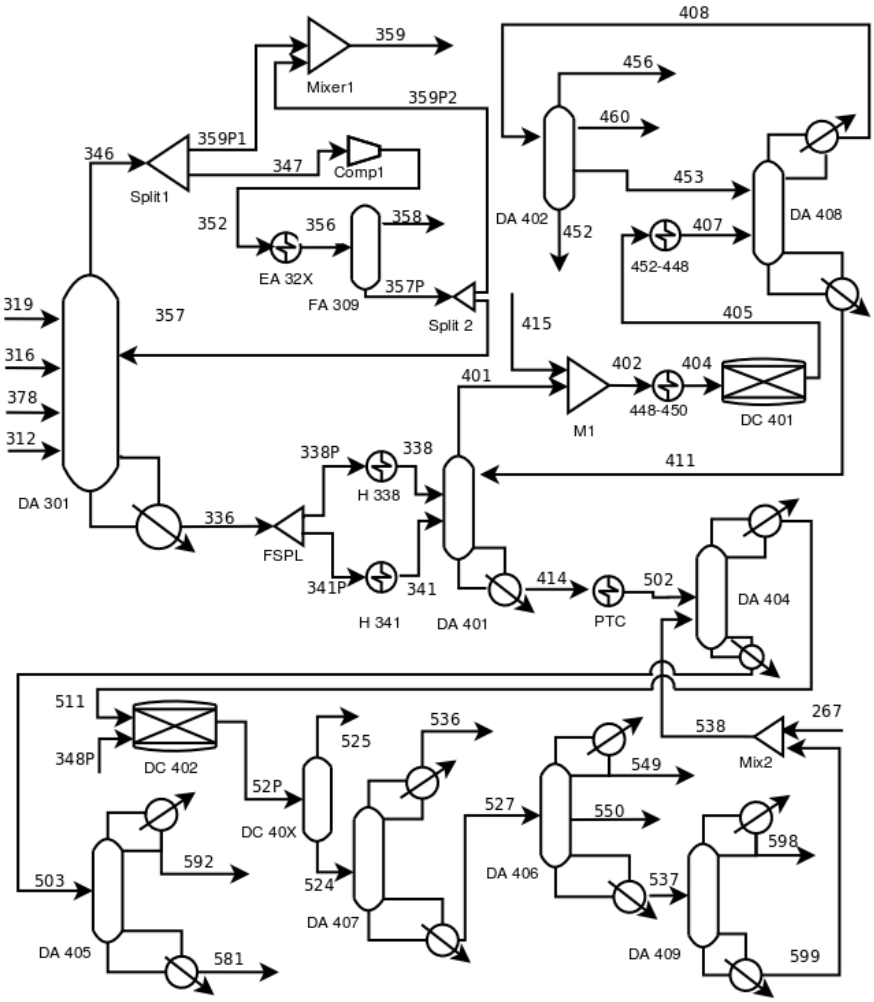


Figure 16: Ethylene plant flowsheet used in GED of Problem 16.

Table 30: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 15.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
1	0.489	0.489	0.603	12.8
2	0.581	0.581	0.878	25.9
3	0.476	0.476	0.628	14.9
4	0.548	0.548	0.715	15.9
5	0.809	0.810	0.938	17.5
6	0.497	0.497	0.730	21.2
7	0.606	0.607	0.692	10.5
8	0.630	0.630	0.682	7.8
9	0.580	0.580	0.687	12.2
10	0.519	0.520	0.624	11.4
11	0.612	0.612	0.850	20.8

[3] Y. Yang, R. Ten, L. Jao, Computers & Chemical Engineering 19 (1995) 217–222.

[4] J. Rosenberg, R. S. H. Mah, C. Iordache, Ind. & Eng. Chem. Proc. Des. Dev, Vol. 26 (1987) 555–564.

[5] S. Narasimhan, R. S. H. Mah, AIChE Journal 33 (1987) 1514–1521.

[6] C. Iordache, R. S. H. Mah, A. C. Tamhane, AIChE Journal 31 (1985) 1187–1201.

[7] M. Sánchez, J. Romagnoli, Q. Jiang, M. Bagajewicz, M. Sa, Computers & Chemical Engineering 23 (1999) 841 – 857.

[8] M. Meyer, B. Koehret, M. Enjalbert, Computers & Chemical Engineering 17 (1993) 807–817.

[9] R. R. Rao, S. Narasimhan, Industrial & Engineering Chemistry Research 35 (1996) 1362–1368.

[10] D. Mandel, A. Abdollahzadeh, D. Maquin, J. Ragot, International Journal of Mineral Processing 53 (1998) 157–169.

[11] M. A. Martins, C. A. Amaro, L. S. Souza, R. A. Kalid, A. Kiperstok, Journal of Cleaner Production (2010) 1–6.

[12] C. L. Mitsas, Measurement 43 (2010) 702–707.

[13] M. Alhaj-Dibo, D. Maquin, J. Ragot, Control Engineering Practice 16 (2008) 159–170.

[14] R. W. Serth, W. A. Heenan, AIChE Journal 32 (1986) 733–747.

Table 31: Results of MT and GLR (testing measurement bias) for measurement bias datasets of Problem 16.

Stream Name	Real Value	Standard Deviation	Adjust.	Detect.	OP MT $z_{a,j}, \alpha$	OP MT $z_{a,j}, \beta$	OP GLR MT $z_{a,j}, \beta$	AEE %
S319	225.45	2.06	0.007	0.117	0	0	0	-
S316	167.89	1.31	0.003	0.074	0	0	0	-
S312	1332.00	11.10	0.223	0.629	0	0	0	-
S378	1332.00	11.20	0.225	0.632	0	0	0	-
S336	2276.90	18.60	0.665	0.942	0.751	0.746	0.748	5.6
S357	137.50	1.22	0.146	0.520	0.211	0.206	0.209	28.9
S346	917.89	7.80	0.556	0.896	0.705	0.700	0.702	7.8
S359P1	532.38	4.74	0.285	0.699	0.515	0.51	0.512	15.4
S347	385.51	3.32	0.627	0.928	0.745	0.742	0.744	6.6
S352	385.51	3.32	0.627	0.928	0.736	0.734	0.735	6.5
S356	385.51	3.32	0.627	0.928	0.748	0.746	0.747	6.5
S358	100.32	0.85	0.061	0.344	0.092	0.0888	0.090	75.9
S357P	285.19	2.26	0.485	0.857	0.674	0.669	0.670	9.5
S359P2	147.69	1.34	0.183	0.576	0.267	0.265	0.266	23
S359	680.07	6.35	0.461	0.842	0.668	0.664	0.666	9.9
S338P	683.07	6.35	0.337	0.749	0.541	0.536	0.539	13.5
S338	683.07	6.24	0.326	0.738	0.522	0.518	0.520	13.4
S341P	1593.80	15.00	0.552	0.894	0.718	0.714	0.716	8.2
S341	1593.8	15.00	0.552	0.894	0.708	0.704	0.706	7.9
S414	582.60	5.28	0.569	0.902	0.71	0.704	0.708	7.7
S502	582.60	5.28	0.569	0.902	0.708	0.704	0.706	7.6
S411	1178.20	9.30	0.179	0.571	0.359	0.352	0.355	23.2
S401	2872.50	24.00	0.658	0.940	0.748	0.745	0.746	5.8
S415	3.84	0.03	0	0.002	0.002	0.002	0.002	30000
S402	2876.30	26.00	0.687	0.950	0.736	0.732	0.733	5.4
S404	2870.00	26.00	0.687	0.950	0.752	0.748	0.75	6.9
S405	2876.30	26.00	0.687	0.950	0.748	0.744	0.745	5.5
S407	2876.30	26.00	0.687	0.950	0.746	0.742	0.744	5.3
S408	3098.20	26.00	0.547	0.892	0	0	0	-
S453	1400.00	11.50	0.081	0.395	0	0	0	-
S460	1337.70	12.20	0.460	0.842	0	0	0	-
S456	11.36	0.10	0	0.007	0	0	0	-
S452	349.07	2.98	0.021	0.206	0	0	0	-
S511	501.44	4.37	0.789	0.978	0.755	0.752	0.753	3.5
S503	392.10	3.34	0.464	0.844	0.647	0.643	0.645	10
S384P	33.80	0.29	0.004	0.088	0.005	0.004	0.005	694
S52P	535.24	1.00	0.118	0.472	0.194	0.190	0.192	37.5
S592	244.46	2.15	0.199	0.599	0	0	0	-
S581	147.64	1.20	0.058	0.334	0	0	0	-
S525	31.13	1.00	0.083	0.4	0.103	0.102	0.103	54
S524	504.11	4.43	0.740	0.966	0.762	0.76	0.761	4.29
S536	7.00	0.06	0	0.022	0.001	0.001	0.001	4100
S527	497.11	4.18	0.724	0.961	0.735	0.732	0.733	4.63
S549	233.72	1.93	0.212	0.616	0	0	0	-
S550	247.55	2.12	0.264	0.677	0	0	0	-
S537	15.84	0.12	0.288	0.703	0.252	0.250	0.251	15.4
S598	13.37	0.12	0.288	0.703	0.163	0.163	0.163	16.4
S599	2.47	0.02	0.007	0.117	0.002	0.002	0.002	576
S267	308.47	2.86	0.374	0.780	0.594	0.584	0.587	12.3
S538	310.94	2.74	0.347	0.757	0.567	0.563	0.565	13.3

Table 32: NT and GLR (testing leakings) results for equipment leakings datasets of Problem 16.

Equipment	OP NT α_{NT}	OP NT β_{NT}	OP GLR NT β_{NTGLR}	AEE %
DA-301	0.999	0.999	1	5.9
Split1	0.990	0.990	0.998	5.9
COMP1	0.970	0.970	0.997	8.7
EA-32X	0.974	0.974	0.998	10.1
FA-309	0.990	0.990	1	8.9
Split2	0.992	0.992	0.996	6.1
Mixer1	0.975	0.975	0.983	4.2
FSPL	0.988	0.988	1	9.5
H338	0.957	0.958	0.982	6.4
H341	0.954	0.954	0.997	11.5
DA-401	0.997	0.997	1	8.64
PTC	0.961	0.961	0.992	8.7
M1	0.970	0.970	1	11.8
448-450	0.951	0.951	0.994	13.6
DC-401	0.974	0.975	0.997	12.2
452-448	0.963	0.963	0.998	12.9
DA-408	0.998	0.998	1	8.8
DA-402	0.994	0.994	0.999	9.1
DA-404	0.999	0.999	1	7.6
DA-405	0.988	0.988	0.995	5.3
DC-402	1	1	1	5.8
DC-40X	0.998	0.998	0.999	7.4
DA-407	0.979	0.979	0.999	11.4
DA-406	0.988	0.989	0.999	8.5
DA-409	0.987	0.987	0.988	0.3
Mix2	0.963	0.963	0.975	4.7

Apêndice B

Estimadores Robustos

Definição da função erro:

$\varepsilon_i = (y_i - x_i) / \sigma_i$

MQP

$\rho(\varepsilon) = {}^1/_2 . \varepsilon_i^2$

Cauchy

$\rho(\varepsilon) = c_C^2 \ln\left(1 + \frac{\varepsilon_i^2}{c_C^2}\right)$

Lorentzian

$\rho(\varepsilon) = -\frac{1}{1+\left(\varepsilon_i^2/2c_{CL}^2\right)}$

Quasi-Weighted

$\rho(\varepsilon) = \frac{\varepsilon_i^2}{2+\beta|\varepsilon_i|}$

onde:c_{CC}, c_{CL} e β são parâmetros do modelo.

Apêndice C

Influência dos Desvios Padrões no Poder Total do Teste

C.1 Problema 1

Tabela C.1: Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,993	0,998	1	1	1
2	0	0,001	0,001	0,002	0,003
3	0	0	0	0	0

Tabela C.2: Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 0,1.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1

Tabela C.3: Poder Total do teste de medição utilizando a incerteza-padrão original.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,59	0,543	0,653	0,814	0,867
2	0,081	0,136	0,215	0,268	0,34
3	0	0	0	0,001	0,001

Tabela C.4: Poder Total do teste nodal utilizando a incerteza-padrão original.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
0,91	0,99	1	1	1

Tabela C.5: Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,579	0,577	0,591	0,54	0,499
2	0,03	0,053	0,15	0,314	0,414
3	0,003	0,002	0,006	0,063	0,217

Tabela C.6: Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 5.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
0,172	0,231	0,31	0,379	0,46

Tabela C.7: Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,395	0,459	0,495	0,506	0,495
2	0,198	0,348	0,446	0,45	0,484
3	0,339	0,458	0,489	0,449	0,466

Tabela C.8: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 0,1.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1

Tabela C.9: Poder Total do teste de medição utilizando a incerteza-padrão unitária.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,231	0,35	0,459	0,549	0,583
2	0,203	0,308	0,405	0,444	0,51
3	0,25	0,376	0,41	0,502	0,505

Tabela C.10: Poder Total do teste nodal utilizando a incerteza-padrão unitária.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
0,175	0,227	0,275	0,339	0,424

Tabela C.11: Poder Total do teste de medida utilizando a incerteza-padrão unitária multiplicada por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,757	0,796	0,82	0,818	0,814
2	0,498	0,531	0,532	0,548	0,512
3	0,538	0,509	0,529	0,471	0,497

Tabela C.12: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 5.

OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
0,102	0,099	0,103	0,102	0,105

C.2 Problema 8

Tabela C.13: Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,946	0,984	0,998	1	1
2	0,7	0,872	0,924	0,978	0,986
3	0	0	0	0,002	0,002
4	0,934	0,978	0,998	0,998	1
5	0,816	0,878	0,916	0,916	0,924
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	0,932	0,974	0,996	0,998	1
10	0	0	0	0	0
11	0,49	0,544	0,626	0,656	0,72

Tabela C.14: Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 0,1.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	1	1	1	1	1

Tabela C.15: Poder Total do teste de medição utilizando a incerteza-padrão original.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,954	0,996	0,996	0,998	0,998
2	0,602	0,672	0,684	0,748	0,704
3	0,072	0,088	0,172	0,182	0,282
4	0,944	0,984	0,994	0,992	0,97
5	0,586	0,556	0,52	0,506	0,558
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	0,934	0,986	0,994	0,982	0,948
10	0	0	0	0	0
11	0,512	0,518	0,498	0,466	0,552
12					

Tabela C.16: Poder Total do teste nodal utilizando a incerteza-padrão original.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
	0,898	0,986	1	1	1
1	0,902	0,996	1	1	1
2	0,972	1	1	1	1
3	0,876	0,986	1	1	1
4	0,85	0,966	0,994	1	1

Tabela C.17: Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,962	0,986	0,998	0,994	0,996
2	0,432	0,544	0,59	0,638	0,66
3	0,294	0,394	0,476	0,58	0,634
4	0,926	0,984	0,994	0,988	0,988
5	0,658	0,67	0,654	0,696	0,702
6	0,032	0,15	0,356	0,476	0,466
7	0,01	0,084	0,236	0,452	0,488
8	0	0	0,02	0,104	0,256
9	0,942	0,982	0,988	0,994	0,992
10	0,012	0,094	0,264	0,446	0,474
11	0,488	0,52	0,534	0,5	0,508

Tabela C.18: Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 5.

Equipamento	OP Vazão Nodal Total 7%	OP Vazão Nodal Total 9%	OP Vazão Nodal Total 11%	OP Vazão Nodal Total 13%	OP Vazão Nodal Total 15%
1	0,064	0,1	0,144	0,198	0,27
2	0,088	0,13	0,19	0,254	0,344
3	0,072	0,094	0,136	0,172	0,222
4	0,058	0,082	0,116	0,17	0,23

Tabela C.20: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 0,1.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1

Tabela C.19: Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,756	0,88	0,94	0,962	0,972
2	0,53	0,62	0,64	0,61	0,624
3	0,082	0,188	0,254	0,298	0,31
4	0,714	0,868	0,948	0,964	0,974
5	0,716	0,882	0,936	0,968	0,982
6	0,228	0,306	0,412	0,46	0,474
7	0,196	0,312	0,382	0,452	0,436
8	0,316	0,42	0,444	0,466	0,504
9	0,342	0,412	0,464	0,446	0,504
10	0,338	0,424	0,454	0,454	0,452
11	0,358	0,426	0,438	0,512	0,492

Tabela C.21: Poder Total do teste de medição utilizando a incerteza-padrão unitária.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,752	0,848	0,836	0,756	0,698
2	0,268	0,282	0,270	0,172	0,096
3	0,422	0,496	0,566	0,588	0,608
4	0,704	0,790	0,790	0,670	0,596
5	0,700	0,794	0,762	0,674	0,626
6	0,200	0,328	0,402	0,468	0,516
7	0,132	0,158	0,132	0,100	0,102
8	0,236	0,278	0,208	0,128	0,110
9	0,420	0,520	0,566	0,616	0,588
10	0,216	0,202	0,230	0,188	0,214
11	0,452	0,534	0,638	0,584	0,556

Tabela C.22: Poder Total do teste nodal utilizando a incerteza-padrão unitária.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1

Tabela C.23: Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,712	0,824	0,846	0,812	0,746
2	0,144	0,162	0,196	0,164	0,144
3	0,464	0,586	0,608	0,558	0,522
4	0,662	0,784	0,784	0,676	0,602
5	0,68	0,798	0,76	0,686	0,614
6	0,362	0,528	0,718	0,87	0,936
7	0,516	0,582	0,6	0,638	0,598
8	0,17	0,212	0,232	0,18	0,138
9	0,464	0,556	0,564	0,54	0,542
10	0,216	0,282	0,318	0,368	0,356
11	0,484	0,604	0,598	0,632	0,576

Tabela C.24: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 5.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1

C.3 Problema 9

Tabela C.25: Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,97	0,994	0,992	0,994	1
2	0	0	0	0	0
3	0,97	0,986	0,992	1	1
4	0	0	0	0	0
5	0,326	0,436	0,592	0,704	0,77
6	0	0	0	0	0
7	0,976	0,986	0,992	0,998	1
8	0	0	0	0	0
9	0,964	0,99	0,998	1	1
10	0,116	0,156	0,148	0,178	0,18
11	0,954	0,99	0,994	0,996	1
12	0,892	0,924	0,938	0,952	0,944

Tabela C.26: Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 0,1.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	1	1	1	1	1
6	1	1	1	1	1

Tabela C.27: Poder Total do teste de medição utilizando a incerteza-padrão original.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,978	0,992	0,994	0,992	1
2	0	0	0	0	0
3	0,606	0,592	0,67	0,686	0,78
4	0,004	0,032	0,05	0,092	0,124
5	0,278	0,408	0,526	0,6	0,632
6	0,002	0,002	0,006	0,002	0,01
7	0,98	0,992	0,996	0,998	1
8	0	0,004	0,002	0,002	0,004
9	0,964	0,982	0,99	0,992	0,996
10	0,356	0,424	0,456	0,456	0,506
11	0,974	0,988	0,992	0,996	0,998
12	0,456	0,504	0,506	0,522	0,562

Tabela C.28: Poder Total do teste nodal utilizando a incerteza-padrão original.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	0,046	0,094	0,16	0,216	0,296
2	1	1	1	1	1
3	0,778	0,962	0,996	1	1
4	0,794	0,96	0,994	1	1
5	0,952	0,996	1	1	1
6	0,202	0,348	0,492	0,66	0,816

Tabela C.29: Poder Total do teste de medição utilizando a incerteza-padrão original multiplicada por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,959	0,981	0,983	0,993	0,996
2	0,407	0,492	0,5	0,474	0,503
3	0,792	0,764	0,75	0,742	0,729
4	0,28	0,418	0,524	0,519	0,535
5	0,318	0,428	0,509	0,592	0,625
6	0,415	0,506	0,498	0,509	0,456
7	0,969	0,993	0,994	0,993	0,998
8	0,018	0,023	0,061	0,099	0,135
9	0,953	0,993	0,988	0,989	0,996
10	0,323	0,446	0,523	0,591	0,599
11	0,965	0,993	0,996	0,997	0,998
12	0,593	0,626	0,619	0,646	0,652

Tabela C.30: Poder Total do teste nodal utilizando a incerteza-padrão original multiplicada por 5.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	0,025	0,025	0,026	0,028	0,03
2	0,099	0,165	0,239	0,32	0,432
3	0,029	0,044	0,07	0,103	0,133
4	0,049	0,065	0,091	0,128	0,178
5	0,064	0,087	0,136	0,198	0,265
6	0,024	0,025	0,03	0,035	0,036

Tabela C.31: Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 0,1.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,337	0,423	0,463	0,487	0,493
2	0,32	0,4067	0,4367	0,52	0,48
3	0,8	0,903	0,96	0,98	0,9967
4	0,703	0,853	0,953	0,963	0,997
5	0,797	0,93	0,987	0,99	0,997
6	0,73	0,897	0,977	0,98	0,997
7	0,79	0,927	0,993	1	1
8	0,76	0,897	0,96	0,997	1
9	0,77	0,92	0,967	0,993	1
10	0,74	0,877	0,963	0,993	1
11	0,78	0,93	0,977	0,993	1
12	0,787	0,93	0,977	0,99	1

Tabela C.32: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 0,1.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	1	1	1	1	1
6	1	1	1	1	1

Tabela C.33: Poder Total do teste de medição utilizando a incerteza-padrão unitária.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,403	0,473	0,487	0,473	0,510
2	0,300	0,453	0,490	0,497	0,503
3	0,763	0,900	0,973	0,977	0,997
4	0,723	0,837	0,947	0,980	1,000
5	0,773	0,927	0,973	1,000	1,000
6	0,737	0,877	0,953	0,993	1,000
7	0,810	0,947	1,000	1,000	1,000
8	0,743	0,893	0,967	0,987	0,993
9	0,760	0,917	0,963	0,990	0,997
10	0,770	0,870	0,953	0,993	1,000
11	0,767	0,920	0,957	1,000	1,000
12	0,777	0,923	0,980	0,983	1,000

Tabela C.34: Poder Total do teste nodal utilizando a incerteza-padrão unitária.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	1	1	1	1	1
6	1	1	1	1	1

Tabela C.35: Poder Total do teste nodal utilizando a incerteza-padrão unitária multiplicada por 5.

Equipamento	OP - Vazão Nodal Total = 7%	OP - Vazão Nodal Total 9%	OP - Vazão Nodal Total 11%	OP - Vazão Nodal Total 13%	OP - Vazão Nodal Total 15%
1	0,917	0,980	0,997	0,997	1,000
2	0,833	0,907	0,960	0,993	0,993
3	0,623	0,730	0,870	0,930	0,973
4	0,840	0,940	0,980	0,990	1,000
5	0,160	0,210	0,310	0,403	0,493
6	0,867	0,943	0,983	0,993	1,000

Tabela C.36: Poder Total do teste de medição utilizando a incerteza-padrão unitária multiplicada por 5.

Corrente	OP δ/u =5	OP δ/u =6	OP δ/u =7	OP δ/u =8	OP δ/u =9
1	0,360	0,437	0,540	0,587	0,590
2	0,080	0,113	0,253	0,353	0,480
3	0,810	0,883	0,933	0,917	0,813
4	0,617	0,780	0,770	0,730	0,773
5	0,773	0,900	0,940	0,920	0,830
6	0,697	0,893	0,977	0,993	0,987
7	0,800	0,940	0,987	1,000	1,000
8	0,750	0,907	0,923	0,837	0,807
9	0,730	0,823	0,813	0,653	0,497
10	0,727	0,893	0,947	0,903	0,890
11	0,750	0,917	0,973	0,950	0,947
12	0,800	0,893	0,920	0,860	0,770

Apêndice D

Resultados Completos

D.1 Erros Únicos

D.1.1 Problema 1

Tabela D.1: Resultados do Problema 1 para erros em medições.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	8,50	0,277	0,846	0	-	1	2,202	0,970
2	4,50	0,145	0,443	0	-	0,978	2,227	0,813
3	4,00	0,097	0,296	0	-	1	2,315	0,890

Tabela D.2: Resultados do Problema 1 para erros em medições.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,997	0,997	0	0,84	-0,040	1

D.1.2 Problema 2

Tabela D.3: Resultados do Problema 2 para erros em medições.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	11	0,179	0,791	0,981	0,090	1,000	0,343	0,993
2	10	0,161	0,404	0,000	-	0,989	1,503	0,998
3	21	0,346	0,869	0,000	-	0,980	0,980	0,996
4	11	0,182	0,798	0,984	0,087	1,000	0,659	1,000
5	5,5	0,228	0,798	0,000	-	0,940	1,045	0,993
6	5,5	0,122	0,429	0,000	-	1,000	1,048	0,993

Tabela D.4: Resultados do Problema 2 para vazamentos.

Equipamento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
	1	1	6,677	1	0,005	1
	1	1	6,810	1	0,0001	1
	1	1	4,848	1	0,001	1

D.1.3 Problema 3

Tabela D.5: Resultados do Problema 3 para erros em medições.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
x1	189,980	2,141	0,900	0,994	0,058	1,000	1,436	0,963
x2	174,600	1,948	0,875	0,000	-6,660	1,000	1,218	0,973
x3	3,139	0,034	0,045	0,017	7,821	1,000	0,980	0,981
x4	32,770	0,345	0,183	0,000	-	-	7,526	-
x5	33,470	0,400	0,212	0,000	-	-	7,406	-
x6	7,250	0,087	0,046	0,000	-	0,793	1,854	1,000
x7	0,316	0,004	0,002	0,000	-	1,000	0,448	1,000
x8	92,376	1,059	0,560	0,000	-	1,000	0,131	1,000
x9	28,629	0,361	0,452	0,000	-	1,000	1,030	0,987
x10	23,800	0,300	0,400	0,405	0,246	1,000	0,821	1,000
x11	18,526	0,200	0,090	0,000	-	1,000	0,246	0,987
x12	55,568	0,645	0,810	0,000	-	1,000	1,299	1,000

Tabela D.6: Resultados do Problema 3 para vazamentos.

Equipmento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1,00	1,000	0,033	1	0,007	1
2	1,00	1,000	0,026	1	0,004	1
3	1,00	1,000	0,019	1	0,001	1

D.1.4 Problema 3 - Incerteza-Padrão = 5 t/h

Tabela D.7: Resultados do Problema 3 para erros de medição com incerteza-padrão padrão igual a 5.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
x1	189,980	2,141	0,900	0,996	0,058	1,000	0,789	1,000
x2	174,600	1,948	0,875	0,000	-	1,000	1,387	1,000
x3	3,139	0,034	0,045	0,015	7,648	1,000	0,654	1,000
x4	32,770	0,345	0,183	0,000	-	1,000	0,249	1,000
x5	33,470	0,400	0,212	0,000	-	1,000	-0,072	1,000
x6	7,250	0,087	0,046	0,000	-	1,000	1,296	1,000
x7	0,316	0,004	0,002	0,000	-	1,000	0,791	1,000
x8	92,376	1,059	0,560	0,000	-	1,000	1,315	0,995
x9	28,629	0,361	0,452	0,000	-	1,000	0,556	1,000
x10	23,800	0,300	0,400	0,386	0,238	1,000	1,272	1,000
x11	18,526	0,200	0,090	0,000	-	1,000	1,654	0,989
x12	55,568	0,645	0,810	0,000	-	1,000	0,243	0,997

Tabela D.8: Resultados do Problema 3 para vazamentos com incerteza-padrão padrão igual a 5.

Equipmento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1,00	1,000	0,033	1	0,007	1
2	1,00	1,000	0,026	1	-0,008	1
3	1,00	1,000	0,019	1	-0,001	1

D.1.5 Problema 4

Tabela D.9: Resultados do Problema 4 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	100	1	0,816	0,985	0,081	1	0,602	0,988
2	64	1	0,816	0,988	0,085	1	0,003	1
3	36	1	0,816	0,989	0,084	1	1,010	1
4	64	1	0,816	0,988	0,083	1	1,037	1
5	36	1	0,816	0,991	0,084	1	0,859	1
6	100	1	0,816	0,990	0,084	1	0,970	1

Tabela D.10: Resultados do Problema 4 para vazamentos.

Equipmento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1	1,000	0,048	1,000	0,0001	1
2	1	1,000	0,061	1,000	-0,001	1
3	0,984	0,995	0,113	1,000	-0,001	1
4	1	1,000	0,047	1	-0,0003	1

D.1.6 Problema 5

Tabela D.11: Resultados do Problema 5 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	98,7	0,993	0,846	0,99	0,074	0,98	1,195	0,993
2	41,1	0,641	0,650	0,93	0,132	1	0,543	1,000
3	78,9	0,888	0,836	0,99	0,081	1	0,840	0,988
4	30,2	0,550	0,424	0,00	-	1	0,348	1,000
5	109,1	1,045	0,806	0,00	-	0,98	1,916	1,000
6	19,8	0,445	0,504	0,74	0,185	1	1,039	0,998
7	57,6	0,759	0,774	0,96	0,095	0,98	1,096	1,000
8	37,8	0,615	0,693	0,94	0,115	1	0,901	0,980

Tabela D.12: Resultados do Problema 5 para vazamentos.

Equipmento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1	1	0,030	1,000	-0,004	1,000
2	1	1	0,038	1,000	-0,004	1,000
3	1	1	0,025	1,000	0,001	1,000
4	1	1	0,041	1,000	-0,023	1,000

D.1.7 Problema 6

Tabela D.13: Resultados do Problema 6 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	5	1	0,787	0,977	0,092	1	0,830	0,998
2	15	1	0,787	0,978	0,093	1	0,984	1,000
3	15	1	0,787	0,976	0,092	1	0,743	1,000
4	5	1	0,690	0,952	0,121	1	1,298	0,996
5	10	1	0,756	0,964	0,102	1	1,183	1,000
6	5	1	0,690	0,943	0,119	1	1,017	0,975
7	5	1	0,787	0,978	0,091	1	0,519	1,000

Tabela D.14: Resultados do Problema 6 para vazamentos.

Equipmento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,29	0,5216	0,2413739	1	-0,00032	1
2	0,524	0,7664	0,2137069	1	0,00008	1
3	0,37	0,6532	0,2491131	1	0,00062	1
4	0,1592	0,2924	0,2078744	1	0,00056	1

D.1.8 Problema 7

Tabela D.15: Resultados do Problema 7 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	1,000	0,810	0,985	0,086	1,000	0,648	1,000
2	75	1,000	0,815	0,984	0,084	1,000	1,066	1,000
3	75	1,000	0,815	0,984	0,084	1,000	0,996	0,971
4	48	1,000	0,802	0,988	0,087	1,000	1,166	0,971
5	30	1,000	0,805	0,983	0,088	1,000	0,978	0,971
6	25	1,000	0,810	0,988	0,087	1,000	1,188	1,000
7	5	0,387	0,806	0,981	0,088	1,000	1,460	0,996
8	5	0,387	0,806	0,982	0,086	1,000	1,303	0,971
9	3	0,316	0,617	0,735	0,140	1,000	1,289	1,000
10	2	0,316	0,623	0,753	0,135	1,000	0,921	0,971

Tabela D.16: Resultados do Problema 7 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1	1,000	0,062	1	0,002	0,971
2	1	1,000	0,052	1	-0,001	0,971
3	1	1,000	0,061	1	-0,0003	0,971
4	0,9336	0,982	0,141	1,000	0,000	0,971
5	0,37	0,512	0,158	1,000	-0,001	0,972
6	0,3252	0,438	0,141	0,000	1,200	0,000

D.1.9 Problema 8

Tabela D.17: Resultados do Problema 8 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	3600	111,210	0,915	0,995	0,052	1,000	1,784	0,984
2	1850	57,000	0,595	0,000	-	1,000	1,393	1,000
3	1750	54,210	0,566	0,000	-	1,000	0,813	1,000
4	2837	87,300	0,859	0,987	0,070	1,000	1,518	1,000
5	730	22,110	0,748	0,802	0,103	1,000	1,886	1,000
6	25	0,780	0,009	0,000	-	0,640	5,508	0,973
7	8	0,228	0,003	0,000	-	0,960	4,510	0,989
8	137	3,150	0,032	0,000	-	1,000	4,938	0,960
9	2700	84,960	0,850	0,000	-	1,000	2,013	0,984
10	58	1,710	0,058	0,000	-	1,000	4,794	0,955
11	672	20,04	0,682	0,000	-	1,000	1,733	0,989

Tabela D.18: Resultados do Problema 8 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,991	0,995	0,066	1,000	-0,040	1,000
2	0,998	0,999	0,086	1,000	0,074	1,000
3	0,988	0,987	0,028	1,000	-0,038	1,000
4	0,981	0,993	0,095	1,000	0,027	1,000

D.1.10 Problema 9

Tabela D.19: Resultados do Problema 9 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	230	37,575	0,998	0,000	-	1,000	1,103	1,000
2	21	1,08	0,029	0,000	-	1,000	1,491	1,000
3	209	5	0,890	0,665	0,060	1,000	1,511	1,000
4	35	1,825	0,327	0,237	0,367	1,000	1,164	1,000
5	174	2	0,522	0,725	0,174	1,000	1,433	0,999
6	15	0,88	0,193	0,095	1,060	1,000	1,877	0,993
7	159	7,245	0,967	0,997	0,031	1,000	1,677	1,000
8	50	1	0,349	0,370	0,320	1,000	1,288	0,999
9	209	5	0,916	0,974	0,052	1,000	1,501	0,999
10	94	2	0,616	0,548	0,145	1,000	0,928	0,991
11	115	18,1	0,991	0,999	0,016	1,000	1,363	1,000
12	44	2,385	0,733	0,638	0,105	1,000	1,144	0,999

Tabela D.20: Resultados do Problema 9 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,159	0,161	0,033	1,000	0,014	1,000
2	1,000	1,000	0,020	1,000	0,040	1,000
3	0,973	0,996	0,135	1,000	0,008	1,000
4	0,976	0,993	0,134	1,000	0,004	1,000
5	0,994	0,994	0,014	1,000	0,033	1,000
6	0,546	0,568	0,065	1,000	0,021	1,000

D.1.11 Problema 10

Tabela D.21: Resultados do Problema 10 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	28	0,275	0,865	0,976	0,070	1,000	0,729	0,959
2	5	0,050	0,753	0,842	0,100	1,000	0,862	0,940
3	5	0,172	0,911	0,995	0,054	1,000	0,821	1,000
4	7	0,145	0,873	0,985	0,067	1,000	0,918	0,954
5	11	0,372	0,935	0,996	0,045	1,000	1,395	1,000
6	4	0,127	0,904	0,991	0,056	1,000	1,027	0,959
7	3	0,136	0,919	0,991	0,050	1,000	0,908	0,998
8	5	0,045	0,690	0,811	0,117	1,000	1,433	0,959
9	5	0,095	0,663	0,916	0,129	1,000	0,725	0,995
10	4	0,073	0,661	0,927	0,126	1,000	1,082	0,959
11	3	0,064	0,538	0,767	0,164	1,000	1,335	1,000
12	14	0,147	0,541	0,754	0,164	1,000	1,292	0,964
13	14	0,131	0,790	0,974	0,089	1,000	1,017	0,959

Tabela D.22: Resultados do Problema 10 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1	1	0,055	1	-0,0006	1
2	1	1	0,012	1	-0,0002	1
3	0,977	0,992	0,080	1	0,0001	0,98
4	0,994	0,999	0,108	1	-0,0001	0,96
5	0,999	0,999	0,081	1	0,0002	1
6	0,988	0,997	0,092	1	1,2300	0
7	0,920	0,970	0,127	1	1,1211	0
8	1	1	0,021	1	0,0004	1

D.1.12 Problema 11

Tabela D.23: Resultados do Problema 11 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	690	20,750	0,862	0,994	0,071	1,000	1,566	1,000
2	725	21,826	0,876	0,995	0,066	1,000	1,615	0,999
3	700	20,981	0,865	0,994	0,069	1,000	1,470	0,977
4	685	20,615	0,860	0,995	0,070	1,000	0,932	1,000
5	35	1,076	0,884	0,559	0,064	1,000	1,728	0,998
6	15	0,375	0,308	0,200	0,410	1,000	2,133	0,983
7	25	0,836	0,877	0,995	0,066	1,000	1,960	0,999
8	20	0,701	0,819	0,987	0,084	1,000	1,777	1,000
9	30	0,680	0,803	0,984	0,089	1,000	1,651	1,000
10	5	0,144	0,438	0,401	0,215	1,000	1,519	1,000
11	5	0,136	0,416	0,386	0,230	1,000	1,430	0,999
12	10	0,279	0,823	0,816	0,079	1,000	1,676	1,000

Tabela D.24: Resultados do Problema 11 para vazamentos.

Equipamento	OP NT beta or	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,978	0,992	0,098	1,000	0,048	1,000
2	0,973	0,994	0,123	1,000	0,012	1,000
3	0,974	0,988	0,097	1,000	0,041	1,000
4	0,984	0,994	0,062	1,000	-0,016	1,000
5	0,996	0,999	0,081	1,000	-0,017	1,000
6	0,988	0,995	0,069	1,000	-0,008	1,000
7	0,994	0,996	0,040	1,000	0,004	1,000

D.1.13 Problema 12

Tabela D.25: Resultados do Problema 12 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	100	1	0,931	0,996	0,046	1,000	1,050	1,000
2	30	0,3	0,687	0,913	0,118	1,000	1,566	0,942
3	40	0,4	0,624	0,900	0,139	1,000	1,855	0,999
4	30	0,3	0,578	0,792	0,150	1,000	1,343	1,000
5	15	0,3	0,652	0,000	-	0,960	2,063	0,998
6	5	0,1	0,217	0,000	-	0,920	2,345	0,997
7	10	0,2	0,435	0,000	-	0,980	1,029	1,000
8	70	0,7	0,889	0,996	0,061	1,000	1,423	1,000
9	10	0,3	0,569	0,756	0,156	1,000	1,460	1,000
10	20	0,4	0,712	0,862	0,114	1,000	1,488	1,000
11	80	0,8	0,887	0,997	0,060	1,000	1,196	1,000
12	100	1	0,931	0,998	0,045	1,000	0,813	1,000

Tabela D.26: Resultados do Problema 12 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	1,000	1,000	0,017	1,000	-0,0002	1,000
2	1,000	1,000	0,029	1,000	0,0013	1,000
3	1,000	1,000	0,034	1,000	0,0008	1,000
4	1,000	1,000	0,030	1,000	0,0010	1,000
5	1,000	1,000	0,039	1,000	0,0005	1,000
6	1,000	1,000	0,030	1,000	0,0003	1,000

D.1.14 Problema 13

Tabela D.27: Resultados do Problema 13 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	25	1	0,913	0,996	0,054	1,000	1,246	1,000
2	27	1,325	0,943	0,987	0,041	1,000	0,856	1,000
3	22	1,460	0,957	0,994	0,036	1,000	1,398	1,000
4	2	0,200	0,184	0,074	1,305	1,000	1,523	1,000
5	20	0,916	0,906	0,993	0,056	1,000	1,080	0,947
6	24	1,101	0,924	0,990	0,050	1,000	1,160	1,000
7	14	1,040	0,934	0,986	0,046	0,980	1,498	1,000
8	10	0,472	0,601	0,815	0,143	1,000	1,256	1,000
9	10	0,401	0,567	0,770	0,150	1,000	1,437	1,000
10	4	0,207	0,243	0,145	0,781	1,000	1,311	0,999
11	5	0,300	0,815	0,973	0,083	1,000	1,449	1,000
12	7	0,328	0,756	0,000	-	0,980	1,474	0,999
13	1	0,052	0,120	0,000	-	0,980	1,327	0,922
14	8	0,369	0,815	0,977	0,083	1,000	0,911	1,000
15	5	0,250	0,719	0,932	0,111	1,000	1,563	1,000
16	3	0,385	0,779	0,972	0,095	0,980	1,811	1,000

Tabela D.28: Resultados do Problema 13 para vazamentos.

Equipamento	OP NT beta	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,799	0,934	0,167	1	0,0015	1
2	0,668	0,936	0,251	1	0,0015	1
3	0,604	0,899	0,263	1	0,0011	1
4	0,802	0,962	0,200	1	0,0012	1
5	0,739	0,955	0,245	1	0,0008	1
6	0,593	0,854	0,200	1	0,0005	1
7	0,663	0,893	0,202	1	-0,0003	1
8	0,812	0,940	0,174	1	-0,0017	1
9	0,641	0,892	0,215	1	0,0002	1

D.1.15 Problema 14

Tabela D.29: Resultados do Problema 14 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	0,533	0,173	0,000	-	1,000	0,057	0,996
2	150	1,460	0,473	0,000	-	1,000	0,416	0,993
3	140	1,541	0,499	0,000	-	0,000	9,000	0,000
4	140	1,451	0,470	0,000	-	0,000	8,000	0,000
5	480	7,611	0,983	0,997	0,022	1,000	2,058	1,000
6	480	7,611	0,983	0,999	0,022	1,000	1,975	0,993
7	480	8,335	0,986	1,000	0,020	1,000	1,918	1,000
8	220	2,136	0,810	0,974	0,085	1,000	2,090	0,994
9	260	2,865	0,949	0,996	0,039	1,000	2,528	0,999
10	260	2,865	0,949	0,998	0,039	1,000	1,976	0,996
11	85,8	0,827	0,745	0,927	0,105	1,000	2,144	0,996
12	15,8	0,441	0,735	0,895	0,103	1,000	3,154	0,988
13	70	0,772	0,701	0,897	0,113	1,000	2,655	0,996
14	15,8	0,441	0,735	0,882	0,104	0,960	3,460	0,987
15	174,2	2,038	0,918	0,996	0,051	1,000	2,711	1,000
16	3,484	0,095	0,708	0,573	0,117	1,000	3,686	0,993
17	3,484	0,095	0,708	0,573	0,113	1,000	3,517	1,000
18	73,484	2,852	0,981	0,997	0,023	1,000	2,508	0,993
19	170,716	1,858	0,900	0,994	0,058	1,000	2,402	0,996
20	170,716	1,999	0,914	0,998	0,053	1,000	2,143	1,000
21	55	0,532	0,292	0,000	-	1,000	2,425	0,996
22	115,716	1,467	0,804	0,000	-	1,000	1,924	0,991
23	200	1,977	0,773	0,000	-	0,960	2,636	0,987
24	20	0,160	0,062	0,000	-	0,860	3,308	1,000

Tabela D.30: Resultados do Problema 14 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
M1	1	1	0,061	0,0	-0,013	0,996
F1	1	1	0,086	0,0	-0,015	0,996
T1	1	1	0,089	0,0	0,003	0,996
S1	1	1	0,069	0,0	0,000	0,996
F2	1	1	0,049	0,0	-0,004	0,996
M2	1	1	0,044	0,0	0,005	0,996
S3	1	1	0,022	0,0	-0,002	0,996
F3	0,9764	0,982	0,041	0,0	0,003	0,996
F4	1	1	0,049	0,0	-0,030	0,996
F5	0,9804	0,9836	0,007	0,0	-0,003	0,996
S5	0,97	0,9756	0,028	0,0	-0,003	0,996
T2	1	1	0,050	0,0	0,007	0,996
S4	1	1	0,040	0,0	-0,003	0,996
S2	1	1	0,024	0,0	0,011	0,996

D.1.16 Problema 14 - Incerteza- Padrão = 2 t/h

Tabela D.31: Resultados do Problema 14 para erros de medição com incerteza-padrão padrão igual a 5 t/h.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	50	0,533	0,173	0,000	-6,660	1,000	1,062	0,997
2	150	1,460	0,473	0,000	-6,660	0,989	0,987	0,988
3	140	1,541	0,499	0,000	-6,660	1,000	1,109	0,990
4	140	1,451	0,470	0,000	-6,660	1,000	0,876	0,991
5	480	7,611	0,983	0,997	0,022	1,000	1,974	0,999
6	480	7,611	0,983	0,999	0,022	1,000	1,716	0,996
7	480	8,335	0,986	1,000	0,020	1,000	1,825	0,988
8	220	2,136	0,810	0,974	0,085	1,000	1,745	0,995
9	260	2,865	0,949	0,996	0,039	1,000	2,291	0,988
10	260	2,865	0,949	0,998	0,039	1,000	1,718	0,992
11	85,8	0,827	0,745	0,927	0,105	1,000	1,343	1,000
12	15,8	0,441	0,735	0,895	0,103	1,000	1,667	0,999
13	70	0,772	0,701	0,897	0,113	1,000	1,703	0,996
14	15,8	0,441	0,735	0,882	0,104	1,000	2,109	0,984
15	174,2	2,038	0,918	0,996	0,051	1,000	2,050	0,988
16	3,484	0,095	0,708	0,573	0,117	1,000	1,619	0,996
17	3,484	0,095	0,708	0,573	0,113	1,000	1,744	1,000
18	73,484	2,852	0,981	0,997	0,023	1,000	1,507	0,990
19	170,716	1,858	0,900	0,994	0,058	1,000	2,001	1,000
20	170,716	1,999	0,914	0,998	0,053	1,000	1,657	0,988
21	55	0,532	0,292	0,000	-6,660	1,000	2,216	0,995
22	115,716	1,467	0,804	0,000	-6,660	1,000	1,735	0,994
23	200	1,977	0,773	0,000	-6,660	1,000	1,747	0,992
24	20	0,160	0,062	0,000	-6,660	0,820	2,227	0,992

Tabela D.32: Resultados do Problema 14 para vazamentos e incerteza-padrão padrão de 5 t/h.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
M1	1	1	0,061	1,0	-0,002	0,992
F1	1	1	0,086	1,0	0,002	0,992
T1	1	1	0,089	1,0	-0,002	0,992
S1	1	1	0,069	1,0	-0,001	0,992
F2	1	1	0,049	1,0	-0,003	0,992
M2	1	1	0,044	1,0	-0,002	0,992
S3	1	1	0,022	1,0	0,000	0,992
F3	0,9764	0,982	0,041	0,0	1,186	0,000
F4	1	1	0,049	1,0	-0,002	0,992
F5	0,9804	0,9836	0,007	0,0	1,100	0,000
S5	0,97	0,9756	0,028	1,0	-0,001	0,992
T2	1	1	0,050	1,0	-0,003	0,992
S4	1	1	0,040	1,0	-0,002	0,992
S2	1	1	0,024	1,0	-0,001	0,992

D.1.17 Problema 15

Tabela D.33: Resultados do Problema 15 para erros de medição.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,86	0,022	0,015	0,0032	32,642	1,000	1,871	1,000
2	1	0,025	0,010	0,0024	47,814	1,000	1,477	1,000
3	111,82	2,796	0,831	0,973	0,076	1,000	2,081	1,000
4	109,96	2,749	0,825	0,971	0,079	1,000	1,536	1,000
5	53,27	1,332	0,829	0,950	0,077	1,000	1,522	1,000
6	112,27	2,807	0,886	0,990	0,062	1,000	1,286	1,000
7	2,32	0,058	0,027	0,0040	17,861	1,000	1,679	1,000
8	164,05	4,101	0,946	0,993	0,041	1,000	1,892	1,000
9	0,83	0,021	0,008	0,00240	53,678	1,000	1,530	1,000
10	52,41	1,310	0,823	0,980	0,079	1,000	1,772	1,000
11	14,86	0,372	0,242	0,118	0,847	1,000	1,830	1,000
12	67,27	1,682	0,887	0,982	0,062	1,000	1,508	0,998
13	111,27	2,782	0,884	0,976	0,063	1,000	1,275	1,000
14	91,86	2,297	0,792	0,942	0,089	1,000	1,357	0,982
15	60	1,500	0,791	0,864	0,091	1,000	1,323	0,986
16	23,64	0,591	0,295	0,177	0,565	1,000	1,624	1,000
17	32,73	0,818	0,255	0,148	0,779	1,000	0,244	0,982
18	16,23	0,406	0,241	0,089	0,801	1,000	2,166	0,999
19	7,85	0,196	0,103	0,010	3,188	1,000	1,195	1,000
20	10,5	0,263	0,121	0,018	2,973	1,000	1,554	1,000
21	87,32	2,183	0,768	0,931	0,096	1,000	1,546	1,000

Continua

Tabela D.34: Resultados do Problema 15 para erros de medição. Continuação

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
22	5,45	0,136	0,092	0,011	3,697	1,000	1,437	1,000
23	2,59	0,065	0,044	0,005	9,511	1,000	1,519	1,000
24	46,63	1,166	0,639	0,685	0,129	1,000	1,146	0,997
25	85,46	2,137	0,905	0,961	0,054	1,000	1,213	1,000
26	81,32	2,033	0,816	0,969	0,083	1,000	1,589	0,998
27	70,79	1,770	0,747	0,920	0,103	1,000	1,022	1,000
28	72,23	1,806	0,647	0,825	0,130	1,000	1,847	0,999

Tabela D.35: Resultados do Problema 15 para vazamentos.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
1	0,994	0,997	0,073	1,000	0.0028	1,000
2	0,996	0,999	0,103	1,000	0.0016	1,000
3	0,993	0,999	0,081	1,000	0.0021	1,000
4	0,996	1,000	0,086	1,000	0.0098	1,000
5	1,000	1,000	0,064	1,000	- 0.0018	1,000
6	0,988	0,998	0,107	1,000	0.0011	1,000
7	0,997	1,000	0,056	1,000	0.0024	1,000
8	0,998	1,000	0,041	1,000	0.0022	1,000
9	0,999	1,000	0,064	1,000	0.0010	1,000
10	0,998	0,999	0,062	1,000	0.0017	1,000
11	0,999	1,000	0,098	1,000	- 0.0005	1,000

D.1.18 Problema 16

Tabela D.36: Resultados de erros grosseiros de medição para o Problema 16.

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
S319	225,45	2,06	0,117	0,000	-	0,469	4,520	0,999
S316	167,89	1,31	0,074	0,000	-	0,921	1,861	0,899
S312	1332,00	11,12	0,629	0,000	-	1,000	0,316	0,996
S378	1332,00	11,16	0,632	0,000	-	1,000	0,017	0,996
S336	2276,90	18,61	0,942	0,995	0,042	1,000	2,463	1,000
S357	137,50	1,22	0,520	0,543	0,179	1,000	1,975	1,000
S346	917,89	7,81	0,896	0,985	0,060	1,000	2,579	1,000
S359P1	532,38	4,74	0,699	0,914	0,112	1,000	2,035	0,995
S347	385,51	3,32	0,928	0,991	0,048	1,000	2,848	1,000
S352	385,51	3,32	0,928	0,997	0,048	0,980	2,965	1,000
S356	385,51	3,32	0,928	0,996	0,047	1,000	2,835	0,991
S358	100,32	0,85	0,344	0,332	0,388	1,000	2,302	0,999
S357P	285,19	2,26	0,857	0,985	0,073	0,960	2,913	0,994
S359P2	147,69	1,34	0,576	0,618	0,154	1,000	2,342	1,000
S359	680,07	6,35	0,842	0,983	0,075	1,000	2,330	0,999
S338P	683,07	6,35	0,749	0,908	0,103	1,000	2,550	1,000
S338	683,07	6,24	0,738	0,897	0,102	0,980	2,815	0,999
S341P	1593,80	14,97	0,894	0,989	0,058	1,000	2,858	1,000
S341	1593,80	14,97	0,894	0,991	0,060	1,000	2,767	1,000
S414	582,60	5,28	0,902	0,991	0,058	1,000	2,776	0,998
S502	582,60	5,28	0,902	0,991	0,058	1,000	2,999	0,999
S411	1178,20	9,27	0,571	0,801	0,145	1,000	2,571	1,000
S401	2872,50	23,76	0,940	0,996	0,043	1,000	2,172	1,000

Continua

Tabela D.37: Resultados de erros grosseiros de medição para o Problema 16. Continuação

Corrente	Valor Nominal	Incerteza-Padrão	Detecta-bilidade	OP GLR MT	AEE GLR	OP Método	AEE Método	Seletividade Método
S415	3,84	0,03	0,002	0,001	225,0	1,000	2,448	1,000
S402	2876,30	25,96	0,950	0,999	0,039	1,000	3,153	0,999
S404	2870,00	25,96	0,950	0,999	0,050	0,940	3,216	0,995
S405	2876,30	25,96	0,950	0,994	0,039	1,000	3,109	1,000
S407	2876,30	25,96	0,950	0,998	0,039	1,000	3,045	1,000
S408	3098,20	25,87	0,892	0,000	-	0,960	2,984	0,997
S453	1400,00	11,46	0,395	0,000	-	0,740	3,557	0,997
S460	1337,70	12,20	0,842	0,000	-	0,880	2,585	0,998
S456	11,36	0,10	0,007	0,000	-	0,920	2,702	0,998
S452	349,07	2,98	0,206	0,000	-	0,735	3,422	0,967
S511	501,44	4,37	0,978	0,992	0,026	1,000	2,289	1,000
S503	392,10	3,34	0,844	0,978	0,071	1,000	2,628	1,000
S384P	33,80	0,29	0,088	0,008	4,534	1,000	2,264	0,999
S52P	535,24	1,00	0,472	0,537	0,199	1,000	2,662	0,996
S592	244,46	2,15	0,599	0,000	-	0,980	2,718	0,994
S581	147,64	1,20	0,334	0,000	-	0,920	2,956	0,996
S525	31,13	1,00	0,400	0,354	0,282	0,920	3,291	0,994
S524	504,11	4,43	0,966	0,995	0,031	1,000	2,717	1,000
S536	7,00	0,06	0,022	0,003	17,359	1,000	2,816	0,999
S527	497,11	4,18	0,961	0,992	0,034	0,900	3,251	0,999
S549	233,72	1,93	0,616	0,000	-	0,980	2,586	0,999
S550	247,55	2,12	0,677	0,000	-	0,920	3,079	0,995
S537	15,84	0,12	0,703	0,438	0,118	1,000	2,207	1,000
S598	13,37	0,12	0,703	0,283	0,114	0,980	2,877	0,999
S599	2,47	0,02	0,117	0,008	2,367	1,000	2,604	1,000
S267	308,47	2,86	0,780	0,951	0,089	0,980	2,240	0,999
S538	310,94	2,74	0,757	0,937	0,097	1,000	2,559	0,999

Tabela D.38: Resultados de vazamentos para o Problema 16.

Equipamento	OP NT	OP GLR NT	AEE GLR	OP Método	AEE Método	Seletividade Método
DA-301	1	1	0,023	1	0,005	1
Split1	1	1	0,022	1	0,005	1
COMP1	1	1	0,033	1	0,005	1
EA-32X	1	1	0,037	1	0,005	1
FA-309	1	1	0,032	1	0,005	1
Split2	1	1	0,023	1	0,005	1
Mixer 1	1	1	0,016	1	0,005	1
FSPL	1	1	0,036	1	0,005	1
H338	1	1	0,025	1	0,005	1
H341	1	1	0,042	1	0,005	1
DA-401	1	1	0,031	1	0,004	1
PTC	1	1	0,034	1	0,005	1
M1	1	1	0,045	1	0,005	1
448-450	1	1	0,046	1	0,004	1
DC-401	1	1	0,045	1	0,002	1
452-448	1	1	0,045	1	0,003	1
DA-408	1	1	0,032	1	0,007	1
DA-402	1	1	0,033	1	0,007	1
DA-404	1	1	0,029	1	0,005	1
DA-405	1	1	0,021	1	0,005	1
DC-402	1	1	0,022	1	0,005	1
DC-40X	1	1	0,027	1	0,007	1
DA-407	1	1	0,041	1	0,006	1
DA-406	1	1	0,030	1	0,005	1
DA-409	1	1	0,001	0	1,500	0
Mix2	1	1	0,018	1	0,005	1

D.2 Erros Múltiplos

Tabela D.39: Cenário de distúrbios com erros múltiplos e os resultados dos ajustes dos modelos.

Problema	Corrente com Erro Medição	Equipamento com Vazamento	R^2 Modelo
P1	1	1	0,97
P2	3	3	0,97
P3	3;10	2	0,96
P4	1	4	0,98
P5	2;4	4	0,95
P6	4;5	2	0,95
P7	6;9	5	0,96
P8	4;11	2	0,94
P9	2;3	5	0,94
P10	3;7;12	8	0,97
P11	6;12	3;7	0,97
P12	3;6;11	2	0,95
P13	2;7;13;15	6	0,97
P14	1;8;9;15;19	4;13;14	0,96
P15	1;13;15;19	1;5;10	0,95
P16	1;6;13;15;19;30;43;46;48	1;12;16;18;26	0,96

D.2.1 Problema 1

Tabela D.40: Resultados dos erros grosseiros múltiplos do Problema 1.

Corrente/Equipamento	OP	AEE	Seletividade
C-1	1	0,382	0,645
E-1	1	0,049	0,609

D.2.2 Problema 2

Tabela D.41: Resultados dos erros grosseiros múltiplos do Problema 2.

Corrente/Equipamento	OP	AEE	Seletividade
C-3	1	0,221	1
E-3	1	0,033	1

D.2.3 Problema 3

Tabela D.42: Resultados dos erros grosseiros múltiplos do Problema 3.

Corrente/Equipamento	OP	AEE	Seletividade
C-3	1	2,49	1
C-10	1	1,94	1
E-2	1	0,024	1

D.2.4 Problema 4

Tabela D.43: Resultados dos erros grosseiros múltiplos do Problema 4.

Corrente/Equipamento	OP	AEE	Seletividade
C-1	1	0,510	0,493
E-4	1	0,078	0,444

D.2.5 Problema 5

Tabela D.44: Resultados dos erros grosseiros múltiplos do Problema 5.

Corrente/Equipamento	OP	AEE	Seletividade
C-2	1	0,259	1
C-4	0,98	0,490	1
E-4	1	0,005	1

D.2.6 Problema 6

Tabela D.45: Resultados dos erros grosseiros múltiplos do Problema 6.

Corrente/Equipamento	OP	AEE	Seletividade
C-4	1	0,647	1
C-5	1	0,216	1
E-2	1	0,019	1

D.2.7 Problema 7

Tabela D.46: Resultados dos erros grosseiros múltiplos do Problema 7.

Corrente/Equipamento	OP	AEE	Seletividade
C-6	1	0,268	1
C-9	1	0,245	1
E-5	1	0,030	1

D.2.8 Problema 8

Tabela D.47: Resultados dos erros grosseiros múltiplos do Problema 8.

Corrente/Equipamento	OP	AEE	Seletividade
C-4	1	0,060	1
C-11	1	0,018	1
E-2	1	0,030	1

D.2.9 Problema 9

Tabela D.48: Resultados dos erros grosseiros múltiplos do Problema 9.

Corrente/Equipamento	OP	AEE	Seletividade
C2	1	0,848	1
C-3	1	0,211	1
E-5	1	0,031	1

D.2.10 Problema 10

Tabela D.49: Resultados dos erros grosseiros múltiplos do Problema 10.

Corrente/Equipamento	OP	AEE	Seletividade
C-3	1	0,080	1
C-7	1	0,210	1
C-12	1	0,484	1
E-8	1	0,020	1

D.2.11 Problema 11

Tabela D.50: Resultados dos erros grosseiros múltiplos do Problema 11.

Corrente/Equipamento	OP	AEE	Seletividade
C-6	1	0,245	1
C-12	1	0,139	1
E-3	1	0,053	1
E-7	1	0,053	1

D.2.12 Problema 12

Tabela D.51: Resultados dos erros grosseiros múltiplos do Problema 12.

Corrente/Equipamento	OP	AEE	Seletividade
C-3	1	0,134	1
C-6	0,91	0,986	1
C-11	1	0,019	1
E-2	1	0,029	1

D.2.13 Problema 13

Tabela D.52: Resultados dos erros grosseiros múltiplos do Problema 13.

Corrente/Equipamento	OP	AEE	Seletividade
C-2	0,978	0,150	1
C-7	0,989	0,157	1
C-13	1	0,332	1
C-15	1	0,206	1
E-6	1	-0,016	1

D.2.14 Problema 14

Tabela D.53: Resultados dos erros grosseiros múltiplos do Problema 14.

Corrente/Equipamento	OP	AEE	Seletividade
C-1	0,98	1,442	0,558
C-8	1	1,971	0,614
C-9	1	-0,443	0,614
C-15	1	1,038	0,614
C-19	1	0,846	0,586
E-4	1	-0,014	0,533
E-13	1	-0,014	0,533
E-14	1	-0,014	0,533

D.2.15 Problema 15

Tabela D.54: Resultados dos erros grosseiros múltiplos do Problema 15.

Corrente/Equipamento	OP	AEE	Seletividade
C-1	0,944	0,771	0,673
C-13	0,967	0,503	0,673
C-15	1	0,227	0,685
C-19	0,989	0,452	0,649
E-1	1	0,035	0,649
E-5	1	0,035	0,649
E-10	1	0,035	0,649

D.2.16 Problema 16

Tabela D.55: Resultados dos erros grosseiros múltiplos do Problema 16.

Corrente/Equipamento	OP	AEE	Seletividade
C-1	0.911	5.441	0.525
C-6	0.595	1.708	0.604
C-13	0.932	0.892	0.604
C-15	0.984	0.740	0.604
C-19	0.968	-0.415	0.604
C-30	0.874	4.608	0.569
C-43	0.805	2.560	0.604
C-46	0.984	1.525	0.604
C-48	0.574	2.263	0.604
E-1	1	-0.016	0.544
E-12	1	-0.016	0.544
E-16	1	-0.016	0.544
E-18	1	-0.016	0.544
E-26	1	-0.016	0.544

Referências Bibliográficas

ABIQUIM, A. a. B. d. I. Q. *A Indústria Química Brasileira*. São Paulo, 2013. 9 p. Disponível em: <http://www.abiquim.org.br/pdf/indQuimica/AIndustriaQuimica-SobreSetor.pdf>.

ADAMS, M. J. *Chemometrics in Analytical Spectroscopy*. 2. ed. Cambridge: Royal Society of Chemistry, 2004.

AGUIRRE, L. A. *Introdução À Identificação de Sistemas: Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. 1. ed. Belo Horizonte: Editora UFMG, 2000.

ALBUQUERQUE, J. a. S.; BIEGLER, L. T. Data reconciliation and gross-error detection for dynamic systems. *AIChE Journal*, v. 42, n. 10, p. 2841–2856, out. 1996.

ALHAJ-DIBO, M.; MAQUIN, D.; RAGOT, J. Data reconciliation: A robust approach using a contaminated distribution. *Control Engineering Practice*, v. 16, n. 2, p. 159–170, fev. 2008.

ALMATO, M.; SANMARTI, E.; ESPUNA, A.; PUIGJANER, L. Rationalizing the water use in the batch process industry. *Computers & Chemical Engineering*, v. 21, n. Supplement 1, p. S971–S976, maio 1997.

AQUIM, P. M. D. *Gestão em Curtumes: Uso Integrado e Eficiente da Água*. Tese (Doutorado), 2009.

ARORA, N.; BIEGLER, L. T. Redescending estimators for data reconciliation and parameter estimation. *Computers and Chemical Engineering*, v. 25, p. 1585–1599, 2001.

BAGAJEWICZ, M. A review of recent design procedures for water networks in refineries and process plants. *Computers & Chemical Engineering*, v. 24, n. 9-10, p. 2093–2113, out. 2000.

BAGAJEWICZ, M.; JIANG, Q.; SÁNCHEZ, M. Performance evaluation of PCA tests for multiple gross error identification. *Computers & Chemical Engineering*, v. 23, n. Supplement 1, p. S589–S592, jun. 1999.

BAGAJEWICZ, M.; RODERA, H.; SAVELSKI, M. Energy efficient water utilization systems in process plants. *Computers & Chemical Engineering*, v. 26, n. 1, p. 59–79, jan. 2002.

BIPM, B. I. d. P. e. M. *JCGM 100:2008 Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement*. [S.l.], 2008. Disponível em: <http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E-.pdf>.

BIPM, B. I. d. P. e. M. *JCGM 104:2009 - Evaluation of measurement data - An introduction to the 'Guide to the expression of uncertainty in measurement' and related documents*. [S.l.], 2009. Disponível em: <http://www.bipm.org/utils-/common/documents/jcgm/JCGM_104_2009_E.pdf>.

STATISTICS CANADA. *Canada - Industrial Water Use*. Otawa, 2009. Disponível em: <http://www.statcan.gc.ca/pub/16-401-x/16-401-x2012001-eng.pdf>.

CHARPENTIER, V.; CHANG, L. J.; SCHWENZER, G.; BARDIN, M. C. An On-line Data Reconciliation System for Crude and Vacuum Units. In: *NPRA Computer Conference*. Houston: [s.n.], 1991.

CNTL, C. N. d. T. L. *Implementação de Programas de Produção mais Limpa*. Porto Alegre, 2003.

EUROSTAT EUROPEAN COMMISION. *Environmental statistics and accounts in Europe*. Luxembourg, 2010. Disponível em: <http://epp.eurostat.ec.europa.eu-/cache/ITY_OFFPUB/KS-32-10-283/EN/KS-32-10-283-EN.PDF>.

CROWE, C. M. Recursive identification of gross errors in linear data reconciliation. *AIChE Journal*, v. 34, p. 541–555, 1988.

CROWE, C. M. Observability and Redundancy of Process Data for Steady State Data Reconciliation. v. 44, n. 12, p. 2909–2917, 1989.

CROWE, C. M. Data reconciliation - progress and challenges. *Journal of Process Control*, v. 6, n. 2/3, p. 89–98, 1996.

CROWE, C. M.; CAMPOS, Y. G.; HRYMAK, A. Reconciliation of Process Flow Rates by Matrix Projection, Part I: Linear Case. *AICHE Journal*, v. 29, n. 6, p. 881–888, 1983.

CUNHA, E.; PESSOA, C. L. P.; COSTA, C. L. V.; SIQUEIRA, R. A.; KALID, R. D. A. Em Busca do Efluente Zero. In: KIPERSTOK, A. (Ed.). *Prata da casa. Construindo a produção limpa na Bahia*. Primeira e. Salvador: Teclim/UFBA, 2008. p. 175–192.

DEVANATHAN, S.; VARDEMAN, S. B.; Rollins Sr., D. K.; ROLLINS, D. K. Likelihood and Bayesian Methods for Accurate Identification of Measurement Biases in Pseudo Steady-State Processes. *Chemical Engineering Research and Design*, v. 83, n. 12, p. 1391–1398, dez. 2005.

EL-HALWAGI, M. *Pollution Prevention through Process Integration: Systematic Design Tools*. 1. ed. London: Academic Press, 1997.

EL-HALWAGI, M.; MANOUSIOUTHAKIS, V. Synthesis of Mass Exchange Networks. *AIChE Journal*, v. 35, n. 8, p. 1233–1244, 1989.

FARENZENA, M.; FERREIRA, L. S.; TRIERWEILER, J. O.; AQUIM, P. M. Curtumes: do desperdício à sustentabilidade. In: *XV Congresso Brasileiro de Engenharia Química*. Curitiba: [s.n.], 2004.

FARIAS, A. C. *Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros*. Tese (Dissertação de Mestrado) — Universidade Federal do Rio Grande do Sul, 2009.

FEPAM. *Diagnóstico da poluição hídrica industrial na região hidrográfica do Guaíba*. Porto Alegre, 2001.

FIESP. *Prêmio FIESP Conservação e Reúso de Água*. 2013. Disponível em: <http://hotsite.fiesp.com.br/premioagua/cases.htm>.

FINKLER, T. F. *Desenvolvimento de uma Ferramenta para Obtenção de Modelos Empíricos*. Tese (Dissertação de Mestrado) — Universidade Federal do Rio Grande do Sul, 2003.

FINKLER, T. F.; SOARES, R. D. P.; VALLE, E. C. D. *VRAnalyst*. Porto Alegre: VRTech Industrial Technologies, 2006. Disponível em: <http://www.vrtech.com-.br/en_us/vranalyst/vranalyst-modules.html>.

FONTANA, D. *Recuperação de águas de processos, desenvolvimento de um problema padrão*. 110 p. Tese (Dissertação (Mestrado em Pesquisa e Desenvolvimento de Processos)) — Universidade Federal do Rio Grande do Sul, 2002.

FONTANA, D.; KALID, R.; SARTORI, I.; KIPERSTOK, A.; SILVA, M.; SALES, E. A.; FILHO, J. G. P.; OLIVEIRA, S.; PERAZZO, C. Balanço hídrico - uma nova sistemática. In: *XV Congresso Brasileiro de Engenharia Química*. Curitiba: [s.n.], 2004.

FONTANA, D.; KALID, R. D. A.; KIPERSTOK, A.; SILVA, M. Methodology for wastewater minimization in industries in the petrochemical complex 1. In: *2nd Mercosur Congress on Chemical Engineering*. Rio de Janeiro: [s.n.], 2004. p. 1–10.

GOMES, J. *Procedimento para minimização de Efluentes Aquosos*. Tese (Mestrado em em Tecnologia de Processos Químicos e Bioquímicos) — Universidade Federal do Rio de Janeiro, 2002.

AUSTRALIA GOVERNMENT. *Australian Water Resources 2005*. Cambera, 2006. Disponível em: <http://www.water.gov.au/publications/index-.aspx?Menu=Level1_9*backslash*level1>.

GRANDE, M. H. D.; SALES, E. Racionalização do Uso de Água na Indústria de Celulose Bahia Pulp S/A. In: KIPERSTOK, A. (Ed.). *Prata da casa. Construindo a produção limpa na Bahia*. Primeira e. Salvador: Teclim/UFBA, 2008. p. 157–173.

HARIKUMAR, P.; NARASIMHAN, S. A Method to Incorporate Bounds in Data Reconciliation and Gross Error Detection - II. Gross Error Detection Strategies. *Computers & Chemical Engineering*, v. 17, n. 11, p. 1121–1128, 1993.

HEYEN, G.; MARECHÁL, E.; KALITVENTZEFF, B. Sensitivity calculations and variance analysis in plant measurement reconciliation. *Computers & Chemical Engineering*, v. 20, n. 96, p. S539–S544, 1996.

HOINKIS, J.; PANTEN, V. Wastewater recycling in laundries-From pilot to large-scale plant. *Chemical Engineering and Processing: Process Intensification*, v. 47, n. 7, p. 1159–1164, jul. 2008.

INMETRO, I. N. d. M. Q. e. T. *VIM - Vocabulário Internacional de Metrologia - 2012*. Duque de Caxias - Rio de Janeiro, 2012. Disponível em: <http://www-.inmetro.gov.br/infotec/publicacoes/vim_2012.pdf>.

INRIA. *Scilab*. Domaine de Voluceau: Claude Gomez, 2010. Disponível em: <http://www.scilab.org/>.

IORDACHE, C.; MAH, R. S. H.; TAMHANE, A. C. Performance Studies of the Measurement Test for Detecting Gross Errors in Process Data. *AIChE Journal*, v. 31, p. 1187–1201, 1985.

JODICKE, G.; FISCHER, U.; HUNGERBUHLER, K. Wastewater reuse: a new approach to screen for designs with minimal total costs. *Computers & Chemical Engineering*, v. 25, n. 2-3, p. 203–215, mar. 2001.

JOLLIFFE, I. T. *Principal Component Analysis*. 2. ed. New York: Springer-Verlag, 2002.

JUNIOR, C. A. V.; MEDEIROS, J. L. D.; O. Q. F. Araújo. Ferramentas de Detecção de Anomalias Baseadas em Classificadores Hierárquicos e Teoria dos Protótipos. In: *XVII Congresso Brasileiro de Engenharia Química*. Recife - Brazil: [s.n.], 2008.

KELLER, J. Y.; DAROUACH, M.; KRZAKALA, G. Fault detection of Multiple Biases or process leaks in Linear Steady State Systems. *Comp. & Chem. Eng*, v. 18, p. 1001–1004, 1994.

KELLY, J. Reconciliation of process data using other projection matrices. *Computers & Chemical Engineering*, v. 23, n. 6, p. 785–789, jun. 1999. ISSN 00981354.

KENNY, J.; BARBER, N.; HUTSON, S.; LINSEY, K.; LOVELACE, J.; MAUPIN, M. *Estimated Use of Water in the United States in 2005*. Reston, 2009. Disponível em: <http://pubs.usgs.gov/circ/1344/pdf/c1344.pdf>.

KHOR, C. S.; SHAH, N.; MAHADZIR, S.; ELKAMEL, A. Optimisation of petroleum refinery water network systems retrofit incorporating reuse, regeneration and recycle strategies. *The Canadian Journal of Chemical Engineering*, v. 90, n. 1, p. 137–143, fev. 2012.

KIPERSTOK, A.; ESQUERRE, K.; KALID, R.; SALES, E.; OLIVEIRA, G. Rationalizing the Use of Water in Industry-Part 2: Instruments Developed by the Clean Technology Network in the State of Bahia. *Journal of Environmental Protection*, v. 4, p. 497–507, 2013.

KIPERSTOK, A.; ESQUERRE, K.; KALID, R. A.; SALES, E.; OLIVEIRA, G. Rationalizing the Use of Water in Industry-Part 1: Summary of the Instruments Developed by the Clean Technology Network in the State of Bahia and Main Results Obtained. *Journal of Environmental Protection*, v. 4, p. 486–496, 2013.

KONGSJAHJU, R.; ROLLINS, D. K.; nANA, M. B. B. Accurate Identification of Biased Measurements Under Serial Correlation. *Trans IChemE*, v. 78, n. Part A, p. 1010–1018, 2000.

KRETSOVALIS, A.; MAH, R. S. H. Effect of redundancy on estimation accuracy in process data reconciliation. *Chemical Engineering Science*, v. 42, n. 9, p. 2115–2121, 1987.

MADRON, F. A New Approach to the Identification of Gross Error in Chemical Engineering Measurements. *Chemica Engineering Science*, v. 40, p. 1855–1860, 1985.

MADRON, F. *Process Plant Performance: Measurement and Data for Optimization and Retrofits*. 1. ed. West Sussex: Ellis Horwood Limited Co, 1992.

MAH, R. S. H.; STANLEY, G. M.; DOWNNING, D. M. Reconciliation of Process Flow and Invetory Data. *Ind. & Eng. Chem. Proc. Des. Dev, Vol.*, v. 15, n. 1, p. 175–183, 1976.

MAH, R. S. H.; TAMHANE, A. C. Detection of gross errors in process data. *AIChE Journal*, v. 28, p. 828–830, 1982.

MANN, J.; LIU, Y. *Industrial water reuse and wastewater minimization*. 1. ed. Ney York: McGraw-Hill Professional, 1999. ISBN 978-0071348553.

MARK, H.; WORKMAN, J. *Chemometrics in Spectroscopy*. 1. ed. London: Academic Press, 2007. ISBN 978-0123740243.

MARONNA, R.; ARCAS, J. Data reconciliation and gross error diagnosis based on regression. *Computers & Chemical Engineering*, v. 33, n. 1, p. 65–71, 2009.

MARTINS, M. A.; AMARO, C. A.; SOUZA, L. S.; KALID, R. A.; KIPERSTOK, A. New objective function for data reconciliation in water balance from industrial processes. *Journal of Cleaner Production*, p. 1–6, mar. 2010.

MEI, C.; SU, H.; CHU, J. An NT-MT Combined Method for Gross Error Detection and Data Reconciliation. *Chinese Journal of Chemical Engineering*, v. 14, n. 5, p. 592–596, out. 2006.

MIRRE, R. C.; YOKOYAMA, L.; PESSOA, F. L. P. Diagrama de Fontes de Água: Uma Ferramenta Sustentável para a P+L na Minimização de Águas e Efluentes Industriais. In: *2nd International Workshop Advances in Cleaner Production*. São Paulo: [s.n.], 2009.

MITSAS, C. L. Data reconciliation and variable classification by null space methods. *Measurement*, ELSEVIER SCI LTD, v. 43, n. 5, p. 702–707, jun. 2010.

NARASIMHAN, S.; JORDACHE, C. *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*. 1. ed. Houston: Gulf Publishing, 2000.

NARASIMHAN, S.; MAH, R. S. H. Generalized Likelihood Ratio Method for Gross Error Identification. *AIChE Journal*, v. 33, p. 1514–1521, 1987.

NARCISO, M. V. A. *Alocação Ótima de Instrumentos de em Sistemas de Medições Aplicados a Plantas Industriais sem Medições.* Tese (Dissertação de Mestrado) — Universidade Federal da Bahia, 2012.

OLIVEIRA, G. L. de. *A Metodologia TECLIM para Uso Racional da Água na Indústria: Uma proposta de sistematização.* Tese (Dissertação de Mestrado) — Universidade Federal da Bahia, 2011.

ÖZYURT, D. B.; PIKE, R. W. Theory and practice of simultaneous data reconciliation and gross error detection for chemical. *processes. Computers and Chemical Engineering*, v. 28, p. 381–402, 2004.

PRATA, D. M.; FREITAS, K. P.; LIMA, E. L. Análise de Clusters para Detecção de Erros Grosseiros e Reconciliação de dados: Aplicação Industrial. In: *XVII Congresso Brasileiro de Engenharia Química.* Recife - Brazil: [s.n.], 2008.

Raghava Rao, J.; CHANDRABABU, N. K.; MURALIDHARAN, C.; NAIR, B. U.; RAO, P. G.; RAMASAMI, T. Recouping the wastewater: a way forward for cleaner leather processing. *Journal of Cleaner Production*, v. 11, n. 5, p. 591–599, ago. 2003.

RAWLINGS, J. O.; PANTULA, S. G.; DICKEY, D. A. *Applied Regression Analysis: A Research Tool.* 2. ed. New York: Springer-Verlag, 1998.

REDDY, V. N.; MAVROVOUNIOTIS, M. L. An Input-Training Neural Network Approach for Gross Error Detection and Sensor Replacement. *Chemical Engineering Research and Design*, v. 76, n. 4, p. 478–489, maio 1998.

ROLLINS, D. K.; CHENG, Y.; DEVANATHAN, S. Intelligent Selection of Hypothesis Tests to Enhance Gross Error Identification. *Comp. & Chem. Engineering'*, v. 20, n. 5, p. 517–530, 1996.

ROLLINS, D. K.; DAVIS, J. F. Unbiased Estimation of Gross Error in Process Measurements. *AIChE Journal*, v. 38, p. 563–572, 1992.

ROMAGNOLI, J.; SÁNCHEZ, M. C. *Process Systems Engineering - Volume 2 - Data Processing and Reconciliation for Chemical Process Operations.* [s.n.], 1999. 270 p.

ROMAGNOLI, J.; STEPHANOPOULOS, G. Rectification of Process Measrement Data in the Presence of Gross Errors. *Chemical Engineering Science*, v. 36, n. 11, p. 1849–1863, 1981.

ROSENBERG, J.; MAH, R. S. H.; IORDACHE, C. Evaluation of Schemes for Detecting and Identifying Gross Errors in Process Data. *Ind. & Eng. Chem. Proc. Des. Dev, Vol.*, v. 26, p. 555–564, 1987.

SÁNCHEZ, M.; ROMAGNOLI, J. Use of orthogonal transformations in data classification-reconciliation. *Computers & Chemical Engineering*, v. 20, n. 5, p. 483–493, 1996.

SÁNCHEZ, M.; ROMAGNOLI, J. J.; JIANG, Q.; BAGAJEWICZ, M.; SA, M. Simultaneous estimation of biases and leaks in process plants. *Computers & Chemical Engineering*, v. 23, n. 7, p. 841 – 857, jul. 1999.

SÁNCHEZ, M. C.; SENTONI, G.; SCHBIB, S.; TONEILI, S.; ROMAGNOLI, J. Gross Measurements Error Detection/Identification for ans Industrial Ethylene Reactor. *Comp. & Chem. Eng*, v. 20, p. S1559–S1564, 1996.

SERTH, R. W.; HEENAN, W. A. Gross error detection and data reconciliation in steam-metering systems. *AIChE Journal*, v. 32, p. 733–747, 1986.

UNITED KINGDON STATISTICAL SERVICE. *Environmental Statistics: Key Facts*. London, 2011. Disponível em: <http://www.defra.gov.uk/statistics/files-/Environmental-key-statistics-Dec-2011.pdf>.

SHUKLA, S. K.; KUMAR, V.; YEOM, I. T.; BANSAL, M. C. Recycling of bleach plant effluent of an Indian paper mill using water cascade analysis technique. *Clean Technologies and Environmental Policy*, v. 14, n. 4, p. 677–685, nov. 2011.

SINGTH, S. R.; MITTAL, N.; SEN, P. A Novel Data Reconciliation and Gross Error Detection Tool for the Mineral Processing Industry. *Minerals Engineering*, v. 14, n. 7, p. 809–814, 2001.

SMITH, R. *Chemical Process: Design and Integration*. First. Southern Gate: John Wiley & Sons, 2005. ISBN 978-0471486817.

SODERSTROM, T. A.; HIMMELBLAU, D. M.; EDGAR, T. F. A mixed integer optimization approach for simultaneous data reconciliation and identification of measurement bias. *Control Engineering Practice*, v. 9, n. 8, p. 869–876, ago. 2001.

SOUZA, L. S. D.; KALID, R.; KIPERSTOK, A. Propagação da incerteza na reconciliação de dados com restrições lineares. 2010.

STANLEY, G. M.; MAH, R. S. H. Observability and redundancy estimation in process data estimation. *Chemical Engineering Science*, v. 36, p. 259–272, 1980.

SWARTZ, C. L. E. Data Reconciliation for Generalized Flowsheet Applications. In: *American Chemical Society National Meeting*. [S.l.: s.n.], 1989.

THEVENDIRARAJ, S.; KLEMES, J.; PAZ, D.; ASO, G.; CARDENAS, G. J. Water and wastewater minimisation study of a citrus plant. *Resources, Conservation and Recycling*, v. 37, p. 227–250, 2003.

TOKOS, H.; PINTARIC, Z. N.; YANG, Y.; KRAVANJA, Z. Multilevel Strategies for the Retrofit of Large- Scale Industrial Water System: A Brewery Case Study. *AIChE Journal*, v. 58, n. 3, p. 884–898, 2012.

TONG, H.; CROWE, C. M. Detection of gross errors in data reconciliation by principal component analysis. *AIChE Journal*, v. 41, p. 1712–1722, 1995.

VALLE, E. C. *Data Reconciliation and Gross Error Detection Benchmark Problems*. 2011. Disponível em: <http://code.google.com/p/dr-ged-benchmarks/>.

VALLE, E. C. D. *Minimização do Uso de Água e Efluentes com Considerações Econômicas e Operacionais via Programação Matemática*. Tese (MSc) — Universidade Federal do Rio Grande do Sul, 2005.

Van Der Leeden, F.; TROISE, F. L.; TODD, D. K. *The Water Encyclopedia*. [S.l.]: Lewis Publishers, 1990. (Geraghty & Miller : Ground-water series).

VARMUZA, K.; FILZMOSER, P. *Introduction to Multivariate Statistical Analysis in Chemometrics*. First edit. Boca Raton: Taylor & Francis, 2009.

VEVERKA, V.; MADRON, F. *Material and Energy Balancing in the Process Industries: From Microscopic Balances to Large Plants*. First edit. Amsterdan: Elsevier B.V., 1997.

WÄCHTER, A.; BIEGLER, L. T. On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming*, v. 106, n. 1, p. 25–57, 2006.

WANG, F.; JIA, X.-p.; ZHENG, S.-q.; YUE, J.-c. An improved MT-NT method for gross error detection and data reconciliation. *Computers & Chemical Engineering*, v. 28, n. 11, p. 2189–2192, out. 2004.

WANG, Y.-P.; SMITH, R. Design of distributed effluent treatment systems. *Chemical Engineering Science*, v. 49, n. 18, p. 3127–3145, set. 1994.

WERKEMA, C.; AGUIAR, S. *Análise de Regresão, Como Entender o Relacionamento entre as Variáveis de um Processo*. 1. ed. Belo Horizonte: Werkema Editora, 2006.

ZHANG, Z.; SHAO, Z.; CHEN, X.; WANG, K.; QIAN, J. Quasi-weighted least squares estimator for data reconciliation. *Computers & Chemical Engineering*, v. 34, n. 2, p. 154–162, fev. 2010.