



MONITORAMENTO DE PROCESSOS DE SEPARAÇÃO BASEADO EM RECONCILIAÇÃO DE DADOS

Diego Queiroz Faria de Menezes

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

Orientadores: José Carlos Costa da Silva Pinto

Argimiro Resende Secchi

Diego Martinez Prata

Rio de Janeiro

Janeiro de 2021

MONITORAMENTO DE PROCESSOS DE SEPARAÇÃO BASEADO EM
RECONCILIAÇÃO DE DADOS

Diego Queiroz Faria de Menezes

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

Orientadores: José Carlos Costa da Silva Pinto

Argimiro Resende Secchi

Diego Martinez Prata

Aprovada por: Prof. José Carlos Costa da Silva Pinto

Prof. Argimiro Resende Secchi

Prof. Diego Martinez Prata

Prof. Príamo Albuquerque Melo Jr.

Profa. Verônica Maria de Araújo Calado

Prof. Lizandro de Sousa Santos

Dr. Thiago de Sá Feital

RIO DE JANEIRO, RJ – BRASIL

JANEIRO DE 2021

de Menezes, Diego Queiroz Faria

Monitoramento de processos de separação baseado em reconciliação de dados/Diego Queiroz Faria de Menezes. – Rio de Janeiro: UFRJ/COPPE, 2021.

XXIII, 409 p.: il.; 29, 7cm.

Orientadores: José Carlos Costa da Silva Pinto

Argimiro Resende Secchi

Diego Martinez Prata

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Química, 2021.

Referências Bibliográficas: p. 266 – 289.

1. Reconciliação de dados. 2. Estatística robusta.
3. Coluna de destilação. 4. Separação por membrana.
5. Monitoramento. I. Pinto, José Carlos Costa da Silva *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química. III. Título.

*A todos os cidadãos brasileiros
contribuintes, pois meu incentivo
através da bolsa de estudos é
fruto desses verdadeiros
criadores de riqueza. Aos
indivíduos que mantêm,
inabaladamente, seu patriotismo,
sua honestidade, moral e ética,
que diante de tantos fatos que
vilipendiam a ordem e o
progresso de nosso país,
continuam a lutar pela
manutenção e seguridade de suas
tradições, liberdade, prosperidade
e paz.*

Agradecimentos

Gostaria de agradecer a minha mãe, que sempre me ajudou de todas as formas possíveis. Seu cuidado, dedicação, carinho e amor me deram base e força para seguir em frente e nunca desistir, mesmo sem entender o porquê de que eu só “estudo” e não “trabalho”. A minha família, por acreditar e investir em mim, mesmo sem entender o que é uma pós-graduação. A meus amigos, que são muitos que não caberiam em uma folha.

Ao professor, orientador e amigo Diego Prata, que sempre acreditou no meu potencial, que de forma paciente e sistemática me mostrou o “caminho das pedras” nesta jornada acadêmica, que desde seu primeiro dia como professor na UFF demonstrou alegria em dar aula, que me empolgava ao falar de sua experiência na indústria química, que me apresentou o fascinante mundo da Reconciliação de Dados e, também, me incentivou nos momentos de recaída emocional. Ao professor e amigo Lizandro, com quem sempre tenho boas horas de conversa sobre qualquer tema, uma pessoa excelente para trabalhar, um grande pesquisador, que me ensinou e sanou muitas de minhas dúvidas. A meus professores e orientadores Zé e Argimiro com quem tive a grande honra de aprender um pouco mais de Engenharia Química, pois bastavam apenas algumas horas de reunião com esses gênios para que minhas dúvidas e problemas de convergência fossem sanados. Agradeço a paciência deles comigo, pois sei que sou uma pessoa atípica e nem um pouco “Caxias”. A liberdade e confiança depositada em mim foram um estímulo para continuar este trabalho.

Aos colegas do PEQ/COPPE/UFRJ que, sem pestanejar, param o que estão fazendo para explicar e sanar qualquer dúvida sua. Admiro todos, pois são excelentes pesquisadores e ótimas pessoas. A todos os demais professores e funcionários da COPPE, que também ajudam esta instituição a manter o mais alto grau de excelência em pesquisa e ensino deste país. Tenho um imenso orgulho de ter feito meu doutoramento na COPPE/UFRJ, instituição *alma mater* de notórios pesquisadores, professores e profissionais da engenharia, que até o presente momento formou 17.347 mestres e doutores dentre os 13 programas de engenharia. A CAPES, CNPq e Petrobras que me concederam bolsas durante o doutoramento.

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

MONITORAMENTO DE PROCESSOS DE SEPARAÇÃO BASEADO EM RECONCILIAÇÃO DE DADOS

Diego Queiroz Faria de Menezes

Janeiro/2021

Orientadores: José Carlos Costa da Silva Pinto

Argimiro Resende Secchi

Diego Martinez Prata

Programa: Engenharia Química

Este trabalho apresenta 4 estudos de caso e uma revisão, com o objetivo de desenvolver uma metodologia para solucionar de maneira robusta o problema de reconciliação de dados dinâmica ou estacionária, em processos de separação para monitoramento em tempo real. Três estudos de casos avaliam o problema de reconciliação de dados dinâmica em colunas de destilação, avaliando também os métodos de otimização, estimadores-M, modelo reduzido de coluna e estratégias para aumentar a velocidade e robustez, baseadas em métricas e modelo preditivo autoregressivo. O quarto caso apresenta uma metodologia baseada em reconciliação de dados, para monitoramento e diagnósticos de erros grosseiros em tempo real, para um processo de separação de CO_2 por membranas. Por fim, uma revisão completa sobre estimadores-M robusto foi realizada, complementando a metodologia desenvolvida para o monitoramento em tempo real e eliminação simultânea de erros grosseiros. Os otimizadores de natureza determinística apresentaram melhores resultados. Os estimadores Welsch e Normal Contaminada apresentaram melhores desempenhos. O modelo reduzido apresentou melhorias na velocidade de otimização sem perder a acurácia. A estratégia apresentou um excelente desempenho, reduzindo os ruídos e erros grosseiros, aumentando a robustez e velocidade. A revisão destacou 50 estimadores-M (2 não robustos e 48 robustos), sintonizando-os para níveis de Eficiência Relativa de 90, 95, 98 e 99%.

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

MONITORING OF SEPARATION PROCESSES BASED ON DATA RECONCILIATION

Diego Queiroz Faria de Menezes

January/2021

Advisors: José Carlos Costa da Silva Pinto

Argimiro Resende Secchi

Diego Martinez Prata

Department: Chemical Engineering

This work presents 4 case studies and a review, with the objective of developing a methodology to robustly solve the problem of dynamic or stationary data reconciliation, in separation processes for real-time monitoring. Three case studies evaluate the problem of dynamic data reconciliation in distillation columns, also evaluating the optimization methods, M-estimators, reduced model of column and strategies to increase speed and robustness, based on metrics and autoregressive predictive model. The fourth case presents a methodology based on data reconciliation, for monitoring and diagnosing gross errors in real time, for a separation process of CO_2 by membranes. Finally, a complete review of robust M-estimators was carried out, complementing the methodology developed for real-time monitoring and simultaneous elimination of gross errors. The deterministic optimizers showed better results. The Welsch and Normal Contaminated estimators showed better performances. The reduced model showed improvements in the optimization speed without losing accuracy. The strategy presented an excellent performance, reducing the noise and gross errors, increasing the robustness and speed. The review highlighted 50 M-estimators (2 not robust and 48 robust), tuning them to Relative Efficiency levels of 90, 95, 98 and 99%.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xviii
Lista de Abreviaturas	xxi
1 Introdução	1
1.1 Contextualização	1
1.2 Motivação	7
1.3 Objetivo	12
1.4 Estrutura da Tese	14
2 Retificação de Dados	16
2.1 Classificação de Variáveis	21
2.2 Matriz de Covariância	24
2.3 Detecção de Erro Grosseiro	28
2.4 Reconciliação de Dados	38
2.4.1 Reconciliação de Dados Estacionária	41
2.4.2 Reconciliação de Dados Dinâmica	62
2.5 Comentários Finais	79
3 Estatística Robusta	80
3.1 Contextualização	80

3.2	Desenvolvimento de estimadores-M robustos	83
3.2.1	Formulação de Máxima Verossimilhança para Estimadores-M Robustos	83
3.2.2	Generalização da função objetivo de máxima verossimilhança .	87
3.2.3	Função de influência	89
3.2.4	Derivada da função de influência	92
3.2.5	Robustez e eficiência relativa	93
3.2.6	Limite de ruptura	94
3.2.7	Exemplos de estimadores-M robustos	95
3.3	Estimador de Correntropia	98
3.3.1	Teoria da Informação e entropia	98
3.3.2	Estimativa de Densidade <i>Kernel</i>	102
3.3.3	Correntropia e suas propriedades	104
3.3.4	Correntropia e aplicações em RD	106
3.4	Critério de avaliação e desempenho	108
3.4.1	Critério Fobj	108
3.4.2	Critério SSE	109
3.4.3	Critério TER	109
3.4.4	Critério SARES	110
3.5	Critério para detecção de valores espúrios	110
3.6	Critério de avaliação da detecção de valores espúrios	111
3.6.1	Critério OP	111
3.6.2	Critério AVTI	112
3.7	Comentários Finais	112
4	Colunas de Destilação	114
4.1	Contextualização	114
4.2	Fundamentos	116

4.2.1	Fundamentos termodinâmicos	120
4.3	O modelo MESH	129
4.3.1	Equações MESH em estado estacionário	129
4.3.2	Modelo dinâmico de coluna de destilação multicomponente . .	134
4.4	Modelos de ordem reduzida	141
4.4.1	Fundamentos teóricos e formulação matemática da técnica de resíduos ponderados pelos momentos	143
4.5	Comentários Finais	152
5	Metodologia	153
5.1	Processo (a) - Coluna de destilação	154
5.1.1	Aquisição de dados e frequência de amostragem	154
5.1.2	Análises de séries temporais	157
5.1.3	Modelagem matemática da coluna e geração dos erros de medida	167
5.1.4	Sintonia dos estimadores-M robustos	168
5.1.5	Comunicação EMSO-Matlab	169
5.1.6	Estratégias na RDDR	172
5.2	Processo (b) - Separação de CO_2 por membrana	179
5.2.1	Introdução	179
5.2.2	Processo de Separação por Membrana	180
5.2.3	Metodologia	182
5.3	Comentários Finais	182
6	Resultados	184
6.1	Apresentação dos estudos de caso	184
6.1.1	Caso 1: Avaliação dos métodos de otimização	184
6.1.2	Caso 2: Avaliação da redução de ordem do modelo de coluna .	195
6.1.3	Caso 3: Avaliação do termo de similaridade com modelo autoregressivo	200

6.1.4	Caso 4: Monitoramento de um processo de separação por membrana	209
6.2	Resultados dos estudos de caso	209
6.2.1	Caso 1: Avaliação dos métodos de otimização	209
6.2.2	Caso 2: Avaliação da redução de ordem do modelo de coluna .	245
6.2.3	Caso 3: Avaliação do termo de similaridade com modelo autoregressivo	250
6.2.4	Caso 4: Monitoramento de um processo de separação por membrana	257
7	Conclusão	258
7.1	Caso 1: Avaliação dos métodos de otimização	258
7.2	Caso 2: Avaliação da redução de ordem do modelo de coluna	259
7.3	Caso 3: Avaliação do termo de similaridade com modelo autoregressivo	260
7.4	Caso 4: Monitoramento de um processo de separação por membrana .	261
7.5	Revisão sobre estimadores-M robustos para análises de regressão . . .	262
7.6	Continuação da pesquisa e desenvolvimento - trabalhos futuros	263
	Referências Bibliográficas	266
A	Artigo Publicado - Modeling of Spiral Wound Membranes for Gas Separations—Part II: Data Reconciliation for Online Monitoring	290
B	Artigo Aceito - A Review on Robust M-Estimators for Regression Analysis	321

Lista de Figuras

1.1	Retificação de dados e aplicações.	3
1.2	Pirâmide de gerenciamento de processo.	5
1.3	Estrutura geralmente utilizada na otimizações de processo em relação às escalas de tempo e variáveis.	6
1.4	Ciclo da Otimização em Tempo Real.	7
1.5	Vantagens da Reconciliação de Dados.	10
2.1	Estrutura resumida do procedimento de retificação de dados.	17
2.2	Dados medidos com erros sistemáticos (<i>bias</i>) e dados reconciliados.	20
2.3	Dados medidos com erros espúrios (<i>outliers</i>) e dados reconciliados.	20
2.4	Dados medidos com erros acumulativos (<i>drifts</i>) e dados reconciliados.	21
2.5	Dados medidos com falha completa e dados reconciliados.	21
2.6	Distribuição marginal de erros de medição.	44
2.7	Bomba centrífuga (problema estacionário linearmente restrito).	59
2.8	Interpretação gráfica da solução do problema de RD linearmente restrito proposto no Estudo de Caso 01.	61
2.9	A abordagem de Horizonte Móvel.	66
2.10	Esquema de solução sequencial para problemas de RDD não linearmente restrito pela abordagem de HM.	68
2.11	Conceito Algoritmo da RDD baseado em predição-correção.	77
2.12	Relação entre os erros de medição, de predição e o valor reconciliado.	77
3.1	Algumas funções de distribuição de probabilidades.	86

3.2	M-Estimators - $\rho(\xi)$	88
3.3	Algumas funções de influência de alguns estimadores (DE MENEZES, 2015).	90
3.4	Algumas propriedades das funções de influência (HAMPEL <i>et al.</i> , 1986).	91
3.5	Uma comparação entre alguns dos estimadores analisados ($\rho(\xi)$).	97
3.6	Diagrama esquemático do sistema de comunicação de SHANNON e WEAVER (1964).	99
3.7	Três sistemas com diferentes níveis de entropia. Fonte: NASCIMENTO (2013b).	101
3.8	Efeito do tamanho da janela de <i>Parzen</i> sobre a estimativa de FDP.	103
3.9	Estimador de Correntropia.	105
3.10	Função de Influência de Correntropia.	106
4.1	Esquema típico de uma coluna de destilação fracionada (PETLYUK, 2004).	120
4.2	Esquema típico de um prato (BUCKLEY <i>et al.</i> , 1985).	120
4.3	Estágio de equilíbrio genérico (KISTER, 1992; SEADER <i>et al.</i> , 2010).	132
4.4	Representação de uma coluna de destilação com múltiplas alimentações, retiradas de líquido e vapor laterais, trocas térmicas, condensador parcial e refeedor (KISTER, 1992).	133
5.1	Discretização de sinal em amplitude e no tempo.	154
5.2	Exemplo de amostragem de sinal contínuo.	155
5.3	Esquema ilustrativo do fenômeno de <i>aliasing</i>	156
5.4	Esquema ilustrativo do processo de escolha do melhor período de amostragem.	157
5.5	Esquema ilustrativo de um processo estocástico $X(t)$. Fonte: GUIMARÃES (2007).	158
5.6	Processo estocástico interpretado como uma família de FDPs. Fonte: MORETTIN e TOLOI (2018).	159
5.7	Processos estocásticos com mesma FDP de 1 ^a ordem. Fonte: YNOGUTI (2011).	160

5.8	Função de autocorrelação para os processos $X(t)$ e $Y(t)$. Fonte: GUIMARÃES (2007).	160
5.9	Diferença na aferição de amostras “através” e “ao longo” de um processo estocástico. Fonte: GUIMARÃES (2007).	162
5.10	Matriz de autocorrelação correspondente de um processo estacionário. Fonte: BOX <i>et al.</i> (2015).	164
5.11	Exemplo de FAC - gráfico gerado no Python com a biblioteca <i>StatsModels</i> (lag=50).	164
5.12	Exemplo de FACP - gráfico gerado no Python com a biblioteca <i>StatsModels</i> ($k=50$).	165
5.13	Sintonização do estimador de Hampel utilizando o <i>software</i> Maple.	169
5.14	Interface EMSO-Matlab.	170
5.15	Comunicação com o Simulink (EMSO-Matlab).	171
5.16	<i>Softwares</i> EMSO e Matlab aplicados ao esquema de solução sequencial para o problema de RDD.	171
5.17	Vetor das estimativas iniciais do otimizador, das condições iniciais do modelo e os limites das variáveis de otimização no Matlab.	176
5.18	Vetor das condições iniciais e de entrada no Simulink.	176
5.19	Vetor das condições iniciais no EMSO.	177
5.20	Vetor das condições iniciais e de entrada no <i>software</i> de conexão EMSO-Matlab.	177
5.21	Termo extra da função objetivo que força a similaridade temporal.	178
5.22	Representação esquemática de uma unidade com folhas de membrana em espirais. Fonte: DIAS <i>et al.</i> (2020).	181
6.1	Comparação dos perfis de temperatura dos dados experimentais e do simulado (KRISHNAMURTHY e TAYLOR, 1985).	188
6.2	Perfis de temperatura dos modelos: completo e resíduo ponderado pelos momentos no estudo de caso 2.	200
6.3	Valores exatos e medidos (corrompidos) - F no estudo de caso 2.	202
6.4	Valor exatos e medidos (corrompidos) - D no estudo de caso 2.	203

6.5	Valores exatos e medidos (corrompidos) - z_1 no estudo de caso 2. . . .	203
6.6	Valores exatos e medidos (corrompidos) - z_2 no estudo de caso 2. . . .	204
6.7	Valores exatos e medidos (corrompidos) - B no estudo de caso 2. . . .	204
6.8	Valores exatos e medidos (corrompidos) - $x_{D,1}$ no estudo de caso 2. . .	205
6.9	Valores exatos e medidos (corrompidos) - $x_{D,2}$ no estudo de caso 2. . .	205
6.10	Valores exatos e medidos (corrompidos) - $x_{B,1}$ no estudo de caso 2. . .	206
6.11	Valores exatos e medidos (corrompidos) - $x_{B,2}$ no estudo de caso 2. . .	206
6.12	Valores exatos e medidos (corrompidos) - T_c no estudo de caso 2. . . .	207
6.13	Valores exatos e medidos (corrompidos) - T_f no estudo de caso 2. . .	207
6.14	Valores exatos e medidos (corrompidos) - T_r no estudo de caso 2. . .	208
6.15	Perfil de temperatura e pressão reduzida dos componentes na coluna no estudo de caso 1.	213
6.16	Comparação entre os critérios TER de avaliação (Nelder-Mead). . . .	222
6.17	Progressão da otimização com o método Nelder-Mead utilizando o estimador NC.	222
6.18	Comparação entre os critérios TER de avaliação (PSO) no estudo de caso 1.	228
6.19	Comparação entre os critérios TER de avaliação (BFGS) no estudo de caso 1.	229
6.20	Comparação entre os critérios TER de avaliação (SQP) no estudo de caso 1.	234
6.21	Comparação entre os critérios TER de avaliação (Ponto Interior) no estudo de caso 1.	238
6.22	Comparação entre os otimizadores com os critérios Fobj e tempo de simulação (MQP).	241
6.23	Comparação entre os otimizadores com os critérios Fobj e tempo de simulação (NC).	242
6.24	Comparação entre os otimizadores com os critérios Fobj e tempo de simulação (Welsch).	243
6.25	Comparação entre os otimizadores com o critério TER (MQP).	243

6.26	Comparação entre os otimizadores com o critério TER (NC).	244
6.27	Comparação entre os otimizadores com o critério TER (Welsch).	244
6.28	Comparação entre os critérios TER no estudo de caso 2.	246
6.29	Perfil de temperatura com modelo completo: dados exatos, medidos e reconciliados no estudo de caso 2.	247
6.30	Perfil de temperatura com modelo reduzido: dados exatos, medidos e reconciliados no estudo de caso 2.	248
6.31	Comparação entre os valores de Função objetivo e número de iterações no estudo de caso 2.	249
6.32	Comparação entre os tempos totais de otimização e os tempos médios de cada iteração no estudo de caso 2.	249
6.33	Resultados - Vazão de alimentação F no estudo de caso 2 parte II.	251
6.34	Resultados - Vazão de destilado D no estudo de caso 2 parte II.	251
6.35	Resultados - Fração molar de propeno na alimentação z_1 no estudo de caso 2 parte II.	252
6.36	Resultados - Fração molar de propano na alimentação z_2 no estudo de caso 2 parte II.	252
6.37	Resultados - Vazão do fundo B no estudo de caso 2 parte II.	253
6.38	Resultados - Fração molar de propeno no destilado $x_{D,1}$ no estudo de caso 2 parte II.	253
6.39	Resultados - Fração molar de propano no destilado $x_{D,2}$ no estudo de caso 2 parte II.	254
6.40	Resultados - Fração molar de propeno no fundo $x_{B,1}$ no estudo de caso 2 parte II.	254
6.41	Resultados - Fração molar de propano no fundo $x_{B,2}$ no estudo de caso 2 parte II.	255
6.42	Resultados - Temperatura do condensador T_c no estudo de caso 2 parte II.	255
6.43	Resultados - Temperatura do prato de alimentação T_f no estudo de caso 2 parte II.	256

6.44 Resultados - Temperatura do referedor T_r no estudo de caso 2 parte	
II.	256

Lista de Tabelas

2.1	Exemplos de estimadores-M robustos aplicados na literatura de RD - parte 1/3	35
2.1	Exemplos de estimadores-M robustos aplicados na literatura de RD - parte 2/3	36
2.1	Exemplos de estimadores-M robustos aplicados na literatura de RD - parte 3/3	37
2.2	Resumo em números dos tipos de sistemas em problemas de RD robusta.	38
2.3	Dados do exemplo motivacional - bomba centrífuga	60
3.1	Alguns estimadores-M: $\rho(\xi)$	95
3.2	Alguns estimadores-M: $\psi(\xi)$	96
3.3	Alguns estimadores-M: $\psi'(\xi)$	96
3.4	Valores das constantes de sintonia para 95% de eficiência relativa em relação à distribuição Normal e referências para os estimadores-M analisados	97
3.5	Conceitos de Entropia derivados das teorias de Shanonn e Boltzmann (PINEDA, 2006).	102
6.1	Configuração e especificações da coluna de recheio (KRISHNAMURTHY e TAYLOR, 1985).	185
6.2	Composição da alimentação F (KRISHNAMURTHY e TAYLOR, 1985).	186
6.3	Resultado da simulação (KRISHNAMURTHY e TAYLOR, 1985). . .	186
6.4	Perfil de temperatura da coluna (KRISHNAMURTHY e TAYLOR, 1985).	189

6.5	Variáveis corrompidas para simular erros grosseiros no estudo de caso 1.	190
6.6	Valores exatos para o estudo de caso 1.	191
6.7	Valores corrompidos para o estudo de caso 1.	192
6.8	Resumo da metodologia para o caso 1 KRISHNAMURTHY e TAYLOR (1985).	194
6.9	Configuração e especificações da coluna de destilação propeno-propano RIBEIRO (2011).	195
6.10	Tempo gasto em segundos para uma iteração.	196
6.11	Comparação entre os modelos completo e reduzido pelos momentos com 51 estágios no estudo de caso 2.	197
6.12	Valores medidos (erro grosseiro em negrito), valores exatos e a magnitude dos erros no estudo de caso 2.	198
6.13	Descrição e parâmetros utilizados no problema de RDD.	202
6.14	Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte I.	210
6.15	Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte II.	211
6.16	Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte III.	212
6.17	Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte IV.	213
6.18	Resultados com o otimizador Nelder-Mead - F no estudo de caso 1.	215
6.19	Resultados com o otimizador Nelder-Mead - D no estudo de caso 1.	216
6.20	Resultados com o otimizador Nelder-Mead - B no estudo de caso 1.	217
6.21	Resultados com o otimizador Nelder-Mead - Temperaturas no estudo de caso 1.	218
6.22	Avaliação de desempenho com o otimizador Nelder-Mead (TER).	221
6.23	Resultados com o otimizador PSO - F no estudo de caso 1.	223
6.24	Resultados com o otimizador PSO - D no estudo de caso 1.	224

6.25	Resultados com o otimizador PSO - B no estudo de caso 1.	225
6.26	Resultados com o otimizador PSO - Temperaturas no estudo de caso 1.	226
6.27	Avaliação de desempenho com o otimizador PSO (TER) no estudo de caso 1.	227
6.28	Resultados com o otimizador BFGS - F no estudo de caso 1.	229
6.29	Resultados com o otimizador BFGS - D no estudo de caso 1.	230
6.30	Resultados com o otimizador BFGS - B no estudo de caso 1.	231
6.31	Resultados com o otimizador BFGS - Temperaturas no estudo de caso 1.	232
6.32	Avaliação de desempenho com o otimizador BFGS (TER) no estudo de caso 1.	232
6.33	Resultados com o otimizador SQP - F no estudo de caso 1.	233
6.34	Resultados com o otimizador SQP - D no estudo de caso 1.	234
6.35	Resultados com o otimizador SQP - B no estudo de caso 1.	235
6.36	Resultados com o otimizador SQP - Temperaturas no estudo de caso 1.	236
6.37	Avaliação de desempenho com o otimizador SQP (TER) no estudo de caso 1.	236
6.38	Resultados com o otimizador Ponto Interior - F no estudo de caso 1.	237
6.39	Resultados com o otimizador Ponto Interior - D no estudo de caso 1.	238
6.40	Resultados com o otimizador Ponto Interior - B no estudo de caso 1.	239
6.41	Resultados com o otimizador Ponto Interior - Temperaturas no estudo de caso 1.	240
6.42	Avaliação de desempenho com o otimizador Ponto Interior (TER) no estudo de caso 1.	240
6.43	Resumo das avaliações de desempenho (TER total) no estudo de caso 1.	240
6.44	Resultados de tempo e número de iterações do caso SEFERLIS e HRYMAK (1994) - no estudo de caso 2 parte II.	250
7.1	Resumo dos objetivos e resultados.	265

Lista de Abreviaturas

BE	Balanço de Energia, p. 38
BM	Balanço de Massa, p. 38
BP	<i>Breakdown Point</i> , p. 95
CLAV	Classificação de Variáveis, p. 2
COV	Matriz de Covariâncias, p. 24
CSTR	<i>Continuous Stirred-Tank Reactor</i> (Reator Tanque Agitado Contínuo), p. 33
DEG	Detecção de Erros Grosseiros, p. 2
DMEG	Detecção de Múltiplos Erros Grosseiros, p. 2
EED	Estimação de Estados Dinâmica, p. 79
EKF	<i>Extended Kalman Filter</i> (Filtro de Kalman Estendido), p. 66
ELV	Equilíbrio Líquido-Vapor, p. 120
EMC	Estimador Máximo de Correntropia, p. 34
EOS	<i>Equation of State</i> (Equação de Estado), p. 126
EP	Estimação de Parâmetros, p. 6
FACP	Função de Autocorrelação Parcial, p. 162
FAC	Função de Autocorrelação, p. 162
FDP	Função Densidade de Probabilidade, p. 29
FI	Função de Influência, p. 89
Fobj	Função Objetivo, p. 13, 38

GT	<i>Global Test</i> (Teste Global), p. 30
HBM	<i>High Breakdown M-estimator</i> (Estimador de Alto Limite de Ruptura), p. 107
HM	Horizonte Móvel, p. 63
IMT	<i>Iterative Measurement Test</i> (Teste de Medida Iterativo), p. 30
KDE	<i>Kernel Density Estimate</i> , p. 43
KF	<i>Kalman Filter</i> (Filtro de Kalman), p. 62
MADn	<i>Normalized Median Absolute Deviation</i> (Desvio Absoluto da Mediana Normalizado), p. 26
MAP	<i>maximum a posteriori</i> , p. 73
MCC	<i>Maximization of Correntropy Criterion</i> (Critério Maximização da Correntropia), p. 104
MHE	<i>Moving Horizon Estimation</i> (Estimação de Horizonte Móvel), p. 79
MIMT	<i>Modified Iterative Measurement Test</i> (Teste de Medida Iterativo Modificado), p. 30
MPC	<i>Model Predictive Control</i> (Controle Preditivo Multivariável), p. 4
MP	<i>Maximum Power Test</i> (Teste da Potência Máxima), p. 30
MQP	Mínimos Quadrados Ponderados, p. 3, 38
MT	<i>Measurement Test</i> (Teste de Medida), p. 30
MWDR	<i>Moving Window Data Reconciliation</i> , p. 71
NC	Normal Contaminada, p. 84
NT	<i>Nodal Test</i> (Teste Nodal), p. 30
PCAT	<i>Principal Component Analysis Test</i> (Teste com Análise de Componentes Principais), p. 30
PCA	<i>Principal Component Analysis</i> (Análise de Componentes Principais), p. 30

PF	Filtro de Partículas, p. 33
PNL	Programação Não-Linear, p. 40
PR	Peng-Robinson, p. 126
PSO	<i>Particle Swarm Optimization</i> (Otimização Enxame de Partículas), p. 11
RDDR	Reconciliação de Dados Dinâmica Robusta, p. 69
RDD	Reconciliação de Dados Dinâmica, p. 62
RDEP	Reconciliação de Dados com Estimção de Parâmetros, p. 39
RDR	Reconciliação de Dados Robusta, p. 2
RD	Reconciliação de Dados, p. 2
RNA	Redes Neurais Artificiais, p. 32
RPG	Razão Probabilística Generalizada, p. 30
RTD	Retificação de Dados, p. 2
RTO	<i>Real Time Optimization</i> (Otimização em Tempo Real), p. 6
SDCD	Sistema Distribuído de Controle Digital, p. 4
SQP	<i>Sequential Quadratic Programming</i> (Programação Quadrática Sucessiva), p. 63
SRK	Soave-Redlich-Kwong, p. 126
SR	<i>Sum-rates method</i> (Método da Soma das Taxas), p. 130
SSDR	<i>Single Set Data Reconciliation</i> , p. 71
TER	<i>Total Error Reduction</i> (Redução Total dos Erros), p. 13
VMNR	Variáveis Medidas Não-Redundantes, p. 22
VMR	Variáveis Medidas Redundantes, p. 22
VNMNO	Variáveis Não-Medidas Não Observáveis, p. 22
VNMO	Variáveis Não-Medidas Observáveis, p. 22
asMHE	<i>Advanced Step Moving Horizon Estimation</i> (Estimção de Horizonte Móvel de Passo Avançado), p. 79

Capítulo 1

Introdução

“A dúvida é o princípio da sabedoria.”

Aristóteles (384 a.C. - 322 a.C.)

1.1 Contextualização

O avanço tecnológico e da computação vem impulsionando o amplo, fácil e rápido acesso aos dados de processos de uma planta industrial. Além disso, os dados são de extrema importância para o monitoramento e a otimização em tempo real de unidades de produção (NICHOLSON *et al.*, 2014). O monitoramento de uma planta, unidade ou equipamento industrial se torna cada vez mais necessário para melhorar a qualidade do produto, aumentar a segurança do processo, reduzir os gastos energéticos e os resíduos do processo; porém, é relevante que as informações sejam confiáveis e validadas junto ao processo, pois a confiabilidade dos dados é de suma importância para qualquer tomada de decisão sobre esse processo (CÂMARA *et al.*, 2017). Entretanto, medidas de processos estão sujeitas a erros e flutuações por conta da imprecisão intrínseca, degradação, mau funcionamento, instalação imprópria, má calibração e falha dos instrumentos de medição, além de erros humanos de operação e aferição, ou até mesmo erros relacionados à operação do processo, o que resulta em dados que não representam fidedignamente o processo. Assim, não se espera que os dados medidos obedeçam às leis de conservação (balanços de massa, momentum e energia) (PRATA *et al.*, 2010). O controlador do processo que recebe esses dados, se não forem tratados adequadamente, pode levar a planta a operar em um ponto inseguro e, dessa forma, tomadas de decisão baseada em dados pouco confiáveis

aumentam a ocorrência de acidentes industriais, queda da qualidade do produto e prejuízos financeiros (FARIAS, 2009). Portanto, um procedimento de Retificação de Dados (RTD) é essencial para tratar estatisticamente a informação contida nos dados e, conseqüentemente, dar uma margem de confiabilidade para o controle e monitoramento em tempo real.

O procedimento de RTD pode ser basicamente dividido em três etapas: Classificação de Variáveis (CLAV); Detecção de Erros Grosseiros (DEG) e Reconciliação de Dados (RD). Entre essas três etapas a RD é a etapa mais estudada da RTD.

A primeira etapa da RTD é a etapa de classificação das variáveis. Essa etapa determina se a informação disponível é suficiente para resolver o problema de RD e identificar os conjuntos de variáveis observáveis (variáveis medidas e não medidas, que podem ser estimadas por meio das demais variáveis medidas e pelas restrições do processo) e não observáveis (variáveis não medidas que não podem ser estimadas). Assim, é possível restringir a quantidade de dados a um número ideal de variáveis relevantes e satisfatórias para o modelo matemático do processo, diminuindo a dimensão do banco de dados do processo (PRATA, 2009).

A segunda etapa da RTD é a DEG, que é executada para identificar um tipo especial de erro (erro grosseiro), que não segue a distribuição estatística de erros admitida. Esses erros podem ser causados por má calibração dos instrumentos de medição, deterioração dos sensores, picos de energia, entre outros já exemplificados anteriormente. Entretanto, para obter uma estimativa acurada de parâmetros e variáveis, o efeito negativo dos erros grosseiros deve ser minimizado ou eliminado. Quando um ou mais erros grosseiros estão presentes no conjunto de dados analisado, é necessária uma estratégia para detectá-los e eliminá-los ou compensá-los, o que é conhecido por Detecção de Múltiplos Erros Grosseiros (DMEG). A literatura tem apontado a utilização de estimadores robustos para a eliminação do efeito negativo dos múltiplos erros grosseiros, de modo simultâneo com a RD, evitando o uso de procedimentos iterativos e computacionalmente intensivos (PRATA *et al.*, 2010). Esse procedimento é conhecido como Reconciliação de Dados Robusta (RDR) e permite reduzir a RTD a duas etapas: CLAV e RDR.

Na RD, terceira etapa da RTD, os dados medidos são ajustados de maneira estatisticamente consistente pelo estimador (função objetivo) resultante da formulação de máxima verossimilhança sobre a distribuição estatística dos erros de medição considerada, de forma a satisfazer às leis de conservação e demais restrições impostas ao sistema, maximizando a probabilidade de ocorrência daquela medida e, simultaneamente, respeitando o modelo matemático do processo. Obtêm-se assim

estimativas confiáveis para as variáveis e parâmetros do processo (DE MENEZES, 2015; PRATA *et al.*, 2009). Tradicionalmente, admite-se como válida a distribuição Normal, que resulta no estimador de Mínimos Quadrados Ponderados (MQP) (BENQLILOU, 2004). Assim, de modo geral as aplicações e o procedimento de RTD podem ser contextualizados e ilustrados conforme mostrado na Figura 1.1.

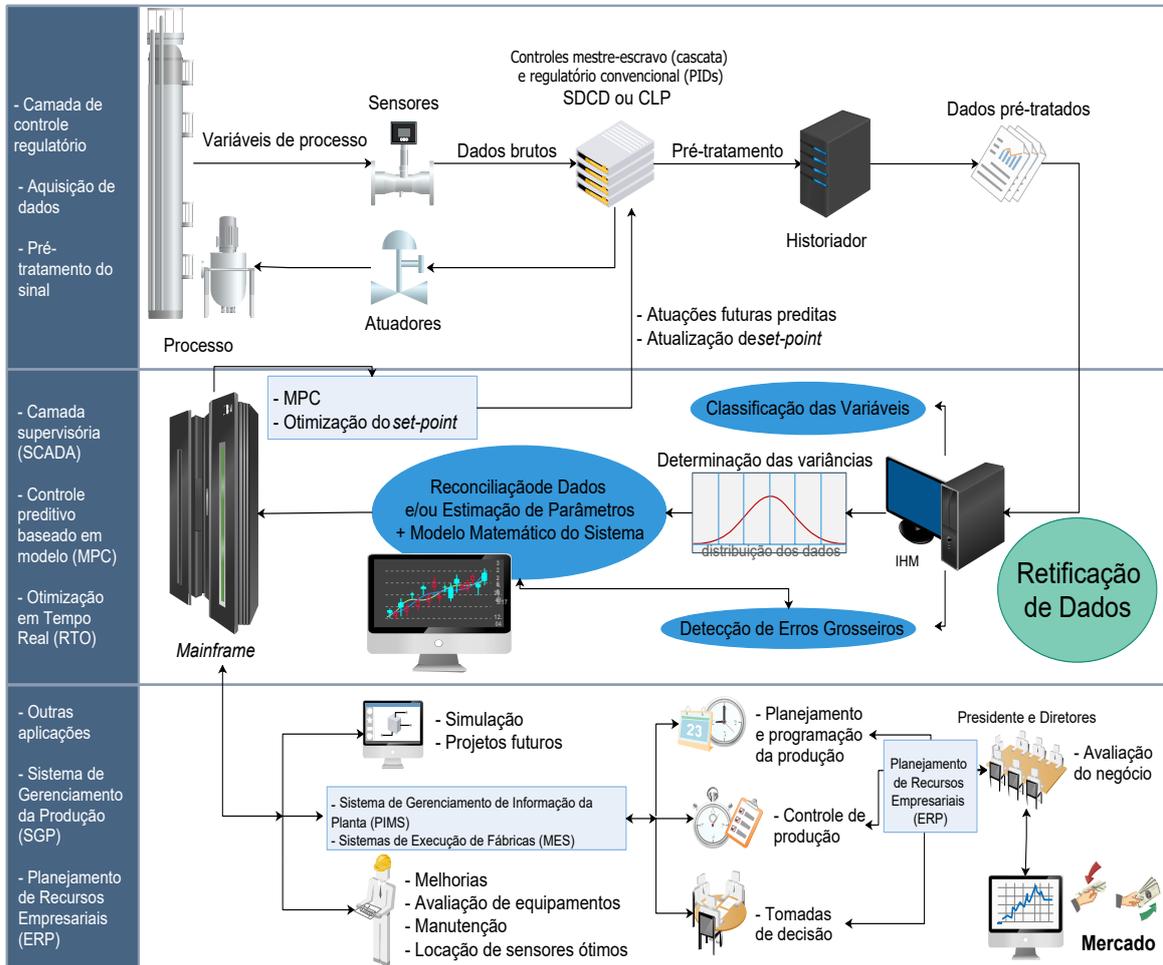


Figura 1.1: Retificação de dados e aplicações (FARIAS, 2009; NARASIMHAN e JORDACHE, 1999).

A maioria dos estudos de RD aplicados a processos químicos está relacionado ao controle do processo. Dessa forma, é comum ver casos que abrangem desde a medida simples do sinal de vazão, passando por problemas que envolvem centenas de variáveis medidas em um só equipamento, até milhares de variáveis medidas e não medidas que se inter-relacionam em um processo que contém múltiplos equipamentos, para dar ao controle de processo dados mais confiáveis. Entretanto, há escassez de estudos de casos que correlacionem a influência do tratamento dos dados de processo às camadas mais elevadas da pirâmide de gerenciamento de processo (Figura 1.2), o que em parte é devido às restrições

de confidencialidade praticadas pelas indústrias (PRATA, 2009). Conjunto de dados pouco confiáveis influencia diretamente a camada inferior da pirâmide, composta pelo Sistema Distribuído de Controle Digital (SDCD), onde o controle regulatório, lógica de proteção e de sequenciamento são executados. Uma camada mais acima faz as otimizações das trajetórias de referências *set-point*, que pode ser feito com um procedimento de Controle Preditivo baseado em Modelo (MPC). Os dados imprecisos influenciam indiretamente o MPC, que por sua vez afeta as análises conduzidas nas camadas superiores da pirâmide. Nas camadas superiores encontram-se a otimização em tempo real, a programação de produção, o planejamento da produção e o gerenciamento geral da empresa. Nas camadas mais elevadas dessa pirâmide, os efeitos dos erros aleatórios e grosseiros são amortecidos por heurísticas e bom senso do engenheiro na hora de fechar as contas; porém, os erros não são corrigidos estatisticamente pelos tratamentos, filtros e otimizações das camadas inferiores. Dessa forma, as camadas mais acima da pirâmide, que estão relacionadas com os processos de tomada de decisão que envolvem grande quantidade de capital, utilizam quase sempre dados de processo sem tratamento estatístico, que podem gerar grandes variações nos lucros corporativos. Portanto, mesmo que a etapa de RD corrija pequenos erros irrelevantes para as primeiras camadas da pirâmide, o tratamento estatístico desses erros pode fazer grande diferença nas tomadas de decisão no topo dessa pirâmide. Além disso, a RDR tem a função de proteger toda a pirâmide da presença dos erros grosseiros, ou de pequenas falhas. Dessa forma, é imprescindível o estudo e aplicação de técnicas de RDR em equipamentos que afetam o desempenho econômico da planta (NARASIMHAN e JORDACHE, 1999; ROMAGNOLI e SANCHEZ, 1999). As Figuras 1.2 e 1.3 ilustram a pirâmide de gerenciamento de processo e as otimizações conduzidas, classificando-as em relação às escalas de tempo e às variáveis de processo.

Indústrias de processos modernas enfrentam uma crescente concorrência no mercado e um aumento de leis ambientais rigorosas, o que exige cada vez mais operações automatizadas e eficientes das plantas, a fim de maximizar a rentabilidade e minimizar impactos ambientais (WANG *et al.*, 2017). Com isso, a partir da década de 1980 a tecnologia Otimização em Tempo Real (RTO) começou a ser estudada para aplicações em indústrias, com o objetivo de maximizar a eficiência energética e o desempenho econômico da planta (BAGAJEWICZ, 2000a). A forma clássica de um projeto de camada de RTO utiliza modelo fenomenológico estacionário para descrever o comportamento da planta, otimizando uma função objetivo econômica sujeita a este modelo (GRACIANO, 2016). QUELHAS *et al.* (2013) descrevem algumas limitações para a implementação prática de RTO: falta de medições de processo, ruído de medição, incompatibilidade estrutural planta/modelo, método

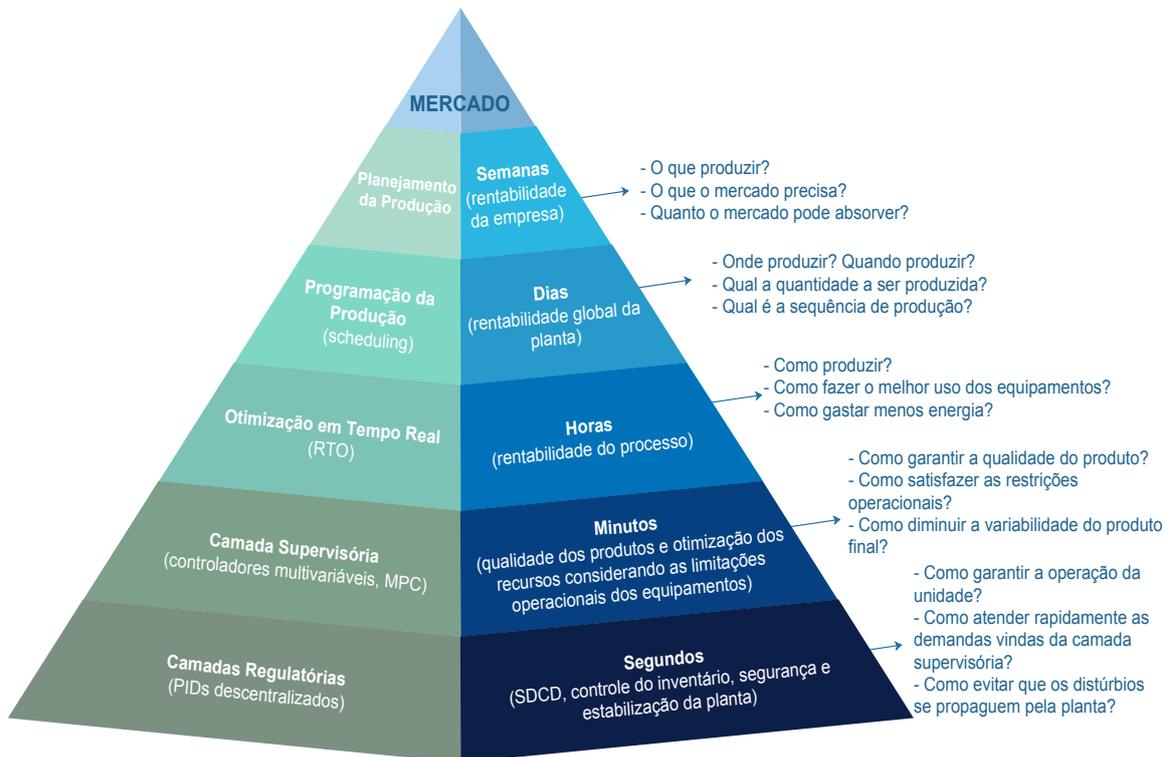


Figura 1.2: Pirâmide de gerenciamento de processo (FARIAS, 2009; ROMAGNOLI e SANCHEZ, 1999).

numérico de otimização e baixa frequência de atualização do *set-point*. A RTO é utilizada entre os níveis de camada supervisória e o de programação da produção, como ilustra a Figura 1.2.

O desempenho econômico da RTO está fortemente relacionado a dois fatores: diferença entre o ótimo verdadeiro do processo e o valor ótimo predito pelo modelo; e a variabilidade do processo (DE JESUS, 2011). O primeiro fator tem como ponto fraco os erros intrínsecos da modelagem matemática determinística e dos parâmetros do próprio modelo. O segundo fator está relacionado à natureza dinâmica do processo e a quanto as perturbações influenciam a dinâmica de outras variáveis do mesmo processo e do próprio ciclo de RTO.

Assim, o procedimento de RD pode ser muito relevante para suavizar os possíveis ruídos, os erros aleatórios e os erros grosseiros do processo, tornando os dados mais confiáveis para a comparação com os dados preditos pelo modelo e o procedimento de estimação dos parâmetros do modelo, conhecido como Estimação de Parâmetros (EP), usado para ajustar o modelo aos dados previamente tratados, tornando o modelo mais fidedigno (PRATA *et al.*, 2009).

Além disso, é impraticável medir todas as variáveis de estados em um processo; contudo, com o procedimento de RD é possível estimar as variáveis não medidas, desde que as outras variáveis medidas satisfaçam o número mínimo de variáveis

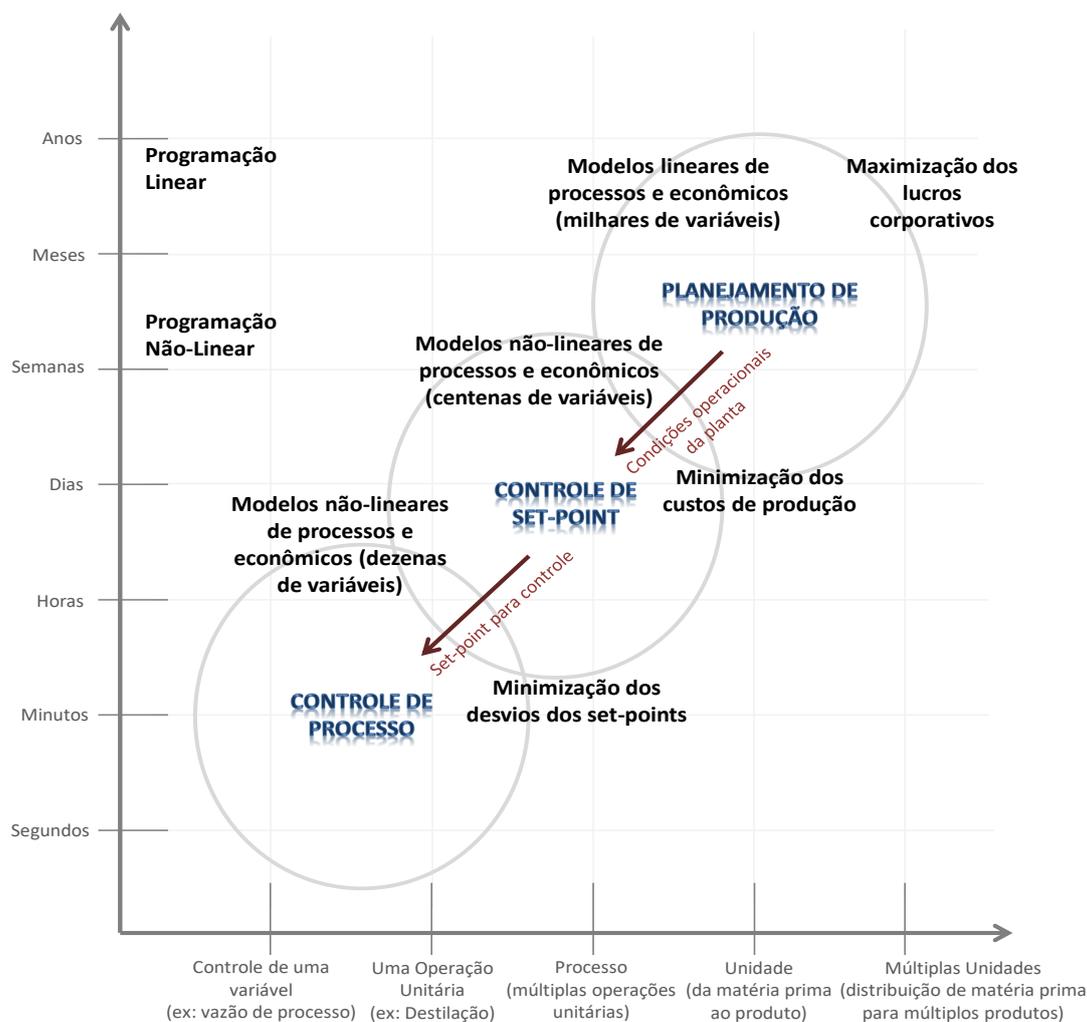


Figura 1.3: Estrutura geralmente utilizada na otimizações de processo em relação às escalas de tempo e variáveis (CHEN, 1998; DE JESUS, 2011).

observáveis do modelo matemático. Por fim, é necessário estudar os algoritmos de otimização aplicados em cada etapa do ciclo da RTO, para averiguar o algoritmo mais adequado (que garanta o ponto ótimo dentro do tempo do ciclo) para cada etapa da RTO (QUELHAS *et al.*, 2013). A Figura 1.4 representa o ciclo da RTO e suas principais etapas.

Conforme a contextualização apresentada previamente, conclui-se que o procedimento de RDR pode ter suma importância em todos os níveis do fluxograma de gerenciamento de processo, constituindo uma ferramenta de segurança que trata estatisticamente das variáveis medidas, tornando mais confiáveis os procedimentos de RTO, de planejamento da produção e da avaliação econômica do negócio.

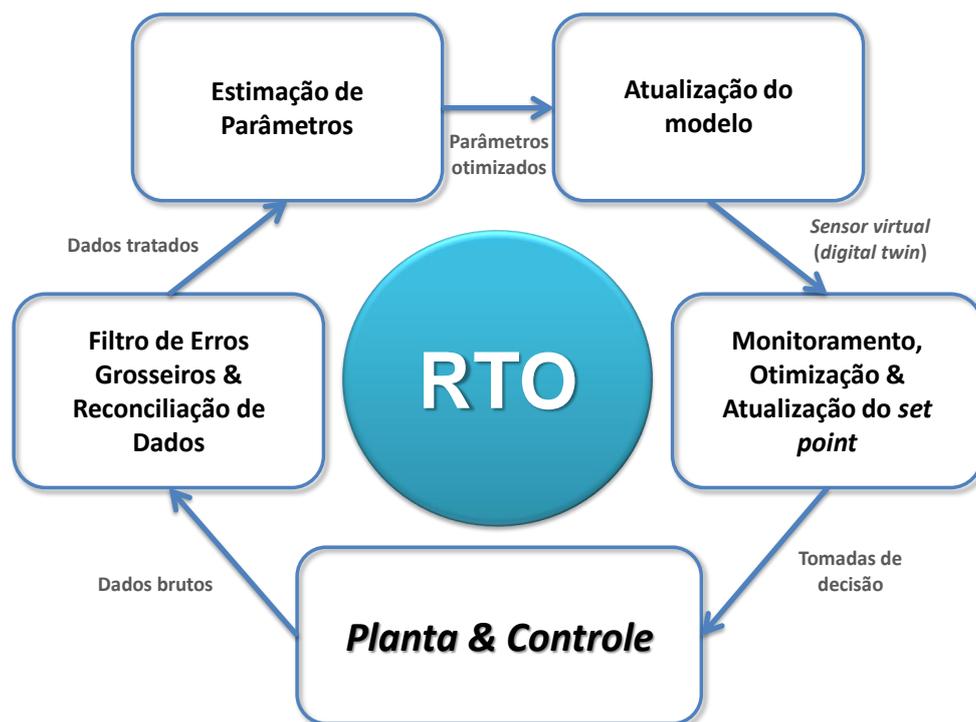


Figura 1.4: Ciclo da Otimização em Tempo Real (BAGAJEWICZ e ROLLINS, 2002).

1.2 Motivação

Um dos processos de maior interesse da indústria química e petroquímica é a separação, sendo a destilação um dos mais comuns. Essa é uma antiga operação unitária, que encontra como principais aplicações o fracionamento do petróleo, a obtenção de álcoois, a extração de essências e purificação de substâncias (KISTER, 1992).

As colunas de destilação em geral são os equipamentos de separação mais empregados na indústria química, petroquímica e de biocombustíveis. O alto custo energético torna o processo muito importante, pois em alguns casos podem concentrar mais de 70% do custo energético total de uma indústria de transformação. Conforme SZKLO e ULLER (2008), o consumo energético em termos absolutos concentra-se em poucos processos que processam grande quantidade de carga, como as unidades de destilação atmosférica e a vácuo, totalizando 35-40% do consumo de energia de uma refinaria. Pois, em geral, processos energo-intensivos (processos que consomem mais energia por barril) não são os que consomem mais energia em termos absolutos. Por outro lado, unidades de separação *topping* recebem todos os barris de petróleo que entram numa refinaria, tornando esses processos os campeões de consumo energético em termos absolutos. Além disso, a destilação é geralmente

a etapa do processo que impede o aumento da produção, constituindo o gargalo de operação (DE ALBUQUERQUE, 2008).

No Brasil em 2019, o consumo de energia por derivados de petróleo representou 39,4% do consumo final, sendo que o petróleo e derivados representaram 34,4% da energia ofertada (EPE, 2020a,b). Segundo ENERGY.GOV (2016), os processos de separação por destilação nos EUA consomem cerca de 2 a 3 Quads (1Quad = 10^{15} BTU) de energia por ano, ao mesmo tempo que representam de 40% a 70% do investimento e custo operacional de uma típica planta de processamento. Conforme LLNL (2020), foi estimado um consumo total de 100,2 Quads de energia em 2019 nos EUA, sendo que a fonte de energia oriunda do petróleo contribuiu com 36,7 Quads e a do gás natural com 32,1 Quads, ou seja, petróleo e gás natural contribuíram com 68,7% da energia produzida nos EUA em 2019. Assim, implica dizer que colunas de destilação consomem de 2% a 3% de toda a energia produzida no ano de 2019 nos EUA. HUMPHREY e SIEBERT (1992) cita que a destilação foi responsável por 11% do gasto de energia em indústrias nos EUA em 1991. Outros dados relevantes também foram apresentados por ENERGY.GOV (2016), que segundo o Departamento de Energia dos EUA: colunas de destilação representam nos EUA 90% a 95% dos equipamentos de separação em plantas de processamento; cerca de 40.000 colunas de destilação em operação nos EUA consomem o equivalente a aproximadamente 1,2 milhões de barris de petróleo por dia; somente a destilação de petróleo bruto nas refinarias dos EUA consomem cerca de 0,4 milhões de barris de petróleo por dia; as recentes reservas de gás de xisto dos EUA esperam produzir ao longo dos anos cerca de 43-108 bilhões de barris de gás natural líquido, exigindo cerca de 3000-8000 quatrilhões de BTU (ou 3000-8000 Quads) de energia somente para o processo de separação.

Um problema comum na produção de petróleo é o excesso de gás CO_2 presente nas correntes de gás natural. O primeiro e notório problema seria a emissão desse gás para o ambiente. Porém, além dos possíveis problemas ambientais, o excesso de CO_2 nessas correntes pode causar problemas na planta de processo, como o congelamento por queda de pressão em seções de compressões e resfriamento da planta e a corrosão de dutos metálicos (MOKHATAB *et al.*, 2018). De acordo com a resolução da ANP (Agência Nacional de Petróleo, Gás Natural e Biocombustíveis), o gás natural comercial deve conter no máximo 3% (mol) de CO_2 (ANP, 2008). Portanto, uma solução possível para lidar com o CO_2 produzido é a reinjeção do CO_2 no poço de petróleo, permitindo o aumento da produtividade do poço. Isso pode certamente minimizar os impactos ambientais e problemas na planta de processamento de gás natural. Portanto, a etapa de separação do CO_2 no tratamento de gás natural é uma operação industrial indispensável na produção de petróleo.

Alguns processos físicos/químicos são utilizados para a separação do CO_2 do gás natural, como a destilação criogênica, absorção ou osmose reversa por membranas (MOKHATAB *et al.*, 2018; SPEIGHT, 2018). Remoções de CO_2 do gás natural por processo de membranas são utilizadas desde 1981 (HENIS e TRIPODI, 1981). No entanto, as aplicações foram inicialmente limitadas devido aos riscos econômicos intrínsecos associados à atividade de produção de petróleo e às restrições operacionais relacionadas às separações por membrana. Entretanto, o cenário vem mudando devido às vantagens relacionadas ao menor consumo de energia, baixo investimento de capital, baixos custos operacionais e à natureza mais compacta desses equipamentos (AL-OBAIDI *et al.*, 2018; KOVVALI *et al.*, 1994; SINGH *et al.*, 2013).

Dada a crescente importância industrial dos processos de separação por membrana, as demandas para o desenvolvimento de modelagem matemática, simulação, otimização, controle, tratamento de dados estatísticos e procedimentos de monitoramento online também aumentaram, pois essas técnicas são fundamentais para o projeto e monitoramento de diversos processos químicos. Como consequência, o desempenho do processo analisado pode ser avaliado de forma mais precisa e monitorada, permitindo a detecção de falhas em linha e em tempo real. Com base nessas tecnologias, os riscos e o tempo necessário para a tomada de decisão podem ser minimizados (DE MENEZES *et al.*, 2020).

Por estes fatos, melhorias no projeto, no arranjo configuracional de processos de separação, no controle de processo, no monitoramento e otimização desses processos exercem grandes impactos na redução de gasto energético e econômico da planta. Assim, é seguro afirmar que os processos de purificação, em especial as colunas de destilação e membranas, apresentam forte correlação com o desempenho econômico da planta. Portanto, conclui-se que estudo de RDR e EP para aplicação de RTO nesses processos de separação podem influenciar positivamente todo o processo industrial.

Inúmeras vantagens do procedimento de RD vêm sendo discutidas desde a década de 1960 (KUEHN e DAVIDSON, 1961) e ainda são discutidas no presente momento (COIMBRA *et al.*, 2017; DE MENEZES, 2015; DE SOUZA, 2015; DO VALLE *et al.*, 2018; SARRUF, 2014; VALLURU *et al.*, 2016). O procedimento de RD e DMEG com estimadores robustos (ou seja, a RDR) ainda não é amplamente reportado na literatura, principalmente no que diz respeito a colunas de destilação. Esses procedimentos são capazes de avaliar a qualidade dos dados medidos e fazer inferências sobre as variáveis não medidas e parâmetros em tempo real, como reporta NICHOLSON *et al.* (2014).

Segundo CORDÃO (2005), o sistema de RD reduziu o índice de perdas de uma refinaria de 1,09%, no período compreendido entre fevereiro de 2001 e janeiro de 2002, anterior à implantação, para 0,34% no período compreendido entre julho e dezembro de 2002, logo após a implantação. A redução no índice equivaleu a uma redução de R\$ 1.609.295,15 por mês nas perdas da refinaria. Esses resultados mostram a importância que este tipo de sistema pode ter para as indústrias de processo.

Portanto, o uso da RD é indubitavelmente de grande interesse prático, uma vez que melhora a confiança e a consistência dos dados, a qualidade das informações, prevê estimativas coerentes e confiáveis para as variáveis não medidas, poupa análises de laboratório, diminui a frequência de calibração dos sensores e representa de maneira mais fidedigna o processo, permitindo melhor monitoramento, controle e otimização em tempo real. A Figura 1.5 ilustra simplificada as muitas vantagens da RD aplicada a equipamentos de processos químicos.

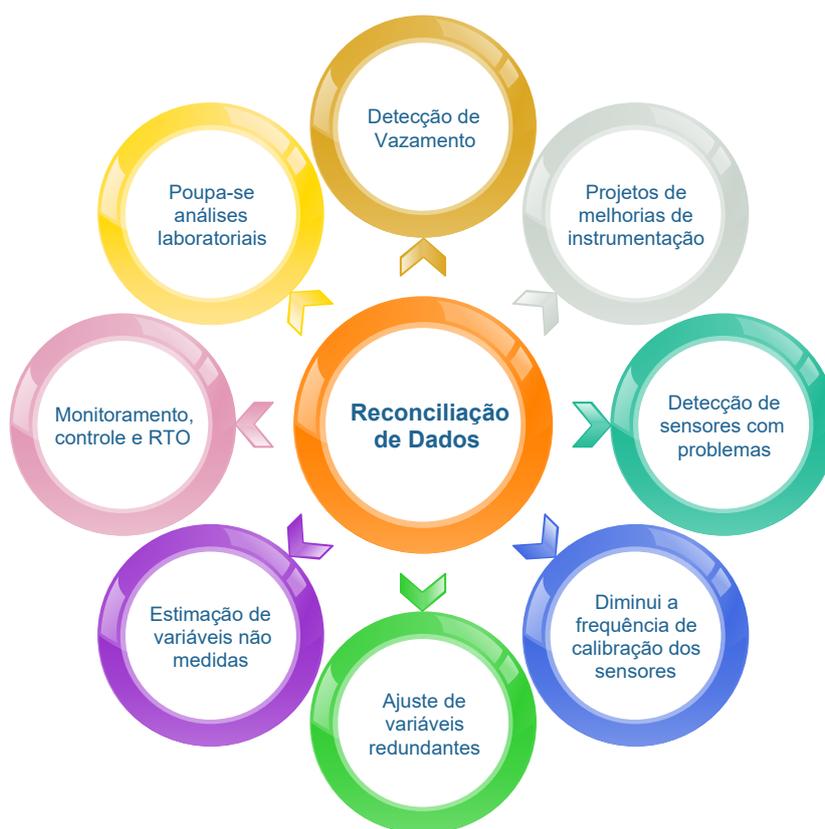


Figura 1.5: Vantagens da Reconciliação de Dados (BENQLILOU, 2004; FARIAS, 2009).

Devido à notória importância e aplicação da RD e DMEG, segundo BAGAJEWICZ e ROLLINS (2002) muitos *softwares* foram desenvolvidos para aplicação em diversos processos, tais como o Aspen Operations Reconciliation and Accounting (ASPENTECH, 2021), o ROMeo Material Balance (AVEVA, 2021), o

IOO (*Interactive On-Line Optimization*) (LSU, 2021), o Massbal (AUREL, 2021), o Production Balance (HONEYWELL, 2021), o RECON (CHEMPLANT, 2021), o Sigmafine (SIGMAFINE, 2021) e o VALI (BELSIM, 2021). A essa lista, pode-se ainda acrescentar os *softwares* nacionais: ESTIMA (NORONHA *et al.*, 1993), ESTIMA-R (PRATA, 2009) e EMSO (SOARES e SECCHI, 2003).

Cabe ressaltar que a maioria destes *softwares* é de utilização comercial e que apenas os *softwares* IOO, ESTIMA-R e EMSO contêm estimadores robustos (definido no Capítulo 3) e são de utilização gratuita para fins acadêmicos. O *software* IOO, voltado para a área de processamento de minérios, apresenta os estimadores Normal Contaminada e Lorenziana. O *software* Estima-R, voltado para processo de polimerização, utiliza o estimador de Welsch, escolhido com base nos resultados apresentados por PRATA *et al.* (2008b). O *software* EMSO, voltado para engenharia de processos, apresenta uma quantidade maior de estimadores robustos: Normal Contaminada, Fair, Cauchy, Logística, Lorenziana e Hampel, com base no trabalho de ÖZYURT e PIKE (2004).

Para resolver o problema resultante da formulação do problema de RD, esses *softwares* utilizam métodos de otimização. O IOO utiliza o pacote de soluções de problemas de otimização GAMS (*General Algebraic Modeling System*). O ESTIMA-R utiliza o método não determinístico enxame de partículas, (*Particle Swarm Optimization*, PSO). O EMSO contém alguns métodos de otimização, como o r-Complex, IPOPT (*Interior-Point Optimization*), PSO, Direct, ARS (*Adaptive Random Search*) e o OPT++ (pacote de algoritmos de otimização não linear escritos na linguagem de programação C++) (DO VALLE *et al.*, 2008).

O *software* EMSO apresenta maior generalidade para aplicações, interface com o *software* Matlab[®] e Scilab, e contém muitos modelos de processos, que inclusive podem ser modificados. Nesse *software*, o problema de RD e DMEG pode ser resolvido nos pacotes “*reconciliation*” e “*optimization*”. Outro *software* nacional capaz de realizar o procedimento de RD e DMEG é o iiSE (VRTECH, 2017), auxiliado pela interface entre o *software* gratuito Scilab, que contém pacotes de otimizadores. Outros *softwares* também fazem interface com o iiSE, como Microsoft[®] Excel, Matlab[®], Java e Python.

Atualmente, aplicativos em ambiente *web* e dados em nuvens vem se tornando uma tendência. Dessa forma, uma plataforma que vem sendo bastante útil para aplicações de monitoramento e gerenciamento de falhas é o Dash (desenvolvido em Python ou R). Essa ferramenta permite criar bonitos e amigáveis gráficos de análise de dados em aplicativo-*web*, sem precisar de JavaScript. A plataforma também permite que a equipe de ciência de dados foque mais tempo nos dados

e modelos, enquanto produzem de forma fácil aplicativos *dashboard* com Python ou R para aplicações on-line e em tempo real. Dessa forma, aplicativo-*web* podem ser feitos apenas com o Dash (Python), o que normalmente exigiria uma equipe de desenvolvedores de *back-end*, *front-end* e TI (Tecnologia da Informação).

Muitos fatores motivaram a escolha dos *softwares* EMSO, Matlab[®] e Python:

- **EMSO**: nacionalidade, gratuidade para fins acadêmicos, comunicação com servidores de dados industriais, muitos modelos matemáticos de interesse da engenharia química (inclusive editáveis), otimizadores confiáveis, capacidade de realizar RD e DMEG com estimadores robustos e disponibilidade de pacotes termodinâmicos (VRTherm);
- **Matlab[®]**: suporte, comunicação com servidores de dados industriais, diversas bibliotecas de análises de dados, de tratamento estatístico e otimizadores confiáveis;
- **Python**: gratuidade, comunicação com servidores de dados industriais, diversas bibliotecas de análises de dados, de tratamento estatístico e otimizadores confiáveis.

1.3 Objetivo

Esta tese apresenta 4 estudos de caso e uma revisão, com o objetivo de desenvolver uma metodologia para solucionar de maneira robusta o problema de reconciliação de dados dinâmica ou estacionária, em processos de separação para monitoramento em tempo real. O problema de RD foi estudado em dois regimes: estacionário (membrana) e dinâmico (destilação). Os dados utilizados no problema de RD em destilação foram gerados artificialmente, com o modelo validado com dados experimentais e outras simulações. Entretanto, para o problema de RD em separação por membrana os dados utilizados são reais e em tempo real. Portanto, os estudos podem ser descritos de forma mais detalhada em quatro casos bases, descritos a seguir.

O primeiro caso avalia os métodos de otimização mais adequados para solucionar o problema de RD e DMEG em coluna de destilação (modelo dinâmico) em regime estacionário, baseado em critérios de desempenho do método de otimização, como velocidade e número de iterações. Esse estudo também visa a averiguar o estimador robusto que melhor atende aos critérios de avaliação de desempenho da RD e DMEG, como por exemplo o valor de Função Objetivo (Fobj)

e o critério TER (*Total Error Reduction*), definido no Capítulo 3. Pois, há uma nova tendência em estudo relacionada aos estimadores-M robustos. A literatura mostra a eficiência desses estimadores em problemas de RD e DMEG, o que mostra a necessidade da análise de desempenho em problemas de RD em colunas de destilação.

O segundo caso estuda a aplicação de modelos reduzidos de coluna de destilação em problemas de RD e DMEG. Há uma grande preocupação com a redução do tamanho de modelos de processo de grandes dimensões, que em geral são bastantes custosos computacionalmente (RIBEIRO *et al.*, 2015). Os modelos mais rigorosos de destilação se enquadram perfeitamente nesse contexto. Entretanto, apesar da busca exaustiva da literatura técnico-científica, não foi encontrado qualquer relato sobre a implementação de modelos reduzidos de coluna em problemas de RD e DMEG em regime dinâmico. Portanto, para uma efetiva aplicação do monitoramento em tempo real é desejável que o modelo apresente um desempenho rápido e acurado.

O terceiro caso estuda a aplicação de RD e DMEG com dados dinâmicos em processo de destilação, utilizando o estimador Normal Contaminada (quasi-robusto). Para sistemas dinâmicos, utilizar a similaridade dos dados no tempo aumenta o grau de liberdade do sistema de RD, dando ainda mais redundância estatística, suavização a ruídos e aumenta a proteção dos dados aos erros grosseiros. Entretanto, não foi encontrado referência de trabalho em sistema de separação, que utiliza modelos de séries temporais em problemas de RD. Dessa forma, a metodologia desenvolvida para o caso inclui modelos de séries temporais, como modelo Autoregressivo de primeira ordem, buscando intensificar a similaridade temporal nos dados das variáveis de entrada do processo. Além disso, estratégias e métricas foram desenvolvida para acelerar o processo de otimização, visando uma efetiva aplicação do monitoramento em tempo real.

O último caso estuda o monitoramento de um processo de separação de CO_2 por membrana *on-line* e em tempo real. O monitoramento é realizado através de aplicativo-*web* desenvolvido em linguagem Python na plataforma Dash, onde os dados podem ser monitorados em *dashboard* iterativos e visualizados em qualquer navegador-*web*. A metodologia desenvolvida para o monitoramento em tempo real consiste em: aquisição de dados, pré-tratamento, RD, detecção de erros grosseiros e monitoramento, ou seja, sensores virtuais (*digital twin*) através da RD e balanço de energia. Uma etapa de estudo *off-line* foi realizada visando o pré-tratamento e caracterização dos dados. A etapa de caracterização fundamenta as escolhas das metodologias aplicadas e desenvolvidas para o caso real.

Para complementar, uma revisão detalhada sobre estimadores-M robusto foi

realizada, pois a literatura tem recomendado o uso de estimadores-M robustos para lidar com problemas de erros grosseiros em RD. Nesse contexto, a revisão apresenta uma pesquisa abrangente sobre estimadores-M robustos, incluindo aplicações em distintos campos do conhecimento e as características intrínsecas dos estimadores, como a função original do estimador, a função de influência, os parâmetros de ajuste, a aplicação original e respectivas referências. Portanto, 48 estimadores-M robustos são apresentados, ilustrados e sintonizados. Além disso, a revisão apresenta novos parâmetros de sintonia, para todos os estimadores-M analisados, nos níveis de eficiência relativa de 90, 95, 98 e 99%, em relação à distribuição normal.

1.4 Estrutura da Tese

Esta tese está estruturada em 7 capítulos e mais 2 apêndices. No **Capítulo 2** é realizada a revisão bibliográfica que aborda os principais fundamentos da pesquisa desenvolvida como: classificação de variáveis; estimação das variâncias e dos desvios-padrão; detecção de erros grosseiros; fundamentos matemáticos relacionado à reconciliação de dados, levando em conta a natureza do processo (estacionário ou dinâmico) e a linearidade do modelo (linear ou não linear); detecção de múltiplos erros grosseiros e formulações matemáticas dos problemas de RD, RD dinâmica e RD robusta.

O **Capítulo 3** apresenta uma revisão simplificada da estatística robusta, em particular dos estimadores-M robustos, com as formulações de máxima verossimilhança das principais funções de probabilidade. Uma descrição da origem do estimador Correntropia é apresentada, assim como suas propriedades e compatibilidade com os estimadores-M robustos. Uma revisão é também feita sobre os principais critérios de avaliação de desempenho do procedimento de RD e DMEG. São apresentados, finalmente os gráficos e as equações de alguns estimadores-M robustos relatados na literatura. Uma revisão mais detalhada sobre estimadores-M robusto está descrito no Apêndice B.

O **Capítulo 4** é dedicado à revisão de modelos de coluna de destilação, da termodinâmica de equilíbrio de fases e do uso de modelos reduzidos de coluna de destilação para fins de simulação. O **Capítulo 5** apresenta a metodologia desenvolvida e implementada para solucionar o problema de reconciliação de dados dinâmica robusta em coluna de destilação. Uma breve descrição do problema de separação de CO_2 por membrana também é descrito neste capítulo. O Apêndice A apresenta o trabalho publicado com os resultados obtidos no monitoramento em tempo real do processo de separação por membrana. No **Capítulo 6** são

apresentados e comentados os resultados obtidos. Finalmente, no **Capítulo 7** são apresentados a conclusão da pesquisa e futuros trabalhos.

Apêndice A apresenta o artigo publicado na revista *Processes*, referente ao estudo sobre o monitoramento em tempo real da separação do CO_2 por membrana, baseado em procedimentos de RD e métricas estatísticas de detecção de erros grosseiros. **Apêndice B** apresenta o trabalho submetido para a revista *Computers & Chemical Engineering*, referente a pesquisa de revisão sobre os estimadores-M da estatística robusta, apresentando 48 estimadores-M robustos e suas formulações funcionais, gráficas e sintonias de 90, 95, 98 e 99% de Eficiência Relativa.

Este trabalho foi desenvolvido no Laboratório de Modelagem, Simulação e Controle de Processos (LMSCP) do Programa de Engenharia Química (PEQ) do Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE) da Universidade Federal do Rio de Janeiro (UFRJ) e no Núcleo de Estudos de Otimização (NEO) do Departamento de Engenharia Química (TEQ) da Universidade Federal Fluminense (UFF).

Capítulo 2

Retificação de Dados

“Se as leis da Matemática referem-se à realidade, elas não estão corretas; e, se estiverem corretas, não se referem à realidade. Mas há outra razão que explica a elevada reputação da Matemática, é que ela leva as ciências naturais exatas a uma faixa de segurança que, sem a Matemática, essas ciências não poderiam obtê-la.”

Albert Einstein (1879 – 1955)

Este capítulo tem objetivo principal de reportar os principais conceitos, estudos, fundamentos e autores da área de reconciliação de dados. Entretanto, é necessário esclarecer que o procedimento de reconciliação de dados está inserido dentro de um conceito mais amplo e que interliga importantes procedimentos, conhecido como **retificação de dados**. Assim, uma breve revisão dos principais conceitos e autores no campo da retificação de dados é indispensável.

Conforme já foi discutido, a Retificação de Dados (RTD) é constituída por três etapas: a Classificação de Variáveis (CLAV), a Detecção de Erros Grosseiros (DEG) e a Reconciliação de Dados (RD) (podendo ter Estimação de Parâmetros (EP) ou não). Segundo EDGAR *et al.* (2001), é comum observar na literatura a existência de certo conflito entre as palavras **retificação** e **reconciliação**, pois muitas vezes são usadas como sinônimos; entretanto, essa consideração não é correta, visto que a retificação de dados deve ser avaliada como um conceito mais amplo. A Figura 2.1 resume a abrangência do procedimento de retificação de dados e suas principais etapas, ilustrando também o sequenciamento e interligação de cada etapa.

Como ilustra a Figura 2.1, a qualidade da informação contida nos dados coletados diretamente da planta influencia todo o procedimento de RTD. É

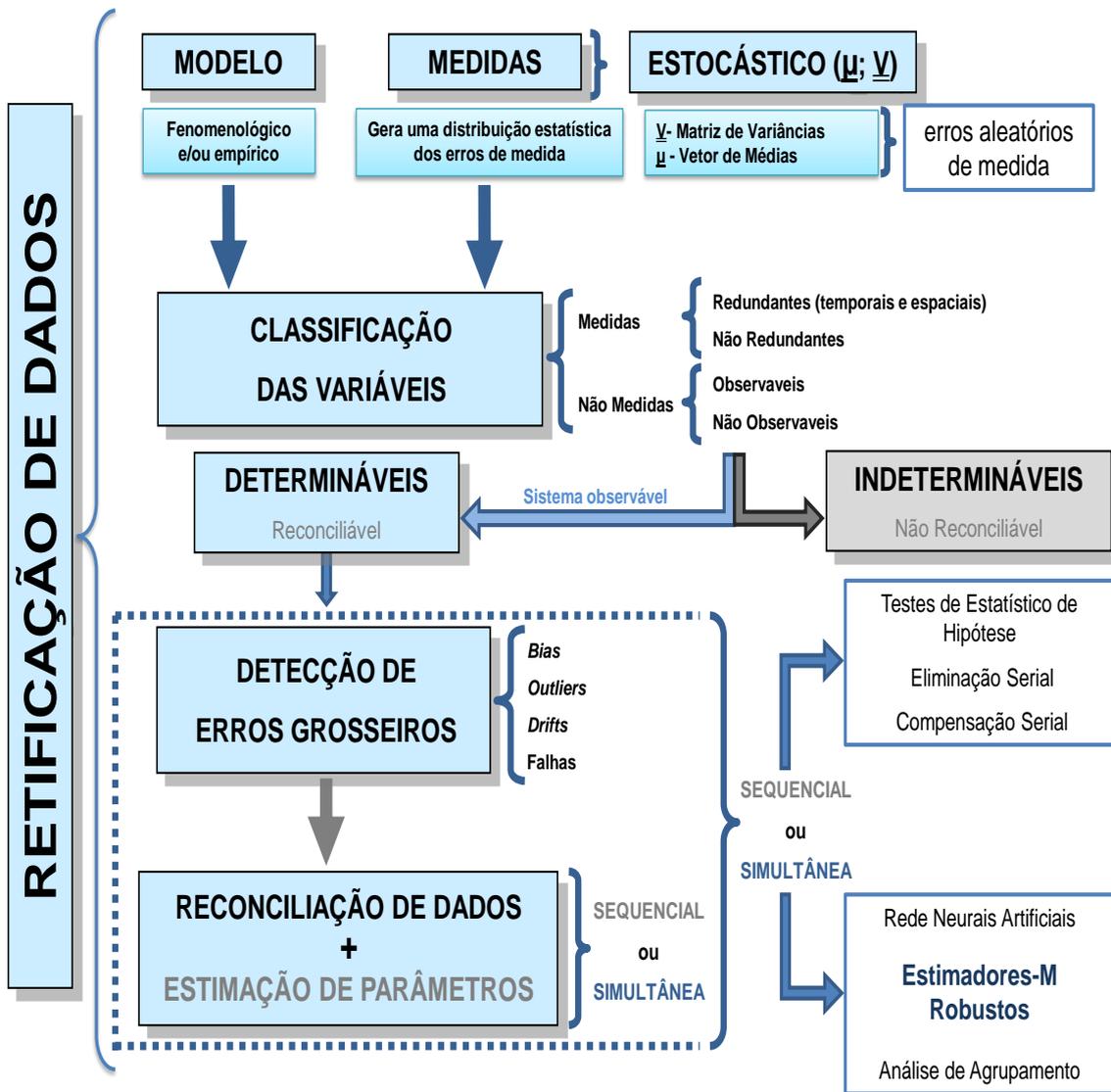


Figura 2.1: Estrutura resumida do procedimento de retificação de dados (SARRUF, 2014).

importante apontar que existe um número mínimo de variáveis medidas que permite determinar o estado de um equipamento (de acordo com o modelo matemático que lhe descreve), necessário para estimar outras variáveis importantes que não são medidas, como especificado pelo CLAV. Como abordado pelo procedimento de CLAV, certas variáveis são indispensáveis para observabilidade do modelo matemático, existindo outras variáveis que podem ser estimadas pelo modelo e dispensam medição. O procedimento de CLAV avalia o número mínimo de medidas para a observabilidade do equipamento, de acordo com o modelo matemático, minimizando custos instrumentais e computacionais. Por outro lado, o uso do número mínimo de medições resulta em perda de redundância de informação, que é fundamental para a etapa da DEG (ou Detecção de Múltiplos Erros Grosseiros,

DMEG) e RD. O estudo de CLAV também pode gerar um resultado fundamental para a análise de Alocação Ótima de Sensores, pois localiza a informação necessária em um processo químico, como vazões, concentrações, temperaturas, níveis, pressões e propriedades finais de interesse dos produtos (PRATA, 2009).

A qualidade da informação está diretamente ligada aos procedimentos de medição, pois estes estão sujeitos a falhas, limitação tecnológicas e resolução intrínseca dos instrumentos de medição (não existe escala com precisão infinita, havendo limites de detecção e zonas mortas de observação). Portanto, as informações contidas nas variáveis medidas (*dados brutos ou crus*) com auxílio de instrumentos físicos ou rotinas de medição apresentam informações combinadas com erros, de forma que as medidas estão sempre corrompidas por erros (em maior ou menor grau). Estes erros são originados por diversas causas, como falhas mecânicas, mau funcionamento de componentes físicos, má calibração dos instrumentos de medição, intervenção humana (amostragem, análise e transcrição), flutuações súbitas de energia, falha na conversão de sinais, deterioração (corrosão) dos equipamentos de medição e incrustação nos sensores (LIEBMAN *et al.*, 1992; MCBRAYER e EDGAR, 1995). Outro problema que contribui ainda mais para a variabilidade dos dados obtidos são os efeitos da não linearidade do processo, a variabilidade natural das variáveis de operação, que pode ser causada por perturbações não controladas do processo, esquemas de controle e histerese (DE JESUS, 2011). Portanto, as preciosas informações contidas nos dados estão sempre corrompidas, em menor ou maior grau. Por isto, não se espera que conjuntos de dados reais medidos obedeçam às leis de conservação (balanços de massa, energia e quantidade de movimento) nem às demais restrições físicas impostas ao sistema, seja em escala industrial ou laboratorial. Desta forma, uma importantíssima etapa do procedimento de RTD, a RD, é crucial para minimizar os ruídos contidos nas informações aferidas, respeitando simultaneamente o modelo matemático do processo. Realizando de forma estatisticamente consistente a recuperação dos valores confiáveis e precisos das variáveis medidas, a estimação das variáveis não medidas e parâmetros do modelo, sempre que possível.

Os erros que afetam os dados medidos podem ser divididos em duas classes principais: **erros aleatórios** e **erros grosseiros**. Erros aleatórios são aqueles que resultam da impossibilidade dos sensores reproduzirem os resultados exatamente, muitas vezes conhecido como ruído branco Gaussiano. A RD, nesse caso, pode ser baseada na consideração de que estes erros são independentes, Normalmente distribuídos (Gaussiana) e que apresentam variância conhecida e média nula. Erros grosseiros são aqueles originários de eventos não aleatórios, tendo pouca ou nenhuma ligação com o valor medido. Podem estar relacionados às medidas (por exemplo, mau funcionamento de instrumentos) ou ao processo (por exemplo, vazamentos). Um fato

demonstrado matematicamente é que uma medida sujeita a infinitos erros aleatórios pequenos apresenta distribuição Normal de probabilidades (Teorema do Limite Central) (FISCHER, 2010). Porém, erros grosseiros invalidam a base estatística clássica dos métodos tradicionais de RD e prejudicam a análise sistemática dos dados, requerendo a etapa de DMEG e de eliminação dos erros grosseiros, que pode preceder a etapa de RD. Os erros grosseiros podem ser ainda subdivididos em, **desvios sistemáticos** (do idioma inglês “*bias*”), **valores espúrios** (do idioma inglês “*outliers*”) e **desvios cumulativos** (do idioma inglês “*drifts*”), como ilustrado nas Figuras 2.2, 2.3 e 2.4, respectivamente (CHEN e ROMAGNOLI, 1998).

Na Figura 2.2, o erro grosseiro sistemático acompanha o processo; ou seja, segue o comportamento do processo, porém com um erro constante, fazendo que os dados sejam desviados sistematicamente. A Figura 2.3 exemplifica erros aleatórios na presença de valores espúrios; ou seja, valores com erros grosseiros completamente distintos do comportamento do processo (falha repentina). Estes erros podem levar a ações de controle equivocadas, caso não sejam retificados *a priori* (PRATA, 2009). Para o tratamento estatístico, o erro do tipo *bias* é interpretado como um grande desvio sistemático, alterando consideravelmente a média móvel da série de dados, enquanto o erro do tipo *outlier* é interpretado como um grande desvio pontual, em relação a variabilidade dos dados. A reconciliação, com ajuda do modelo matemático e dos estimadores robustos, “negligencia” os erros sistemáticos e espúrios durante o procedimento matemático de otimização, de forma que os erros são identificados e os dados são corrigidos, respeitando o modelo do processo. Para isso, é necessário que existam redundâncias dos valores medidos, um bom modelo matemático, um estimador robusto bem sintonizado para a variância dos dados, um intervalo de amostragem e uma janela de pontos adequada para descrever a dinâmica do sistema (BAGAJEWICZ e ROLLINS, 2002).

A Figura 2.4 ilustra erros aleatórios que se desvia cumulativamente do valor real ao longo do tempo. Erros do tipo *drift* podem ser causados por desgastes ou incrustações nos sensores, agravando-se gradualmente com o tempo, fazendo com que a magnitude do erro grosseiro aumente suavemente com o passar do tempo. Estes erros acumulados ao longo do tempo levam a ações de controle equivocadas e a perda do *set-point* ótimo de produção, caso não sejam detectados e tratados consistentemente (ZHANG e CHEN, 2015).

Existe ainda outro tipo de erro não mencionado anteriormente, que são os erros que não seguem o comportamento do processo. Essa classe de erros é conhecida como falha completa, segundo NARASIMHAN e JORDACHE (1999). A falha completa não se encaixa como erro tipo *bias*, pois não acompanha o comportamento do sistema; entretanto, este tipo de falha também é caracterizado pela leitura de

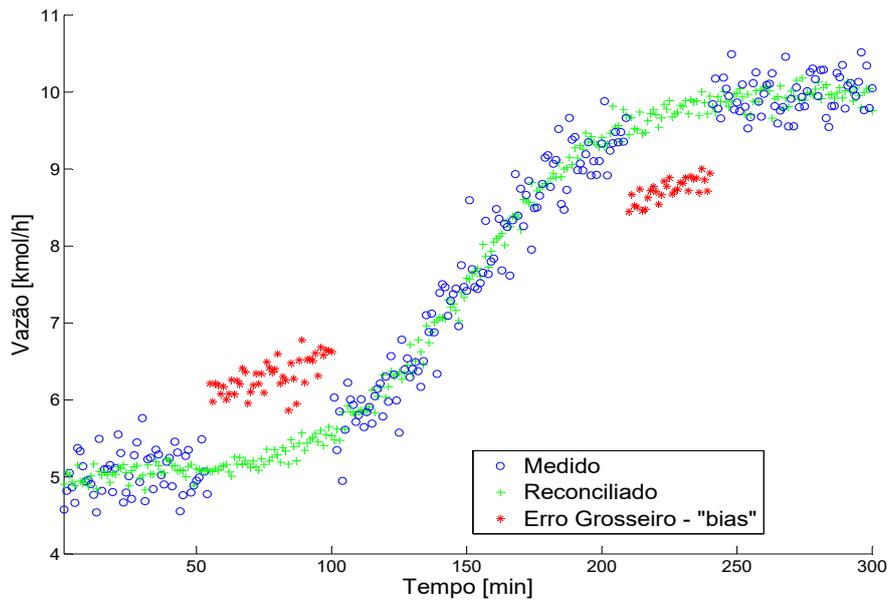


Figura 2.2: Dados medidos com erros sistemáticos (*bias*) e dados reconciliados. Adaptado: (BAGAJEWICZ, 2000a).

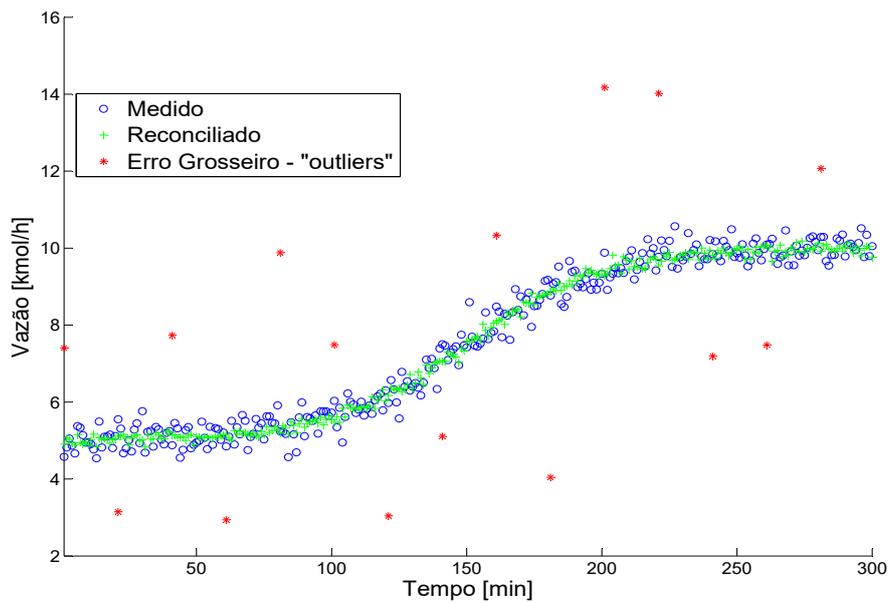


Figura 2.3: Dados medidos com erros espúrios (*outliers*) e dados reconciliados. Adaptado: (BAGAJEWICZ, 2000a).

valores sistematicamente mais altos ou mais baixos que os valores “verdadeiros”, como aqueles gerados por uma falha na comunicação (levando à repetição do último sinal) ou do sinal padrão da instrumentação. A Figura 2.5 ilustra um caso de falha completa.

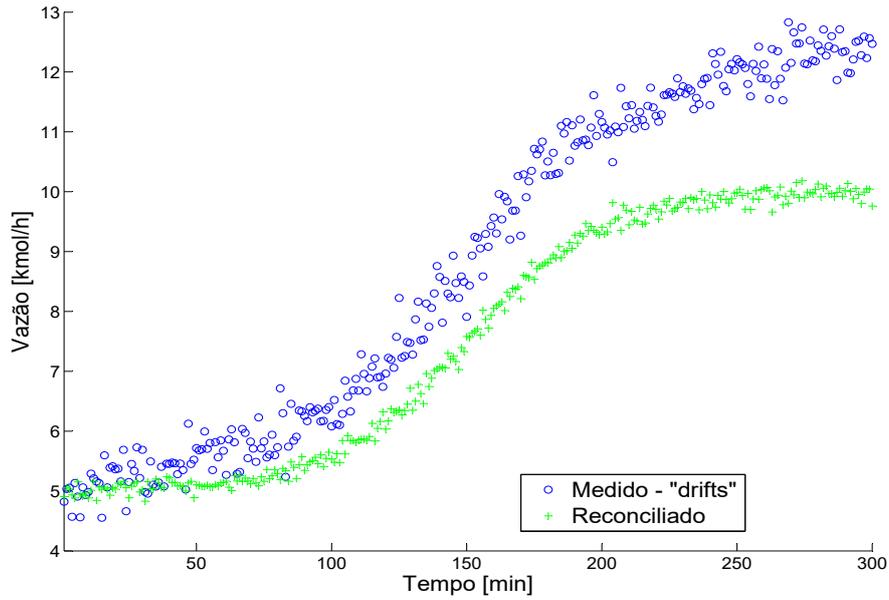


Figura 2.4: Dados medidos com erros cumulativos (*drifts*) e dados reconciliados. Adaptado: (ZHANG e CHEN, 2015).

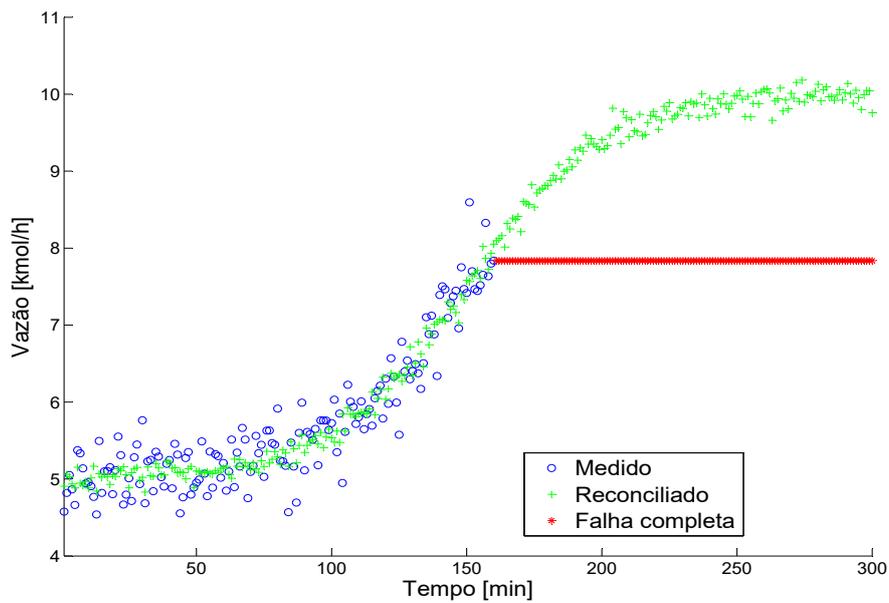


Figura 2.5: Dados medidos com falha completa e dados reconciliados. Adaptado: (BAGAJEWICZ, 2000a).

2.1 Classificação de Variáveis

A técnica de classificação de variáveis consiste basicamente em determinar a observabilidade e redundância do sistema, de acordo com a dimensão do modelo de processo e com as variáveis medidas disponíveis. A técnica permite avaliar se é possível ou não prever o comportamento das variáveis não medidas com o conjunto de dados e o modelo do processo disponíveis. Uma consequência importante é

que, se o conjunto de dados e o modelo de processo não forem suficientes para a observabilidade, o sistema é não observável, e o procedimento de Reconciliação de Dados é inviável.

Segundo ROMAGNOLI e SANCHEZ (1999), a CLAV pode classificar dois tipos de variáveis: **Variáveis Medidas** e **Variáveis Não-Medidas**. Além disso, esses dois tipos podem ser subdivididos em **Variáveis Medidas Redundantes** (VMR), **Variáveis Medidas Não-Redundantes** (VMNR), **Variáveis Não-Medidas Observáveis** (VNMO) e **Variáveis Não-Medidas Não Observáveis** (VNMNO). É importante também definir dois tipos de redundância das variáveis medidas. NARASIMHAN e JORDACHE (1999) define que a **Redundância Espacial** é aquela em que uma variável se relaciona com outra(s) por meio de um modelo matemático determinístico (fenomenológico ou não), mesmo em estado estacionário. A **Redundância Temporal** é a que aparece nos problemas de RD dinâmica, em que a variável medida é discretizada no tempo e está relacionada a medidas deslocadas no tempo. A falta de Redundância Temporal não afeta necessariamente a observabilidade do sistema, visto que o modelo pode ser observado em um dado tempo. Entretanto, a redundância temporal pode ser extremamente importante quando no conjunto de dados existem erros grosseiros, pois a Redundância Temporal pode ajudar as etapas de DEG/DMEG e RD dinâmica (BAI e THIBAUT, 2010; FARIAS, 2009; PRATA, 2009).

Na etapa de RD somente é possível reconciliar as **VMR**; ou seja, variáveis que podem ser observáveis (calculáveis) e que não afetam a observabilidade do sistema mesmo quando são excluídas (STANLEY e MAH, 1981b). As variáveis medidas do tipo **VMNR**, que não são calculáveis com o auxílio das VMR, requerem que o valor medido permaneça inalterado. O sistema só se torna observável quando o conjunto mínimo de variáveis medidas disponível (VMR e VMNR) é suficiente para determinar as **VNMO** (determináveis) por meio do modelo do processo. Entretanto, se o sistema apresentar **VNMNO** (indetermináveis), o procedimento de RD pode se tornar impossível, pois as VNMNO tornam o sistema globalmente não observável (não determinável); ou seja, globalmente não reconciliável (CROWE *et al.*, 1983; PRATA, 2009; STANLEY e MAH, 1981a,b). Do ponto de vista prático, no entanto, pode ser que a existência de VNMNO não seja determinante para interpretação dos dados de processo e que a parte observável do modelo seja suficiente para uso na operação.

A redundância nas medidas é essencial para a DEG, porque quando uma VMNR é eliminada, ela não é reconciliada no procedimento de RD. Assim, apenas as VMR devem estar presentes na função objetivo, de forma que os testes estatísticos de hipóteses utilizados na etapa de DEG/DMEG devem ser feitos somente com VMR

(FARIAS, 2009).

Para obter um sistema observável, pode ser necessário reestruturar o modelo de processo ou adicionar novas variáveis medidas (VMR ou VMNR) ao conjunto de dados. A viabilização de novas variáveis medidas requer a instalação de novos instrumentos de medição para determinar as variáveis não medidas indetermináveis. Este procedimento é conhecido como “alocação ótima de sensores”; ou seja, é o procedimento que prova a viabilidade técnica e utilidade da aquisição de um novo instrumento de medição (NARASIMHAN e JORDACHE, 1999).

Pesquisadores vêm realizando estudos de CLAV desde 1969 (SARRUF, 2014). Existem, basicamente, duas abordagens para a etapa da CLAV: abordagem baseada na **Teoria dos Grafos** (MAH *et al.*, 1976) e abordagem baseada na **Teoria Matricial** (CROWE *et al.*, 1983; DAROUACH *et al.*, 1986; OLIVEIRA JR., 2006).

Teoria dos Grafos pressupõe que o processo estudado pode ser organizado na forma de fluxograma, tratando cada unidade ou equipamento como um nó e ligando as saídas e entradas dos equipamentos com setas. Além das unidades e equipamentos de processo, um nó extra é criado, para ligar as correntes que saem e entram do ambiente externo ao processo. Assim, todas as setas têm um começo e um fim, formando o grafo representativo do processo.

A abordagem matricial usada para o estudo da CLAV pode ser dividida em quatro subgrupos: **Matriz de Projeção** (CROWE *et al.*, 1983), **Fatorização QR** (ROMAGNOLI e SANCHEZ, 1999), **Matriz de Incidência** (DAROUACH *et al.*, 1986) e **Matriz de Ocorrência** (OLIVEIRA JR., 2006).

A abordagem da Matriz de Projeção de Crowe consiste na eliminação direta de variáveis não medidas (observáveis) em problemas em que o modelo de processo é linear. CROWE *et al.* (1983) propuseram uma elegante estratégia para decompor as variáveis não medidas em variáveis medidas, formando um sistema linear. O procedimento da Matriz de Projeção pode ser estendido para modelos bilineares, fazendo a substituição do produto de duas variáveis por uma nova variável, tornando o problema linear.

A Fatoração QR tem o mesmo fundamento da Matriz de Projeção; entretanto, a solução matemática requer a dissociação da matriz que representa o modelo de processo em matrizes de variáveis medidas e de variáveis não medidas. Em seguida, aplica-se a transformação ortogonal QR na matriz de variáveis não medidas. Após uma organização algébrica matricial, o problema pode ser analisado e verificado no qual diz respeito à observabilidade. Este método pode ser usado para modelos de processos não-lineares utilizando a solução por linearização sucessiva. O

procedimento desta abordagem será detalhado na subseção de Reconciliação de Dados.

As outras abordagens usadas para solucionar a CLAV utilizam técnicas de ordenamento matricial, fazendo atribuições como o valor (1) para variável presente na equação do modelo de processo e (0) para variável ausente no modelo. A Matriz de Incidência, além de numerar as variáveis presentes e ausentes, sinaliza as entradas (+) e saídas (-) com sinais. A grande vantagem das abordagens baseadas em Matriz de Incidência e de Ocorrência é que a abordagem de Matriz de Ocorrência pode ser aplicada a modelos de processo não-lineares e dinâmicos, incluindo aqueles que contêm parâmetros desconhecidos (OLIVEIRA JR., 2006).

2.2 Matriz de Covariância

A estimação da **Matriz de Covariâncias** (COV) não é considerada de fato como uma etapa do procedimento de RTD. Entretanto, é sempre importante entender a distribuição dos erros associados aos dados; ou seja, a caracterização da variabilidade dos erros de medição que está associada à etapa de RD. A COV é de fundamental importância no procedimento de RTD, pois as variáveis envolvidas no processo apresentam grandezas e desvios distintos, obtidas por instrumentos de medição com diferentes graus de amplitude e resolução.

O inverso desta matriz pondera consistentemente o peso dado às variáveis na função objetivo, dando margem para um ajuste adequado em relação à variabilidade e à resolução do instrumento (DE MENEZES, 2015). Sabendo que a variância mede a dispersão em torno da média, as quantidades que têm maiores variâncias impõem maior flexibilização numérica no procedimento de minimização dos erros, estando conseqüentemente sujeitas a maiores ajustes (PRATA, 2005).

A forma da matriz de covariâncias depende se os erros de medida são **correlacionados** ou **independentes**. Para erros correlacionados, a matriz é cheia. Quando as covariâncias são negativas, as variáveis são inversamente proporcionais; quando são positivas, as variáveis são diretamente proporcionais. A variância é sempre uma grandeza positiva, o que torna a Matriz de Covariâncias positiva definida (NARASIMHAN e JORDACHE, 1999). A matriz de covariâncias é também uma matriz simétrica. Entretanto, quando se admite que os erros são independentes e não correlacionados todas as covariâncias são nulas, de forma que a matriz de covariâncias é diagonal. Admitindo a forma diagonal, a matriz é chamada de **Matriz de Variâncias** ou **Matriz de Dispersão** (DE MENEZES, 2015). Nem todos os dados são suficientemente disponibilizados em processos industriais, dificultando a

análise dos erros de medida nesses processos. Desta forma, alguns métodos são usados para a estimação da matriz de covariâncias. Esses métodos podem ser divididos em **Métodos Diretos** e **Métodos Indiretos**.

Nos métodos diretos, a estimação das covariâncias utiliza dados medidos em intervalos temporais, representados dentro de uma janela (horizonte) e admitindo que o processo esteja em regime estacionário (NARASIMHAN e JORDACHE, 1999; ROMAGNOLI e SANCHEZ, 1999). Várias técnicas podem ser usadas para detectar o estado estacionário de um certo período de tempo, facilitando a aplicação automática e em linha de métodos diretos para a estimação de covariâncias (ZHANG *et al.*, 2014). As Equações (2.1), (2.2) e (2.3) podem ser utilizadas para calcular a Média ($\hat{\mu}$), Variância ($\hat{\sigma}_i^2$) e Covariância ($\hat{\sigma}_{ij}$), respectivamente, no intervalo de tempo considerado.

$$\hat{\mu}_i(\underline{z}_i) = \frac{1}{n} \sum_{t=1}^n z_{ti} \quad (2.1)$$

$$\hat{\sigma}_i^2(\underline{z}_i) = \frac{1}{n-1} \sum_{t=1}^n (z_{ti} - \hat{\mu}_i(\underline{z}_i))^2 \quad (2.2)$$

$$\hat{\sigma}_{ij}(\underline{z}_i, \underline{z}_j) = \frac{1}{n-1} \sum_{t=1}^n (z_{ti} - \hat{\mu}_i(\underline{z}_i))(z_{tj} - \hat{\mu}_j(\underline{z}_j)) \quad \forall i \neq j; \quad (2.3)$$

em que z_{ti} é a variável aleatória discreta i medida no instante de tempo t discreto e \underline{z}_i é o vetor da janela de tempo da variável i , sendo $i = j = 1 \dots m$, em que m é o número de variáveis, e $t = 1 \dots n$, em que n é o número de pontos na janela de tempo. Com isso, dentro da mesma janela de tempo definem-se que $\hat{\mu}_i()$ é a média amostral da variável i , $\hat{\sigma}_i^2()$ é variância amostral da variável i e $\hat{\sigma}_{ij}()$ é a covariância amostral entre duas variáveis aleatórias discretas i e j , sendo $i \neq j$.

Porém, a hipótese de estado estacionário nem sempre está correta ou não é possível manter uma planta operando em regime permanente. Assim, as estimativas obtidas pelo método direto podem estar incorretas na presença de oscilações e perturbações na dinâmica do processo. Por isso, métodos indiretos foram propostos para tentar contornar esse problema. Nos métodos indiretos, adiciona-se a informação do modelo de processo (balanço de massa, por exemplo) para acompanhar as mudanças temporais das variáveis medidas. Contudo, o método indireto apresenta algumas limitações, pois o modelo de processo pode não ser completamente fidedigno (SARRUF, 2014).

Com isso, CHEN *et al.* (1997) propôs um método indireto, chamado de Método

Robusto Indireto, que não incorpora informação do modelo de processo, mas o método se limita a dados em regime permanente. O método utiliza a mediana como base de cálculo, ao invés da média. Isto ocorre porque a média torna-se mais sensível à contaminação por valores espúrios. Para compensar a influência dos erros grosseiros, a mediana é utilizada, por ser mais robusta na presença destes erros.

A Equação (2.4) representa a mediana dos valores absolutos dos desvios em relação à mediana da série, que é chamada de *Normalized Median Absolute Deviation* (MADn). A MADn é um estimador robusto do desvio-padrão ($\hat{\sigma}_i$) muito utilizado por diversos autores da estatística robusta (PRATA, 2009).

$$\hat{\sigma}_i = MADn(\underline{z}_i) = 1,4826 \cdot Q_{2/4} (|\underline{z}_i - \underline{1} \cdot Q_{2/4}(\underline{z}_i)|) \quad (2.4)$$

em que $\underline{1}$ é um vetor unitário com a mesma dimensão do vetor \underline{z}_i e $Q_{2/4}()$ é o 2º quartil das distâncias ou a mediana amostral.

ROUSSEEUW e CROUX (1992) propuseram dois estimadores alternativos ao MADn motivados por dois pontos fracos de técnica: a baixa eficiência relativa (definida no capítulo de Estatística Robusta) do estimador MADn em relação à distribuição Normal (37%) e o fato do MADn admitir uma distribuição simétrica. As Equações (2.5) e (2.6) representam os estimadores Sn e Qn, respectivamente. Os estimadores Sn e Qn apresentam eficiência relativa à distribuição normal de 58% e 82%, respectivamente.

$$\hat{\sigma}_i = Sn(\underline{z}_i) = 1,1926 \cdot Q_{2/4} \left[\underline{Q}_{2/4} (|\underline{1} \cdot z_{ti} - \underline{z}_i|) \right] : t = 1 \dots n \quad (2.5)$$

$$\hat{\sigma}_i = Qn(\underline{z}_i) = 2,2219 \cdot Q_{1/4} (|z_{ti} - z_{si}|) : t < s; t = 1 \dots n - 1 \text{ e } s = 2 \dots n \quad (2.6)$$

em que $\underline{Q}_{2/4}$ é o vetor de medianas produzido pelas medianas de n vetores, e $Q_{1/4}()$ é o 1º quartil das distâncias, sendo $\frac{n(n-1)}{2}$ a dimensão do vetor no argumento do operador amostral. Uma revisão mais abrangente e criteriosa desses métodos de estimação de variância pode ser encontrada em PRATA (2009).

Uma outra forma de estimar desvio-padrão e médias é utilizando os estimadores-M robustos. Na literatura da estatística robusta os parâmetros são usualmente chamados de localização (média - $\hat{\mu}$) e escala (desvio-padrão - $\hat{\sigma}$). Portanto, uma estimação-M de localização e escala pode ser feita iterativamente utilizando a função peso $w(u)$ de um estimador-M robusto. Os Algoritmos 1 e 2

ilustram os processos iterativos para estimar a localização e escala, através de um estimador-M robusto, respectivamente (ZOUBIR *et al.*, 2018).

Algoritmo 1: Estimação-M da Localização com o cálculo *a priori* da Escala

Dados: Amostras $\mathbf{y} \in \mathfrak{R}^N$, parâmetro de sintonia $c \in \mathfrak{R}$, estimativa de escala $\hat{\sigma} = MADn(\mathbf{y})$.

Saída: Estimativa-M da localização $\hat{\mu}$.

Entrada: Estimativa inicial da localização $\hat{\mu}^{(0)} = med(\mathbf{y})$, índice de iteração $n = 1$, nível de tolerância $\delta \in \mathfrak{R}$.

enquanto $\frac{|\hat{\mu}^{(n+1)} - \hat{\mu}^{(n)}|}{\hat{\sigma}} > \delta$ **faça**

 Cálculo dos pesos;

$$w_i^{(n)} = w\left(\frac{y_i - \hat{\mu}^{(n)}}{\hat{\sigma}}\right);$$

 Cálculo da estimativa da localização;

$$\hat{\mu}^{(n+1)} = \frac{\sum_{i=1}^N w_i^{(n)} y_i}{\sum_{i=1}^N w_i^{(n)}};$$

$n \leftarrow n + 1$

fim

Resultado: $\hat{\mu} \leftarrow \hat{\mu}^{(n+1)}$

No Algoritmo 2 a constante positiva b deve satisfazer $0 < b < \rho(\infty)$ e é escolhida como $\mathbf{E}[\rho(u)]$, onde u é uma variável aleatória normal padrão para atingir consistência com a distribuição Gaussiana, $\rho()$ é o estimador-M robusto, detalhado no Capítulo 3 e no Apêndice B, e $\mathbf{E}[\cdot]$ é o operador esperança, detalhado na Seção 2.4.1.

Algoritmo 2: Estimação-M da Escala com o cálculo *a priori* da Localização

Dados: Amostras $\mathbf{y} \in \mathfrak{R}^N$, parâmetro de sintonia $c \in \mathfrak{R}$, estimativa da localização $\hat{\mu}^{(0)} = med(\mathbf{y})$.

Saída: Estimativa-M da escala $\hat{\sigma}$.

Entrada: Estimativa inicial da escala $\hat{\sigma} = MADn(\mathbf{y})$, índice de iteração $n = 1$, nível de tolerância $\delta \in \mathfrak{R}$.

enquanto $|\frac{\hat{\sigma}^{(n+1)}}{\hat{\sigma}^{(n)}} - 1| > \delta$ **faça**

 Cálculo dos pesos;

$$w_i^{(n)} = w\left(\frac{y_i - \hat{\mu}^{(n)}}{\hat{\sigma}}\right);$$

 Cálculo da estimativa da escala;

$$\hat{\sigma}^{(n+1)} = \sqrt{\frac{1}{Nb} \sum_{i=1}^N w_i^{(n)} (y_i - \hat{\mu})^2};$$

$n \leftarrow n + 1$

fim

Resultado: $\hat{\sigma} \leftarrow \hat{\sigma}^{(n+1)}$

Um método chamado de Espectro de Variâncias vem sendo utilizado para analisar dinâmicas de processo e estimar incertezas de medições em DE MENEZES *et al.* (2020); FEITAL e PINTO (2015); FEITAL *et al.* (2014). Esta análise mostra como a variância do processo depende do tamanho da janela de amostragem. Este tipo de espectro fornece informações sobre as várias fontes que contribuem para o sinal de uma variável, incluindo ruído/erros de medição (tamanhos de janela curtos) e variações intrínsecas do processo (tamanhos de janela grandes). O espectro de variâncias pode ser definido como um conjunto de variâncias calculadas enquanto alguma variável relacionada a elas evolui (FEITAL e PINTO, 2015). O espectro de variações para janelas de amostragem curtas é controlado pelas variações do instrumento de medição. Dessa forma, a melhor estimativa para a variância do dispositivo de medição pode ser calculada usando o espectro de variância com janelas de amostragem curtas. No entanto, o uso de janelas de amostragem muito curtas pode não revelar a variabilidade real dos dados, devido à baixa qualidade da medição. O espectro para janelas de amostragem suficientemente grandes captura a variabilidade de todo o processo, incluindo mudanças operacionais. Mais detalhes sobre a utilidade desta técnica para caracterização de dados de processo são fornecidos por FEITAL e PINTO (2015).

2.3 Detecção de Erro Grosseiro

A ocorrência de erros grosseiros durante o monitoramento de processo advém de diversas casualidades, como já discutido. Portanto, há a necessidade de entender melhor a distribuição dos valores espúrios e estudar as formas de tratá-lo ou eliminá-lo do conjunto de dados.

A Detecção de Erro Grosseiro (DEG) é uma etapa da RTD necessária para obter maior confiança nos dados, tornando o tratamento estatístico mais eficiente. A DEG identifica os valores que não seguem a tendência de distribuição estatística adotada (modelo estatístico da distribuição dos erros), podendo eliminá-los do conjunto de dados (BAGAJEWICZ, 2000b). Para isso, foram desenvolvidas diversas técnicas, com a capacidade de lidar com um único erro grosseiro ou com múltiplos erros, como na Detecção de Múltiplos Erros Grosseiros (DMEG). Entretanto, técnicas de DEG podem acusar algum tipo de erro durante o teste estatístico de hipóteses. NARASIMHAN e JORDACHE (1999) citam que existem dois tipos de erros (ou falhas) associados a testes estatísticos de hipóteses de DEG:

- **Erro do Tipo I:** Um erro grosseiro é detectado no teste; entretanto, **não é** de fato um **erro grosseiro**. A falsa detecção pode prejudicar a precisão da

análise e a observabilidade do sistema, se a remoção resultar na redução do grau de redundância.

- **Erro do Tipo II:** Um erro grosseiro não é detectado no teste; entretanto, é de fato um **erro grosseiro**. A falsa impressão de que o sistema está livre de erros grosseiros. Faz com que os erros grosseiros afetem as estimativas, corrompendo os resultados e causando o efeito “*smearing*” (o erro grosseiro é compensado numericamente por desvios em outras variáveis no procedimento de RD).

Ambos os erros são prejudiciais ao tratamento do conjunto de dados, uma vez que o erro do Tipo I proporciona a redução indevida do grau de redundância do sistema, o que diminui a precisão da análise, enquanto o erro do Tipo II possibilita que erros grosseiros influenciem negativamente as estimativas dos dados reconciliados.

O teste estatístico de hipóteses padrão para DEG é bastante utilizado e simples, e também constitui a base para todos os demais procedimentos de DEG clássicos. O teste é aplicado ao conjunto de dados para determinar se as medições seguem uma distribuição com média zero; ou seja, distribuição simétrica dos erros de medição. Assim, o teste é aplicado ao erro médio ($\mu = \hat{E}(y_i - x_i^*)$) dos dados, conforme as seguintes hipóteses:

- H_0 : Hipótese nula: $\mu = 0$.
- H_1 : Hipótese alternativa: $\mu \neq 0$.

$$T = \frac{\hat{\mu}(y_i - x_i^*) - 0}{\sqrt{\hat{V}(y_i - x_i^*)}} \quad (2.7)$$

em que y_i é o valor medido, x_i^* é o valor ótimo estimado (reconciliado), $\hat{\mu}()$ representa a média amostral do erro de medição, (ou seja, o erro médio estimado) e $\hat{V}()$ é a variância amostral dos erros de medição (variância estimada). Admitindo uma Função Densidade de Probabilidade (FDP) com média nula para T , então grandes valores de T devem ser interpretados como valores inesperados na série de dados; conseqüentemente, prova que a hipótese H_1 é verdadeira, ou seja, que existe o erro grosseiro (*outlier*) (ÖZYURT e PIKE, 2004). Admite-se usualmente a distribuição Normal para o erro, embora esta hipótese não seja de fato necessária para conduzir o teste.

REILLY e CARPANI (1963) foram os primeiros a estudar a detecção de erros grosseiros em engenharia de processos. Os autores propuseram um teste estatístico do tipo χ^2 (qui-quadrado) baseado nos resíduos do modelo de processo, que foi chamado de Teste Global (GT). Um outro teste que utiliza restrições lineares e bilineares foi chamado de Teste Nodal (NT), proposto por REILLY e CARPANI (1963) e depois por MAH *et al.* (1976).

ALMASY e SZTANO (1975) propuseram um teste chamado de Teste da Máxima Potência (MP) para sistemas sujeitos a restrições lineares e com todas as variáveis medidas; entretanto, o teste só detecta um único erro grosseiro quando a variância é conhecida. O teste conhecido como Teste de Medida (MT) foi desenvolvido por MAH e TAMHANE (1982), sendo um dos mais utilizados na DEG. O teste MT requer que os dados sejam reconciliados previamente.

Outro tipo de método foi proposto por TONG e CROWE (1995). O Teste com Análise de Componentes Principais (PCAT) utiliza a técnica de (*Principal Component Analysis*, PCA) no teste de DEG, utilizando a hipótese de que os erros de medição seguem uma distribuição Normal.

Os testes propostos para DEG não são utilizados quando no conjunto de dados há presença de múltiplos erros grosseiros. Para isso, outras estratégias foram propostas para contornar essa dificuldade (BAGAJEWICZ e ROLLINS, 2002). Para o teste de DMEG, é comum a utilização de testes iterativos que visam a eliminar ou compensar os erros grosseiros. Destacam-se o método da Razão Probabilística Generalizada (RPG), proposto por NARASIMHAN e MAH (1987), e os métodos Teste de Medida Iterativo (IMT) e Teste de Medida Iterativo Modificado (MIMT), propostos por SERTH e HEENAN (1986). Também é possível a utilização de dois testes de DEG combinados; ou seja, testes podem ser combinados para formar uma estratégia iterativa que detecta, elimina ou compensa os erros grosseiros. Destacam-se o método proposto por YANG *et al.* (1995), MT-NT, que combina os métodos MT e NT, e o método proposto por CONGLI *et al.* (2006), NT-MT, combinando os métodos NT e MT, nesta ordem.

A DMEG pode ser realizada de duas formas em relação ao procedimento de RD: **sequencial** ou **simultânea**. Na DMEG realizada de forma **sequencial**, os erros grosseiros são identificados um de cada vez, sequencialmente. Essa estratégia sequencial ainda pode ser subdividida em dois métodos: método da eliminação serial e método da compensação serial, como discutido adiante. A estratégia de DMEG realizada de forma **simultânea** com a RD trata todos os dados simultaneamente; ou seja, sem iteração ou em uma única iteração de testes estatísticos, minimizando todos os tipos de erros (aleatórios e grosseiros) ao mesmo tempo (NARASIMHAN

e JORDACHE, 1999).

Assim, NARASIMHAN e JORDACHE (1999) destacaram três estratégias para DMEG:

- Eliminação Serial;
- Compensação Serial;
- Compensação Simultânea ou Coletiva.

Na estratégia de Eliminação Serial os erros grosseiros são detectados um por um, sequencialmente. A cada eliminação da variável com erro grosseiro no conjunto de dados, o procedimento de RD deve ser realizado em sequência. Ao término da RD, realiza-se nova busca e eliminação de erro grosseiro, realizando este procedimento até que nenhum erro grosseiro seja detectado no conjunto de dados. A desvantagem da técnica é que as eliminações das variáveis contaminadas com valores espúrios podem tornar o sistema matemático não observável e, conseqüentemente, o procedimento de RD inviável (PRATA, 2009).

A Compensação Serial também detecta os erros grosseiros um por um, sequencialmente. Entretanto, uma vez que o erro grosseiro é detectado, a magnitude dele é estimada. Assim, o erro grosseiro é corrigido (compensado), ao invés de eliminado do sistema de variáveis medidas. O procedimento realiza o passo de detecção e compensação do erro em sequência com a RD, até que não se identifique mais nenhum erro grosseiro. Este procedimento elimina a desvantagem de perder redundância espacial (ou temporal no caso de RD dinâmica) no problema de RD, mas os resultados são dependentes da precisão da estimativa da magnitude do erro (ROLLINS e DAVIS, 1992).

Segundo PRATA *et al.* (2010), todas essas estratégias são baseadas na hipótese que os erros de medida seguem uma distribuição Normal. Entretanto, erro grosseiro presente nos dados, com alta probabilidade de ocorrência, inviabiliza severamente a hipótese de distribuição Normal. Assim, um esquema iterativo é necessário durante a DMEG, aumentando o esforço computacional e, dependendo do problema de RD, inviabilizando a aplicação do procedimento em tempo real.

FARIAS (2009) desenvolveu uma nova estratégia de DMEG, o Teste de Medida Iterativo Robusto (IMT-robusto), em que utilizou a estratégia IMT juntamente com estimadores-M robusto como hipótese da distribuição dos erros de medida. O estudo indicou que a estratégia IMT + estimadores-M robusto ajudou a reduzir o efeito “*smearing*”. Além disso, o estudo avaliou 11 diferentes testes de DMEG, comparando-os com a estratégia que utilizou o estimador-M robusto.

Conforme descrito por PRATA (2009), estas estratégias foram amplamente utilizadas em sistemas que opera em estado estacionário ou pseudo-estacionário (considerado estacionário por um certo período de tempo). Entretanto, os processos industriais reais estão frequentemente sendo manipulados, a fim de ajustar o sistema às condições ótimas e atender às especificações de mercado, segurança do processo e índices ambientais. Portanto, as abordagens aplicáveis a sistemas dinâmicos e baseadas em modelos de processo não-lineares estão ganhando bastante notoriedade. Seguindo essa nova tendência, três abordagens alternativas para realização do procedimento de RD e DMEG simultaneamente se destacam:

- Análise de Agrupamento (*clusters*);
- Redes Neuronais Artificiais (RNA);
- Estimadores-M Robustos (estatística robusta).

A abordagem baseada em Análise de Agrupamento identifica grupos de dados que apresentam similaridade (na redundância temporal) e os divide. Assim, dentro de um mesmo grupo os dados são mais similares entre si que quando comparados com dados de outros grupos. Para cada agrupamento, é estabelecida uma ponderação que relaciona as distâncias médias entre os grupos de dados. Desta forma, um peso é atribuído para cada agrupamento e, conseqüentemente, para cada variável nele contida e computada na função objetivo. Assim, a influência negativa de um erro grosseiro pode ser reduzida. Esta abordagem foi desenvolvida por CHEN e ROMAGNOLI (1998) para realizar simultaneamente DMEG e RD.

Na abordagem de RNA, as redes devem ser “treinadas” com dados simulados ou reais dentro de um período de tempo e uma faixa de operação limitada, para então selecionar o melhor modelo algébrico que interpola os dados do “treinamento”. Sendo o dado selecionado para “treinamento” o mais confiável e estável da série temporal selecionada. O modelo é então usado para detectar os erros grosseiros dos dados de “testes”, que se desviam do modelo “treinado”. A desvantagem é que sistemas reais podem estar sujeitos a grandes perturbações, manipulações e mudanças das condições operacionais, fazendo com que o modelo “treinado” em certa condição de operação passa ser inadequado para simular o processo em outra condição de operação.

GUPTA e NARASIMHAN (1993) e TERRY e HIMMELBLAU (1993) realizaram estudos de DMEG e RD simultaneamente via RNA, utilizando como exemplo uma de rede de vapor em estado estacionário, como originalmente proposto por SERTH e HEENAN (1986). VACHHANI *et al.* (2001) utilizaram a abordagem de RNA para DMEG em um sistema dinâmico sujeito a restrição não-lineares para

representar um reator (*Continuous Stirred-Tank Reactor*) (CSTR) não isotérmico. Todos esses estudos utilizaram dados corrompidos com erros, para simular o comportamento real de medida.

ZHANG e CHEN (2014) utilizaram uma estratégia de Filtro de Partículas (PF) combinada com MT para solucionar simultaneamente o problema de RD e DMEG em sistemas dinâmicos não-lineares. Os autores realizaram o estudo com um sistema de polimerização em larga escala.

Estimadores-M robustos originam-se da estatística robusta, que admite que os dados não seguem a distribuição Normal. Quando há uma probabilidade maior de ocorrência de valores espúrios, que contaminam a hipótese de distribuição Normal, a modelagem dos erros de medida pela hipótese de Normalidade se torna ruim. Assim, os estimadores-M robustos modelam melhor os dados na presença de erros grosseiros. Além disso, durante o procedimento de otimização (minimização dos desvios), o estimador-M robusto aumenta a flexibilidade matemática para obter melhores ajustes nas variáveis sujeitas a erros grosseiros, sem prejudicar com efeito “*smearing*” as demais variáveis no procedimento de RD (PRATA, 2009).

TJOA e BIEGLER (1991) mostraram que o estimador-M permite reconciliar variáveis sujeita a erros grosseiros sem necessidade de eliminar o erro do conjunto de dados. ARORA e BIEGLER (2001) avaliaram estimadores robustos num problema de Reconciliação de Dados e Estimação de Parâmetros em sistemas dinâmicos com dados simulados. Assim, utilizando como exemplo um sistema de dois CSTR em série, demonstraram melhor desempenho dos estimadores robustos frente aos estimadores Fair e MQP, na presença de erros grosseiros e parâmetros a serem estimados. PRATA *et al.* (2010) foram os primeiros a resolver a RD dinâmica com o estimador-M robusto de Welsch em um reator de polimerização, na presença de dados industriais. Resultados satisfatório foram demonstrados utilizando o estimador de Welsch. ÖZYURT e PIKE (2004) fizeram um estudo comparativo com estimadores robustos em um problema de RD estacionário, demonstrando o melhor desempenho do procedimento de RD e DMEG simultâneo em relação ao procedimento sequencial.

PRATA *et al.* (2008b) realizaram o primeiro estudo comparativo entre estimadores robustos em um problema de RD dinâmico sujeito a restrições não-lineares, concluindo que os estimadores de Welsch e Lorenziano resultaram nos melhores desempenhos. ZHANG *et al.* (2010) e JIN *et al.* (2012) desenvolveram e avaliaram seus próprios estimadores robustos, mostrando o enorme potencial da estatística robusta. DE MENEZES (2015) realizou um estudo de RD e DMEG com estimadores robustos em colunas de destilação, usando modelos lineares, bilineares e

não-lineares em estado estacionário, com variáveis não medidas (a serem estimadas). Três estimadores-M robustos (Collins, Tangente hiperbólica e Alamgir) exclusivos da estatística robusta e não anteriormente reportados em problemas de RD e DMEG foram estudados. Os três estimadores robustos demonstraram melhor desempenho em relação ao estimador de MQP.

VALDETARO e SCHIRRU (2011) apresentaram um novo método para executar simultaneamente a sintonia (estimação) dos três parâmetros do estimador-M robusto de Hampel, a reconciliação robusta de dados e a detecção de erros grosseiros. O procedimento de sintonização automática foi baseado na minimização dos critérios de Akaike robusto. Para isso, o algoritmo de PSO foi usado como um método de otimização. O sistema estudado pelos autores foi um reator térmico nuclear. Os resultados mostraram o potencial de uso do método para resolver RD e DMEG, evitando procedimentos em que a sintonia dos parâmetros do estimador deve ser feita *a priori*.

MUNOZ e CHEN (2012) propuseram um novo estimador chamado de Estimador Máximo de Correntropia (EMC), que apresenta as mesmas propriedades de um estimador-M robusto, e removendo os efeitos de valores espúrios. As vantagens do novo estimador foram comprovadas em um reator batelada. CHEN *et al.* (2013) sintonizou um parâmetro fundamental do novo estimador de Correntropia proposto, a largura do *kernel* Gaussiano, utilizando a minimização do critério Akaike.

A Tabela 2.1 resume algumas aplicações de estimadores-M robustos, utilizados no procedimento de RD em problemas de engenharia.

Tabela 2.1: Exemplos de estimadores-M robustos aplicados na literatura de RD - parte 1/3

Referência	Sistema*	Estimadores-M	Processo(s)
TJOA e BIEGLER (1991)	S-E	Contaminated Normal	Trocador de calor
JOHNSTON e KRAMER (1995)	S-E	Contaminated Normal and Lorentzian	Trocador de calor e rede de fluxo
ZHANG <i>et al.</i> (1995)	R-E	Contaminated Normal	Planta de ácido sulfúrico
ALBUQUERQUE e BIEGLER (1996)	S-D	Contaminated Normal and Fair	Trocador de calor, tanques e hidrólises
CHEN <i>et al.</i> (1998)	R-E	Fair and Lorentzian	Reator químico
BOUROUIS <i>et al.</i> (1998)	R-E	Contaminated Normal	Plantas de dessalinização flash multiestágios
MINGFANG <i>et al.</i> (2000)	S-E	Kong	CSTR adiabático
ARORA e BIEGLER (2001)	S-E/D	Fair and Hampel	Medição de vapor e tanques Reator químico, medição de vapor,
ÖZYURT e PIKE (2004)	S/R-E	Contaminated Normal, Cauchy, Fair, Hampel, Logistic e Lorentzian	moagem metalúrgica, trocador de calor, ácido sulfúrico e alquilação
RAGOT <i>et al.</i> (2005)	S-E	Contaminated Normal	Balanço de massa por componente
WONGRAT <i>et al.</i> (2005)	S-E	Hampel	Problema de PAI e FISHER (1988)
LINGKE <i>et al.</i> (2006)	S-E	Huber and Kong	Problema de PAI e FISHER (1988)
FABER <i>et al.</i> (2006, 2007)	R-E	Kong	Purificação de gás de forno de coque
SCHLADT e HU (2007)	R-E	Contaminated Normal	Reação química e coluna de destilação

***S** - dados simulados; **R** - dados reais; **E** - estado estacionário e **D** - dinâmico.

Tabela 2.1: Exemplos de estimadores-M robustos aplicados na literatura de RD - parte 2/3

Referência	Sistema*	Estimadores-M	Processo(s)
ALHAJ-DIBO <i>et al.</i> (2008)	S-E	Contaminated Normal	Planta de processamento mineral
LID e SKOGESTAD (2008a,b)	S/R-E	Contaminated Normal	Reformador de nafta catalítica, modelo de tubo simples e flash
PRATA <i>et al.</i> (2008b)	S-D	Contaminated Normal, Cauchy, Fair, Hampel, Logistic, Lorentzian e Welsch	CSTR - Problema de LIEBMAN <i>et al.</i> (1992)
PRATA <i>et al.</i> (2010)	R-D	Welsch	Reator industrial de polipropileno Torre atmosférica, separação de etileno, separação de ar e medição de vapor
ZHANG <i>et al.</i> (2010)	S-E	Zhang	Potência do reator térmico
VALDETARO e SCHIRRU (2011)	S/R-E	Hampel	Rede de medição e Problema de PAI e FISHER (1988)
JIN <i>et al.</i> (2012)	S-E	Jin, Cauchy and Huber	Torre de destilação atmosférica e medição de vapor
CHEN <i>et al.</i> (2013)	S-E	Zhang, Fair, Hampel and Correntropy	
ZHANG <i>et al.</i> (2014)	R-E	Correntropy	Separação de ar
ZHANG e CHEN (2015)	S-D	Correntropy	Polimerização de estireno
LLANOS <i>et al.</i> (2015)	S-E	Zhang, Biweight, Welsch and Correntropy	Medição de vapor e Problema de PAI e FISHER (1988)
KORPELA <i>et al.</i> (2016)	S/R-E	Welsch	Caldeiras industriais multi-combustível

***S** - dados simulados; **R** - dados reais; **E** - estado estacionário e **D** - dinâmico.

Tabela 2.1: Exemplos de estimadores-M robustos aplicados na literatura de RD - parte 3/3

Referência	Sistema*	Estimadores-M	Processo(s)
COIMBRA <i>et al.</i> (2017)	R-D	Welsch	Polimerizações de metacrilato de metila
DA CUNHA <i>et al.</i> (2017)	S-E	Contaminated Normal, Bell, Huber, Lorentzian, Hampel, Welsch, Andrews, Smith, Jin, Biweight and Correntropy	Reator de Van de Vusse (VAN DE VUSSE, 1964)
WU <i>et al.</i> (2017)	S-E	Wu	Conjunto de destilação de óleo cru
LLANOS <i>et al.</i> (2017)	S-E	Biweight, Huber, Hampel, Welsch and Correntropy	Rede de medição de vapor e trocador de calor
VALLURU <i>et al.</i> (2018)	S/R-D	Fair and Hampel	Reator de Williams-Otto (WILLIAMS e OTTO, 1960) e destilação reativa
DO VALLE <i>et al.</i> (2018)	S-E/D	Contaminated Normal, Cauchy, Fair, Hampel, Logistic, Lorentzian and Zhang (quasi-Weighted)	Coleção de problemas teste de benchmark
LIU <i>et al.</i> (2018)	R-D	Fair, Logistic, Lorentzian and Welsch	Manufatura farmacêutica
SU <i>et al.</i> (2019)	R-D	Welsch	Manufatura farmacêutica
XIE <i>et al.</i> (2019)	S/R-E	Fair, Welsch, Cauchy e Xie	Rede de medição e processo de evaporação industrial
DA CUNHA <i>et al.</i> (2020)	S-E	16 estimadores-M robustos	Problemas de reatores químicos

***S** - dados simulados; **R** - dados reais; **E** - estado estacionário e **D** - dinâmico.

Observa-se na Tabela 2.1 que a maioria das aplicações de estimadores robustos foi feita em processos operando em estado estacionário (32 casos). Apenas 5 trabalhos aplicaram estimadores-M robustos em um processo dinâmico e usando dados reais. Apesar disso, a literatura tem mostrado as amplas aplicações possíveis com os estimadores-M robustos; em particular, essas estratégias vêm se destacando cada vez mais dentre os métodos de DMEG simultâneo que podem lidar com sistemas estacionários ou dinâmicos, lineares ou não-lineares. A Tabela 2.2 resume em números o resultado da Tabela 2.1 em relação ao tipo de sistema.

Tabela 2.2: Resumo em números dos tipos de sistemas em problemas de RD robusta.

Sistema	Estado Estacionário	Dinâmico	Total
Dado Simulado	22	6	(28)
Dado Real	11	5	(16)
Total	(33)	(11)	44

2.4 Reconciliação de Dados

Reconciliação de Dados pode ser definido como o procedimento de estimação dos “verdadeiros” valores das variáveis medidas (sempre corrompidas por erros), respeitando as restrições fenomenológicas do sistema. Portanto, o problema de RD é um problema de otimização em que se deve minimizar os erros de medida descritos por uma Função Objetivo (Fobj), de modo a satisfazer os Balanços de Massa (BM), Balanços de Energia (BE) e outras restrições impostas ao equipamento(s) ou unidade(s) do processo em estudo. A Fobj pode ser modelada por alguma FDP ou um estimador-M (usualmente de Mínimos Quadrados Ponderados, MQP), que descreve ponderadamente pelas variâncias das medidas, os desvios observados entre os valores medidos e reconciliados (preditos ou estimados). Assim, a estrutura genérica do problema de RD pode ser representada pela Equação (2.8) e suas restrições, Equações (2.9) a (2.12).

$$\hat{\underline{\epsilon}} = \arg \left[\min_{\underline{\hat{\epsilon}}} \rho (\underline{\hat{\epsilon}}, \underline{\underline{V}}, \underline{\kappa}) \right] \quad (2.8)$$

sujeito a:

$$\hat{\underline{z}} = \underline{z} - \hat{\underline{\epsilon}} \quad (2.9)$$

$$\underline{f} (\hat{\underline{z}}, \hat{\underline{u}}, \underline{\theta}) = \underline{0} \quad (2.10)$$

$$\underline{h} (\hat{\underline{z}}, \hat{\underline{u}}, \underline{\theta}) = \underline{0} \quad (2.11)$$

$$\underline{g} (\hat{\underline{z}}, \hat{\underline{u}}, \underline{\theta}) \geq \underline{0} \quad (2.12)$$

em que $\rho()$ representa a forma funcional de um estimador-M (robusto ou não-robusto), que é em função dos erros de medida, das variâncias dos erros de medida e dos parâmetros de sintonia do estimador; $\hat{\underline{\epsilon}}$ é o vetor dos erros de medida estimados (dependente do tempo ou não); $\hat{\underline{z}}$ é o vetor das variáveis reconciliadas (dependente do tempo ou não); \underline{z} é o vetor das variáveis medidas (discretizadas no tempo ou não); \underline{u} é o vetor das variáveis estimadas não medidas (dependente do tempo ou não); $\underline{\theta}$ é o vetor dos parâmetros do modelo (caso tenha e pré-estimados); $\underline{\underline{V}}$ é a matriz de variâncias das medidas disponíveis; $\underline{\kappa}$ é o vetor dos parâmetros de sintonia do estimador (caso tenha e pré-sintonizados); $\underline{f}()$ é o vetor das equações diferenciais (temporais e/ou espaciais); $\underline{h}()$ é o vetor das equações algébricas (lineares e/ou não-lineares); $\underline{g}()$ é o vetor das inequações algébricas (incluindo os limites físicos superiores e inferiores das variáveis).

Sempre que possível, o procedimento deve estimar os valores das variáveis não medidas, $\underline{u}(t, \hat{\underline{\theta}})$, e parâmetros físicos do problema, $\hat{\underline{\theta}}$, preferencialmente em tempo real. A RD associada ao problema de estimação de parâmetros é usualmente conhecida como Reconciliação de Dados com Estimação de Parâmetros (RDEP). A estimação de parâmetros na RD pode ser realizada de forma sequencial ou simultânea, sendo esta última majoritariamente utilizada (PRATA, 2009).

Para entender melhor os diferentes problemas de RD que resultam principalmente dos tipos de restrições impostas ao sistema, é necessário esclarecer os tipos de restrições mais comuns que descrevem os modelos de processos. De forma geral, os modelos de processos podem ser **Lineares**, **Bilineares** e **Não-lineares**. O modelo pode ainda conter restrições na forma de **Equações** ou **Inequações** algébricas. Além disso, os modelos de processos podem apresentar variações temporais ou não, representando variáveis em regime permanente ou dinâmico, e sendo **Estacionários** ou **Dinâmicos**, sendo constituídos por sistemas algébricos, diferenciais ou algébricos-diferenciais, contendo inequações algébricas ou não. Para cada tipo de restrição diversas abordagens podem ser usada para

solucionar o problema de otimização restrito.

A maioria das técnicas de RD encontradas na literatura admite a hipótese de que os modelos são lineares e operam em regime permanente (MAH, 1990). Isso provavelmente decorre do fato de que, na época em que surgiram as técnicas de RD, existiam limitações tecnológicas e poucos algoritmos desenvolvidos para Programação Não-Linear (PNL).

O problema de otimizar uma função linear sujeita a restrições lineares teve origem nos estudos de Jean-Baptiste Joseph Fourier (1768 – 1830), matemático francês, sobre sistemas lineares de inequações em 1826. Quase um século mais tarde, em 1939, o matemático russo Leonid Vitaliyevich Kantorovich (1912 – 1986) publicou uma extensa obra denominada de “*Métodos Matemáticos de Organização e Planificação da Produção*”. No entanto, esse trabalho apenas ficou conhecido vinte anos mais tarde. Nessa época, os problemas logísticos eram resolvidos por tentativa e erro e por cálculo manual, justificando a preferência por soluções analíticas e o grande desenvolvimento da programação linear no início do século XX (NEVES, 2011).

Segundo UEHARA (2011), durante a Segunda Guerra Mundial, os avanços tecnológicos motivados pelas forças armadas resultaram na criação do microprocessador, na forma de válvulas termiônicas, desenvolvidas por Thomas Edison. A velocidade com que um microprocessador realizava os cálculos ultrapassava a velocidade dos cálculos manuais, sem contar a eliminação de erros humanos, constituindo um marco da nova revolução industrial. As válvulas termiônicas deram origem à válvula eletrônica e foram utilizadas nos primeiros computadores constituídos na década de 40, como o célebre ENIAC (*Electronic Numerical Integrator Analyzer and Computer*), desenvolvido a pedido do exército dos EUA para o laboratório de pesquisa balística, construído em 1945 e constituído por 18.000 válvulas, com capacidade de processar 5.000 adições, 357 multiplicações e 38 divisões por segundo. O ENIAC funcionou do período da Guerra Fria até 1955 e durante toda sua vida fez mais cálculos matemáticos do que toda a humanidade havia realizado anteriormente (UEHARA, 2011).

NEVES (2011) relatou que, com o fim da Segunda Guerra Mundial, a Pesquisa Operacional continuou bem estabelecida, principalmente nos EUA. No ano de 1947, as técnicas de Programação Linear (PL) se consolidaram quando George Dantzig, conhecido como o “Pai da Programação”, criou e desenvolveu o algoritmo conhecido como “Simplex”. Este algoritmo permitia operar em termos práticos a resolução de problemas de otimização, surgindo pela primeira vez o termo PL para designar o método criado e que ainda hoje é aplicado nas mais variadas situações para encontrar

a melhor solução (solução ótima) dentre várias disponíveis.

Com o passar do tempo e com a evolução dos microprocessadores, as inovações fomentaram e continuam a fomentar a eficiência computacional e o desenvolvimento de novos algoritmos de PL e PNL, que constituem casos particulares dos problemas de RD. Os algoritmos que surgiram para resolver os problemas de PL e PNL são cada vez mais rápidos e eficazes, facilitando a resolução de problemas de RD mais complexos, como RD dinâmica e DMEG simultâneo com restrições dinâmica não-lineares.

2.4.1 Reconciliação de Dados Estacionária

KUEHN e DAVIDSON (1961) foram os primeiros a utilizar procedimentos de RD como ferramenta de monitoramento de processo em problemas de engenharia química, formulado para sistemas de balanço de massa e energia lineares em estado estacionário e com todas as variáveis medidas (sem variáveis a serem estimadas, apresentando máxima redundância espacial para o modelo de processo). Este problema, por ser linear e em estado estacionário, admite solução analítica. A solução pode ser obtida aplicando o método de Multiplicadores de Lagrange numa Fobj quadrática, baseada no estimador MQP sujeito a restrições lineares. Assim, os primeiros benefícios da RD foram observados e relatados. MAH *et al.* (1976) estudaram a solução do problema de RD com restrições lineares e em estado estacionário com variáveis não medidas, desde que a etapa de CLAV indique que o sistema é globalmente observável.

HLAVÁČEK (1977) estudou um problema de RD em uma planta química operando em estado estacionário e sujeita a restrições lineares e não-lineares, representadas por balanços de massa e energia. O autor propôs cinco alternativas para solucionar o problema de RD linearmente restrito e não linearmente restrito: 1) substituição algébrica direta de variáveis das equações de restrição (BM e BE) na Fobj; 2) minimização com auxílio de multiplicadores de Lagrange; 3) minimização através da PNL; 4) minimização através da PL e 5) uso do critério “minimax” de Chebyshev, minimização do desvio máximo. A primeira alternativa mostrou o melhor ganho computacional; a segunda e a terceira alternativa permitem a utilização da etapa CLAV, que pode admitir variáveis não medidas; a quarta alternativa é recomendada para sistemas multicomponentes que geram equações de balanço de massa bilineares e a última alternativa é utilizada quando problemas de mal condicionamento afetam nas abordagens de estimadores quadráticos.

KNEPPER e GORMAN (1980) utilizaram o método de linearizações sucessivas

para resolver o problema de RD não linearmente restrito, já que para restrições não-lineares não é possível obter solução analítica. Combinando um processo de linearização com iterações sucessivas é possível obter a solução numérica aproximada do problema. Restrições na forma de inequações algébricas não foram consideradas neste método de solução.

Outro método iterativo parecido utiliza a Matriz de Projeção de CROWE *et al.* (1983), baseado no algoritmo de fatoração QR de matrizes. PAI e FISHER (1988) propuseram a utilização do método da Matriz de Projeção de Crowe para solucionar problemas de RD com variáveis não medidas, não linearmente restrito e em estado estacionário. Este método também não considera limites nas variáveis, na forma de inequações algébricas.

O método de PNL foi introduzido para solucionar os problemas de RD por LIEBMAN e EDGAR (1988). Os autores solucionaram um problema de RD com restrições não-lineares, limites nas variáveis e em regime permanente. Comprovou-se que o método é eficaz e que pode constituir uma alternativa para os métodos de linearização sucessiva e da Matriz de Projeção de Crowe. RAMAMURTHI e BEQUETTE (1990) desenvolveram uma solução para um problema parecido utilizando um método de PNL, resolvendo o problema de RD na presença de variáveis não medidas, restrições não-lineares e em estado estacionário.

A partir do anos 90, observou-se clara mudança no foco dos problemas de RD, pois os pesquisadores começaram a olhar mais atentamente para sistemas reais em estado transiente. Problemas de RD dinâmica, que usam modelos dinâmicos de processo, requerem outros tipos de técnicas e métodos numéricos para solucionar o problema, como detalhada adiante.

O princípio da máxima verossimilhança

O procedimento de RD consiste essencialmente em resolver um problema de otimização caracterizado por uma Fobj a ser minimizada/maximizada, respeitando certas restrições (modelo).

A Fobj do problema é o estimador de **máxima verossimilhança** decorrente de uma FDP dos erros de medição, sendo comumente adotada a distribuição Normal (Gaussiana). O princípio da máxima verossimilhança aplicado à distribuição Normal resulta no estimador de MQP (ALBUQUERQUE e BIEGLER, 1995).

Para entender este princípio precisamos descrever primeiro o modelo de erros dos dados brutos medidos. Sabemos que medidas de processos inevitavelmente contêm informações imprecisas e, as causas são, por exemplo, devido a(s):

- Instrumentações imperfeitas, mal instaladas e/ou mal calibradas;
- Faixa e definição do sinal limitada;
- Estocasticidade na transmissão do sinal (entropia da informação - ver subseção 3.3.1);
- Flutuações da corrente elétrica (impacto no sinal transmitido);
- Outras diversas fontes de erros (falhas humanas e acidentes).

Admitindo que o dado medido é composto de um valor verdadeiro e um erro. Com isso, amostras de um conjunto de M variáveis aleatórias medidas podem ser modeladas como:

$$\underline{z} = \underline{x} + \underline{\varepsilon} \quad (2.13)$$

em que \underline{z} é o vetor das variáveis medidas (corrompidas por erro); \underline{x} é o vetor dos valores “verdadeiros” (livres de erro) e $\underline{\varepsilon}$ é o vetor de erros aleatórios de medida (livres de erros grosseiros). Portanto, as estimativas resultante para \underline{x} é denotada por $\hat{\underline{z}}$ e expressa por:

$$\hat{\underline{z}} = \rho(\underline{z}_1, \underline{z}_2, \dots, \underline{z}_n) \quad (2.14)$$

onde ρ representa a forma funcional de um estimador. Entretanto, para estimar $\hat{\underline{z}}$ outras informações sobre os dados são requeridas. De fato, informações sobre \underline{x} ou $\underline{\varepsilon}$ são desejáveis. Com isso, é comum admitir que $\underline{\varepsilon} \sim N(\underline{0}, \underline{\Sigma})$ para os n dados amostrados. Assim, a FDP de $\underline{\varepsilon}$ pode ser escrita como:

$$f(\underline{\varepsilon}) = \frac{1}{(2\pi)^{M/2} |\underline{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} \underline{\varepsilon}^T \underline{\Sigma}^{-1} \underline{\varepsilon}\right) \quad (2.15)$$

onde $|\underline{\Sigma}|$ é o determinante da matriz das variâncias-covariâncias dos erros de medida. A Figura 2.6 ilustra um exemplo de distribuição marginal de uma série de erros associado a medição. O exemplo pode ser interpretado como uma série temporal de erros de medição, para uma variável de processo em estado estacionário. A função densidade de probabilidade dos erros de medição pode ser modelada, de forma não paramétrica como o histograma ou por Estimador Densidade de Núcleo (*Kernel Density Estimate* - KDE). Entretanto, ainda é mais usual admitir um modelo de distribuição dos erros *a priori*.

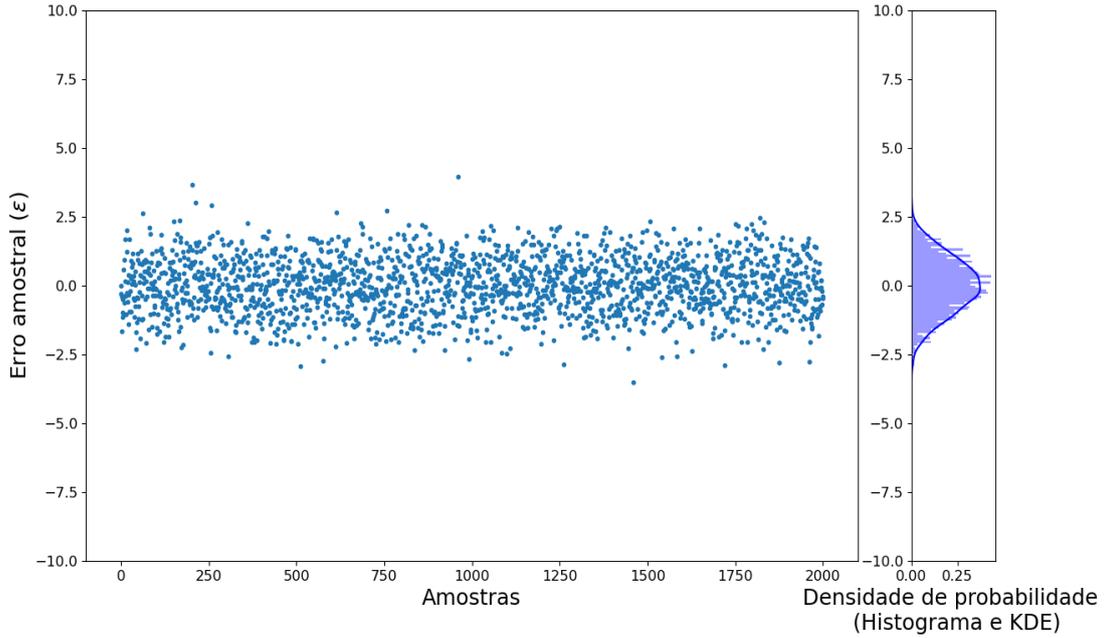


Figura 2.6: Distribuição marginal de erros de medição.

Aplicando a Equação (2.13) na Equação (2.15) e admitindo que \underline{x} e $\underline{\varepsilon}$ são independentes, então a função densidade condicional de medida \underline{z} dada a condição \underline{x} é expressa por:

$$f(\underline{z} | \underline{x}) = f(\underline{z} - \underline{x}) = f(\underline{\varepsilon}) \quad (2.16)$$

Portanto, a FDP de \underline{z} pode ser interpretada como Função de Verossimilhança de \underline{x} , dado as medidas \underline{z} , sendo denotada como:

$$L(\underline{x} | \underline{z}) = \frac{1}{(2\pi)^{M/2} |\underline{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} (\underline{z} - \underline{x})^T \underline{\Sigma}^{-1} (\underline{z} - \underline{x}) \right] \quad (2.17)$$

Dessa forma, a maximização de $L(\underline{x} | \underline{z})$ significar estimar o valor mais semelhante de \underline{x} . Essas estimativas são conhecidas como Estimativas de Máxima Verossimilhança. Assim, as estimativas para \underline{x} são meramente $\hat{\underline{x}} = \underline{z}$, o que significa que as estimativas ótimas dos valores verdadeiros das variáveis do processo são simplesmente as medições brutas quando a única informação disponível sobre o ruído de medição é $\underline{\varepsilon} \sim N(\underline{0}, \underline{\Sigma})$ (BAI e THIBAUT, 2010). Por isso o princípio da máxima verossimilhança (originário da estatística Bayesiana) admite que as máximas probabilidades de ocorrência dos eventos representam otimamente o comportamento de um sistema. A máxima probabilidade é obtida com o operador

Máximo, que retorna as variáveis que maximizam a função:

$$\hat{\underline{z}} = \max_{\hat{\underline{z}}} L(\hat{\underline{z}} | \underline{z}) \quad (2.18)$$

A maximização da Equação (2.17) é equivalente a minimizar o argumento positivo do termo exponencial. Dessa forma, o problema formulado originalmente por KUEHN e DAVIDSON (1961) pode ser escrito na forma (NARASIMHAN e JORDACHE, 1999).

$$\hat{\underline{z}} = \min_{\hat{\underline{z}}} \frac{1}{2} [\underline{z} - \hat{\underline{z}}]^T \underline{\underline{\Sigma}}^{-1} [\underline{z} - \hat{\underline{z}}] \quad (2.19)$$

sujeito a:

$$\underline{h}(\hat{\underline{z}}, \underline{u}) = \underline{0} \quad (2.20)$$

$$\underline{g}(\hat{\underline{z}}, \underline{u}) \geq \underline{0} \quad (2.21)$$

$$\hat{\underline{z}}^L \leq \hat{\underline{z}} \leq \hat{\underline{z}}^U \quad (2.22)$$

$$\underline{u}^L \leq \underline{u} \leq \underline{u}^U \quad (2.23)$$

em que $\hat{\underline{z}}$ é o vetor das variáveis medidas reconciliadas; \underline{z} é o vetor das variáveis medidas; $\underline{\underline{\Sigma}}$ é a matriz das variâncias-covariâncias dos erros de medida; \underline{u} é o vetor das variáveis não medidas estimadas (observáveis); $\underline{h}()$ é o vetor das equações de restrições algébricas lineares ou não-lineares; $\underline{g}()$ é o vetor das inequações de restrições algébricas lineares ou não-lineares; $\hat{\underline{z}}^L, \hat{\underline{z}}^U$ são os vetores de parâmetros inferiores e superiores do vetor $\hat{\underline{z}}$ e $\underline{u}^L, \underline{u}^U$ são os vetores de parâmetros inferiores e superiores do vetor \underline{u} .

RD estacionária linearmente restrita

KUEHN e DAVIDSON (1961) formularam o problema de RD, resultando no estimador dos mínimos quadrados ponderados. Este estimador deve satisfazer as restrições impostas pelas equações de conservação que modelam um equipamento ou um processo. As equações do modelo podem ser fenomenológicas ou empíricas sendo comumente utilizados os balanços materiais, balanços energéticos e balanços de quantidade de movimento. Outros tipos de restrições são as restrições físicas das variáveis, como limites mínimos e máximos para as variáveis reconciliadas. Entretanto, a maioria dos casos estudados na literatura utiliza restrições lineares e em estado estacionário. As restrições lineares podem ser representadas genericamente pela Equação (2.24):

$$\underline{\underline{A}} \hat{\underline{z}} = \underline{b} \quad (2.24)$$

em que $\underline{\underline{A}}$ é a matriz dos coeficientes das equações de restrições lineares; $\hat{\underline{z}}$ é o vetor das variáveis medidas reconciliadas e \underline{b} é o vetor dos coeficientes lineares das restrições (ou resíduo \underline{e} do modelo). A variável “verdadeira” pode ser representada hipoteticamente como a soma de variável medida, e de um erro intrínseco de medição, na forma:

$$\underline{x} = \underline{z} - \underline{\varepsilon} \quad (2.25)$$

Como os valores verdadeiros não são conhecidos, pode-se inferi-los a partir dos valores reconciliados $\hat{\underline{z}}$:

$$\underline{x} \simeq \hat{\underline{z}} = \underline{z} - \underline{\varepsilon} \quad (2.26)$$

Assim, isolando o vetor dos erros aleatórios, $\underline{\varepsilon}$, chega-se a:

$$\underline{\varepsilon} = \underline{z} - \hat{\underline{z}} \quad (2.27)$$

Contudo, é possível rescrever a equação do modelo como:

$$\underline{\underline{A}} (\underline{z} - \underline{\varepsilon}) = \underline{b} \quad (2.28)$$

Com a consideração de que o vetor de erros $\underline{\varepsilon}$ é livre de erros grosseiros, pode-se admitir que o erro tenha média nula, ou seja:

$$\mathbf{E} [\underline{\varepsilon}] = \underline{0} \quad (2.29)$$

em que $\mathbf{E}[\underline{u}] = \int_{D^n} \dots \int \underline{u} f(\underline{u}) (d\underline{u})^n = \underline{\mu}$ é o operador esperança de $f(\underline{x})$. Admite-se ainda que a matriz covariância-variância (covariância) $\underline{\underline{\Sigma}}$ das medidas seja conhecida e definida, aplicando o operador covariância $\mathbf{cov}(\underline{u}) = \mathbf{E} \left[(\underline{u} - \mathbf{E}[\underline{u}]) (\underline{u} - \mathbf{E}[\underline{u}])^T \right]$ no erro de medição, obtemos a matriz na forma :

$$\underline{\underline{\Sigma}} [\underline{\varepsilon}] = \mathbf{cov} (\underline{\varepsilon}) = \mathbf{E} \left[(\underline{\varepsilon} - \mathbf{E} [\underline{\varepsilon}]) (\underline{\varepsilon} - \mathbf{E} [\underline{\varepsilon}])^T \right] = \mathbf{E} [\underline{\varepsilon} \underline{\varepsilon}^T] \quad (2.30)$$

Sabe-se que a matriz de covariância é simétrica, com diagonal principal

composta pelas variâncias das medidas, enquanto os demais elementos são as covariâncias. Por ser simétrica e diagonal dominante, a matriz $\underline{\underline{\Sigma}}$ tem determinante não nulo e é inversível. A matriz de covariância é ainda positiva-definida (LIEBMAN, 1991). Quando os erros das variáveis medidas são independentes uns dos outros (não correlacionados), as covariâncias se tornam nulas e a matriz se torna diagonal passando a ser denominada matriz de variâncias $\underline{\underline{V}} = \text{diag}[\underline{\underline{\sigma}}^2]$.

$$\underline{\underline{V}} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \quad (2.31)$$

em que σ_i^2 é a variância associada à variável medida z_i e σ_i é o desvio-padrão associado à mesma variável. A grande maioria dos problemas de RD não utiliza dados de covariâncias, admitindo a natureza diagonal de $\underline{\underline{V}}$.

Dessa forma, o problema de RD descrito pelas Equações (2.19) e (2.24) pode ser escrito da seguinte maneira:

$$\underline{\underline{\varepsilon}} = \min_{\underline{\underline{\varepsilon}}} \frac{1}{2} \underline{\underline{\varepsilon}}^T \underline{\underline{V}}^{-1} \underline{\underline{\varepsilon}} \quad (2.32)$$

sujeito a:

$$\underline{\underline{A}}(z - \underline{\underline{\varepsilon}}) = \underline{\underline{b}} \quad (2.33)$$

Desde que todas as variáveis sejam medidas, a solução analítica do problema de RD apresentado pode ser feita por meio da técnica de *Multiplicadores de Lagrange*, que reduz o problema de otimização restrito para não restrito (HIMMELBLAU, 1978). Para este tipo de problema, a função de *Lagrange* \mathcal{L} pode ser definida pela seguinte equação:

$$\mathcal{L}(\underline{\underline{x}}, \underline{\underline{\lambda}}) = f(\underline{\underline{x}}) - \underline{\underline{\lambda}}^T \underline{\underline{h}}(\underline{\underline{x}}) \quad (2.34)$$

em que $f(\underline{\underline{x}})$ é a função a ser minimizada (Fobj); $\underline{\underline{h}}(\underline{\underline{x}})$ é o vetor das equações algébricas de restrições lineares (modelo matemático) e $\underline{\underline{\lambda}}$ é o vetor dos multiplicadores de Lagrange.

Para o caso em questão, a função de *Lagrange* assume a seguinte forma:

$$\mathcal{L}(\underline{\varepsilon}, \underline{\lambda}) = \frac{1}{2} \underline{\varepsilon}^T \underline{V}^{-1} \underline{\varepsilon} - \underline{\lambda}^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}] \quad (2.35)$$

Como a matriz \underline{V} é positiva-definida e as restrições são lineares, as condições necessárias e suficientes para a minimização são dadas pelas seguintes equações (BENQLILOU, 2004; LIEBMAN, 1991; ROMAGNOLI e SANCHEZ, 1999; VEVERKA e MADRON, 1997):

$$\underline{\nabla}_{\underline{\lambda}} \mathcal{L}(\underline{\varepsilon}, \underline{\lambda}) = \underline{0} \quad (2.36)$$

$$\underline{\nabla}_{\underline{\varepsilon}} \mathcal{L}(\underline{\varepsilon}, \underline{\lambda}) = \underline{0} \quad (2.37)$$

$$\begin{aligned} \underline{\nabla}_{\underline{\lambda}} \left\{ \frac{1}{2} \underline{\varepsilon}^T \underline{V}^{-1} \underline{\varepsilon} - \underline{\lambda}^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}] \right\} &= \underline{0} \\ \frac{1}{2} \underline{\nabla}_{\underline{\lambda}} [\underline{\varepsilon}^T \underline{V}^{-1} \underline{\varepsilon}] - \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T \underline{A} z] + \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T \underline{A} \underline{\varepsilon}] + \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T \underline{b}] &= \underline{0} \\ \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T \underline{A} \underline{\varepsilon}] - \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T \underline{A} z] + \underline{\nabla}_{\underline{\lambda}} [\underline{\varepsilon}^T \underline{A}^T \underline{\lambda}] - \underline{\nabla}_{\underline{\lambda}} [z^T \underline{A}^T \underline{\lambda}] + \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T] \underline{b} &= \underline{0} \\ \underline{\nabla}_{\underline{\lambda}} [\underline{\lambda}^T] \{ \underline{A} [\underline{\varepsilon} - z] + \underline{b} \} &= \underline{0} \\ \underline{I} \{ \underline{A} [\underline{\varepsilon} - z] + \underline{b} \} &= \underline{0} \\ \underline{A} [\underline{\varepsilon} - z] + \underline{b} &= \underline{0} \end{aligned} \quad (2.38)$$

$$\begin{aligned} \underline{\nabla}_{\underline{\varepsilon}} \left\{ \frac{1}{2} \underline{\varepsilon}^T \underline{V}^{-1} \underline{\varepsilon} - \underline{\lambda}^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}] \right\} &= \underline{0} \\ \frac{1}{2} \underline{\nabla}_{\underline{\varepsilon}} [\underline{\varepsilon}^T \underline{V}^{-1} \underline{\varepsilon}] - \underline{\nabla}_{\underline{\varepsilon}} [\underline{\lambda}^T \underline{A} z] + \underline{\nabla}_{\underline{\varepsilon}} [\underline{\lambda}^T \underline{A} \underline{\varepsilon}] + \underline{\nabla}_{\underline{\varepsilon}} [\underline{\lambda}^T \underline{b}] &= \underline{0} \\ \frac{1}{2} 2 \underline{V}^{-1} \underline{\varepsilon} + \underline{\nabla}_{\underline{\varepsilon}} [\underline{\lambda}^T \underline{A} \underline{\varepsilon}] + \underline{\nabla}_{\underline{\varepsilon}} [\underline{\varepsilon}^T \underline{A}^T \underline{\lambda}] &= \underline{0} \\ \underline{V}^{-1} \underline{\varepsilon} + \underline{I} \underline{A}^T \underline{\lambda} &= \underline{0} \\ \underline{V} \underline{V}^{-1} \underline{\varepsilon} + \underline{V} \underline{A}^T \underline{\lambda} &= \underline{0} \\ \underline{I} \underline{\varepsilon} + \underline{V} \underline{A}^T \underline{\lambda} &= \underline{0} \\ \underline{\varepsilon} &= -\underline{V} \underline{A}^T \underline{\lambda} \end{aligned} \quad (2.39)$$

Explicitando $\underline{\lambda}$ a partir da substituição da Equação (2.39) na Equação (2.38).

$$\begin{aligned}
\underline{A} [-\underline{V} \underline{A}^T \lambda - \underline{z}] + \underline{b} &= \underline{0} \\
-\underline{A} \underline{V} \underline{A}^T \lambda &= \underline{A} \underline{z} - \underline{b} \\
\lambda &= - [\underline{A} \underline{V} \underline{A}^T]^{-1} [\underline{A} \underline{z} - \underline{b}]
\end{aligned} \tag{2.40}$$

Substituindo a Equação (2.40) na Equação (2.39), pode-se explicitar o vetor $\underline{\varepsilon}$, eliminando-se os *Multiplicadores de Lagrange* da equação. Entretanto, a Equação (2.41) representa agora o erro da reconciliação, ou seja, do resultado final da minimização do erros de medida satisfazendo simultaneamente as restrições. Então, o erro de medição passa a ser o erro estimado da RD, $\underline{\varepsilon} \rightarrow \hat{\underline{\varepsilon}}$, pois agora o erro está relacionado ao ajuste da reconciliação e não somente ao erro de medição.

$$\hat{\underline{\varepsilon}} = \underline{V} \underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} [\underline{A} \underline{z} - \underline{b}] \tag{2.41}$$

Substituindo a Equação (2.41) na Equação (2.26), é possível obter o vetor das variáveis reconciliadas $\hat{\underline{z}}$:

$$\hat{\underline{z}} = \underline{z} - \underline{V} \underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} [\underline{A} \underline{z} - \underline{b}] \tag{2.42}$$

A Equação (2.42) consiste na solução do problema de RD quando todas as variáveis são medidas, as restrições são lineares e o processo opera em estado estacionário, sem que existam limites nas variáveis.

O valor esperado de uma variável reconciliada pode ser demonstrada aplicando o operador esperança na Equação (2.42), em ambos os lados. Além disso, a variável \underline{z} deve ser substituída pela Equação (2.25).

$$\mathbf{E} [\hat{\underline{z}}] = \mathbf{E} [\underline{x}] + \mathbf{E} [\underline{\varepsilon}] - \mathbf{E} \left[\underline{V} \underline{A}^T (\underline{A} \underline{V} \underline{A}^T)^{-1} (\underline{A} \underline{x} - \underline{b}) \right] - \mathbf{E} \left[\underline{V} \underline{A}^T (\underline{A} \underline{V} \underline{A}^T)^{-1} (\underline{A} \underline{\varepsilon}) \right] \tag{2.43}$$

Entretanto, \underline{V} e \underline{A} são matrizes determinísticas, \underline{x} é um vetor determinístico e, com isso, $\mathbf{E} [\underline{x}] = \underline{x}$ e $\mathbf{E} [\underline{\varepsilon}] = \underline{0}$. Assim, a Equação (2.43) reduz para:

$$\mathbf{E} [\hat{\underline{z}}] = \underline{x} - \mathbf{E} \left[\underline{V} \underline{A}^T (\underline{A} \underline{V} \underline{A}^T)^{-1} (\underline{A} \underline{x} - \underline{b}) \right] \tag{2.44}$$

Substituindo a Equação (2.28) de restrições, ou seja, $\underline{A} \underline{x} - \underline{b} = \underline{0}$ na Equação

(2.44) obtém:

$$\mathbf{E} [\hat{z}] = \underline{x} \quad (2.45)$$

O que mostra que os valores reconciliados são estimativas imparciais para as medições brutas.

A matriz de covariâncias dos erros estimados da reconciliado, $\underline{\underline{\Sigma}}_{\hat{\varepsilon}}$, podem ser demonstrados analiticamente. Aplicando o operador esperança na forma da Equação (2.30) na Equação (2.41), temos que:

$$\begin{aligned} \underline{\underline{\Sigma}}_{\hat{\varepsilon}} &= \mathbf{E} [\hat{\varepsilon}\hat{\varepsilon}^T] = \mathbf{E} \left[\underline{V}\underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} [\underline{A} z - \underline{b}] [\underline{A} z - \underline{b}]^T [\underline{A} \underline{V} \underline{A}^T]^{-1} \underline{A}\underline{V}^T \right] \\ \underline{\underline{\Sigma}}_{\hat{\varepsilon}} &= \mathbf{E} [\hat{\varepsilon}\hat{\varepsilon}^T] = \mathbf{E} \left[\underline{V}\underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} [\underline{A} z - \underline{b}] [z^T \underline{A}^T - \underline{b}^T] [\underline{A} \underline{V} \underline{A}^T]^{-1} \underline{A}\underline{V}^T \right] \\ \underline{\underline{\Sigma}}_{\hat{\varepsilon}} &= \mathbf{E} [\hat{\varepsilon}\hat{\varepsilon}^T] = \underline{V}\underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} \mathbf{E} \left[[\underline{A} z - \underline{b}] [z^T \underline{A}^T - \underline{b}^T] \right] [\underline{A} \underline{V} \underline{A}^T]^{-1} \underline{A}\underline{V}^T \end{aligned} \quad (2.46)$$

Essa demonstração é encontrada em alguns livros e teses, porém sempre resumida e na grande maioria dos casos o termo linear \underline{b} do modelo é nulo, o que facilita as demonstrações. Por exemplo, no livro do ROMAGNOLI e SANCHEZ (1999) essa demonstração é bastante breve, sem muitos detalhes. Entretanto, para uma demonstração mais detalhada vamos expandir o termo dentro do operador esperança da Equação (2.46). Deve-se notar que $\mathbf{E} [z] = \mathbf{E} [x + \varepsilon] = \underline{x}$, $\underline{A}x = \underline{b}$, $x^T \underline{A}^T = \underline{b}^T$ e $\mathbf{E} [\varepsilon\varepsilon^T] = \underline{V}$. Com isso, temos que:

$$\begin{aligned} \mathbf{E} \left[[\underline{A} z - \underline{b}] [z^T \underline{A}^T - \underline{b}^T] \right] &= \mathbf{E} \left[\underline{A}z z^T \underline{A}^T - \underline{b} z^T \underline{A}^T - \underline{A} z \underline{b}^T + \underline{b} \underline{b}^T \right] \\ &= \underline{A} \mathbf{E} [z z^T] \underline{A}^T - \underline{b} \mathbf{E} [z^T] \underline{A}^T - \underline{A} \mathbf{E} [z] \underline{b}^T + \underline{b} \underline{b}^T \\ &= \underline{A} \mathbf{E} \left[(x + \varepsilon) (x + \varepsilon)^T \right] \underline{A}^T - \underline{b} x^T \underline{A}^T - \underline{A} x \underline{b}^T + \underline{b} \underline{b}^T \\ &= \underline{A} \mathbf{E} \left[\begin{array}{c} xx^T + \cancel{x\varepsilon^T} + \cancel{\varepsilon x^T} + \varepsilon\varepsilon^T \end{array} \right] \underline{A}^T - \underline{b} \underline{b}^T - \underline{b} \underline{b}^T + \underline{b} \underline{b}^T \\ &= \underline{A} x x^T \underline{A}^T + \underline{A} \mathbf{E} [\varepsilon\varepsilon^T] \underline{A}^T - \underline{b} \underline{b}^T \\ &= \underline{b} \underline{b}^T + \underline{A} \underline{V} \underline{A}^T - \underline{b} \underline{b}^T = \underline{A} \underline{V} \underline{A}^T \end{aligned} \quad (2.47)$$

Com isso, podemos concluir a Equação (2.46) e expressar a matriz de covariâncias dos erros de medida estimados.

$$\begin{aligned}
\underline{\underline{\Sigma}}_{\hat{\varepsilon}} &= \mathbf{E} [\underline{\varepsilon}\underline{\varepsilon}^T] = \underline{V}\underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} \underline{A}\underline{V}\underline{A}^T \underline{A}\underline{V}\underline{A}^T [\underline{A}\underline{V} \underline{A}^T]^{-1} \underline{A}\underline{V}^T \\
\underline{\underline{\Sigma}}_{\hat{\varepsilon}} &= \mathbf{E} [\underline{\varepsilon}\underline{\varepsilon}^T] = \underline{V}\underline{A}^T [\underline{A} \underline{V} \underline{A}^T]^{-1} \underline{A}\underline{V}^T
\end{aligned} \tag{2.48}$$

Também podemos calcular analiticamente a matriz de covariâncias sobre a reconciliação $\underline{\underline{z}}$, ou seja, sobre o ajuste do valor medido. Aplicando o operador covariância na Equação(2.42), temos que:

$$\begin{aligned}
\underline{\underline{\Sigma}}_{\hat{z}} &= \mathbf{cov}(\underline{z}) - \mathbf{cov}(\hat{\varepsilon}) = \mathbf{cov}(\underline{z}) - \mathbf{E} [\hat{\varepsilon}\hat{\varepsilon}^T] \\
\underline{\underline{\Sigma}}_{\hat{z}} &= \mathbf{E} \left[(\underline{z} - \mathbf{E}[\underline{z}]) (\underline{z} - \mathbf{E}[\underline{z}])^T \right] - \underline{\underline{\Sigma}}_{\hat{\varepsilon}} \\
\underline{\underline{\Sigma}}_{\hat{z}} &= \mathbf{E} \left[(\underline{x} + \underline{\varepsilon} - \mathbf{E}[\underline{x} + \underline{\varepsilon}]) (\underline{x} + \underline{\varepsilon} - \mathbf{E}[\underline{x} + \underline{\varepsilon}])^T \right] - \underline{\underline{\Sigma}}_{\hat{\varepsilon}} \\
\underline{\underline{\Sigma}}_{\hat{z}} &= \mathbf{E} \left[(\underline{x} + \underline{\varepsilon} - \underline{x}) (\underline{x} + \underline{\varepsilon} - \underline{x})^T \right] - \underline{\underline{\Sigma}}_{\hat{\varepsilon}} \\
\underline{\underline{\Sigma}}_{\hat{z}} &= \mathbf{E} [\underline{\varepsilon}\underline{\varepsilon}^T] - \underline{\underline{\Sigma}}_{\hat{\varepsilon}} \\
\underline{\underline{\Sigma}}_{\hat{z}} &= \underline{V} - \underline{\underline{\Sigma}}_{\hat{\varepsilon}}
\end{aligned} \tag{2.49}$$

A Equação (2.49) mostra que a variabilidade do dado reconciliado será sempre menor que o dados medido. Pois, o ajuste do dado é proporcional ao erro estimado da reconciliação. De fato, se os valores da matriz de covariância dos erros estimados de reconciliação, $\underline{\underline{\Sigma}}_{\hat{\varepsilon}}$, for significativamente maior que os valores da matriz de covariância dos erros de medição, \underline{V} , então o ajuste será grande, dando uma variabilidade ao dado reconciliado bem menor que a variabilidade do dado bruto.

RD estacionária linearmente restrita: Generalizado para Estimadores-M

Claramente, a função objetivo (estimador) de mínimos quadrados ponderado na formulação anterior vem do pressuposto de que todas as medidas são normalmente distribuídas, sem levar em conta erros grosseiros que podem estar presentes (ALHAJ-DIBO *et al.*, 2008). A influência desses erros grosseiros nas estimativas pode ser minimizada utilizando estimadores de máxima verossimilhança (estimadores-M) oriundo da estatística robusta, detalhada no Capítulo 3 e no Apêndice B. Por exemplo, vamos considerar um estimador-M generalizado $\rho(\underline{\varepsilon}, \underline{V})$:

$$\underline{\varepsilon} = \min_{\underline{\varepsilon}} \rho(\underline{\varepsilon}, \underline{V}) \quad (2.50)$$

sujeito a:

$$\underline{A}(z - \underline{\varepsilon}) = \underline{b} \quad (2.51)$$

O problema descrito pelas Equações (2.50) e (2.51) pode ser solucionado aplicando a Equação (2.34). Para este caso, a função de *Lagrange* assume a seguinte forma:

$$\mathcal{L}(\underline{\varepsilon}, \underline{\lambda}) = \rho(\underline{\varepsilon}, \underline{V}) - \underline{\lambda}^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}] \quad (2.52)$$

Como a matriz \underline{V} é constante e as restrições são lineares, as condições necessárias e suficientes para a minimização são dadas pelas Equações (2.36) e (2.37). Entretanto, antes de aplicar as condições necessárias e suficientes para a minimização é importante definir algumas propriedades dos estimadores-M robustos, detalhada no Capítulo 3 e no Apêndice B. Para solucionar este problema precisamos definir duas funções importantes dos estimadores-M, a Função de Influência $\psi(u)$ e a Função Peso $w(u)$:

$$\psi(u) = \frac{\partial \rho(u)}{\partial u} \quad (2.53)$$

$$w(u) = \frac{\psi(u)}{u} \quad (2.54)$$

Com isso, aplicando as condições necessárias e suficiente, temos que:

$$\begin{aligned} \nabla_{\underline{\lambda}} [\rho(\underline{\varepsilon}, \underline{V}) - \underline{\lambda}^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}]] &= \underline{0} \\ \cancel{\nabla_{\underline{\lambda}} [\rho(\underline{\varepsilon}, \underline{V})]} \xrightarrow{0} - \cancel{\nabla_{\underline{\lambda}} [\underline{\lambda}^T]} \xrightarrow{I} [\underline{A}(z - \underline{\varepsilon}) - \underline{b}] &= \underline{0} \\ \underline{A}[\underline{\varepsilon} - z] + \underline{b} &= \underline{0} \end{aligned} \quad (2.55)$$

$$\begin{aligned}
\nabla_{\underline{\varepsilon}} [\rho(\underline{\varepsilon}, \underline{V}) - \lambda^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}]] &= \underline{0} \\
\nabla_{\underline{\varepsilon}} [\rho(\underline{\varepsilon}, \underline{V})] - \nabla_{\underline{\varepsilon}} [\lambda^T [\underline{A}(z - \underline{\varepsilon}) - \underline{b}]] &= \underline{0} \\
\underline{W}_{\underline{\varepsilon}} \underline{\varepsilon} - \nabla_{\underline{\varepsilon}} \left[\lambda^T \underline{A} z - \lambda^T \underline{A} \underline{\varepsilon} - \lambda^T \underline{b} \right] &= \underline{0} \\
\underline{W}_{\underline{\varepsilon}} \underline{\varepsilon} + \nabla_{\underline{\varepsilon}} \left[\lambda^T \underline{A} \underline{\varepsilon} + \lambda^T \underline{A}^T \lambda \right] &= \underline{0} \\
\underline{W}_{\underline{\varepsilon}} \underline{\varepsilon} + \underline{A}^T \lambda &= \underline{0}
\end{aligned} \tag{2.56}$$

Substituindo a Equação (2.26) na Equação (2.56)

$$\begin{aligned}
\underline{W}_{\underline{\varepsilon}} (z - \hat{z}) + \underline{A}^T \lambda &= \underline{0} \\
\underline{W}_{\underline{\varepsilon}} z - \underline{W}_{\underline{\varepsilon}} \hat{z} + \underline{A}^T \lambda &= \underline{0} \\
\hat{z} &= z + \underline{W}_{\underline{\varepsilon}}^{-1} \underline{A}^T \lambda
\end{aligned} \tag{2.57}$$

Substituindo a Equação (2.57) e a Equação (2.26) na Equação (2.55), então temos:

$$\begin{aligned}
-\underline{A} \left[z + \underline{W}_{\underline{\varepsilon}}^{-1} \underline{A}^T \lambda \right] + \underline{b} &= \underline{0} \\
\underline{A} z + \underline{A} \underline{W}_{\underline{\varepsilon}}^{-1} \underline{A}^T \lambda &= \underline{b} \\
\lambda &= \left[\underline{A} \underline{W}_{\underline{\varepsilon}}^{-1} \underline{A}^T \right]^{-1} (\underline{b} - \underline{A} z)
\end{aligned} \tag{2.58}$$

Substituindo a Equação (2.58) na Equação (2.57), obtemos:

$$\hat{z} = z - \underline{W}_{\underline{\varepsilon}}^{-1} \underline{A}^T \left[\underline{A} \underline{W}_{\underline{\varepsilon}}^{-1} \underline{A}^T \right]^{-1} (\underline{A} z - \underline{b}) \tag{2.59}$$

Deve-se notar que a matriz de pesos $\underline{W}_{\underline{\varepsilon}} = w(\underline{\varepsilon}, \underline{V})$ é diagonal e depende do ajuste do erro $\underline{\varepsilon} = z - \hat{z}$, ou seja, depende da variável reconciliada \hat{z} , visto que a variável medida z e a matriz de variâncias \underline{V} são constantes. Portanto, a expressão da Equação (2.59) é implícita. Entretanto, a solução pode ser obtida de forma iterativa, atualizando a cada passo a matriz de pesos $\underline{W}_{\underline{\varepsilon}}$. Para a k -ésima etapa deste cálculo iterativo, vamos denotar $\hat{z}^{(k)}$ e $\underline{W}_{\underline{\varepsilon}}^{(k)}$ como o valor reconciliado e a matriz de pesos, respectivamente, na iteração k . A estimativa inicial da iteração é

$\hat{z}^{(0)} = z$ e a solução expressa por passo de iteração:

$$\hat{z}^{(k+1)} = z - \left[\frac{W^{(k)}}{\underline{\underline{\varepsilon}}} \right]^{-1} \underline{\underline{A}}^T \left[\underline{\underline{A}} \left[\frac{W^{(k)}}{\underline{\underline{\varepsilon}}} \right]^{-1} \underline{\underline{A}}^T \right]^{-1} (\underline{\underline{A}}z - \underline{\underline{b}}) \quad (2.60)$$

O critério de parada do procedimento iterativo pode ser condicionada por um teste de norma, como a diferença entre as variáveis reconciliadas consecutivas em cada passo da iteração, $\|\hat{z}^{(k+1)} - \hat{z}^{(k)}\| \leq \epsilon$, onde ϵ é o critério de tolerância desejado.

RD estacionária não linearmente restrita: Linearização Sucessiva

O problema de RD não linearmente restrito não apresenta necessariamente uma solução analítica; portanto, métodos de PNL ou de linearizações sucessivas são utilizados para solucionar o problema. Nesta subseção será abordada o método de linearização sucessiva para solucionar o problema de RD não linearmente restrito e fazer a classificação de variáveis do modelo de processo, com a obtenção da matriz de projeção por fatoração QR.

Em uma planta química, correntes de processo apresentam normalmente muitos componentes químicos. Assim, apenas algumas vazões (ou composições) são medidas e analisadas (em laboratório, normalmente, pois os cromatógrafos *on-line* são mais caros e algumas vezes não podem ser usados). No entanto, medidas podem apresentar erros, de maneira que nem o balanço global, nem o balanço por componente, são satisfeitos com os dados medidos. Portanto, é necessário reconciliar tanto a vazão global quanto as vazões por componente simultaneamente. Dessa forma, modelos de balanço de massa por componente são não-lineares, necessitando de uma estratégia particular de solução (NARASIMHAN e JORDACHE, 1999).

As não linearidades surgem nos modelos de processos representados por balanços de massa por componente e energia, por conta das reações químicas, das correlações termodinâmicas e das propriedades físicas. Para a RD simultânea das vazões globais e das vazões por componente, o modelo matemático usado para representar o processo utiliza o produto de duas das variáveis de interesse (vazão global \times composição do componente i). Além disso, a soma das frações molares normalizadas de todos os componentes presentes na corrente deve ser igual a 1, impondo uma restrição algébrica aos valores de composição reconciliados. Assim, o modelo constitui um sistema bilinear. Considerando que não há limites nas variáveis, é possível resolver analiticamente esse problema não linear utilizando técnicas de RD com base em transformações de variáveis (LIEBMAN, 1991). Entretanto, na grande maioria das vezes não é possível desenvolver uma solução analítica para o

problema de RD em estado estacionário e não linearmente restrito. Neste caso, para resolver este problema, geralmente, são utilizadas duas abordagens: (i) a técnica de **linearização sucessiva** proposta por KNEPPER e GORMAN (1980), que lineariza as restrições e resolve o problema linear com a solução analítica em sucessivas iterações; (ii) o uso de métodos de PNL, que pode lidar explicitamente com restrições de igualdade e desigualdade não-lineares (PRATA *et al.*, 2009). As duas abordagens conseguem lidar com o problema de ter **variáveis não medidas** \underline{u} no processo, desde que o modelo tenha observabilidade global.

A abordagem de linearização sucessiva é desenvolvida a seguir para permitir um melhor entendimento da solução do problema de RD em estado estacionário, não linearmente restrito e com variáveis não medidas. A técnica se baseia no método desenvolvido para restrições lineares, que então se estende para restrições não-lineares. O método se baseia na ideia de que as restrições são representadas pela Equação (2.61), sem considerar a ocorrência de inequações ou limites nas variáveis.

$$\underline{h}(\underline{w}) = \underline{0} \quad (2.61)$$

$$\underline{w} = \begin{bmatrix} \hat{\underline{z}} \\ \underline{u} \end{bmatrix} \quad (2.62)$$

O primeiro procedimento consiste em linearizar a Equação (2.61) por expansão em Série de Taylor em torno do valor estimado. Em geral, o valor medido é usado como estimativa inicial. As equações das restrições linearizadas são escritas da forma matricial:

$$\underline{\underline{A}}_{\hat{\underline{z}}} \hat{\underline{z}} + \underline{\underline{A}}_{\underline{u}} \underline{u} = \underline{b} \quad (2.63)$$

em que

$$\underline{\underline{A}}_{\hat{\underline{z}}}^i = \nabla_{\hat{\underline{z}}}^T \underline{h}(\hat{\underline{z}}^i, \underline{u}^i) = \begin{bmatrix} \frac{\partial h_1}{\partial \hat{z}_1} & \frac{\partial h_1}{\partial \hat{z}_2} & \dots & \frac{\partial h_1}{\partial \hat{z}_m} \\ \frac{\partial h_2}{\partial \hat{z}_1} & \frac{\partial h_2}{\partial \hat{z}_2} & \dots & \frac{\partial h_2}{\partial \hat{z}_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_k}{\partial \hat{z}_1} & \frac{\partial h_k}{\partial \hat{z}_2} & \dots & \frac{\partial h_k}{\partial \hat{z}_m} \end{bmatrix} \quad (2.64)$$

$$\underline{\underline{A}}_u^i = \nabla_{\underline{u}}^T \underline{h}(\hat{\underline{z}}^i, \underline{u}^i) = \begin{bmatrix} \frac{\partial h_1}{\partial u_1} & \frac{\partial h_1}{\partial u_2} & \dots & \frac{\partial h_1}{\partial u_l} \\ \frac{\partial h_2}{\partial u_1} & \frac{\partial h_2}{\partial u_2} & \dots & \frac{\partial h_2}{\partial u_l} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_k}{\partial u_1} & \frac{\partial h_k}{\partial u_2} & \dots & \frac{\partial h_k}{\partial u_l} \end{bmatrix} \quad (2.65)$$

$$\underline{b}^i = \underline{\underline{A}}_{\hat{\underline{z}}}^i \hat{\underline{z}}^i + \underline{\underline{A}}_u^i \underline{u}^i - \underline{h}(\hat{\underline{z}}^i, \underline{u}^i) \quad (2.66)$$

Na Equação (2.63), os vetores \underline{b} , $\hat{\underline{z}}$ e \underline{u} representam: o vetor dos coeficientes lineares das restrições linearizadas, o vetor das variáveis medidas e o vetor das variáveis não medidas, respectivamente. Nas Equações (2.64) e (2.65), m representa o número de variáveis medidas, l representa o número de variáveis não medidas, k é o número total de restrições e i é a i -ésima iteração do procedimento de linearização sucessiva. Então, para $i = 1$, $\hat{\underline{z}}^{i=1} = \underline{z}$ e $\underline{u}^{i=1} = \underline{u}_{guess}$; ou seja, para o vetor das variáveis medidas a suposição inicial é o próprio valor medido, enquanto para o vetor das variáveis não medidas a suposição pode ser calculada por balanço, utilizando os valores medidos não reconciliados. As matrizes $\underline{\underline{A}}_{\hat{\underline{z}}}$ e $\underline{\underline{A}}_u$ são as matrizes Jacobianas em relação às variáveis medidas e não medidas, respectivamente.

O problema escrito na forma da Equação (2.63) impõe o uso de uma estratégia para eliminar a matriz das variáveis não medidas. Essa estratégia tem como base o uso de transformação ortogonal. CROWE (1986) propôs uma elegante estratégia para decompor as variáveis não medidas em termos das variáveis medidas usando o modelo do processo, gerando um sistema linear, que ficou conhecida como Matriz de Projeção de Crowe. Um outra abordagem alternativa faz o uso da fatoração QR das matrizes, dissociando as variáveis não medidas das variáveis medidas. Este método de fatoração QR pode ser usado na abordagem de linearizações sucessivas e, também, para a classificação de variáveis. Uma matriz $\underline{\underline{A}}_u$ pode ser fatorada em matrizes QR na seguinte forma:

$$\underline{\underline{A}}_u = \underline{\underline{Q}}_u \underline{\underline{R}}_u \quad (2.67)$$

$$\underline{\underline{A}}_{u(k \times l)} \underline{\underline{\Pi}}_{u(l \times l)} = \underline{\underline{Q}}_{u(k \times k)} \underline{\underline{R}}_{u(k \times l)} \quad (2.68)$$

$$\underline{\underline{Q}}_{u(k \times k)} = \begin{bmatrix} \underline{\underline{Q}}_{u1(k \times l)} & \underline{\underline{Q}}_{u2(k \times k-l)} \end{bmatrix} \quad (2.69)$$

$$\underline{\underline{R}}_{u(k \times l)} = \begin{bmatrix} \underline{\underline{R}}_{u1(r \times r)} & \underline{\underline{R}}_{u2(r \times l-r)} \\ \underline{\underline{0}}_{1(k-r \times r)} & \underline{\underline{0}}_{2(k-r \times l-r)} \end{bmatrix} \quad (2.70)$$

em que k , l e r são o número de restrições, número de variáveis não medidas e o posto da matriz $rank(\underline{\underline{A}}_u) = rank(\underline{\underline{R}}_{u1}) = r$, respectivamente. $\underline{\underline{Q}}_u$ é uma matriz ortogonal, $\underline{\underline{R}}_{u1}$ é uma matriz triangular superior não singular de dimensão r e $\underline{\underline{\Pi}}_u$ é uma matriz de permutação, obtida a partir das colunas da matriz $\underline{\underline{I}}_{(l \times l)}$, permutadas para ordenar e classificar as variáveis não medidas em dois tipos (ROMAGNOLI e SANCHEZ, 1999). O vetor das variáveis não medidas ainda pode ser particionado em dois subvetores: vetor das **variáveis não medida observáveis** $\underline{u}_{1(r \times 1)}$ e vetor das **variáveis não medidas não observáveis** $\underline{u}_{2(l-r \times 1)}$:

$$\underline{\underline{\Pi}}_u^T \underline{u} = \begin{bmatrix} \underline{u}_{1(r \times 1)} \\ \underline{u}_{2(l-r \times 1)} \end{bmatrix} \quad (2.71)$$

Substituindo-se a Equação (2.67) na Equação (2.63) e, em seguida, pré-multiplicando-se toda a equação por $\underline{\underline{Q}}_u^T = \underline{\underline{Q}}_u^{-1}$, pois $\underline{\underline{Q}}_u^T \underline{\underline{A}}_u \underline{\underline{\Pi}}_u = \underline{\underline{Q}}_u^T \underline{\underline{Q}}_u \underline{\underline{R}}_u = \underline{\underline{I}} \underline{\underline{R}}_u = \underline{\underline{R}}_u$, então se obtém:

$$\underline{\underline{Q}}_{u1}^T \underline{\underline{A}}_{\hat{z}} \hat{z} + \underline{\underline{R}}_{u1} \underline{u}_1 + \underline{\underline{R}}_{u2} \underline{u}_2 = \underline{\underline{Q}}_{u1}^T \underline{b} \quad (2.72)$$

$$\begin{aligned} \underline{\underline{Q}}_{u2}^T \underline{\underline{A}}_{\hat{z}} \hat{z} + \underline{\underline{0}}_{u1} \underline{u}_1 + \underline{\underline{0}}_{u2} \underline{u}_2 &= \underline{\underline{Q}}_{u2}^T \underline{b} \\ \underline{\underline{Q}}_{u2}^T \underline{\underline{A}}_{\hat{z}} \hat{z} &= \underline{\underline{Q}}_{u2}^T \underline{b} \end{aligned} \quad (2.73)$$

Assim, a restrição do problema passa a ser escrita conforme a Equação (2.73), que substituído na Equação (2.42), resulta na solução:

$$\hat{\underline{z}} = \underline{z} - \underline{\underline{V}} \left(\underline{\underline{Q}}_{u2}^T \underline{\underline{A}}_{\hat{z}} \right)^T \left[\left(\underline{\underline{Q}}_{u2}^T \underline{\underline{A}}_{\hat{z}} \right) \underline{\underline{V}} \left(\underline{\underline{Q}}_{u2}^T \underline{\underline{A}}_{\hat{z}} \right)^T \right]^{-1} \left[\left(\underline{\underline{Q}}_{u2}^T \underline{\underline{A}}_{\hat{z}} \right) \underline{z} - \underline{\underline{Q}}_{u2}^T \underline{b} \right] \quad (2.74)$$

Explicitando \underline{u}_1 na Equação (2.72), obtém-se a base para estimar a variável não medida e, também, para o procedimento de classificação de variáveis do modelo de processo.

$$\underline{u}_1 = \underline{\underline{R}}_{u1}^{-1} \underline{\underline{Q}}_{u1}^T \underline{b} - \underline{\underline{R}}_{u1}^{-1} \underline{\underline{Q}}_{u1}^T \underline{\underline{A}}_{u1} \underline{\underline{z}} - \underline{\underline{R}}_{u1}^{-1} \underline{\underline{R}}_{u2} \underline{u}_2 \quad (2.75)$$

A análise da classificação de variáveis por fatoração QR consiste na eliminação do vetor \underline{u}_2 da Equação (2.75). Entretanto, a única possibilidade de eliminação do vetor da equação é se ele for $\underline{u}_2 = \underline{0}$; ou seja, não podem existir variáveis não medidas não observáveis no sistema. Para isso, a matriz das variáveis não medidas $\underline{\underline{A}}_u$ deve satisfazer as seguintes propriedades (VEVERKA e MADRON, 1997):

$$r = l = \text{rank}(\underline{\underline{A}}_u) = \text{rank}(\underline{\underline{R}}_{u1}) \quad (2.76)$$

$$\underline{\underline{R}}_{u(k \times r=l)} = \begin{bmatrix} \underline{\underline{R}}_{u1(r \times r)} & \underline{\underline{R}}_{u2(r \times l-r)} \\ \underline{\underline{0}}_{1(k-r \times r)} & \underline{\underline{0}}_{2(k-r \times l-r)} \end{bmatrix} \quad (2.77)$$

Dessa forma, todas as variáveis não medidas são observáveis e podem ser estimadas, possibilitando a solução do problema de RD. A classificação de variáveis identifica o número mínimo de variáveis a serem medidas. Pela ordenação da matriz de permutação, identificam-se e separam-se as variáveis não observáveis do sistema. Após a classificação das variáveis o sistema pode ser reconciliado com as Equações (2.64), (2.65), (2.66), (2.74) e (2.78) em iterações sucessivas.

$$\underline{u}_1 = \underline{\underline{R}}_{u1}^{-1} \underline{\underline{Q}}_{u1}^T \underline{b} - \underline{\underline{R}}_{u1}^{-1} \underline{\underline{Q}}_{u1}^T \underline{\underline{A}}_{u1} \underline{\underline{z}} \quad (2.78)$$

O método de linearizações sucessivas utiliza as variáveis calculadas e estimadas a partir da primeira iteração para atualizar a iteração seguinte, $i = 1, 2, \dots$, repetindo-se o procedimento até que se atinja o critério de tolerância desejado.

O método de linearização sucessiva também pode ser acoplado com o método de RD em estado estacionário generalizado para estimadores-M robustos. Assim, a solução deste problema de RD robusta e não linearmente restrita em estado estacionário pode ser expressa por:

$$\hat{z}^{(k+1)} = \underline{z} - \left\{ \underline{W}_\varepsilon^{-1} \left(\underline{Q}_{\underline{u}_2}^T \underline{A}_{\underline{z}} \right)^T \left[\underline{Q}_{\underline{u}_2}^T \underline{A}_{\underline{z}} \underline{W}_\varepsilon^{-1} \left(\underline{Q}_{\underline{u}_2}^T \underline{A}_{\underline{z}} \right)^T \right]^{-1} \right\}^{(k)} \left[\left(\underline{Q}_{\underline{u}_2}^T \underline{A}_{\underline{z}} \right)^{(k)} \underline{z} - \underline{b}^{(k)} \right] \quad (2.79)$$

$$\underline{u}_1^{(k+1)} = \left(\underline{R}_{\underline{u}_1}^{-1} \underline{Q}_{\underline{u}_1}^T \underline{b} \right)^{(k)} - \left(\underline{R}_{\underline{u}_1}^{-1} \underline{Q}_{\underline{u}_1}^T \underline{A}_{\underline{z}} \right)^{(k)} \hat{z}^{(k+1)} \quad (2.80)$$

Entretanto, o procedimento iterativo da linearização ocorre tanto na matriz das restrições $\underline{A}_{\underline{z}}, \underline{A}_{\underline{u}}$, quanto na matriz de pesos $\underline{W}_\varepsilon$, oriunda de estimares-M não quadráticos (robustos). Com isso, deve-se atentar com a possibilidade de problemas de convergência, pois pode aumentar a dificuldade de convergência com a não linearidade de modelos de restrição juntamente com a não convexidade de estimadores-M robustos.

Estudo de Caso: Medidores de vazão de uma bomba

Um problema simples, porém didático, pode ser esquematizado em termos de uma bomba centrífuga e dois medidores de vazão, um posicionado na entrada e outro posicionado na saída da bomba, como ilustra a Figura 2.7. Para este problema, admite-se que os medidores tenham precisões diferentes. A Tabela 2.3 indica as vazões que entram e saem da bomba, assim como as variâncias dos erros de medida de cada medidor. O exemplo simula a realidade, observando-se a presença de erros, que fazem com que o balanço de massa global não feche, como indica o resíduo da restrição ($z_1 - z_2 = e = 0,5 \text{ kmol/s}$), conforme impõe o modelo na Equação (2.81).

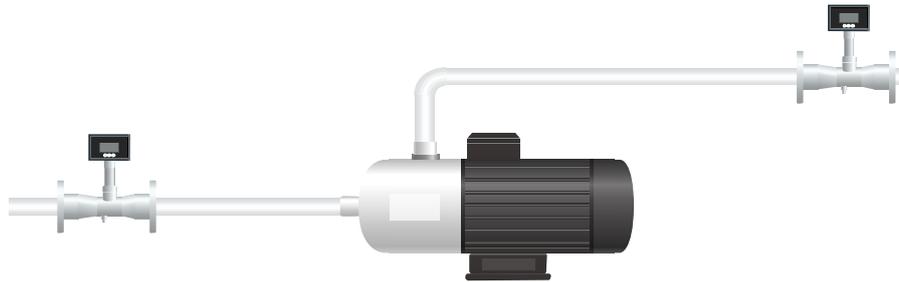


Figura 2.7: Bomba centrífuga (problema estacionário linearmente restrito).

$$\hat{z}_1 - \hat{z}_2 = 0 \quad (2.81)$$

Admite-se ainda, que os medidores apresentem correlação entre si. Assim, o coeficiente de correlação de Pearson, ρ_{ij} , definido pela Equação (2.82), deve ser

Tabela 2.3: Dados do exemplo motivacional - bomba centrífuga

Variáveis/parâmetros	Valor medido [$kmol/s$]	Variância [$kmol/s$] ²
Entrada (1)	1,5 (z_1)	0,36 (σ_1^2)
Saída (2)	2,0 (z_2)	0,64 (σ_2^2)

apresentado. Para este problema, admite-se que a correlação seja diretamente igual a $\rho_{12} = 0,70$.

$$\begin{aligned} \rho_{ij} = \rho_{ji} &= \frac{\sum_{t=1}^n (z_{ti} - \text{mean}_i(\underline{z}_i))(z_{tj} - \text{mean}_j(\underline{z}_j))}{\sqrt{\sum_{t=1}^n (z_{ti} - \text{mean}_i(\underline{z}_i))^2} \sqrt{\sum_{t=1}^n (z_{tj} - \text{mean}_j(\underline{z}_j))^2}} = \\ &= \frac{\text{cov}(\underline{z}_i, \underline{z}_j)}{\sqrt{\text{var}(\underline{z}_i)\text{var}(\underline{z}_j)}} \quad \forall i \neq j \end{aligned} \quad (2.82)$$

Assim, é possível também definir as covariâncias entre as duas variáveis medidas:

$$\sigma_{12}^2 = \sigma_{21}^2 = \rho_{12} \sqrt{\sigma_1^2 \sigma_2^2} = 0,70 \sqrt{0,36 \cdot 0,64} = 0,336 \text{ [} kmol/s \text{]}^2 \quad (2.83)$$

Desta forma, a matriz das variâncias-covariâncias $\underline{\underline{\Sigma}}$, a matriz das restrições $\underline{\underline{A}}$, conforme mostra a Equação (2.81), e o vetor das variáveis medidas \underline{z} podem ser definidas na forma:

$$\underline{\underline{\Sigma}} = \begin{bmatrix} 0,36 & 0,336 \\ 0,336 & 0,64 \end{bmatrix} \quad (2.84)$$

$$\underline{\underline{A}} = \begin{bmatrix} 1 & -1 \end{bmatrix} \quad \text{e} \quad \underline{b} = \begin{bmatrix} 0 \end{bmatrix} \quad (2.85)$$

$$\underline{z} = \begin{bmatrix} 1,5 \\ 2,0 \end{bmatrix} \quad (2.86)$$

Aplicando os dados à Equação (2.42):

$$\hat{\underline{z}} = \begin{bmatrix} 1,5 \\ 2,0 \end{bmatrix} + \begin{bmatrix} 0,36 & 0,336 \\ 0,336 & 0,64 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \left\{ \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 0,36 & 0,336 \\ 0,336 & 0,64 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}^{-1} \cdot \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1,5 \\ 2,0 \end{bmatrix} \quad (2.87)$$

Então, os valores reconciliados para este caso são:

$$\hat{\underline{z}} = \begin{bmatrix} 1,5366 \\ 1,5366 \end{bmatrix}, \quad (2.88)$$

com a restrição satisfeita $e = \hat{z}_1 - \hat{z}_2 = 0,000 \text{ kmol/s}$

Assim, o balanço de massa global é respeitado e, ao mesmo tempo, a probabilidade de ocorrência daquela medida ser “verdadeira” é maximizada. A solução pode ser interpretada graficamente por ter apenas 2 variáveis medidas, como mostrado na Figura 2.8.

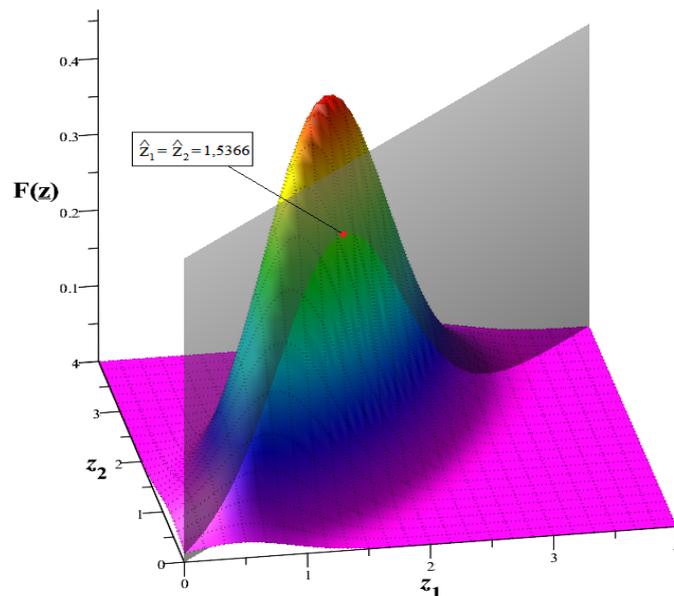


Figura 2.8: Interpretação gráfica da solução do problema de RD linearmente restrito proposto no Estudo de Caso 01.

2.4.2 Reconciliação de Dados Dinâmica

Modelos estacionários são bastante úteis para prever informações de um processo em regime permanente, para projetar equipamentos, para estimar os custos de um empreendimento, verificar condições ótimas de operação e outras aplicações (DOUGLAS, 1988). Entretanto, os processos estão frequentemente sujeitos a oscilações e perturbações por muitos motivos. Para cada perturbação em uma variável de entrada do sistema, as variáveis de saída do processo responde dinamicamente às perturbações, podendo apresentar relações não-lineares entre essas variáveis. Por este motivo, modelos estacionários são incapazes de prever as relações dinâmica entre essas variáveis do sistema.

A dinâmica de um processo pode ser modelada classicamente por sistemas de equações algébrico-diferenciais, ou por equações algébricas de recorrência, após discretização apropriada. De maneira geral, sistemas dinâmicos são representações matemáticas das dependências temporais das variáveis de processo, sendo muitas vezes não-lineares e contendo parâmetros e condições iniciais desconhecidas. Modelos dinâmicos são mais representativos do comportamento real do processo que modelos estacionários; entretanto, o preço da representatividade matemática está usualmente associada ao custo numérico/analítico mais elevado para a solução (BAGAJEWICZ e JIANG, 1997). Portanto, novas técnicas e estratégias matemáticas têm sido propostas para solucionar o problema de Reconciliação de Dados Dinâmica (RDD), em que o modelo do processo é dinâmico.

Os pioneiros no estudo de RDD foram GERTLER e ALMASY (1973), usando sistemas lineares dinâmicos resultantes de balanços de massa. Filtro de Kalman (KF) (KALMAN, 1960) foi utilizado por STANLEY e MAH (1977) para solucionar o problema de RDD com condições de estado quasi-estacionário. Os autores demonstraram com sucesso a aplicação da técnica para reconciliar vazões e temperaturas de uma unidade de destilação de petróleo, utilizando como modelo do processo equações de balanço de massa e energia das correntes.

O KF foi proposto para estimar os estados e a matriz de variâncias-covariâncias de um sistema dinâmico linear a partir de uma série temporal de medições ruidosas. O filtro é aplicado a cada passo de tempo, com o novo estado estimado e suas respectivas variâncias-covariâncias estimadas, gerando-se novas predições para a próxima iteração. Portanto, o filtro de Kalman é um estimador que requer apenas a última estimativa, e não o histórico completo do dados, para estimar os estados de um sistema e avaliar calcular o próximo ponto (KALMAN, 1960).

Entretanto, o KF não estima bem os estados quando o processo opera

em regiões significativamente não-lineares, como na grande maioria dos casos de engenharia química. A linearização do sistema pode levar à obtenção das estimativas imprecisas dos estados. Além disso, o KF é restrito a estimativas dos erros de medida Normalmente distribuídos, o que pode levar a estimativas imprecisas na presença de erros grosseiros (LIEBMAN *et al.*, 1992).

A abordagem de Horizonte Móvel (HM) ou janela móvel foi introduzida por KIM *et al.* (1991) para solucionar um problema de otimização dinâmica com restrições algébrico-diferenciais, utilizando PNL e integração numérica para resolver o problema de RDD. O estudo de caso foi um reator CSTR adiabático. Além de reconciliar as variáveis medidas, alguns parâmetros do processo também foram estimados simultaneamente.

LIEBMAN *et al.* (1992) utilizou uma estratégia para solucionar problemas de RDEP em sistemas dinâmicos não-lineares com HM. A estratégia consistiu em utilizar técnica de Programação Quadrática Sucessiva (SQP) para minimizar o estimador MQP sujeito as restrições. Para reduzir o sistema algébrico-diferencial a um sistema algébrico, os autores discretizaram as equações diferenciais por colocação ortogonal.

Variantes da abordagem que utiliza o HM foram estudadas por ALBUQUERQUE e BIEGLER (1995); entretanto, as maiores diferenças foram associadas à técnica de redução do sistema algébrico-diferencial, já que os autores utilizaram o método padrão de Runge-Kutta implícito para a discretização. Técnicas baseadas em RNA foram utilizadas para modelar o sistema usando dados históricos. Com um modelo mais simples, sugere-se que a solução do problema de RDD fica menos custosa computacionalmente (PRATA, 2009). Nesta mesma linha BENQLILOU (2004) reformulou o problema de RDD para usar modelos baseados em análise de *Wavelets*, também aplicando a abordagem de HM.

PRATA (2009) desenvolveu um estudo de RDEP utilizando o conceito de HM e estimadores-M robustos de Welsch, para a detecção e eliminação de erros grosseiros simultaneamente. O caso estudado pelo autor foi um reator de polimerização isotérmico. O autor não seguiu a abordagem tradicional de reduzir o modelo diferencial. Para cada iteração do otimizador, o modelo era solucionado, fazendo com que os balanços de massa e energia fossem sempre respeitados a cada iteração. Para minimizar o estimador robusto, foi utilizado o método estocástico (não determinístico) de PSO, que não quer o cômputo de derivadas direcionais e dá um caráter mais global para a solução.

Processos industriais reais que operam em múltiplas condições operacionais estão sujeitos a muitos efeitos, como erros de medição dos dados, variações não-

lineares dos parâmetros do modelo a cada condição de operação e diferentes tipos de erros grosseiros. Todos esses efeitos aumentam a dificuldade de estimar os parâmetros do modelo de processo. ZHANG *et al.* (2014) abordou a influência desses efeitos sobre os problemas RDEP. Os autores propuseram uma metodologia de RDEP para um sistema de múltiplas condições operacionais, com o objetivo de diminuir o impacto desses efeitos sobre os resultados. A metodologia utilizou PCA para detectar o estado estacionário, análise de agrupamentos para identificar as diferentes condições operacionais e um estimador de Correntropia para solucionar o problema de RDEP, em diferentes condições operacionais. A metodologia foi aplicada a um processo de fracionamento de ar.

ZHANG e CHEN (2015) apresentaram a solução do problema de RD para um sistema dinâmico não linear juntamente com DMEG. Os autores utilizaram o estimador de Correntropia para a RD e DMEG. Os autores mostraram que, tanto para reconciliação de dados quanto para a estimação de parâmetros, o estimador de Correntropia apresentou bons resultados, apesar da presença de erros grosseiros. Um teste de hipóteses (distâncias entre as variáveis no tempo) foi usado para identificar tipos de erros grosseiros específicos de sistemas dinâmicos, como *bias* e *drifts*. Utilizaram para o estudo de caso um reator de polimerização por radicais livres de estireno em um processo dinâmico. A estratégia proposta diminuiu consideravelmente a influência dos erros grosseiros nos resultados reconciliados e permitiu classificar com sucesso os tipos de erros grosseiros dos dados medidos.

CÂMARA *et al.* (2017) estudaram vários aspectos dos problemas de RD. Os aspectos discutido foram: a formulação do problema, em relação a presença ou não de parâmetros a serem estimados na função objetivo; um teste de hipótese global baseado na distribuição χ^2 , com o propósito de indicar que a otimização atingiu o valor mínimo satisfatório; a abordagem da solução quando técnicas de PNL são utilizadas; métodos para estimar gradientes de função objetivo; formas de estimativas iniciais para a variável reconciliada e algoritmos de otimização. Além disso, os autores apresentaram uma revisão bibliográfica de aplicações e estudos de RD em indústrias ou com dados reais de processo.

COIMBRA *et al.* (2017) apresentaram um estudo e aplicação de RDD em um processo de polimerização real. Os autores relataram que a RDD permitiu com sucesso o monitoramento do processo em tempo real, o que levou a entender melhor o processo de polimerização.

DO VALLE *et al.* (2018) apresentaram uma coleção de estudos de caso referência para problemas de RD e DMEG. Avaliaram a RD e DMEG em diferentes tipos de fluxogramas, apresentam os desafios e dificuldades relacionados com o

problema de RD e DMEG e apresentam um resumo dos principais estudos de casos referência de RD e DMEG.

SU *et al.* (2019) estudaram um problema de RDD robusta em um processo piloto industrial farmacêutico de fabricação de comprimidos. Uma estratégia de RD foi proposta e avaliada com dados experimentais e em linha com o sistema de controle. O estimador Welsch foi utilizado para garantir robustez sobre os valores espúrios. A estratégia obteve bom desempenho em reconciliar o peso do comprimido, garantindo o controle de qualidade do peso do comprimido em produção piloto.

A abordagem de Horizonte Móvel

A abordagem de HM (ou janela móvel) pode ser aplicada tanto a problemas de RDD linearmente restrita quanto a RDD não linearmente restrita. Um dos pioneiros no estudo foram LIEBMAN *et al.* (1992), que tinha como objetivo a redução do tamanho do problema de RDD. Os autores fizeram testes com o tamanho do horizonte móvel, chegando à conclusão de que o tamanho da janela é um parâmetro que deve ser estimado, pois uma janela com poucos pontos torna o procedimento mais rápido computacionalmente, embora, janelas pequenas resultem em respostas oscilatórias e em baixa redundância temporal. Por outro lado, muita informação na janela torna o procedimento muito custoso computacionalmente, podendo inviabilizar a aplicação em tempo real.

Com o intuito de diminuir a dimensão do problema de RDD, a técnica de HM propõe um número fixo de pontos, representando as medidas no tempo de forma discreta. O tamanho da janela H é definido como a distância entre o primeiro ponto da janela (“o mais antigo” t_{c-H}) e o último ponto (“o mais recente” t_c); ou seja, a RDD utiliza informação do passado para reconciliar também o tempo corrente, representado pelo último ponto da janela. Então, para cada variável medida i no tempo, usa-se uma janela de amostragem com H pontos. Quando uma nova medida é aferida pela instrumentação, a janela move 1 ponto, deixando o ponto mais antigo para trás e incorporando a informação mais recente. Portanto, o problema de otimização de RDD é resolvido para cada janela; ou seja, para cada medida nova aferida o problema deve ser resolvido, atualizando-se o ponto mais recente com o valor reconciliado. A Figura 2.9 exemplifica a abordagem de HM para uma variável genérica.

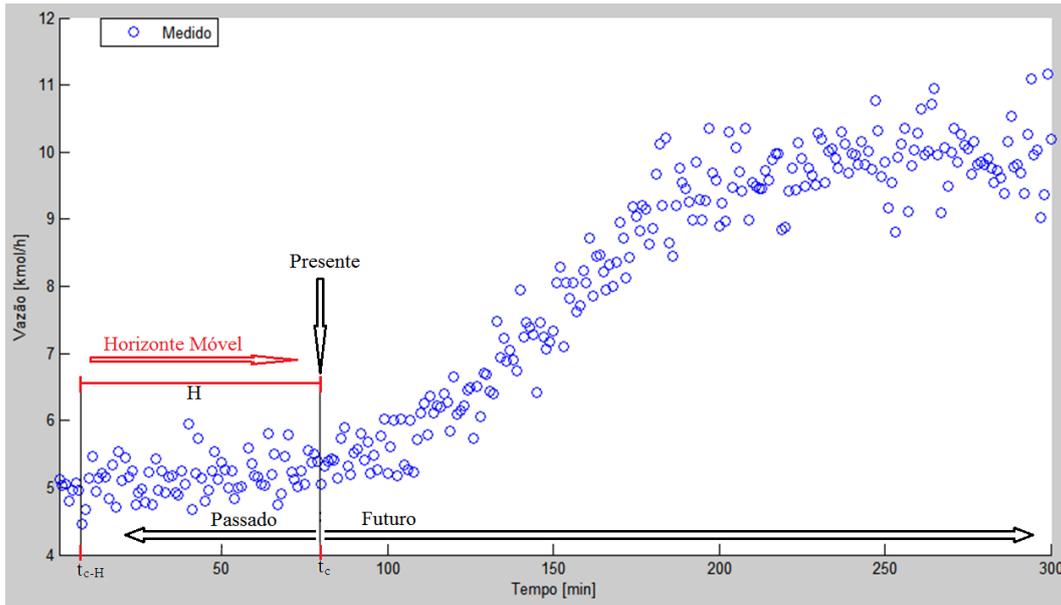


Figura 2.9: A abordagem de Horizonte Móvel.

RD dinâmica não linearmente restrita

A maioria dos modelos dinâmicos de processos é representada por sistemas algébrico-diferenciais (lineares ou não), com inequações que limitam as variáveis e com parâmetros a serem estimados. O problema pode ser escrito como descrito nas Equações (2.9) a (2.12). Por outro lado, o estimador utilizado para modelar a distribuição dos erros de medida pode ser o tradicional MQP ou os estimadores-M robustos. De maneira geral, esses estimadores apresentam parâmetros que devem ser sintonizados (estimados *a priori* ou simultaneamente), para prover um ajuste robusto. Com isso, a função objetivo pode ser escrita na forma mais abrangente da Equação (2.8).

O problema de RDD não linearmente restrito pode ser solucionado classicamente por 3 estratégias (PRATA, 2009): Filtro de Kalman Extendido (EKF), RNA e PNL. O EKF é a versão do KF para sistemas não-lineares, que utiliza discretização linearizada do modelo dinâmico e resolve o problema também de forma recursiva. O procedimento é mais rápido computacionalmente, porém a solução obtida é apenas uma aproximação do problema, pois a técnica só atinge o ponto de ótimo quando o modelo de medida e de transição de estado são ambos lineares. Além disso, EKF pode apresentar divergência nas estimativas, se a estimativa inicial do estado for errada ou se o modelo de processo for incorreto (CHUI *et al.*, 1999).

O uso de RNA tem a vantagem de não usar sistemas de equações diferenciais para representar o processo, evitando linearização e discretização de modelos, fazendo com que a solução do problema de RDD seja computacionalmente mais

rápida. Entretanto, o modelo obtido por treinamento da RNA deve ser feito *a priori*. Por isso, para casos em que múltiplas variações, perturbações e oscilações nas condições operacionais são possíveis, a estratégia de RNA pode ser muito ineficiente (ANDRADE, 2004).

A estratégia de PNL é muito mais abrangente quando aplicada aos problemas de RDD; entretanto, pode apresentar mais dificuldades em aplicações em tempo real devido ao custo computacional mais elevado em relação às demais estratégias. Por outro lado, a utilização de modelos dinâmicos fenomenológicos pode dar maior precisão e representar melhor a realidade do processo químico. O problema de RDD não linearmente restrito pode ser usualmente resolvido por 3 técnicas diferentes de PNL (PRATA, 2009): **linearização sucessiva**, **solução sequencial** e **solução simultânea**.

Na **linearização sucessiva**, como o nome diz, o sistema é linearizado e solucionado como um problema de RD estacionária não linearmente restrita. Entretanto, é preciso aproximar o problema dinâmico para um sistema de equações algébricas. Admite-se que o termo derivativo temporal pode ser substituído pela discretização da variável do processo, gerando um sistema de equações algébricas não-lineares. Dessa forma, reconciliar as variáveis medidas em cada ponto no tempo acaba sendo similar a reconciliar a média dessas variáveis medidas. BENQLILOU (2004) chamou essa metodologia de Média Móvel e Reconciliação de Dados em Estado Estacionário. Para contornar a existência de inequações e limites nas variáveis, alguns autores utilizam funções penalidade para limitar a região de busca na otimização (LIEBMAN, 1991).

A **solução sequencial**, conhecida também como algoritmo de otimização *feasible path*, não lineariza nem reduz as equações diferenciais a sistemas algébricos. Então, as restrições (o ponto de partida requer estimativas de condições iniciais e parâmetros do modelo) e a Fobj são resolvidas em sequência, repetindo os passos iterativos do algoritmo de otimização até convergência. Garantindo assim, que as restrições sejam sempre satisfeitas. O sistema algébrico-diferencial deve ser sempre resolvido em uma camada interna de programação com um método de integração numérica sempre que o algoritmo de otimização, em uma camada externa de programação, requerer o cálculo do modelo para avaliar a Fobj. Isso pode resultar em esforço computacional elevado, pois o método numérico de integração é repedido todas as vezes que o otimizador executar os passos de iteração.

Algumas vantagens do procedimento são que o modelo fica mais preciso e flexível a mudanças de condições operacionais, sendo as restrições do problema sempre respeitadas. Por outro lado, pode perder velocidade computacional devido

ao esforço do algoritmo de otimização, que depende de alguns fatores relacionados ao problema de RDD, como relatado por CÂMARA *et al.* (2017). Apesar disso, PRATA *et al.* (2010) utilizaram esta técnica para solucionar um problema de RD robusta dinâmica não linear e conseguiram implementar o método *on-line* em um sistema real de polimerização. CHEN *et al.* (1998), OLIVEIRA JR. (2006), DE JESUS (2011), CÂMARA *et al.* (2017) e COIMBRA *et al.* (2017) também utilizaram a solução sequencial para resolver o problema de RDD não linearmente restrito. O esquema da solução sequencial está ilustrado pela Figura 2.10.

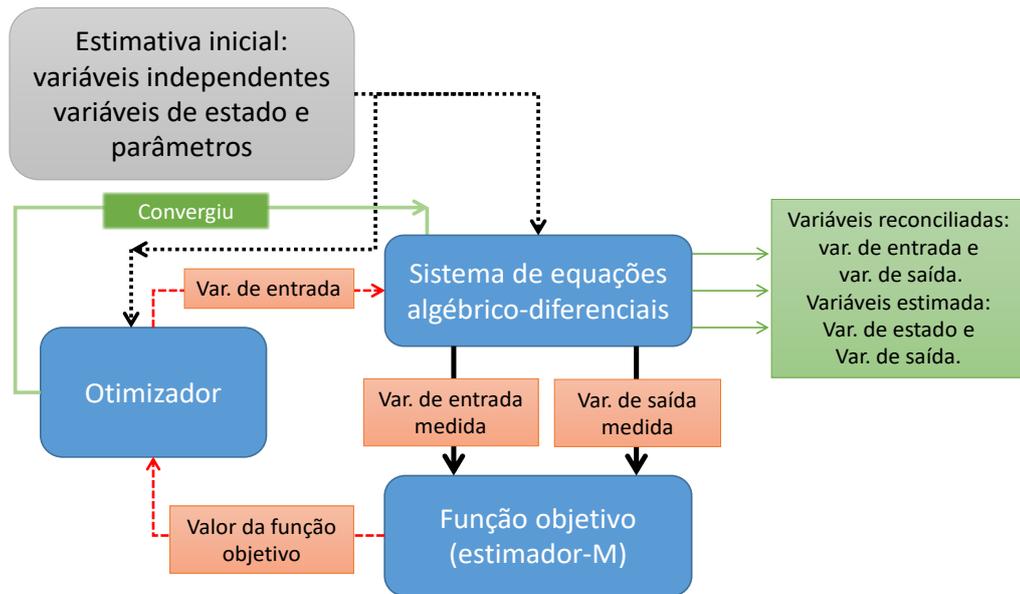


Figura 2.10: Esquema de solução sequencial para problemas de RDD não linearmente restrito pela abordagem de HM.

A **solução simultânea**, conhecida também como algoritmo de otimização *infeasible path*, é uma técnica bastante utilizada para solucionar problemas de RDD não linearmente restrita. LIEBMAN (1991), ALBUQUERQUE e BIEGLER (1995) e ARORA e BIEGLER (2004) utilizaram esta técnica para resolver diferentes problemas de RDD. A solução consiste em reduzir o modelo algébrico-diferencial a um sistema algébrico por meio de discretização com de técnicas como colocação ortogonal, aproximação de Euler de primeira ordem ou Runge-Kutta implícito (BENQLILOU, 2004). Assim, torna-se possível usar o sistema algébrico não linear para aplicar métodos de solução similares aos de RD estacionária não linearmente restrita. Os métodos de solução de RD estacionária não linear são os mesmos que resultam em problema de otimização não linear, incluindo linearizações sucessivas, algoritmos de otimização baseados em gradientes e SQP. A vantagem desta técnica é que ela elimina as inúmeras integrações numéricas das equações diferenciais, pois tanto as restrições quanto a minimização da Fobj convergem

simultaneamente (LIEBMAN, 1991). Isso quer dizer que, durante o processo de otimização, as restrições não são obedecidas, apenas quando convergem. Por outro lado, a discretização aumenta muito a dimensão do problema de otimização, o que pode tornar o esforço computacional maior que no método sequencial (CÂMARA *et al.*, 2017). Outra desvantagem é que o desempenho de técnicas baseadas em discretização depende da seleção do tempo de integração. Na prática, o tempo de amostragem para obter dados de medição é relativamente maior do que o tempo de integração, o que viabiliza a técnica de solução sequencial (BENQLILOU, 2004). No entanto, a representação do modelo pode ser muito inferior que a obtida no método sequencial, em função da necessidade de usar maiores intervalos de integração para implementar a estratégia.

Formulação matemática da RDD robusta não linearmente restrita

Este problema tem a mesma formulação generalizada já descrita pelas Equações (2.8) a (2.12). Entretanto, deve-se estender a formulação da função objetivo para estimadores-M robustos, como proposto por HUBER (1964), com a inclusão da abordagem de horizonte móvel, Equação (2.89). Assim, o problema de Reconciliação de Dados Dinâmica Robusta (RDDR) pode apresentar a formulação completa das Equações (2.90) a (2.92).

$$\min_{\hat{\underline{z}}(t, \hat{\underline{\theta}})} \sum_{t=t_c-t_H}^{t_c} \sum_{i=1}^{NV} \rho_r \left(\frac{z_i(t) - \hat{z}_i(t, \hat{\theta})}{\sigma_i} \right) = \min_{\hat{\underline{z}}(t, \hat{\underline{\theta}})} \sum_{t=t_c-t_H}^{t_c} \sum_{i=1}^{NV} \rho_r(\xi_{ti}) \quad (2.89)$$

sujeito a:

$$\underline{f} \left[\frac{d\hat{\underline{z}}(t, \hat{\underline{\theta}})}{dt}, \hat{\underline{z}}(t, \hat{\underline{\theta}}), \underline{u}(t, \hat{\underline{\theta}}) \right] = \underline{0} \quad (2.90)$$

$$\underline{h} \left[\hat{\underline{z}}(t, \hat{\underline{\theta}}), \underline{u}(t, \hat{\underline{\theta}}) \right] = \underline{0} \quad (2.91)$$

$$\underline{g} \left[\hat{\underline{z}}(t, \hat{\underline{\theta}}), \underline{u}(t, \hat{\underline{\theta}}) \right] \geq \underline{0} \quad (2.92)$$

em que ρ_r é o valor escalar do estimador estimador-M robusto; $\hat{\underline{z}}(t, \hat{\underline{\theta}})$ é o vetor das variáveis reconciliadas; $\underline{z}(t)$ é o vetor das variáveis medidas; $\underline{u}(t, \hat{\underline{\theta}})$ é o vetor das variáveis estimadas não medidas; $\hat{\underline{\theta}}$ é o vetor dos parâmetros estimados do modelo; σ_i é o desvio-padrão da medição da variável i ; ξ_{ti} é o erro padronizado da variável i no tempo t ; t_c é o tempo corrente de amostragem da janela móvel; t_H é o primeiro tempo de amostragem (ou o mais antigo) da janela móvel; H é o tamanho da janela móvel; NV é o número de variáveis medidas; $\underline{f}()$ é o vetor das equações diferenciais;

$\underline{h}()$ é o vetor das equações algébricas e $\underline{g}()$ é o vetor das inequações algébricas.

Formulação matemática da RDD com erro do modelo

Embora as abordagens para solucionar o problema DDR formulado pelas Equações (2.90) a (2.92) tenham sido bem estabelecidas na literatura. Conforme BAI e THIBAUT (2010) estratégias de soluções de RDD sofrem de várias deficiências, notavelmente:

1. Admite que o modelo representa perfeitamente a dinâmica de um processo e os dados reconciliados devem satisfazer perfeitamente essas restrições impostas. Entretanto nenhum modelo matemático representa perfeitamente uma planta e modelos apresentam sempre algum grau de incerteza. Portanto, dados reconciliados devem ser obtidos considerando simultaneamente os erros de medida e do modelo;
2. Modelos de primeiro princípio, como taxa de reação e equações termodinâmicas, foram usados como restrições na reconciliação de medições (LIEBMAN *et al.*, 1992). Infelizmente, esses modelos fenomenológicos raramente estão disponíveis na prática em fábricas de produtos químicos. Conseqüentemente, a dificuldade de obtenção de tais modelos fenomenológicos é um fator que tem impedido o uso do RDD para uma ampla gama de aplicações;
3. O algoritmo RDD requer manipulações matemáticas complexas e sofisticadas. Como resultado, os cálculos online para resolver iterativamente o problema de otimização não linear para reconciliação dinâmica de dados podem ser significativos.

Portanto, visando superar algumas dessas dificuldades MAKNI *et al.* (1995) e HODOUIN e MAKNI (1996) modificaram o algoritmo de reconciliação de dados de restrições em estado estacionária para realizar em tempo real, considerando o processo dinâmico. Assim, a Equação (2.90) passa a ser tratada com incertezas, $\underline{f}[\underline{\hat{z}}, \underline{u}] = \underline{\hat{\delta}}$, onde o vetor $\underline{\hat{\delta}}$ representa a imprecisão residual estimada do modelo. Com essa modificação o problema de reconciliação de dados passa a tratar tanto os erros dos dados medidos quanto os erros das equações de restrições, simultaneamente e de forma consistente dentro da função objetivo. Segundo BAI *et al.* (2005) para um problema de RD não robusta (MQP) e um sistema discretamente amostrado, em cada tempo de amostragem t , a Equação (2.89), conseqüentemente, pode ser escrita como:

$$\min_{\hat{z}_t} J(\hat{z}_t, \underline{u}_t) = (z_t - \hat{z}_t)^T \underline{\Sigma}^{-1} (z_t - \hat{z}_t) + \underline{f}^T [\hat{z}_t, \underline{u}_t] \underline{\Omega}^{-1} \underline{f} [\hat{z}_t, \underline{u}_t] \quad (2.93)$$

sujeito a:

$$\underline{g} [\hat{z}_t, \underline{u}_t] \geq \underline{0} \quad (2.94)$$

onde $\underline{\Omega}$ é a matriz de covariâncias dos erros randômicos do modelo, sendo o erro admitido como normalmente distribuído, ou seja, $\underline{f} [\hat{z}_t, \underline{u}_t] = \delta_t \sim N(\underline{0}, \underline{\Omega})$.

Segundo BAI e THIBAUT (2010) o problema formulado pelas Equações (2.93) e (2.94) pode ser chamada de Reconciliação de Dados de Conjunto Único (*Single Set Data Reconciliation* - SDR), porque apenas um conjunto amostral é reconciliado em um instante de tempo. A estratégia SDR pode ser aplicada utilizando modelos empíricos ou modelos fenomenológicos, com variáveis medidas e não medidas. Entretanto, a classificação de variáveis deve satisfazer a observabilidade do sistema.

Essa estratégia pode ser estendida para a estratégia de média móvel, adicionando informações passadas da série temporal para reduzir o nível de ruído. Portanto, a estratégia SDR em conjunto com média móvel é conhecida como Reconciliação de Dados de Horizonte Móvel (*Moving Window Data Reconciliation* - MWDR) (BAI *et al.*, 2005). A formulação dessa estratégia pode ser escrita como:

$$\min_{\hat{z}_t} J(\hat{z}_t, \underline{u}_t) = \sum_{k=0}^{H-1} \left[(z_{t-k} - \hat{z}_t)^T \underline{\Sigma}^{-1} (z_{t-k} - \hat{z}_t) + \underline{f}^T [\hat{z}_t, \underline{u}_t] \underline{\Omega}^{-1} \underline{f} [\hat{z}_t, \underline{u}_t] \right] \quad (2.95)$$

sujeito a:

$$\underline{g} [\hat{z}_t, \underline{u}_t] \geq \underline{0} \quad (2.96)$$

onde H é o tamanho da janela da série temporal mais recente. A estratégia MWDR considera que o processo esteja em regime estacionário dentro da janela amostral, o que significa que o tamanho da janela deve ser considerado um parâmetro de ajuste da estratégia. Quanto maior o parâmetro H maior será o “amortecimento” dos ruídos da série temporal, por outro lado, o atraso na resposta do sistema dinâmico será maior. Outra observação é sobre a variância do erro do modelo, pois se o erro do modelo for significativamente maior que o erro de medição, ou seja, a diagonal da matriz $\underline{\Omega}$ for de números muito elevados, a estratégia MWDR passa a se comportar

como um simples filtro de média móvel (BAI e THIBAUT, 2010).

Análise da RDD com erro do modelo

Para analisarmos a relação entre o erro de medição e o erro do modelo, devemos apresentar a solução de um problema mais simples de RDD com erro de modelo. Essa estratégia é conhecida como “Algoritmo de RDD baseado em predição-correção” e apresentada por BAI *et al.* (2006). A solução analítica desse problema é viável, porém o modelo preditivo deve ser linear. Portanto, admite que o modelo de predição $\hat{\underline{x}}_t^*$ pode ser expresso com adição de um ruído na forma:

$$\hat{\underline{x}}_t^* = \underline{x}_t - \underline{\delta}_t \quad (2.97)$$

onde $\underline{\delta}_t$ representa o vetor do erro do modelo de predição, o que pode ser uma função complexa de vários fatores, como a estrutura de modelo incorreta, incerteza dos parâmetros do modelo e incerteza das variáveis de entrada do modelo. Entretanto, para simplificar, admite-se que $\underline{\delta}_t \sim N(\underline{0}, \underline{\Omega})$. Assim, o problema para a melhor estimativa de \underline{x}_t é ponderado simultaneamente pelo o erro dos dados medidos e do erro do modelo de predição. Utilizando a fórmula de Bayes (BOX e TIAO, 2011) nos dados medidos, como na Equação (2.98).

$$f(\underline{x}_t | \underline{z}_t) \propto L(\underline{x}_t | \underline{z}_t) f(\underline{x}_t) \quad (2.98)$$

onde $f(\underline{x}_t | \underline{z}_t)$ é a distribuição de \underline{x}_t *a posteriori*, condicionado ao dado medido \underline{z}_t ; $L(\underline{x}_t | \underline{z}_t)$ é a Função de Verossimilhança apresentada na Equação (2.17) e $f(\underline{x}_t)$ é a distribuição *a priori*, o que expressa o conhecimento prévio sobre \underline{x}_t , antes mesmo de ter os dados.

Agora, devemos aplicar novamente a fórmula de Bayes na outra informação independente de \underline{x}_t , representada pelo modelo preditivo $\hat{\underline{x}}_t^*$.

$$f(\underline{x}_t | \hat{\underline{x}}_t^*, \underline{z}_t) \propto L(\underline{x}_t | \hat{\underline{x}}_t^*) f(\underline{x}_t | \underline{z}_t) \propto L(\underline{x}_t | \hat{\underline{x}}_t^*) L(\underline{x}_t | \underline{z}_t) f(\underline{x}_t) \quad (2.99)$$

onde $f(\underline{x}_t | \hat{\underline{x}}_t^*, \underline{z}_t)$ representa o conhecimento *a posteriori* de \underline{x}_t , condicionado simultaneamente em \underline{z}_t e $\hat{\underline{x}}_t^*$. Baseado nessas suposições e aplicando a Função de Verossimilhança na informação do modelo preditivo, temos que:

$$L(\underline{x}_t | \hat{\underline{x}}_t^*) = \frac{1}{(2\pi)^{M/2} |\underline{\underline{\Omega}}|^{1/2}} \exp \left[-\frac{1}{2} (\hat{\underline{x}}_t^* - \underline{x}_t)^T \underline{\underline{\Omega}}^{-1} (\hat{\underline{x}}_t^* - \underline{x}_t) \right] \quad (2.100)$$

Além disso, devemos admitir inicialmente que a informação *a priori* de \underline{x}_t é desconhecida, ou seja, $f(\underline{x}_t)$ é uma constante por toda a faixa de \underline{x}_t . Assim, a distribuição *a posteriori* é expressa por:

$$f(\underline{x}_t | \hat{\underline{x}}_t^*, \underline{z}_t) = \frac{k}{(2\pi)^{M/2} |\underline{\underline{\Sigma}} \underline{\underline{\Omega}}|^{1/2}} \exp \left\{ -\frac{1}{2} \left[(\underline{z}_t - \underline{x}_t)^T \underline{\underline{\Sigma}}^{-1} (\underline{z}_t - \underline{x}_t) + (\hat{\underline{x}}_t^* - \underline{x}_t)^T \underline{\underline{\Omega}}^{-1} (\hat{\underline{x}}_t^* - \underline{x}_t) \right] \right\} \quad (2.101)$$

onde k é uma constante. O que indica que a maximização da Equação (2.101) é a melhor estimativa do valor verdadeiro da variável aleatória de processo. Com isso, a maximização da Equação (2.101) é equivalente a minimizar o argumento positivo do termo exponencial, e é conhecida como estimador de *maximum a posteriori* (MAP).

O estimador MAP denotado por $\hat{\underline{z}}_t$ é obtido minimizando o termo:

$$J(\hat{\underline{z}}_t) = \frac{1}{2} \left[(\underline{z}_t - \hat{\underline{z}}_t)^T \underline{\underline{\Sigma}}^{-1} (\underline{z}_t - \hat{\underline{z}}_t) + (\hat{\underline{x}}_t^* - \hat{\underline{z}}_t)^T \underline{\underline{\Omega}}^{-1} (\hat{\underline{x}}_t^* - \hat{\underline{z}}_t) \right] \quad (2.102)$$

Aplicando a derivada parcial na Equação (2.102) em respeito a $\hat{\underline{z}}_t$ e igualando a zero, temos que:

$$\frac{\partial J}{\partial \hat{\underline{z}}_t} = \frac{1}{2} \left[\underline{\underline{\Sigma}}^{-1} (\underline{z}_t - \hat{\underline{z}}_t) + \underline{\underline{\Omega}}^{-1} (\hat{\underline{x}}_t^* - \hat{\underline{z}}_t) \right] = \underline{0} \quad (2.103)$$

Isolando $\hat{\underline{z}}_t$ de Equação (2.103), temos que:

$$\begin{aligned} \hat{\underline{z}}_t &= (\underline{\underline{\Sigma}}^{-1} + \underline{\underline{\Omega}}^{-1})^{-1} (\underline{\underline{\Sigma}}^{-1} \underline{z}_t + \underline{\underline{\Omega}}^{-1} \hat{\underline{x}}_t^*) \\ \hat{\underline{z}}_t &= (\underline{\underline{\Sigma}}^{-1} + \underline{\underline{\Omega}}^{-1})^{-1} \underline{\underline{\Sigma}}^{-1} \underline{z}_t + (\underline{\underline{\Sigma}}^{-1} + \underline{\underline{\Omega}}^{-1})^{-1} \underline{\underline{\Omega}}^{-1} \hat{\underline{x}}_t^* \end{aligned} \quad (2.104)$$

Adicionando e subtraindo o termo $(\underline{\underline{\Sigma}}^{-1} + \underline{\underline{\Omega}}^{-1})^{-1} \underline{\underline{\Sigma}}^{-1} \hat{\underline{x}}_t^*$ do lado direito da

Equação (2.104), temos que:

$$\begin{aligned}
\hat{z}_t &= (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \underline{\Sigma}^{-1} z_t + (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \left[(\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} - \underline{\Sigma}^{-1} \right] \hat{x}_t^* \\
\hat{z}_t &= (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \underline{\Sigma}^{-1} z_t + \hat{x}_t^* - (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \underline{\Sigma}^{-1} \hat{x}_t^* \\
\hat{z}_t &= \hat{x}_t^* + (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \underline{\Sigma}^{-1} (z_t - \hat{x}_t^*) \\
\hat{z}_t &= \hat{x}_t^* + \underline{\mathbf{K}} (z_t - \hat{x}_t^*)
\end{aligned} \tag{2.105}$$

onde \hat{z}_t é o vetor da melhor estimativa de x_t , ou seja, o vetor das variáveis reconciliadas das medidas brutas z_t ; \hat{x}_t^* é o vetor de predição por modelo de x_t , ou seja, predição um passo a frente do tempo $t - 1$; $\underline{\mathbf{K}} = (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \underline{\Sigma}^{-1} = (\underline{\mathbf{I}} + \underline{\Sigma} \underline{\Omega}^{-1})^{-1}$ no qual $\underline{\Sigma}$ é a matriz de covariâncias dos erros de medição e $\underline{\Omega}$ é a matriz de covariâncias dos erros de predição do modelo.

A Equação (2.105) representa o Algoritmo de RDD baseado em predição-correção, indicando que a variável reconciliada \hat{z}_t é formada por duas informações: o valor predito pelo modelo \hat{x}_t^* e a correção do valor medido $\underline{\mathbf{K}} (z_t - \hat{x}_t^*)$. $\underline{\mathbf{K}}$ representa o ganho da RDD baseado em corretor-preditor, ponderando simultaneamente os erros de medição e erros de predição do modelo na variável reconciliada.

Analisando o termo de ganho da RDD, podemos observar que quando o erro de medição é significativamente maior que o erro de predição do modelo, $\underline{\Sigma} \gg \underline{\Omega}$, $\underline{\mathbf{K}}$ tende a valores muito pequenos. Consequentemente, os valores reconciliados se aproximam dos valores preditos pelo modelo, $\hat{z}_t \rightarrow \hat{x}_t^*$. Por outro lado, quando o erro de predição do modelo é significativamente maior que o erro de medição, $\underline{\Omega} \gg \underline{\Sigma}$, $\underline{\mathbf{K}}$ tende a $\underline{\mathbf{I}}$, e como consequência, o valor reconciliado se aproxima do valor medido, $\hat{z}_t \rightarrow z_t$.

O erro da reconciliação para este problema é dado pela Equação (2.106).

$$\gamma_t = \hat{z}_t - x_t \tag{2.106}$$

Substituindo a Equação (2.105) na Equação (2.106) e usando as Equações (2.13) e (2.97), resulta em:

$$\begin{aligned}
\gamma_t &= \delta_t + \underline{\mathbf{K}} (\varepsilon_t - \delta_t) \\
\gamma_t &= \underline{\mathbf{K}} \varepsilon_t + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \delta_t
\end{aligned} \tag{2.107}$$

Aplicando o operador esperança em ambos os lados da Equação (2.107), temos que:

$$\mathbf{E} [\gamma_t] = \underline{\mathbf{K}} \mathbf{E} [\varepsilon_t] + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \mathbf{E} [\delta_t] \quad (2.108)$$

Entretanto, $\mathbf{E} [\varepsilon_t] = \underline{\mathbf{0}}$ e $\mathbf{E} [\delta_t] = \underline{\mathbf{0}}$, logo $\mathbf{E} [\gamma_t] = \underline{\mathbf{0}}$. O que significa que a RDD baseado em corretor-preditor é um estimador imparcial.

A matriz de covariâncias sobre os erros de reconciliação da RDD é dado pela aplicação do operador covariâncias sobre a Equação(2.107). Além disso, como os erros são independentes, a matriz de covariância entre os erros descorrelacionados será nula, $\mathbf{E} [\varepsilon_t \delta_t^T] = \mathbf{E} [\delta_t \varepsilon_t^T] = \underline{\mathbf{0}}$. Com isso, temos que:

$$\begin{aligned} \mathit{cov} [\gamma_t] &= \mathbf{E} [\gamma_t \gamma_t^T] = \mathbf{E} \left\{ [\underline{\mathbf{K}} \varepsilon_t + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \delta_t] [\underline{\mathbf{K}} \varepsilon_t + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \delta_t]^T \right\} \\ &= \mathbf{E} \left\{ [\underline{\mathbf{K}} \varepsilon_t + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \delta_t] \left[\varepsilon_t^T \underline{\mathbf{K}}^T + \delta_t^T (\underline{\mathbf{I}} - \underline{\mathbf{K}})^T \right] \right\} \\ &= \mathbf{E} \left[\underline{\mathbf{K}} \varepsilon_t \varepsilon_t^T \underline{\mathbf{K}}^T + \cancel{\underline{\mathbf{K}} \varepsilon_t \delta_t^T (\underline{\mathbf{I}} - \underline{\mathbf{K}})^T}^{\underline{\mathbf{0}}} + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \cancel{\delta_t \varepsilon_t^T \underline{\mathbf{K}}^T}^{\underline{\mathbf{0}}} + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \delta_t \delta_t^T (\underline{\mathbf{I}} - \underline{\mathbf{K}})^T \right] \\ \mathit{cov} [\gamma_t] &= \underline{\mathbf{K}} \mathbf{E} [\varepsilon_t \varepsilon_t^T] \underline{\mathbf{K}}^T + (\underline{\mathbf{I}} - \underline{\mathbf{K}}) \mathbf{E} [\delta_t \delta_t^T] (\underline{\mathbf{I}} - \underline{\mathbf{K}})^T \end{aligned} \quad (2.109)$$

Substituindo $\mathbf{E} [\varepsilon_t \varepsilon_t^T] = \underline{\Sigma}$, $\mathbf{E} [\delta_t \delta_t^T] = \underline{\Omega}$ e $\underline{\mathbf{K}} = \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1}$ na Equação (2.109), onde $\underline{\mathbf{X}}^{-1} = (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} = \underline{\Omega} (\underline{\mathbf{I}} + \underline{\Sigma}^{-1} \underline{\Omega})^{-1} = (\underline{\mathbf{I}} + \underline{\Sigma} \underline{\Omega}^{-1})^{-1} \underline{\Sigma}^{-1}$ é simétrica como $\underline{\Sigma}$ e $\underline{\Omega}$, temos que:

$$\begin{aligned}
\mathbf{cov} [\underline{\gamma}_t] &= \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Sigma} (\underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1})^T + (\underline{I} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1}) \underline{\Omega} (\underline{I} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1})^T \\
&= \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} + (\underline{I} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1}) \underline{\Omega} (\underline{I} - \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1}) \\
&= \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} + \underline{\Omega} - \underline{\Omega} \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} + \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Omega} \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Omega} \\
&= (\underline{\mathbf{X}}^{-1} - \underline{\Omega} + \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Omega}) \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} + \underline{\Omega} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Omega} \\
&= \begin{bmatrix} \underline{\Omega} (\underline{I} + \underline{\Sigma}^{-1} \underline{\Omega})^{-1} \\ \underline{\mathbf{X}}^{-1} (\underline{I} + \underline{\Sigma}^{-1} \underline{\Omega}) - \underline{\Omega} \end{bmatrix} \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} + \underline{\Omega} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Omega} \\
&= \underline{\Omega} - \underline{\mathbf{X}}^{-1} \underline{\Sigma}^{-1} \underline{\Omega} = \underline{\Omega} - (\underline{I} + \underline{\Sigma} \underline{\Omega}^{-1})^{-1} \underline{\Omega} \\
&= \underline{\Omega} - (\underline{\Omega}^{-1} + \underline{\Omega}^{-1} \underline{\Sigma} \underline{\Omega}^{-1})^{-1} = \underline{\Omega} - \underline{\Omega} (\underline{I} + \underline{\Omega}^{-1} \underline{\Sigma})^{-1} \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1} \\
&= \underline{\Omega} (\underline{I} - \underline{\Sigma}^{-1} \underline{\mathbf{X}}^{-1}) = \underline{\Omega} (\underline{\mathbf{X}} - \underline{\Sigma}^{-1}) \underline{\mathbf{X}}^{-1} \\
&= [\underline{\Omega} (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1}) - \underline{\Omega} \underline{\Sigma}^{-1}] \underline{\mathbf{X}}^{-1} = (\underline{\Omega} \underline{\Sigma}^{-1} + \underline{I} - \underline{\Omega} \underline{\Sigma}^{-1}) \underline{\mathbf{X}}^{-1} = \underline{\mathbf{X}}^{-1}
\end{aligned}$$

$$\mathbf{cov} [\underline{\gamma}_t] = \mathbf{cov} [\underline{\hat{z}}_t] = (\underline{\Sigma}^{-1} + \underline{\Omega}^{-1})^{-1} \quad (2.110)$$

A Equação (2.110) mostra que a matriz de covariâncias dos erros de reconciliação da RDD é função de $\underline{\Sigma}$ e $\underline{\Omega}$. De fato, $\underline{\Sigma}$ e $\underline{\Omega}$ são matrizes simétrica positiva definida, com os elementos de $\mathbf{cov} [\underline{\gamma}_t]$ menores que os elementos de $\underline{\Sigma}$ e $\underline{\Omega}$. Isso significa que a combinação das incertezas do erro de medição mais o erro de predição do modelo, resulta no erro de reconciliação mais preciso do que os erros dessas incertezas individualmente. Importante lembrar que o erro admitido da predição do modelo é um ruído branco Gaussiano. Entretanto, essa suposição estatística não se aplica aos problemas reais, pois na prática o cálculo da estimativa dessa matriz de covariâncias $\underline{\Omega}$ é complexo. Consequentemente, os elementos de $\underline{\Omega}$ são usualmente tratado como parâmetros de sintonia. As Figuras 2.11 e 2.12 ilustram o conceito e análise dos erros da RDD baseado em corretor-preditor.

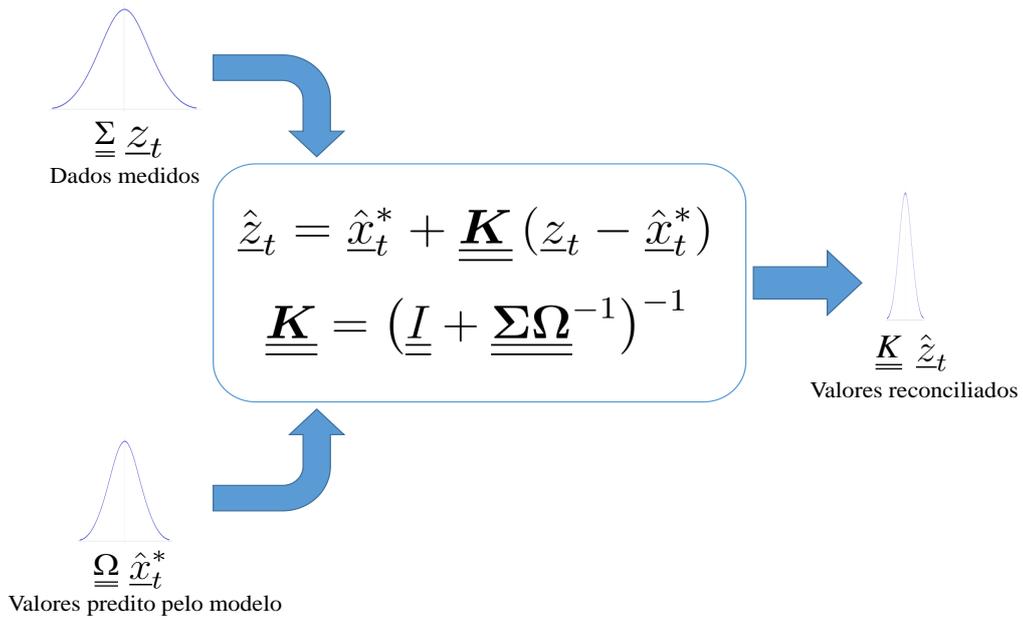


Figura 2.11: Conceito Algoritmo da RDD baseado em predição-correção
 Fonte: Adaptado de BAI e THIBAUT (2010).

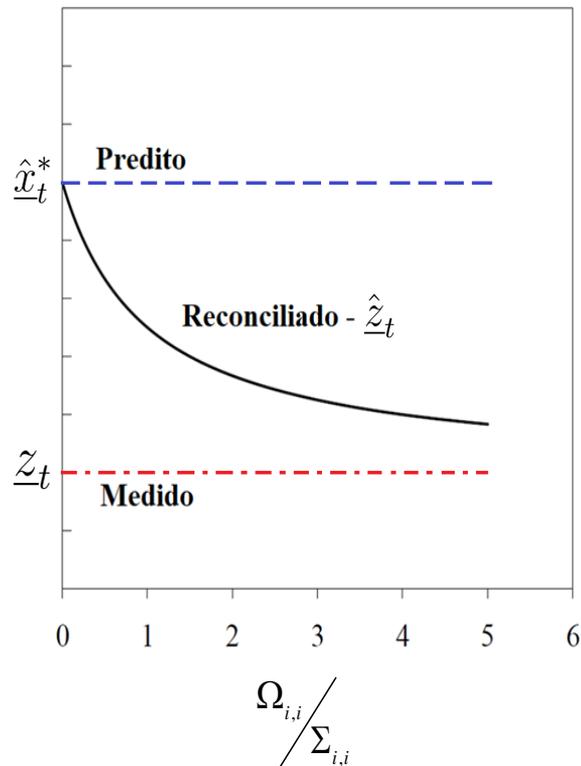


Figura 2.12: Relação entre os erros de medição, de predição e o valor reconciliado
 Fonte: Adaptado de BAI e THIBAUT (2010).

RDD em colunas de destilação

Os pioneiros em problemas de RD não linearmente restrita para modelos de coluna de destilação foram NOORAI e ROMAGNOLI (1998). Neste trabalho, os autores implementaram um sistema de controle distribuído em uma coluna de destilação em escala piloto, com detecção, identificação e estimação de erros grosseiros, reconciliação de dados estacionária, simulação, modificação da eficiência e otimização do *set-point*. O problema de otimização não linear gerado pelo modelo utilizado (balanço de massa, energia e componente da coluna) foi resolvido pelo pacote de otimização GAMS. O modelo utilizado na simulação e a otimização do *set-point* foram implementados no simulador PROCESS. Para o controle do processo, detalhes mais rigorosos foram simulados no simulador dinâmico SPEEDUP.

SODERSTROM *et al.* (2000) implementaram a RDD em uma planta química da ExxonMobil, utilizando um sistema algébrico-diferencial para modelar os balanços de massa nos equipamentos do processo. Os autores utilizaram a abordagem de HM com tempo de amostragem de 2 minutos e horizonte de tamanho de 60 minutos. O sistema algébrico-diferencial foi convertido em equações algébricas por colocação ortogonal em elementos finitos, de maneira que os autores utilizaram a técnica de solução simultânea. O problema de otimização não linear foi solucionado pelo sistema de otimização NOVA, um otimizador comercial que utiliza o algoritmo SQP de otimização.

BAI *et al.* (2006) compararam RDD com KF em uma coluna de destilação binária, utilizando modelo dinâmico de coluna (balanço de massa e energia, equilíbrio-líquido-vapor e comportamento hidráulico dos pratos), além de 4 controles proporcional-integral (PI). BAI *et al.* (2007b) desenvolveram um algoritmo para fazer correções simultânea de erros tipo *bias* em conjunto com RDD, utilizando a mesma coluna do trabalho anterior. BAI *et al.* (2007a) examinaram o impacto da estrutura do modelo dinâmico de uma coluna no desempenho da RDD. Modelo linear e modelo empírico baseado em rede neural foram utilizados em conjunto com a técnica de RDD.

FARZI *et al.* (2008) desenvolveram um software de RDD com a arquitetura de programação orientada a objeto. Os autores exemplificam com dois casos: EKF e rede neural foram utilizados na RDD de uma coluna binária. FARZI *et al.* (2009) utilizam EKF na RDD e exemplificam com dois casos: coluna de destilação com medidas de temperatura e CSTR com medidas de concentração.

NICHOLSON *et al.* (2014) utilizaram para reduzir o tempo computacional a Estimação de Horizonte Móvel (MHE) de passo avançado (asMHE) como técnica

de Estimação de Estados Dinâmica (EED), problema muito parecido com RDD. Os autores utilizaram estimadores robustos Fair e Hampel para amenizar os erros grosseiros. O método de otimização IPOPT foi utilizado para resolver o problema de “RDD”. Modelo rigoroso de coluna foi utilizado e discretizado no tempo, gerando um sistema algébrico não linear, resultando em abordagem simultânea, e utilizaram as temperaturas e líquidos *holdup* como medida.

VALLURU *et al.* (2016) utilizaram o método de MHE em conjunto com EKF para EED “RDD”. O sistema estudado pelos autores foi uma coluna de destilação reativa. Os autores compararam 3 opções de tamanho do HM ($H = 5, 10$ e 20 pontos) com o custo computacional e o desempenho da RDD, concluindo que janela de $H = 10$ pontos se saiu com o melhor custo/desempenho.

2.5 Comentários Finais

Com base no que foi descrito, observa-se que a aplicação de RD tem um grande potencial em monitorar os dados de maneira confiável, detectar falhas e possivelmente reduzir gastos energéticos em processos de separação, principalmente quando aplicada juntamente com a técnica de RTO, já que se trata de equipamentos amplamente utilizados e em muitos casos demandando uma energia extremamente alta. Entretanto, há uma escassez no estudo de RDD aplicada em colunas de destilação e separação de CO_2 por membrana, principalmente no uso de abordagem de solução sequencial. Além disso, não foi encontrado estudo de RDD em colunas de destilação com o uso de modelos reduzidos juntamente com estimadores-M robustos e, estudos de RD em processo de separação por membrana aplicado em linha e em tempo real.

Portanto, é observado que existe um espaço na literatura para pesquisar e desenvolver novas metodologias de RD e RDD não linearmente restrita, com foco no processo de separação. Dessa forma, a tese apresenta um estudo aplicando um modelo preditivo de séries temporais em problemas de RDD, em coluna de destilação para detecção e eliminação simultânea dos erros grosseiros. Além disso, apresenta um estudo aplicando modelo de redução de ordem para coluna de destilação, utilizando a técnica de solução sequencial com algoritmos de otimização determinísticos e estocásticos para solucionar o problema de RDD não-linear. Outro estudo inovador está apresentado no Apêndice A, que trata do monitoramento em tempo real de um sistema de separação por membrana, o qual utiliza estratégias de RD e detecção de erros grosseiros para analisar e diagnosticar o processo em tempo real.

Capítulo 3

Estatística Robusta

“A teoria da probabilidade é, no fundo, nada mais do que o senso comum reduzido ao cálculo. O peso da evidência de uma afirmação extraordinária deve ser proporcional a sua estranheza.”

Pierre Simon de Laplace (1749 – 1827)

Este capítulo tem como objetivo apresentar um panorama sobre a Estatística Robusta, focando principalmente nos estimadores-**M** robustos – generalizações de um estimador de Máxima Verossimilhança – e suas características. São ilustrados 7 estimadores-**M**, incluindo os estimadores de Correntropia, até então nunca aplicados a problemas de colunas de destilação, além do tradicional estimador de Mínimos Quadrados Ponderados (MQP), resultante da hipótese de que os erros de medição seguem uma distribuição Normal. Por fim, são apresentados os critérios alternativos para detecção de valores espúrios, incluindo aqueles que utilizam a forma da curva do estimador, e critérios de avaliação de desempenho para problemas de regressão, incluindo valores espúrios.

3.1 Contextualização

Uma ferramenta estatística amplamente utilizada em quase todas as áreas de conhecimento, quando se busca ajustar um modelo matemático a um conjunto de dados experimentais, é a Análise de Regressão. O termo “Análise de Regressão” usualmente inclui os problemas de Reconciliação de Dados, Estimação de Parâmetros ou ambos em conjunto. Embora vários métodos possam ser usados para obter

o valor estatisticamente consistente das variáveis e estimar os parâmetros de um modelo matemático, o método dos mínimos quadrados ponderados, proveniente da distribuição Normal, é utilizado na maioria das vezes por causa de sua ampla aceitação, propriedades estatísticas elegantes e facilidade de cálculo. Os principais motivos para utilização da distribuição Normal são (BARD, 1974):

- apresenta um comportamento próximo ao de muitas medidas experimentais;
- com o aumento do número de dados medidos, muitas distribuições aproximam-se da distribuição Normal (Teorema do Limite Central);
- simplicidade, já que utiliza apenas dois parâmetros de sintonia: a média e a variância;
- fácil tratamento matemático, por ser simétrica em torno da média e ser convexa, o que facilita o ajuste da distribuição aos dados experimentais.

A adoção de técnicas de mínimos quadrados ponderados foi considerada por muitos como a melhor abordagem para o problema de regressão (REY, 1983). Entretanto, esse método clássico e popular depende de uma série de pressupostos que são muitas vezes irrealistas ou não aceitáveis. A suposição da distribuição Normal pode ser severamente violada se um ou mais valores espúrios (*outliers*) estiverem presentes no conjunto de dados medidos, mesmo que a maioria dos dados esteja em conformidade com uma distribuição Normal, resultando em estimativas ruins ou desviadas para as variáveis e parâmetros, conforme apresentado por TUKEY (1960). Esse autor mostrou que muitos estimadores que têm um desempenho ótimo para dados amostrados a partir de uma distribuição Normal apresentam desempenho ruim se pequenas alterações são feitas na distribuição (contaminação). Ressalta-se que inferências estatísticas são sempre baseadas em observações parciais de um fenômeno real (cuja realidade é sempre desconhecida, pois nunca é possível se obter infinitas observações). Portanto, sempre haverá dúvidas quanto ao modelo mais adequado para representar a curva de distribuição probabilística atribuída às finitas observações da realidade desconhecida. Logo, o estimador de mínimos quadrados ponderados pode ser excessivamente sensível a pequenas incertezas sobre as hipóteses idealizadas usualmente admitidas (TUKEY, 1960).

Na década de 1960, três trabalhos foram de suma importância para o desenvolvimento de um novo campo da estatística matemática chamada de estatística robusta (HAMPEL, 1968; HUBER, 1964; TUKEY, 1960). O termo “robustez”, de modo qualitativo, pode ser entendido como “insensibilidade para pequenas diferenças em relação às hipóteses idealizadas”, para as quais o estimador

é otimizado (HUBER e RONCHETTI, 2009). A palavra “pequena” pode ter duas interpretações diferentes: pode ser entendida como um número pequeno de diferenças consideráveis (valores espúrios) ou como um número apreciável de diferenças pequenas (distribuição não-Normal) (PRATA, 2009). A estatística robusta visa a corrigir métodos estatísticos convencionais, de forma a melhorar a representação da curva de distribuição de probabilidade com valores espúrios presentes, suavizando ou eliminando pequenos desvios existente no modelo estocástico clássico. Os estimadores robustos são ferramentas de crescente potencial para uso em problemas de regressão de dados laboratoriais ou industriais (estimação de parâmetros, reconciliação de dados, ou ambos conjuntamente), em que hipóteses idealizadas são violadas, pois estes estimadores são capazes de ignorar valores atípicos (valores espúrios) por conta da estrutura matemática PRATA *et al.* (2008a).

Conforme descrito por PRATA (2009), as três principais classes de estimadores robustos existentes são:

- Estimadores-**L** (combinações lineares de observações de ordem);
- Estimadores-**R** (estimador baseado no ranqueamento dos resíduos);
- Estimadores-**M** (generalizações de estimadores de Máxima Verossimilhança).

PRATA (2009) ainda comenta sobre a existência de outras classes de estimadores robustos (D, S, τ , CM, GM, GS, MM, RM, LMS, L e TS), embora estes estimadores não constituam o escopo do presente trabalho.

Os estimadores-L geralmente são estatísticas extremamente simples e robustas, como por exemplo a mediana da amostra e média aparada- α (*α -trimmed mean*). Astrônomos do antigo Egito e de Roma já descartavam dados extremos para calcular a média das observações, o que indica que foram os primeiros a usarem o estimador-L *α -trimmed mean* (VICHARE, 1993). Os estimadores-R envolvem a classificação de resíduos. A classificação de uma amostra é um mapeamento de n números reais, admitindo uma ordenação dos dados. Assim, o cálculo dos pesos e interpretações são facilitadas e geralmente resistem a valores discrepantes. Os estimadores-L e -R são, portanto, úteis em estatísticas robustas, como estatística descritiva, em educação estatística e quando o cômputo do problema é muito custoso. Os estimadores-L e -R desempenham um papel fundamental em muitas abordagens na estatística não-paramétrica. No entanto, na atualidade estimadores-M são preferidos, embora sejam computacionalmente mais custosos (HUBER, 1981; JUREČKOVÁ, 1984).

Segundo SARRUF (2014), a estatística robusta vem sendo aplicada em diversos campos, tais como astronomia; econometria; finanças; processamento de

imagens; engenharia elétrica; engenharia mecânica; engenharia química (ÖZYURT e PIKE, 2004); engenharia petroquímica (PRATA *et al.*, 2010); engenharia nuclear (VALDETARO e SCHIRRU, 2011), dentre outros. Vale ressaltar que os métodos robustos para regressão ainda não são extensamente utilizados, quando comparados aos métodos não robustos convencionais. Várias razões podem ajudar a explicar a não popularidade dos métodos robustos:

- para muitos estimadores robustos, e mais especificamente aqueles classificados como “descendentes” (*redescending*), existe a necessidade de métodos de usar otimização do tipo global, uma vez que a estrutura do problema é não convexa (BASELGA, 2007);
- alguns pacotes computacionais comerciais populares não apresentam estimadores robustos como opções (STROMBERG *et al.*, 2004);
- a dificuldade computacional aumenta com o uso de estimadores diferencialmente descontínuos, especialmente aqueles classificados como *hard redescending*, cuja segunda derivada é uma função descontínua dos parâmetros e variáveis analisadas, o que dificulta o uso de métodos determinísticos baseados em derivadas para otimização (PRATA, 2009).

3.2 Desenvolvimento de estimadores-M robustos

A classe de estimadores robustos mais importante para o campo de retificação de dados, segundo descrito por ALBUQUERQUE e BIEGLER (1995), ÖZYURT e PIKE (2004) e PRATA *et al.* (2010), é a classe de estimadores-M, que compreende generalizações de estimadores de Máxima Verossimilhança. Conforme descrito por PRATA (2009), a classe de estimadores-M compreende a grande maioria dos estimadores robustos utilizados no campo de retificação de dados.

3.2.1 Formulação de Máxima Verossimilhança para Estimadores-M Robustos

Fisher introduziu a formulação de Máxima Verossimilhança (ALDRICH *et al.*, 1997), procedimento mais utilizado para obter estimadores em problemas de regressão de dados amostrais (experimentais, laboratoriais ou industriais). Para tanto, considere uma amostragem e uma variável aleatória discreta z_i , relacionada a essa amostragem, com FDP f_i . Se os dados não estão correlacionados (ou seja, são

independentes, de forma que a matriz de covariâncias é diagonal) e distribuídos segundo uma mesma distribuição, a FDP conjunta f pode ser obtida como o produto das funções de probabilidade individuais f_i . Assim, a estimativa de máxima verossimilhança é aquela que maximiza a probabilidade de observação dos dados experimentais f , conforme mostra a Equação (3.1):

$$\max_{\hat{\underline{z}}} f = \max_{\hat{\underline{z}}} \prod_{i=1}^n f_i \quad (3.1)$$

Entretanto, na prática é mais conveniente trabalhar com a minimização da função de verossimilhança (REY, 1983), conforme mostra a Equação (3.2):

$$\max_{\hat{\underline{z}}} \prod_{i=1}^n f_i = - \min_{\hat{\underline{z}}} \prod_{i=1}^n f_i \quad (3.2)$$

O desenvolvimento da estimativa de máxima verossimilhança frequentemente faz uso da aplicação do logaritmo Neperiano sobre o produtório. Assim, a Equação (3.2) pode ser escrita na forma da Equação (3.3):

$$\min_{\hat{\underline{z}}} - \ln \left[\prod_{i=1}^n f_i \right] = \min_{\hat{\underline{z}}} - \sum_{i=1}^n \ln [f_i] \quad (3.3)$$

O estimador MQP, também conhecido como estimador L_2 (REY, 1983) e que advém da distribuição Normal, já foi desenvolvido no Capítulo 2. Outros estimadores para regressão de dados podem ser construídos, conforme mostrado nos exemplos seguintes. Para a **Distribuição Normal Contaminada** (NC) (JEFFREYS, 1932):

$$\max_{\hat{z}} \prod_{i=1}^n \left\{ (1-p) \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] + p \frac{1}{b\sigma_i \sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{b^2 \sigma_i^2} \right] \right\} \quad (3.4)$$

$$\min_{\hat{z}} - \sum_{i=1}^n \ln \left\{ (1-p) \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] + \frac{p}{b} \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{b^2 \sigma_i^2} \right] \right\} + \sum_{i=1}^n \ln(\sigma_i \sqrt{2\pi}) \quad (3.5)$$

$$\min_{\hat{z}} - \sum_{i=1}^n \ln \left\{ (1-p) \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] + \frac{p}{b} \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{b^2 \sigma_i^2} \right] \right\} \quad (3.6)$$

$$\min_{\hat{z}} - \sum_{i=1}^n \ln \left[(1-p) \exp \left(-\frac{1}{2} \xi_i^2 \right) + \frac{p}{b} \exp \left(-\frac{1}{2b^2} \xi_i^2 \right) \right] \quad (3.7)$$

O resíduo padronizado ξ_i corresponde à diferença, ponderada pelo desvio padrão σ_i , entre o valor medido z_i e o valor estimado (reconciliado) pelo modelo \hat{z}_i , para a variável i . A Equação (3.7) apresenta o estimador com base na distribuição Normal Contaminada, também chamada bivariada (TJOA e BIEGLER, 1991), em que ξ_i^2 é o resíduo padronizado; p é a probabilidade de obter valores espúrios (com $p < 0,5$); e $b^2 \sigma_i^2$ é o grau de contaminação da variância (com $b > 1$).

Para a **Distribuição Laplaciana** (DE LAPLACE, 1774):

$$\max_{\hat{z}} \prod_{i=1}^n \left\{ \frac{1}{2\sigma_i} \exp \left[-\frac{|z_i - \hat{z}_i|}{\sigma_i} \right] \right\} \quad (3.8)$$

$$\min_{\hat{z}} - \sum_{i=1}^n \ln \left[\exp \left(-\frac{|z_i - \hat{z}_i|}{\sigma_i} \right) \right] + \sum_{i=1}^n \ln(2\sigma_i) \quad (3.9)$$

$$\min_{\hat{z}} \sum_{i=1}^n \frac{|z_i - \hat{z}_i|}{\sigma_i} \quad (3.10)$$

$$\min_{\hat{z}} \sum_{i=1}^n |\xi_i| \quad (3.11)$$

A distribuição Laplaciana, também conhecida como exponencial dupla, resulta no estimador de mínimos valores absolutos - MVA, apresentado na Equação (3.11), também conhecido como estimador L_1 (REY, 1983).

Para a **Distribuição de Cauchy** (CAUCHY, 1853):

$$\max_{\hat{z}} \prod_{i=1}^n \left\{ \frac{1}{\sigma_i \pi \left[1 + \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right]} \right\} \quad (3.12)$$

$$\min_{\hat{z}} - \sum_{i=1}^n \ln \left\{ \sigma_i \pi \left[1 + \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] \right\}^{-1} \quad (3.13)$$

$$\min_{\hat{z}} \sum_{i=1}^n \ln \left\{ \sigma_i \pi \left[1 + \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] \right\} \quad (3.14)$$

$$\min_{\hat{z}} \sum_{i=1}^n \ln \left[1 + \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] + \sum_{i=1}^n \ln(\sigma_i \pi) \quad (3.15)$$

$$\min_{\hat{z}} \sum_{i=1}^n \ln (1 + \xi_i^2) \quad (3.16)$$

A Figura 3.1 apresenta graficamente as funções de distribuição de probabilidades: Normal, Normal Contaminada, Laplaciana e Cauchy. Na forma apresentada, observa-se que a principal diferença entre esses estimadores diz respeito à tolerância por desvios elevados.

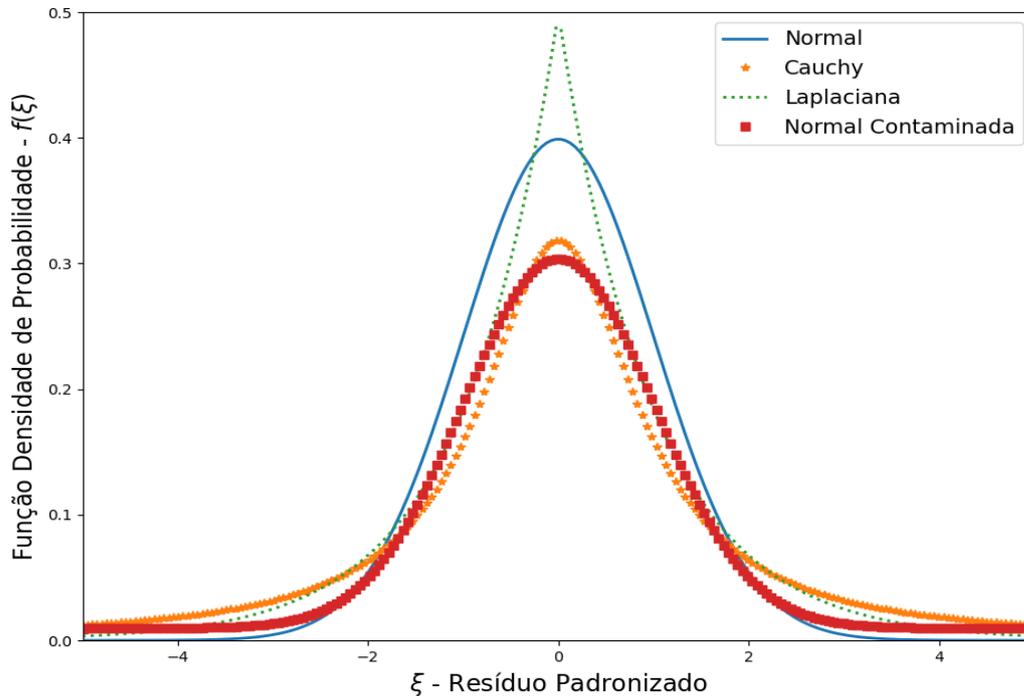


Figura 3.1: Algumas funções de distribuição de probabilidades.

Alguns estimadores-M podem ser apresentados com base na derivada do estimador; ou seja, baseado na função de influência, e não em uma distribuição

de probabilidades de origem conhecida. Conforme discutido por REY (1983), alguns estimadores-M têm sido construídos com base em uma estrutura matemática conveniente, como por exemplo, o estimador Fair (FAIR, 1974), descrito adiante.

3.2.2 Generalização da função objetivo de máxima verossimilhança

A generalização da função objetivo de máxima verossimilhança foi proposta por HUBER (1964). Em 1973, o mesmo autor estendeu a ideia de estimador-M para regressão, obtido por meio da minimização de uma função suave, simétrica e razoavelmente monotônica dos resíduos $-\rho$ em relação às estimativas ótimas (PRATA, 2009), conforme ilustrado na Equação (3.17).

$$\hat{z} = \min_{\hat{z}} \sum_{i=1}^n \rho \left(\frac{z_i - \hat{z}_i}{\sigma_i^2} \right) = \min_{\hat{z}} \sum_{i=1}^n \rho(\xi_i) \quad (3.17)$$

Nenhum estimador-M é livre da escala (desvio-padrão - σ), dado pela medida de dispersão. Estimadores-M também podem estimar a localização (média - μ), que é a média do conjunto amostral da variável. O procedimento é iterativo como demonstrado nos Algoritmos 1 e 2 (ZOUBIR *et al.*, 2018). Entretanto, para estimar a média de forma simples e robusta utiliza-se a mediana, $Q_{2/4}(z)$. Na Equação (3.17) é admitido que os erros de medição não são correlacionados, ou seja, independentes, sendo a matriz de covariância-variância diagonal.

A estimativa da escala pode ser obtida pela estatística tradicional ou de maneira robusta, usando os próprios estimadores-M com algoritmo iterativo de ponderação ou com os estimadores-L robustos MADn (HAMPEL, 1974; ZOUBIR *et al.*, 2018), Tn (ROUSSEEUW e CROUX, 1992) e ainda, Sn e Qn ROUSSEEUW e CROUX (1993), explicados na seção 2.2, pelos métodos direto ou indireto. Encontrar um ótimo global para a Localização e Escala de funções de Máxima Verossimilhança não-Gaussianas ainda é um problema em aberto (ZOUBIR *et al.*, 2018). A Figura 3.2 ilustra os Estimadores-M - $\rho(\xi)$ - para as funções: Normal, Normal Contaminada, Laplaciana e Cauchy.

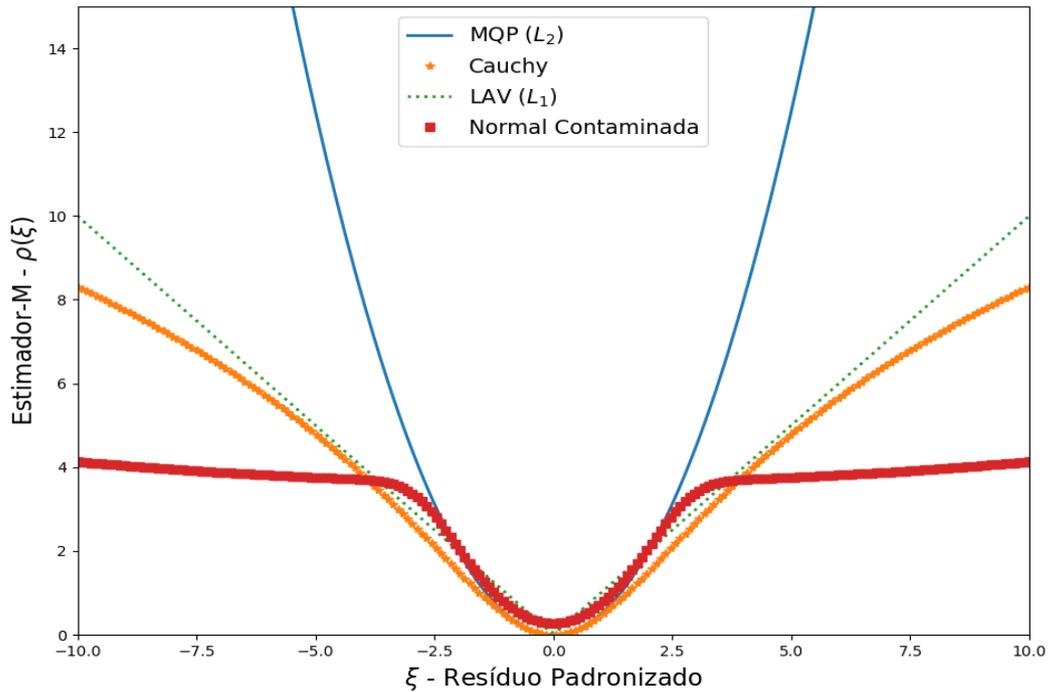


Figura 3.2: M-Estimators - $\rho(\xi)$.

Algumas propriedades desejáveis para ρ são (HOAGLIN *et al.*, 1983):

- ρ deve ser contínua;
- ρ deve ser simétrica: $\rho(\xi_i) = \rho(-\xi_i)$;
- $\rho(\xi_i) \geq 0$;
- $\rho(\xi_i) \geq \rho(\xi_j)$, para $|\xi_i| > |\xi_j|$;
- $\rho(\xi_i = 0) = 0$;
- ρ deve ser convexa, preferencialmente.

A convexidade de ρ garante unicidade de solução (ótimo global) para problemas descritos somente por modelos lineares (HUBER e RONCHETTI, 2009), ou problemas não lineares que possam ser escritos por restrições lineares, após linearização. Deve-se perceber que as propriedades desejáveis de ρ podem ser modificadas convenientemente em diferentes problemas, dado que não constituem requisito matemático fundamental.

3.2.3 Função de influência

HAMPEL (1968) introduziu o conceito de Função (ou curva) de Influência (FI) como medida de robustez. Para um estimador ser considerado robusto, a função de influência deve ser limitada. Pode-se interpretar que a função de influência ψ corresponde ao peso (influência ou impacto) da magnitude de um valor espúrio (quase sempre mensurado em termos de múltiplos do resíduo padronizado ξ) sobre as estimativas obtidas. Esta função pode ser definida na forma da Equação (3.18), caracterizada pela primeira derivada da função objetivo, ρ , em relação ao resíduo padronizado, ξ (HAMPEL *et al.*, 1986).

$$\psi(\xi) = \frac{\partial \rho(\xi)}{\partial \xi} \quad (3.18)$$

O estimador MQP não é robusto, pois a respectiva função de influência é ξ_i , de maneira que a influência de valores espúrios sobre as estimativas é ilimitada (falta de robustez) e aumenta proporcionalmente com o aumento da magnitude do valor espúrio. A Equação (3.19) apresenta a FI do Estimador-M NC, que também não apresenta comportamento robusto na forma considerada.

$$\psi_{NC}(\xi) = \frac{\partial \rho_{NC}(\xi)}{\partial \xi} = \frac{\xi_i}{b^2} \left[\frac{b^3(p-1) \exp\left(-\frac{\xi_i^2}{2}\right) - p \exp\left(-\frac{\xi_i^2}{2b^2}\right)}{b(p-1) \exp\left(-\frac{\xi_i^2}{2}\right) - p \exp\left(-\frac{\xi_i^2}{2b^2}\right)} \right] \quad (3.19)$$

Os estimadores de Welsch (REY, 1983) e Hampel (HAMPEL *et al.*, 1986) apresentam função de influência que decresce de forma limitada e tendendo a zero, anulando o efeito negativo dos valores espúrios sobre as estimativas com o aumento da magnitude, sendo por isso classificados como **descendentes** (*redescending*). Cabe ressaltar que a função ρ destes estimadores não é convexa. Segundo (ROUSSEEUW e YOHAI, 1984), os estimadores robustos podem ser divididos como: **descendentes suaves** (*soft-redescending*), em que a função de influência se aproxima de zero assintoticamente (como Welsch), e **descendentes bruscos** (*hard-redescending*), em que a função de influência torna-se nula para valores espúrios de magnitude elevada (como Hampel). O estimador de Fair está sujeito à influência dos valores espúrios, mesmo que de forma limitada, sendo por isso classificado como **monótono**. Entretanto, esta influência é bem menor do que no caso do estimador de MQP, como ilustra a Figura 3.3. Cabe ressaltar que a função ρ de um estimador monótono é convexa. Dessa forma, os estimadores-M podem ser classificados genericamente em termos da estrutura da FI: Não Robusto (como MQP); quasi-

Robusto (como NC); Robusto monótono (como Fair); Robusto *soft-redescending* (como Welsch); Robusto *hard-redescending* (como Hampel). Uma comparação entre diversas funções de influência é apresentada na Figura 3.3.

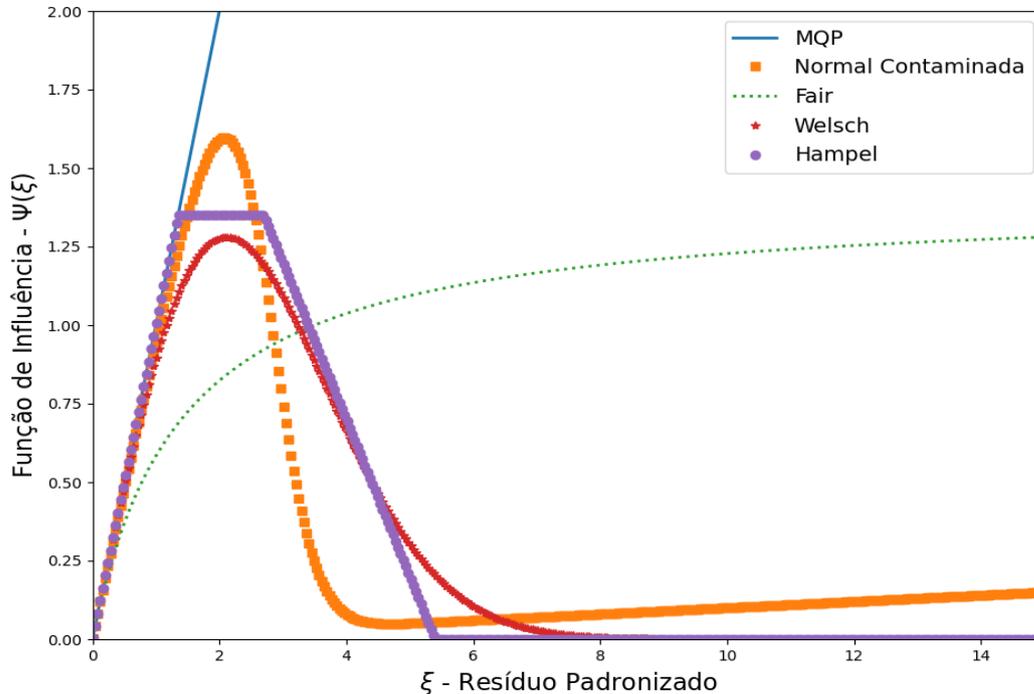


Figura 3.3: Algumas funções de influência de alguns estimadores (DE MENEZES, 2015).

Na Figura 3.3 observa-se que o estimador Normal Contaminada apresenta redução parcial de influência; entretanto, a FI é ilimitada, à medida que aumenta o resíduo padronizado, o que caracteriza a não robustez. A FI dos estimadores de Hampel tornam-se identicamente nulas em $\xi = 5, 4$; a partir daí, o peso atribuído pelo erro à função objetivo é nulo, sendo por isso chamados de *hard-redescending* (descendente brusco). Já o estimador de Welsch apresentam FI com uma assíntota em zero, sendo classificados como *soft-redescending* (descendente suave). Todos os estimadores-M foram sintonizados para eficiência relativa de 95% em relação à distribuição Normal, conceito explicado na Seção 3.2.5.

Com base na estrutura da função de influência, três outras propriedades podem ser também derivadas (HAMPEL *et al.*, 1986):

- Sensibilidade a erros grosseiros, γ (*gross error sensitivity*), definida como o supremo do valor absoluto da função de influência, medindo o efeito máximo que a contaminação do valor medido pode exercer sobre o estimador;

- Sensibilidade em relação ao desvio de medida, λ (*local-shift sensitivity*), definida como a inclinação da função de influência no ponto considerado, medindo a influência do estimador sobre o resíduo padronizado quando a medida se afasta do erro zero.
- Ponto de Rejeição, c_P (*rejection point*), definido como o ponto a partir do qual a função de influência é identicamente nula, relacionado à anulação completa do efeito (negativo) do valor espúrio sobre as estimativas dos parâmetros e variáveis. Nos estimadores tipo descendente suave, o ponto de rejeição é um valor de caráter apenas aproximado.

A Figura 3.4 ilustra algumas das propriedades de função de influência.

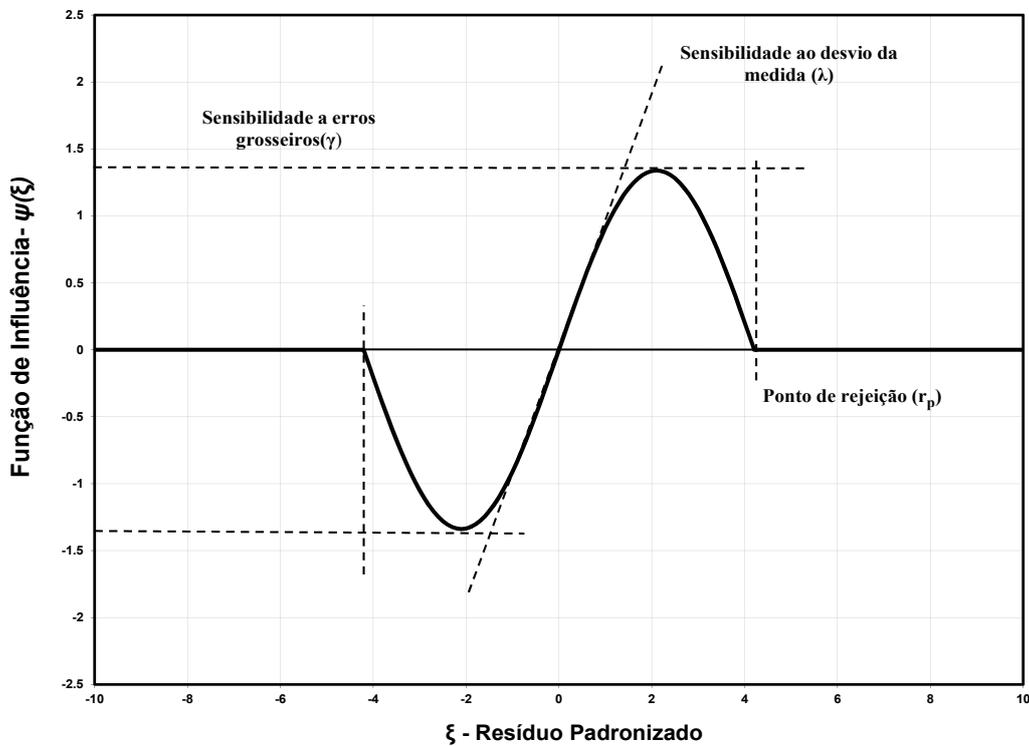


Figura 3.4: Algumas propriedades das funções de influência (HAMPEL *et al.*, 1986).

Algumas propriedades desejáveis para Ψ são (HOAGLIN *et al.*, 1983):

- ψ deve ser limitado, caracterizando a robustez (sensibilidade a erros grosseiros γ finita);
- ψ deve ser moderadamente contínua ou contínua por partes (sensibilidade em relação ao desvio de medida λ finita);
- ψ deve ser uma função ímpar: $\psi(-\xi) = -\psi(\xi)$, caracterizando simetria no trato dos erros;

- ψ deve ser aproximadamente linear na origem: $\psi(\xi) \approx k\xi$, com $k \neq 0$, para ξ pequeno, aproximando um comportamento Normal dos erros pequenos;
- ψ deve possuir um ponto de rejeição c_P finito (se a resistência a valores espúrios de grande magnitude é desejada), matematicamente similar com as rejeições dos teste de hipóteses no tratamento de erros grosseiros.

Como dito anteriormente, as propriedades desejáveis de ψ podem ser convenientemente modificadas, dado que não constitui um requisito matemático fundamental para a aplicação.

3.2.4 Derivada da função de influência

Outra função muito importante pode ser definida em termos da derivada da função de influência ψ , em relação ao resíduo padronizado ξ . Esta função corresponde à derivada segunda de ρ em relação ao resíduo padronizado ξ , conforme apresentado na Equação (3.20).

$$\psi'(\xi) = \frac{\partial\psi(\xi)}{\partial\xi} = \frac{\partial^2\rho(\xi)}{\partial\xi^2} \quad (3.20)$$

$\psi'(\xi)$ corresponde à matriz Hessiana; ou seja, a matriz de derivadas de segunda ordem da função objetivo ρ em relação às variáveis de decisão. Segundo HOAGLIN *et al.* (1983), $\psi'(\xi)$ tem sido pouco estudada no campo da estatística robusta, já que a solução do problema de regressão é usualmente realizada por métodos iterativos que não usam derivada ou baseados na matriz para a matriz de derivadas de primeira ordem, como métodos de Gauss-Newton, Gauss-Marquardt ou Gradiente Reduzido Generalizado (BENQLILOU, 2004). O estimador NC, por exemplo, apresenta $\psi'(\xi)$ definida como na Equação (3.21).

$$\psi'_{NC}(\xi) = \frac{\partial\psi_{NC}(\xi)}{\partial\xi} = \frac{\left\{ b^4(p-1)^2 \exp(-\xi_i^2) - bp(p-1) \left(b^2 - b^2\xi_i^2 + 2\xi_i^2 - \frac{\xi_i^2}{b^2} + 1 \right) \exp\left[-\frac{\xi_i^2}{2} \left(1 + \frac{1}{b^2} \right)\right] + p^2 \exp\left(-\frac{\xi_i^2}{b^2}\right) \right\}}{b^2 \left[b(p-1) \exp\left(-\frac{\xi_i^2}{2}\right) - p \exp\left(-\frac{\xi_i^2}{2b^2}\right) \right]^2} \quad (3.21)$$

Quando $\psi'(\xi)$ é descontínua, pode ocorrer falta de robustez no que diz respeito ao desempenho dos procedimentos clássicos de otimização (determinísticos). Para lidar com possíveis descontinuidades de $\psi'(\xi)$ é possível (ARORA e BIEGLER, 2001): modificar a forma descontínua de $\psi'(\xi)$ com o uso de funções de suavização,

para tornar viável o uso de procedimentos de otimização determinísticos ou utilizar procedimentos de otimização não determinísticos.

3.2.5 Robustez e eficiência relativa

Os conceitos de eficiência e robustez estão associadas com o uso de estimadores-M robustos em problema de regressão de dados (ALBUQUERQUE e BIEGLER, 1995). Por exemplo, eficiência se refere à qualidade do ajuste realizado com um estimador, quando os erros seguem uma outra distribuição, chamada de referência, quase sempre a distribuição Normal. A eficiência indica o quanto o estimador pode se aproximar da distribuição de referência. O conceito de robustez se refere ao desempenho do estimador em relação à presença de uma variedade de distribuições de erros não normais. A robustez indica o quanto o estimador depende de uma distribuição de referência.

A eficiência e a robustez podem ser geralmente manipuladas, especificando-se uma constante de sintonia. Deve-se atentar que, quanto mais robusto é um estimador-M, menos eficiente ele usualmente é (ALBUQUERQUE e BIEGLER, 1995). Estas medidas são importantes para uma comparação apropriada entre estimadores robustos, quando especificadas uma eficiência e uma distribuição de referência, permitindo comparação de regras de rejeição para supostos valores espúrios. Geralmente, as constantes de sintonia são calculadas para o nível de 95% de eficiência relativa em relação à distribuição Normal. Segundo PRATA (2009), os estimadores robustos são projetados para um desempenho muito bom quando a distribuição está contaminada; ou seja, corrompida por valores espúrios, com uma pequena perda de eficiência quando a distribuição é realmente Normal. Cabe ressaltar que a comparação de desempenho entre estimadores robustos é comumente feita com base na mesma eficiência relativa (PRATA *et al.*, 2008b; REY, 1983; ÖZYURT e PIKE, 2004). A Equação (3.22) apresenta a definição matemática de eficiência relativa.

$$E_{ff}[\psi(\xi), f(\xi)] = \frac{V_f[\psi_f(\xi), f(\xi)]}{V[\psi(\xi), f(\xi)]} \quad (3.22)$$

em que V_f é a variância assintótica do estimador de referência (para MQP como referência $V_f = 1$) e V é a variância assintótica atribuída a um estimador-M escolhido (ALBUQUERQUE e BIEGLER, 1995). Os estimadores-M tem variância assintótica V definida pela Equação (3.23) (SHEVLYAKOV e VILCHEVSKI, 2001):

$$V[\psi(\xi), f(\xi)] = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[\int_{-\infty}^{+\infty} \psi'(\xi) f(\xi) d\xi \right]^2} \quad (3.23)$$

em que f é a verdadeira distribuição de erros, ψ é a função de influência e ψ' é a respectiva derivada.

Algumas propriedades devem ser seguidas para a correta sintonia da eficiência, envolvendo relações entre ψ e f e suas derivadas ψ' e f' (HAMPEL *et al.*, 1986; SHEVLYAKOV e VILCHEVSKI, 2001):

- f deve ser simétrica e unimodal;
- f deve ser duas vezes continuamente diferenciável e satisfazer $f(\xi) > 0 \forall \xi \in \mathfrak{R}$;
- $\int_{-\infty}^{+\infty} \psi(\xi) f(\xi) d\xi = 0$ (ou seja, ψ deve ser ímpar);
- $\int_{-\infty}^{+\infty} \psi^2(\xi) d\xi < \infty$ (ou seja, ψ deve ser assintoticamente ou identicamente nula quando tende para $\pm\infty$);
- $0 < \int_{-\infty}^{+\infty} \psi'(\xi) f(\xi) d\xi = - \int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi$ (ou seja, ψ' e f' deve ser limitadas).

3.2.6 Limite de ruptura

O conceito de Limite de Ruptura (*Breakdown Point* - BP) também foi introduzido por HAMPEL (1968) e é outra medida importante de robustez, sendo geralmente classificada como robustez quantitativa (HUBER, 1981). O BP pode ser definido como a proporção de observações incorretas (erros grosseiros) em um conjunto de dados para que a técnica de regressão robusta ainda possa ser bem-sucedida. Quanto mais alto é o BP de um estimador, mais robusto ele é. Teoricamente, o BP mais alto possível é 0,5 (50%) (ROUSSEEUW e LEROY, 1987; ZOUBIR *et al.*, 2018). Pois, se mais da metade das medidas estiverem corrompidas, não é possível distinguir entre distribuição verdadeira e distribuição contaminada por valores espúrios. O BP de um estimador-M não depende de sua distribuição de probabilidade (ROUSSEEUW e LEROY, 1987, p. 10), (HOAGLIN *et al.*, 1983, p. 370). É importante ressaltar que o estimador MQP tem BP próximo de zero; portanto, um único valor espúrio causa desvios em suas estimativas. Isso novamente reflete a extrema sensibilidade do estimador MQPS para valores espúrios (ROUSSEEUW e LEROY, 1987). Espera-se que o BP de um estimador-M robusto

seja 0,5 (HUBER, 1984) - podendo lidar com até 50% de erros grosseiros nos dados. Isso foi comprovado assintoticamente por dados de simulação para os estimadores-M Biweight, Hampel, Andrews e Tangente Hiperbólica (ZHANG *et al.*, 1998).

3.2.7 Exemplos de estimadores-M robustos

Muitos estimadores-M foram propostos na literatura e alguns destes já foram utilizados em problemas de regressão de dados, visando a minimizar a influência negativa dos menos frequentes valores espúrios. Uma revisão completa sobre os estimadores-M robustos é apresentada no Apêndice B. Entretanto, nesta seção são apresentados apenas alguns estimadores. Portanto, as Tabelas 3.1, 3.2 e 3.3 apresentam as funções ρ , ψ , ψ' como função do resíduo padronizado ξ e de suas respectivas constantes de sintonia. Assim, todos os estimadores-M apresentados e citados nesta seção são:

- Normal Contaminada (quasi-robusto)
- Fair (monótono)
- Cauchy (*soft-redescending*)
- Welsch (*soft-redescending*)
- Collins (*hard-redescending*)
- Hampel (*hard-redescending*)

Tabela 3.1: Alguns estimadores-M: $\rho(\xi)$

Estimador-M	$\rho(\xi_i)$	Faixa
Normal Contaminada	Equação (3.7)	
Fair	$c_F^2 \left[\frac{ \xi_i }{c_F} - \ln \left(1 + \frac{ \xi_i }{c_F} \right) \right]$	$ \xi_i \leq \infty$
Cauchy	$\frac{c_C^2}{2} \ln \left(1 + \frac{\xi_i^2}{c_C^2} \right)$	$ \xi_i \leq \infty$
Welsch	$\frac{c_W^2}{2} \left[1 - \exp \left(-\frac{\xi_i^2}{c_W^2} \right) \right]$	$ \xi_i \leq \infty$
Collins	$\begin{cases} \frac{\xi_i^2}{2} & \xi_i \leq qc_l \\ -2 \left\{ \log \left[\cosh \left(\frac{ac_l}{2} (\xi_i - rc_l) \right) \right] \right\} + dc_l & qc_l < \xi_i \leq rc_l \\ dc_l & \xi_i > rc_l \end{cases}$	$qc_l < \xi_i \leq rc_l$
Hampel	$\begin{cases} \frac{\xi_i^2}{2} & \xi_i \leq a_H \\ a_H \xi_i - \frac{a_H^2}{2} & a_H < \xi_i \leq b_H \\ a_H b_H - \frac{a_H^2}{2} + \frac{a_H(c_H - b_H)}{2} \left[1 - \left(\frac{c_H - \xi_i }{c_H - b_H} \right)^2 \right] & b_H < \xi_i \leq c_H \\ a_H b_H - \frac{a_H^2}{2} + \frac{a_H(c_H - b_H)}{2} & \xi_i > c_H \end{cases}$	$b_H < \xi_i \leq c_H$

A Figura 3.5 ilustra as funções ρ dos estimadores-M robustos apresentados, conjuntamente com o estimador clássico de MQP. Os estimadores-M robustos foram

Tabela 3.2: Alguns estimadores-M: $\psi(\xi)$

Estimador-M	$\psi(\xi_i)$	Faixa
Normal Contaminada	Equação (3.19)	$ \xi_i \leq \infty$
Fair	$\frac{\xi_i}{1 + \frac{ \xi_i }{c_F}}$	$ \xi_i \leq \infty$
Cauchy	$\frac{\xi_i}{1 + \frac{\xi_i^2}{c_C^2}}$	$ \xi_i \leq \infty$
Welsch	$\xi_i \exp\left(-\frac{\xi_i^2}{c_W}\right)$	$ \xi_i \leq \infty$
Collins	$\begin{cases} \xi_i^2 & \xi_i \leq qc_l \\ a_{cl} \tanh\left[\frac{a_{cl}}{2}(qc_l - \xi_i)\right] \text{sgn}(\xi_i) & qc_l < \xi_i \leq rc_l \\ 0 & \xi_i > rc_l \end{cases}$	$ \xi_i \leq qc_l$ $qc_l < \xi_i \leq rc_l$ $ \xi_i > rc_l$
Hampel	$\begin{cases} \xi_i^2 & \xi_i \leq a_H \\ a_H \text{sgn}(\xi_i) & a_H < \xi_i \leq b_H \\ \frac{a_H(c_H - \xi_i)}{c_H - b_H} \text{sgn}(\xi_i) & b_H < \xi_i \leq c_H \\ 0 & \xi_i > c_H \end{cases}$	$ \xi_i \leq a_H$ $a_H < \xi_i \leq b_H$ $b_H < \xi_i \leq c_H$ $ \xi_i > c_H$

Tabela 3.3: Alguns estimadores-M: $\psi'(\xi)$

Estimador-M	$\psi'(\xi_i)$	Faixa
Normal Contaminada	Equação (3.21)	$ \xi_i \leq \infty$
Fair	$\frac{1}{\left(1 + \frac{ \xi_i }{c_F}\right)^2}$	$ \xi_i \leq \infty$
Cauchy	$\frac{1}{\left(1 + \frac{\xi_i^2}{c_C^2}\right)^2}$	$ \xi_i \leq \infty$
Welsch	$\left(1 - \frac{2\xi_i^2}{c_W}\right) \exp\left(-\frac{\xi_i^2}{c_W}\right)$	$ \xi_i \leq \infty$
Collins	$\begin{cases} 1 & \xi_i \leq qc_l \\ -\frac{a_{cl}^2}{2} \cosh\left[\frac{a_{cl}}{2}(\xi_i - qc_l)\right]^{-2} & qc_l < \xi_i \leq rc_l \\ 0 & \xi_i > rc_l \end{cases}$	$ \xi_i \leq qc_l$ $qc_l < \xi_i \leq rc_l$ $ \xi_i > rc_l$
Hampel	$\begin{cases} 1 & \xi_i \leq a_H \\ 0 & a_H < \xi_i \leq b_H \\ -\frac{a_H}{c_H - b_H} & b_H < \xi_i \leq c_H \\ 0 & \xi_i > c_H \end{cases}$	$ \xi_i \leq a_H$ $a_H < \xi_i \leq b_H$ $b_H < \xi_i \leq c_H$ $ \xi_i > c_H$

elaboradas considerando a eficiência relativa padrão de 95% em relação à distribuição Normal, cujas constantes de sintonia são apresentadas na Tabela 3.4.

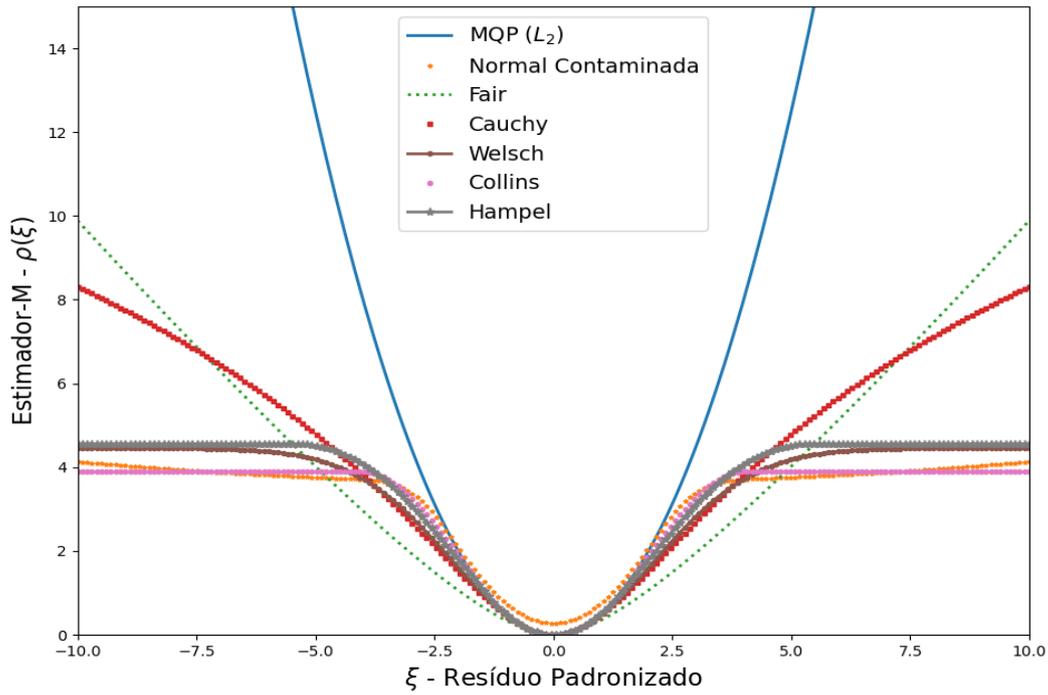


Figura 3.5: Uma comparação entre alguns dos estimadores analisados ($\rho(\xi)$).

Tabela 3.4: Valores das constantes de sintonia para 95% de eficiência relativa em relação à distribuição Normal e referências para os estimadores-M analisados

Estimador-M	Constantes	Referência (constante)
Normal Contaminada	$\begin{cases} b = 10 \\ p = 0,235 \end{cases}$	(ÖZYURT e PIKE, 2004)
Fair	$c_F = 1,3998$	(ÖZYURT e PIKE, 2004)
Cauchy	$c_C = 2,3849$	(ÖZYURT e PIKE, 2004)
Welsch	$c_W = 2,9846$	(PRATA, 2009)
Collins	$\begin{cases} a_{Cl} = 1,65145901 \\ q_{Cl} = 1,590796619 \\ r_{Cl} = 4 \\ d_{Cl} = 3,894797324 \end{cases}$	(DE MENEZES, 2015)
Hampel	$\begin{cases} a_H = 1,35 \\ b_H = 2,70 \\ c_H = 5,40 \end{cases}$	(PRATA, 2009)

3.3 Estimador de Correntropia

Esta seção introduz os conceitos, as propriedades e a formulação do estimador de Correntropia, visto que o estimador não é oriundo da estatística robusta e do princípio de Máxima Verossimilhança. Entretanto, recentemente ele vem sendo utilizado em problemas de RD e DMEG por diversos autores e em diversas áreas da engenharia. Dito isto, devemos definir sua origem, teoria e fundamentos que deram base para a formulação do estimador de Correntropia. Com isso, esta seção aborda a teoria da informação, o conceito de entropia da informação, o conceito do estimador de densidade *kernel*, o potencial de informação e, por fim, a correntropia, suas propriedades, sua similaridade com os estimadores-M e aplicações.

3.3.1 Teoria da Informação e entropia

A Teoria da Informação é um ramo da matemática que estuda a quantificação da informação, inicialmente destinada a resolver problemas de engenharia de telecomunicações. Essa teoria teve seus pilares estabelecidos por SHANNON (1948), que formalizou vários conceitos em aplicações relacionadas à Teoria da Comunicação e da Estatística.

A palavra *comunicação* é entendida por SHANNON e WEAVER (1964) em um sentido filosófico amplo, para incluir todos os procedimentos pelos quais uma mente pode afetar outra. Portanto, para os autores, o significado da palavra envolve não apenas a fala escrita e oral, mas também a música, as artes pictóricas, o teatro, o balé e, de fato, todo o comportamento humano. Com isso, os autores definem três níveis de problemas compatíveis com esse amplo conceito de *comunicação*:

- Nível A: Com qual precisão os símbolos de comunicação podem ser transmitidos? (O problema técnico);
- Nível B: Quão precisamente os símbolos transmitidos transmitem o significado desejado? (O problema semântico);
- Nível C: Quão efetivamente o significado recebido afeta a conduta da maneira desejada? (O problema de eficácia).

O problema técnico se preocupa com a precisão da transferência do remetente para o receptor de conjuntos de símbolos; o discurso escrito envolve a transmissão de um conjunto finito de símbolos discretos; a transmissão telefônica, rádio, de voz ou música transmite uma função contínua no tempo, enquanto a televisão transmite

muitas funções contínuas no tempo e dependente de duas coordenadas espaciais. Os problemas semânticos dizem respeito à identidade satisfatória, para interpretação do significado pelo receptor, em comparação com o significado pretendido do remetente. Esta é uma situação muito relevante, mesmo em problemas simples, como o de comunicação através da fala. O problema de eficácia se preocupa com o sucesso com o qual o significado transmitido ao receptor leva à conduta desejada de sua parte. Em suma, esses problemas todos estão intimamente relacionados e se sobrepõem de forma bastante vaga (SHANNON e WEAVER, 1964).

Com base no exposto, Claude Shannon desenvolveu uma teoria matemática admitindo inicialmente apenas o problema de nível A; ou seja, o problema técnico de precisão da transferência de vários tipos de sinais do transmissor para o receptor. O diagrama esquemático do sistema de comunicação proposto por Shannon é ilustrado na Figura 3.6.

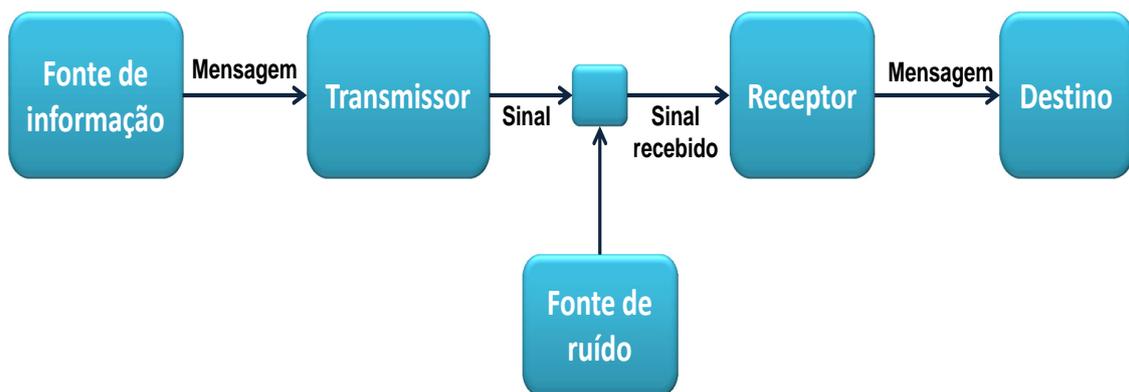


Figura 3.6: Diagrama esquemático de um sistema de comunicação de SHANNON e WEAVER (1964).

Com base neste modelo, Shannon definiu a capacidade de transmissão de um canal como a quantidade de sinais transmitidos, durante um intervalo de tempo T , definido na Equação (3.24) (PINEDA, 2006).

$$C = \lim_{T \rightarrow \infty} \frac{\log N(T)}{T} \quad (3.24)$$

em que C é a capacidade de transmissão de um canal (usualmente medida em bits por segundo), $N(T)$ é a quantidade de símbolos permitidos durante o período de amostragem T .

Diante de fatos experimentais, Shannon afirmou que o processo de geração de informação é um processo não-determinístico discreto e, portanto, deve ser analisado

de maneira probabilística.

Um sistema que produz uma sequência aleatória de símbolos tem caráter de **processo estocástico** (PINEDA, 2006). Um tipo especial de processo estocástico, o cujas características estatísticas não variam com o tempo é classificado como um **processo estocástico estacionário**. Um tipo especial de processo estocástico estacionário, para os quais as médias de um conjunto são iguais às médias temporais de qualquer função amostra, é classificado como um **processo ergódico** (YNOGUTI, 2011). Para mais detalhes sobre processos estocásticos ver seção 5.2.

Assim, Shannon propôs que a Quantidade de Informação ou Incerteza Probabilística gerada por um processo ergódico deve considerar todas as probabilidades p_i de ocorrência de todos os n eventos i possíveis, multiplicado pelo seu próprio logaritmo, como descrito na Equação (3.25) (BESSA, 2009; NASCIMENTO, 2013b).

$$H = - \sum_{i=1}^n p_i \log p_i \quad (3.25)$$

onde $\sum_{i=1}^n p_i = 1$. Nesse tratamento um certo evento i representa um certo símbolo emitido pela fonte, com probabilidade p_i .

A Equação (3.25) tem papel fundamental na Teoria da Informação e a forma de Quantidade média de Informação H foi reconhecida, por analogia, como a Entropia da Mecânica Estatística, como no Teorema-H de Boltzmann da Teoria dos Gases (PINEDA, 2006). Assim, o termo Entropia passou a ser utilizado para designar a Quantidade de Informação, ou seja, a Equação (3.25) tratar estatisticamente uma mensagem considerando o conjunto de símbolos que a compõe, da mesma forma que a mecânica estatística trata uma medida de probabilidade sobre um macro-estado, levando em consideração todas as configurações possíveis dos micro-estados (NASCIMENTO, 2013b).

Para variáveis aleatórias contínuas com FDP $f_y(z)$, pode-se obter a versão contínua de entropia pela definição de entropia de *Renyi* (BESSA, 2009), como definido na Equação (3.26).

$$H_{R\alpha} = \frac{1}{1-\alpha} \log \left(\int_{-\infty}^{+\infty} f_y(z)^\alpha dz \right) \quad (3.26)$$

Quando $\alpha = 2$, obtém-se a entropia quadrática de *Renyi*, como definido na Equação (3.27).

$$H_{R2} = -\log \left(\int_{-\infty}^{+\infty} f_y(z)^2 dz \right) \quad (3.27)$$

A Figura 3.7 ilustra uma comparação de 3 diferentes sistemas e suas distribuições. No sistema *A* a ocorrência dos símbolos *i* são equiprováveis, onde ocorre a Entropia máxima, maior variabilidade possível dos símbolos. No sistema *C* temos a concentração de probabilidade em torno de um grupo de símbolos, o que faz com que esse sistema tenha uma variabilidade de símbolos menor. O sistema *B* é um caso intermediário entre os sistemas *A* e *C* de ocorrência de símbolos (NASCIMENTO, 2013b).

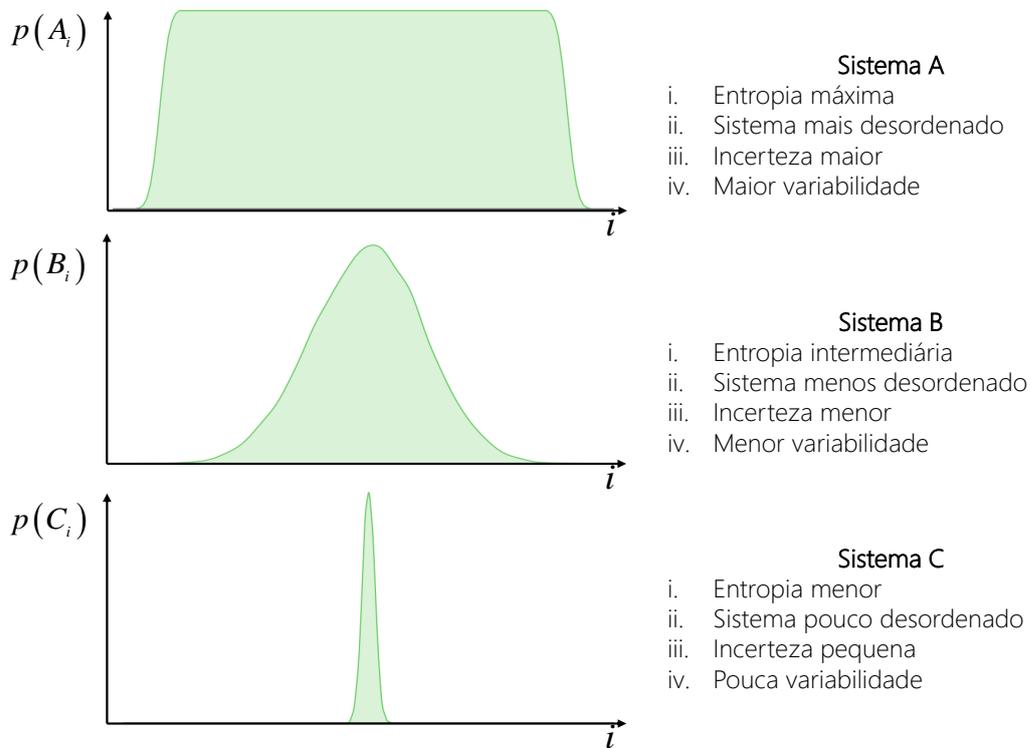


Figura 3.7: Três sistemas com diferentes níveis de entropia.
Adaptado: NASCIMENTO (2013b).

A Tabela 3.5 compara os conceitos de Entropia conforme as ideias de Boltzmann e Shannon (PINEDA, 2006).

Tabela 3.5: Conceitos de Entropia derivados das teorias de Shanonn e Boltzmann (PINEDA, 2006).

Entropia		
<i>Autor</i>	Boltzmann	Shannon
<i>Objeto</i>	Gases	Informação
<i>Conceito</i>	Desordem molecular	Quantidade de informação
<i>Medida</i>	Distribuição das velocidades das moléculas	Probabilidade de ocorrência de símbolos
<i>Variação</i>	Aumenta, à medida que o movimento das moléculas se torna aleatório	Aumenta, à medida que a ocorrência dos símbolos se torna aleatória (a distribuição se aproxima à equiprobabilidade)
<i>Tendência</i>	Num sistema fechado aumenta, até atingir o equilíbrio (distribuição de Maxwell-Boltzmann)	Numa língua natural, a frequência de símbolos tende à distribuição Medida de Gibbs (generalização do conjunto canônico para sistemas infinitos)

3.3.2 Estimativa de Densidade *Kernel*

A técnica Estimativa de Densidade *Kernel* (KDE - *Kernel Density Estimation*) ou Janela de *Parzen* é uma técnica que estima a FDP de uma variável aleatória contínua a partir de uma amostra discreta de N pontos, $i = 1, 2, \dots, N$ em um espaço M -dimensional. Esta técnica usa uma função núcleo (*Kernel*) centrada em cada ponto da amostra, atribuindo a cada ponto uma FDP de Dirac (PARZEN, 1962). Assim, a expressão da estimativa de um FDP real $f_y(z)$ pode ser descrita a partir da soma das contribuições individuais do *Kernel* aplicado em cada ponto, conforme a Equação (3.28).

$$f_y(z) = \frac{1}{N} \sum_{i=1}^N K_{\sigma_{co}}(z - y_i) \quad (3.28)$$

em que o vetor $y_i \in \mathfrak{R}^M$ representa amostras da variável aleatória, $K_{\sigma_{co}}(\cdot)$ é a função *Kernel* e σ_{co} é o tamanho da janela de *Parzen* ou largura de banda do *Kernel*. O *Kernel* Gaussiano simétrico é o mais utilizado, pois é derivável, contínuo, não-negativo e permite simplificações (BESSA, 2009). A propriedade mais interessante e importante entre duas variáveis aleatórias independentes $y_i \in \mathfrak{R}^M$ e $y_j \in \mathfrak{R}^M$ do

Kernel Gaussiano é a relação dada pela Equação (3.29).

$$\int_{-\infty}^{+\infty} G(z - y_i, \underline{\underline{\Sigma}}_1) G(z - y_j, \underline{\underline{\Sigma}}_2) dz = G(y_i - y_j, \underline{\underline{\Sigma}}_1 + \underline{\underline{\Sigma}}_2) \quad (3.29)$$

em que G é o *Kernel* Gaussiano, $\underline{\underline{\Sigma}}_1$ e $\underline{\underline{\Sigma}}_2$ são duas matrizes de covariâncias. A equação representa a convolução de duas funções Gaussianas, que gera outra função Gaussiana, cuja variância é a soma das variâncias iniciais. O *Kernel* Gaussiano simétrico tem a seguinte forma (BESSA, 2009):

$$K_{\sigma_{co}}(z - y_j, \sigma_{co}^2) = G(z - y_j, \sigma_{co}^2) = \frac{1}{\sigma_{co}\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_{co}^2}(z - y_j)^2\right] \quad (3.30)$$

O tamanho da janela de *Parzen* na estimativa da FDP tem grande importância. Valores altos de σ_{co} levam a estimativas mais suaves, enquanto valores menores geram funções mais irregulares e com mais picos, como ilustra a Figura 3.8. A Figura 3.8 foi gerada no *software* MATLAB®.

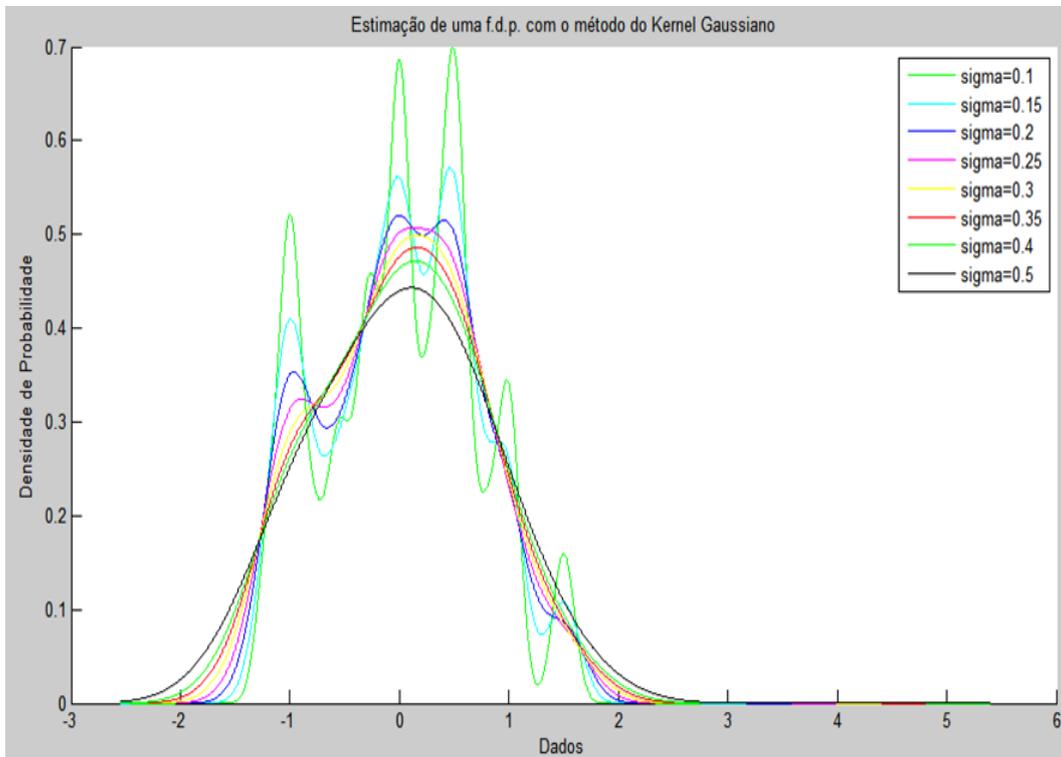


Figura 3.8: Efeito do tamanho da janela de *Parzen* sobre a estimativa de FDP.

3.3.3 Correntropia e suas propriedades

A Correntropia é uma medida estatística que apresenta características de autocorrelação, muitas vezes citada como autocorrentropia (LIU *et al.*, 2007). Uma propriedade da Correntropia é a autocorrelação e a similaridade entre instantes temporais, embora o valor médio ao longo dos instantes convirja assintoticamente para o Potencial de Informação estimado com as janelas de *Parzen* (SANTAMARÍA *et al.*, 2006). A propriedade de similaridade está diretamente relacionada com quão similares duas variáveis aleatórias são na vizinhança do espaço de junção, controlado pela largura da janela de *Parzen*. A possibilidade de ajustar o tamanho da janela de *Parzen* fornece um mecanismo eficaz que permite eliminar o efeito de *outliers*. A Correntropia tem muitas propriedades interessantes, como estabelecer conexões entre treinar algoritmos de aprendizagem baseada nos conceitos da Teoria da Informação, estimar uma FDP pelo método da densidade de *Kernel* e estimador-M robusto da estatística robusta (LIU *et al.*, 2006).

Admitindo dois vetores de variáveis aleatórias $\underline{X} = (x_1, x_2, \dots, x_N)$ e $\underline{Y} = (y_1, y_2, \dots, y_N)$ e definindo o erro como $\underline{\varepsilon} = \underline{X} - \underline{Y} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)$, então pode-se definir o Critério Maximização da Correntropia (MCC) entre duas variáveis aleatórias escalares como:

$$\hat{V}(\underline{X}, \underline{Y}) = \mathbf{E}[K_{\sigma_{co}}(\underline{X} - \underline{Y})] \cong \frac{1}{N} \sum_{i=1}^N K_{\sigma_{co}}(\varepsilon_i) \quad (3.31)$$

LIU *et al.* (2007) e SANTAMARÍA *et al.* (2006) descreveram algumas das propriedades da Correntropia. Três propriedades devem ser ressaltadas: 1) a Correntropia é positiva, limitada e atinge o valor máximo se $\underline{X} = \underline{Y}$; 2) permite quantificar a probabilidade de dois eventos serem iguais; 3) induz uma métrica no espaço amostral.

O *Kernel* usualmente utilizado para o estimador de Correntropia é o Gaussiano, conforme descrito na Equação (3.29). Para definir o estimador de Correntropia dentro do aspecto geral dos estimadores-M robustos é preciso primeiro dividir os erros de medida pelos respectivos desvios-padrões, na forma de resíduos padronizados $\underline{\xi} = (\frac{\varepsilon_1}{\sigma_1}, \frac{\varepsilon_2}{\sigma_2}, \dots, \frac{\varepsilon_N}{\sigma_N})$, e depois defini-la da seguinte forma (LIU *et al.*, 2006; ZHANG e CHEN, 2015):

$$\rho_{co}(\xi) = \frac{1}{\sigma_{co}\sqrt{2\pi}} - K_{\sigma_{co}}(\xi) = \frac{1}{\sigma_{co}\sqrt{2\pi}} \left[1 - \exp\left(-\frac{\xi^2}{2\sigma_{co}^2}\right) \right] \quad (3.32)$$

Com esta forma, a Correntropia consegue atender às propriedades desejáveis

de um estimador-M robusto:

- $\rho_{co}(\xi)$ é contínua;
- $\rho_{co}(\xi)$ é simétrica: $\rho_{co}(\xi) = \rho_{co}(-\xi)$;
- $\rho_{co}(\xi) \geq 0$;
- $\rho_{co}(\xi_i) \geq \rho_{co}(\xi_j)$, para $|\xi_i| > |\xi_j|$;
- $\rho_{co}(0) = 0$;
- ρ é robusta *soft-redescending*.

Dessa forma, o critério de MCC equivale a minimizar o estimador de Correntropia $\rho_{co}(\xi)$:

$$\min \sum_{i=1}^N \rho_{co}(\xi_i) \quad (3.33)$$

As Figuras 3.9 e 3.10 ilustram a função $\rho_{co}(\xi)$ e sua derivada $FI_{co}(\xi)$, respectivamente. A Equação (3.34) apresenta a Função de Influência do estimador Correntropia.

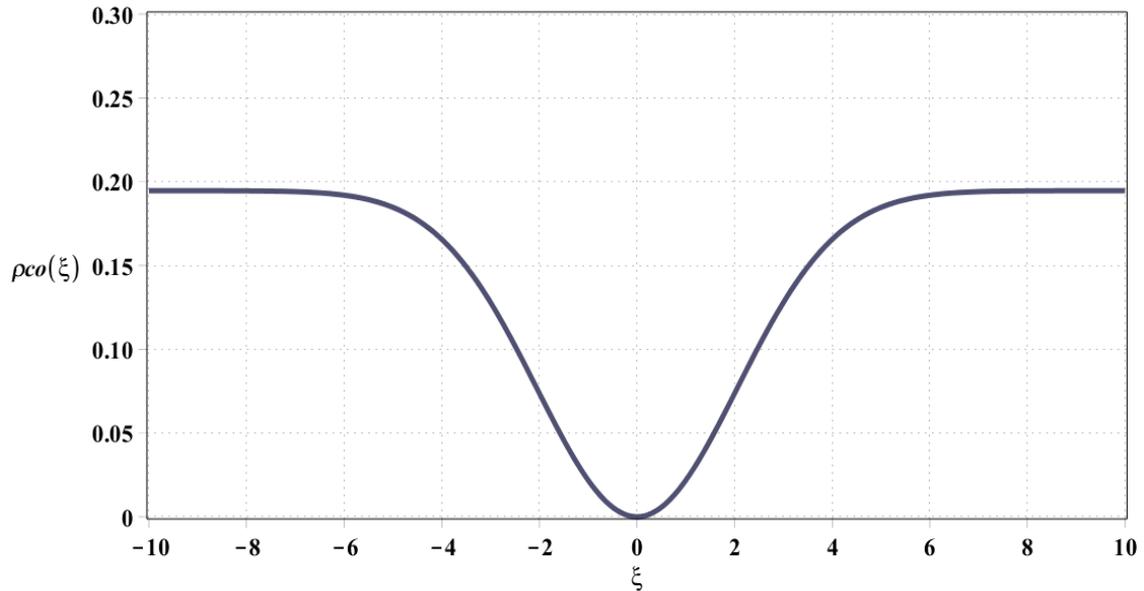


Figura 3.9: Estimador de Correntropia.

A Figura 3.10 indica que o estimador de Correntropia se classifica como um estimador *soft-redescending* robusto. Segundo LLANOS *et al.* (2015), a sintonia do

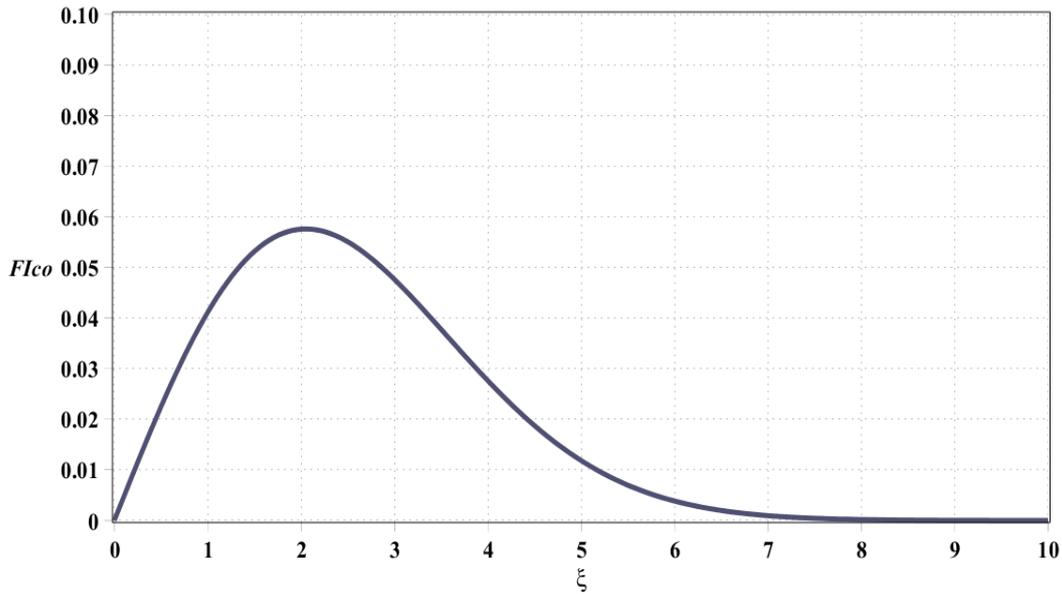


Figura 3.10: Função de Influência de Correntropia.

tamanho da janela de *Parzen* σ_{co} do estimador de Correntropia para uma eficiência relativa E_{ff} de 95% é $\sigma_{co} = 2,05$. Os autores utilizaram a minimização do Critério de Informação Akaike para sintonizar este parâmetro. No Apêndice B apresentam as sintonias de 90, 95, 98 e 99% da Correntropia, utilizando o conceito de Eficiência Relativa, apresentada na Equação (3.22).

$$FI_{co}(\xi) = \psi_{co}(\xi) = \frac{\xi}{\sigma_{co}^4 \sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2\sigma_{co}^2}\right) \quad (3.34)$$

3.3.4 Correntropia e aplicações em RD

Um dos pioneiros na utilização do estimador de Correntropia em problemas de regressão de dados robusta foram LIU *et al.* (2007) e HOSEINNEZHAD e BAB-HADIASHAR (2011). LIU *et al.* (2007) abordaram no artigo o significado probabilístico e geométrico da função de Correntropia como uma medida de similaridade localizada, apresentam uma relação estreita entre Correntropia e estimadores-M, conexões e diferenças entre os métodos de Correntropia e método de estimador *Kernel* e propriedades de Correntropia que podem ser muito úteis no processamento de sinais não gaussianos. HOSEINNEZHAD e BAB-HADIASHAR (2011) abordaram um problema de estimação de erro utilizando o método de estimador *Kernel* Gaussiano como função objetivo, chamando o estimador de *High Breakdown M-estimator* (HBM), com alto limite de ruptura. Os autores utilizaram o estimador HBM para tratamento de imagem e demonstraram que o novo estimador é computacionalmente mais rápido, sem perder eficácia no tratamento dos dados.

CHEN *et al.* (2013) foram os pioneiros em utilizar o estimador de Correntropia em problemas de RD e DMEG em engenharia química. Os autores utilizaram a minimização do Critério de Informação de Akaike para sintonizar a largura do *Kernel* ideal. Dois estudos de casos demonstraram as vantagens de usar o estimador de Correntropia: o primeiro estudo foi um conjunto de colunas atmosféricas de destilação, utilizando apenas o balanço de massa global em estado estacionário e usando dados temporais como redundância de similaridade; o segundo problema foi uma rede de medidores de vapor em estado estacionário e sujeita a restrições lineares. A eficácia do estimador proposto foi comparada com vários métodos convencionais, especialmente com o estimador MQP, demonstrando bons resultados.

ZHANG *et al.* (2014) estudaram a influência de vários fatores, como efeitos de dados medidos contaminados, as características não lineares dos parâmetros do modelo com diferentes condições de operação e erros grosseiros, sobre os problemas RDEP. Os autores propuseram uma metodologia de RDEP prática para um sistema com múltiplas condições operacionais, com o objetivo de diminuir o impacto desses efeitos. A metodologia utiliza PCA para a detecção do estado estacionário, análise de agrupamento para identificar as diferentes condições operacionais e o estimador de Correntropia para solucionar o problema de RDEP, para diferentes condições operacionais. A metodologia foi aplicada a um processo de separação de ar com múltiplas condições operacionais. A metodologia proposta apresentou eficácia, demonstrada através do problema de DRPE.

LLANOS *et al.* (2015) analisaram o desempenho entre estratégias de reconciliação de dados robustos. O estudo envolveu dois procedimentos baseados em função bponderada e três técnicas de estimação baseadas em estimadores-M: Welsh, quasi-MQP e Correntropia. Todos os procedimentos foram corretamente sintonizados para ter a mesma eficiência relativa e capacidades de detecção e identificação de erros grosseiros em relação a distribuição ideal. Os resultados foram analisados considerando os critérios de avaliação da detecção de valores espúrios: AVTI, OP, erro quadrático médio e tempo computacional. Os autores mostraram que a estratégia simples e robusta fornece um bom resultado nos casos *benchmarks* lineares e não lineares utilizados. Os critérios AVTI e OP estão detalhados adiante.

YAO e WANG (2015) propuseram o método *Kernel* Aditivo Generalizado e Análise de Componentes Principais (GAKPCA) para o monitoramento *on-line* de processos em batelada. O método proposto é um método não-linear especial de PCA, que pode lidar com as relações não lineares entre diferentes variáveis e em intervalos de tempo. Quando o *Kernel* Gaussiano for usado como a função de *Kernel* em cada ponto amostral no tempo, o método proposto foi conectado com o conceito de Correntropia que pode trazer robustez para a aplicação. Os

autores mostraram a validade dos resultados com o método de monitoramento on-line baseado no GAKPCA, com dados experimentais de um processo de fermentação de penicilina.

ZHANG e CHEN (2015) utilizaram dois algoritmos de processamento de dados, incluindo o estimador de Correntropia em estratégia de horizonte móvel para solucionar problemas de RDD Não linear, chamado de *Correntropy based Nonlinear Dynamic Data Reconciliation* (CNDDR), bem como Detecção e Identificação de Erros Grosseiros (GEDI). Além de CNDDR, a GEDI com um teste de hipóteses e um critério sequencial de distância-tempo identificou os tipos de erros grosseiros em sistemas dinâmicos, como os erros tipo *bias* e *drifts*. Os autores mostraram que CNDDR diminuiu consideravelmente a influência do erros grosseiros nos resultados reconciliados e que a GEDI classificou com sucesso os tipos de erros grosseiros dos dados medidos, usando como estudo de caso a polimerização por radicais livres de estireno.

3.4 Critério de avaliação e desempenho

Seguindo a formulação do problema de RD, deve-se minimizar uma função objetivo e satisfazer as restrições, que computacionalmente corresponde a minimizar uma função dos respectivos resíduos. Para uma avaliação honesta dos resultados dos problemas de RD, na presença de valores espúrios ou não, em que diferentes estimadores-M são utilizados, são apresentados nesta seção alguns critérios utilizados na literatura, tais como:

- Critério Fobj (Função Objetivo);
- Critério SSE (Soma dos Erros Quadráticos);
- Critério TER (Redução de Erros Totais);
- Critério SARES (Soma Absoluta dos RESÍDUOS).

3.4.1 Critério Fobj

Quando uma mesma função objetivo é utilizada no procedimento de RD, independentemente das estratégias utilizadas, quanto menor é o seu valor, provavelmente melhor terá sido o resultado, desde que as restrições estejam satisfeitas. Essa hipótese advém do fato de que o problema de RD é definido como

um problema de minimização de função objetivo, resultado alcançado com maior eficiência quando a função objetivo é menor.

3.4.2 Critério SSE

O critério SSE foi proposto por WANG (2000) apud (LINGKE *et al.*, 2006) e é caracterizado pelo somatório do quadrado da diferença entre os valores regredidos (reconciliados/estimados) e exatos (base da simulação dos dados medidos, isentos de erros). Este critério avalia o quanto os valores reconciliados \hat{z}_i diferem dos valores exatos z_i^* e como as variáveis não medidas (observáveis) estimadas \hat{u}_j diferem dos valores exatos u_j^* , sendo m o número de variáveis medidas e l o número de variáveis não medidas observáveis, conforme apresenta a Equação (3.35).

$$SSE = \sum_{i=1}^m (\hat{z}_i - z_i^*)^2 + \sum_{j=1}^l (\hat{u}_j - u_j^*)^2 \quad (3.35)$$

Quanto menor é a SSE obtida, provavelmente melhor terá sido o resultado, indicando menor diferença entre os valores reconciliados /estimados e os exatos. Uma crítica é que este critério não leva em consideração o desvio-padrão para as variáveis medidas (resíduo padronizado).

O critério SSE pode ser utilizado independentemente da Fobj utilizada (estimador). Desta forma, este critério pode ser utilizado para comparação de estimadores robustos e outras estratégias nos procedimentos de RD e DMEG. Este critério deve ser utilizado, preferencialmente, quando existirem variáveis não medidas observáveis. A Equação (3.35) pode ser ampliada para incluir parâmetros, no caso de problemas de RDEP. Cabe ressaltar que a utilização é limitada aos problemas de simulação, para os quais a solução exata é conhecida, não representando o problema real de estimação. Além disso, a Equação (3.35) pode ser questionada, dado que ela impõe uma métrica quadrática não ponderada ao problema de estimação.

3.4.3 Critério TER

O critério TER foi proposto por SERTH e HEENAN (1986) e é bastante utilizado para a avaliação de desempenho do procedimento de RD e DMEG, quando todos os valores exatos (base da simulação) são fornecidos (ÖZYURT e PIKE, 2004), conforme apresenta a Equação (3.36).

$$TER = \frac{\sqrt{\sum_{i=1}^m \frac{(z_i - z_i^*)^2}{\sigma_i^2}} - \sqrt{\sum_{i=1}^m \frac{(\hat{z}_i - z_i^*)^2}{\sigma_i^2}}}{\sqrt{\sum_{i=1}^m \frac{(z_i - z_i^*)^2}{\sigma_i^2}}} \quad (3.36)$$

O critério TER expressa de forma quantitativa a redução total dos erros (valores espúrios e aleatórios), avaliando o quanto os erros foram reduzidos entre os valores medidos z_i e os regredidos (reconciliados) em relação aos exatos. Quanto maior for a TER obtida, ou mais próximo de 1, provavelmente melhor terá sido o resultado. Uma crítica é que este critério não leva em consideração variáveis não medidas (observáveis) ou parâmetros. O critério TER pode ser utilizado independentemente da função objetivo (estimador) utilizada, permitindo assim a comparação de estimadores-M robustos e outras estratégias nos procedimentos de RD e DMEG. De novo, a técnica pressupõe o conhecimento dos valores reais e impõe uma norma de avaliação distinta da usada para estimação.

3.4.4 Critério SARES

O critério SARES consiste na soma do valor absoluto de cada resíduo das equações de restrição do modelo matemático, sendo k o número de equações de restrição. Quanto menor é o valor, provavelmente melhor terá sido o resultado, conforme apresenta a Equação (3.37) (SARRUF, 2014).

$$SARES = \sum_{i=1}^k |e|_i \quad (3.37)$$

O critério SARES pode ser utilizado independentemente do estimador. Desta forma, este critério pode ser utilizado para comparação de estimadores robustos e outras estratégias nos procedimentos de RD e DMEG. Uma crítica é que a maioria dos resultados reportados na literatura não indica as tolerâncias utilizadas durante o procedimento de otimização.

3.5 Critério para detecção de valores espúrios

O tradicional conceito de nível de confiança em teste de hipóteses também se aplica para rejeição (detecção) de valores espúrios (erros grosseiros), o que torna uma vantagem conhecer a distribuição dos erros de medida. Um valor medido, cuja probabilidade de ocorrência é menor do que o nível de confiança adotado, pode ser

considerado como valor espúrio. Segundo ÖZYURT e PIKE (2004), o desvio da variável além do fator de corte pode ser considerado como valor espúrio (*cut point*), mesmo que seja usado outro critério alternativo que não considere a distribuição *a priori*.

Para a distribuição Normal, com média zero e variância um $[N(0, 1)]$, têm-se:

- 95% de confiança (clássico): Fator de corte $\xi = 1,96$;
- 97% de confiança: Fator de corte $\xi = 2,16$ (ÖZYURT e PIKE, 2004).

Cabe ressaltar que a rejeição de valores espúrios por meio da utilização da abordagem de teste de hipótese pode gerar resultados errôneos, se a hipótese considerada para a distribuição do erro de medida admitida for violada. Conforme descrito por PRATA (2009), quando se utiliza os estimadores-M robustos não é necessário detectar o valor espúrio durante o procedimento de regressão ou procedimento iterativo, como feito quando o estimador não robusto MQP é utilizado. Teoricamente, a influência negativa do valor espúrio é minimizada ou eliminada pela estrutura matemática pouco sensível do estimador, com base na função de influência. Entretanto, estes valores espúrios podem ser assinalados após a regressão robusta para averiguação de falsas detecções ou detecção de valores espúrios consecutivos, o que pode caracterizar um desvio permanente (*bias*).

3.6 Critério de avaliação da detecção de valores espúrios

Para uma avaliação honesta dos resultados dos problemas de regressão robusta, na presença de valores espúrios e com diferentes estimadores-M, esta seção introduzirá alguns critérios utilizados na literatura, tais como:

- Critério OP (Potência Global);
- Critério AVTI (Erro Médio do Tipo I);

3.6.1 Critério OP

O critério OP foi proposto por NARASIMHAN e MAH (1987), conforme mostra a Equação (3.38):

$$OP = \frac{\text{Número de erros grosseiros corretamente detectados}}{\text{Número de erros grosseiros simulados}} \quad (3.38)$$

A Equação (3.38) expressa a razão entre o número de valores espúrios corretamente identificados e o número de valores espúrios totais (simulados); ou seja, avalia os Erros Tipo II indiretamente. Quanto maior for o valor de OP obtido, provavelmente melhor terá sido o resultado, indicando uma melhor detecção dos valores espúrios (erros grosseiros).

3.6.2 Critério AVTI

O critério AVTI foi proposto por NARASIMHAN e MAH (1987), conforme mostra a Equação (3.39):

$$AVTI = \frac{\text{Número de erros grosseiros erroneamente detectados}}{\text{Número de ensaios de simulação}} \quad (3.39)$$

A Equação (3.39) expressa a razão entre o número de valores espúrios erroneamente identificados (erro tipo I, ou seja, um valor espúrio é detectado, quando na verdade não é) pelo número de diferentes simulações realizadas em um mesmo caso, avaliado para cada simulação separadamente, mesmo que não exista nos dados um valor espúrio simulado. Quanto menor for a AVTI obtida, provavelmente melhor terá sido o resultado, indicando uma melhor detecção dos valores espúrios (erros grosseiros).

3.7 Comentários Finais

Com base no que foi discutido, observa-se que os estimadores-M robustos e o estimador Correntropia (robusto, mas não oriundo do princípio da Máxima Verossimilhança e da estatística robusta), vem apresentando bons resultados até o presente trabalho. Observa-se, também, que os estimadores oriundos da estatística robusta estão se consolidando cada vez mais nos estudos e aplicações de RD, que os tornam indispensáveis de qualquer estudo de RD e DMEG. Entretanto, é oportuno uma análise dos problemas de RD e DMEG em sistemas de separação de forma robusta, aplicando os estimadores-M robustos. Portanto, uma abrangente revisão sobre os estimadores-M robustos foi feita e apresentada no Apêndice B, onde são detalhados 50 estimadores-M (48 robustos, MQP e Normal Contaminada), suas

formas funcionais e gráficas das famílias de funções (ρ , ψ e ψ'), os parâmetros de sintonia para as eficiências relativas de 90, 95, 98 e 99% e uma tabela guia para o leitor, detalhando as principais características e propriedades desses estimadores.

Capítulo 4

Colunas de Destilação

“Pensar é o trabalho mais difícil que existe. Talvez por isso tão poucos se dediquem a ele.”

Henry Ford (1863 – 1947)

Este capítulo tem como objetivo contextualizar uma grande e antiga área da engenharia química: a destilação. Para isso, apresenta-se uma revisão bibliográfica dos fundamentos teóricos para modelos fenomenológicos de colunas de destilação, incluindo a termodinâmica de equilíbrio de fases, uma revisão dos principais modelos dinâmicos e, também, dos modelos de redução de ordem usados, como aqueles baseados no método de colocação ortogonal e de resíduos ponderados.

4.1 Contextualização

A destilação é uma antiga operação unitária que vem sendo praticada por milhares de anos. As primeiras aplicações eram bem rudimentares, com aparatos usados para concentrar o teor alcoólico das bebidas (KISTER, 1992). Evidências do processo de destilação são muito antigas. Aristóteles mencionou que água pura podia ser obtida a partir da água do mar por evaporação, mas ele não explicou como condensar o vapor. Segundo YOUNG (1903), um método primitivo de condensação foi descrito por Dioscorides e Pliny, que mostraram que um óleo pode ser obtido aquecendo um recipiente cônico comprido, com líquido que contém o óleo. Na parte de cima do recipiente foi colocado um pouco de lã; assim, o óleo se condensa na lã, sendo retirado quando se espreme a lã. Os químicos Alexandrinos adicionavam

um segundo recipiente, chamados pelos Árabes de *alembic*, para destilar bebidas alcoólicas (YOUNG, 1903).

Em 1813, na França, Cellier-Blumenthal desenvolveu a primeira coluna vertical de destilação simples. Os primeiros recheios foram usados em 1820 por um tecnólogo chamado Clement, que utilizou bolas de vidro para conduzir uma destilação simples de álcool. Perrier introduziu os primeiros pratos com borbulhadores na Inglaterra em 1822. Coffey desenvolveu a primeira destilação com pratos perfurados em 1830. O primeiro livro de fundamentos de destilação foi o *La Rectification de l'alcohol* por Ernest Sorel, em 1893. Nos primeiros 25 anos do século XX, a aplicação da destilação passou de uma ferramenta de concentração do teor de álcool em bebidas para uma sofisticada técnica de separação na indústria química e petroquímica. Uma vez que a destilação foi reconhecida, como técnica eficaz para separar óleo bruto em vários componentes, uma expansão acelerada de sua utilização ocorreu. A partir daí, essa aplicação se transformou no principal processo da indústria química e petroquímica (KISTER, 1992).

Antes da década de 1950, cálculos dos modelos de colunas eram realizados manualmente. Por isso, cálculos com modelos rigorosos eram muito difíceis de serem avaliados e procedimentos como os métodos de McCabe-Thiele, Ponchon-Savarit e *shortcut*, eram normalmente utilizados (SEADER *et al.*, 2010). Com a invenção dos computadores, os cálculos rigorosos começaram a se desenvolver. Inicialmente, gastavam-se muitos dias resolvendo-se o problema, mas a evolução dos processadores permitiu que os mesmos cálculos atualmente sejam feitos em alguns segundos com muita eficiência, eliminando as chances de erros manuais. Atualmente, simuladores de processos de destilação utilizam métodos e modelos rigorosos como ferramenta para projetos, otimização e controle supervisionado (DE MENEZES, 2015). Os antigos métodos mais simplificados ainda são utilizados para fins didáticos e estimativas iniciais de projetos.

Muitos programas comerciais e acadêmicos, como Aspen Plus (PLUS, 2003), Hysys (MANUAL, 1999), Unisim Design (HONEYWELL, 2018), Pro II (ELECTRIC'S, 2018), iiSE (VRTECH, 2018), EMSO (SOARES e SECCHI, 2003) e até sub-rotinas públicas com modelos matemáticos e termodinâmicos já foram desenvolvidas para simular diversos processos químicos. Para o estudo de processos de separação, em particular a destilação, fica evidente que os parâmetros estimados dos modelos termodinâmicos são tão importantes quanto a precisão do modelo de coluna utilizado. É fundamental contar com dados de boa qualidade para uma adequada representação das condições do equilíbrio, o que pode resultar em substancial economia de energia e gastos com equipamentos.

Como já discutido no Capítulo 1, a destilação é um dos equipamentos de separação mais utilizados na indústria química, apesar da sua alta demanda por energia devido a baixa eficiência energética (GARY, 2016). Com isso, inúmeros estudos foram feitos para reduzir o custo do equipamento, o consumo energético, o custo operacional e o custo de manutenção do processo. Estudos vem mostrando que existe um potencial significativo para a redução do consumo de energia do processo (HALVORSEN e SKOGESTAD, 2011). Um desses estudos é o Método de Grande Composição da Coluna (CGCC) que analisa os gastos energéticos através dos parâmetros e variáveis da coluna, com o objetivo de melhorar a eficiência ajustando as variáveis e parâmetros (AKSHAYA, 2016). Estudos com diferentes arranjos de projeto de destilação também foram propostos, como Coluna com Parede Divisória (DWC), Colunas com Integração Energética Internamente (HIDiC) e a aplicação de uma tecnologia que utiliza bombas de calor mecânica (ReCompressão de Vapor, VRC) (HALVORSEN e SKOGESTAD, 2011; MODLA e LÁNG, 2017). Uma tecnologia importante para a eficiência energética do processo é o controle e monitoramento em tempo real do equipamento. Assim, um bom controle pode ser feito em conjunto com uma camada de RTO do processo; com isso, a técnica de RTO pode ser tão relevante quanto os arranjos de projeto e análises energéticas. Por estes fatos, melhorias no processo de controle e monitoramento de destilação podem exercer grandes impactos na redução de gasto energético e econômico de uma indústria.

4.2 Fundamentos

O objetivo geral da destilação é separar componentes que tenham pressões de vapor diferentes em uma dada temperatura. A pressão de vapor de uma substância aumenta com a temperatura, pois nas temperaturas mais elevadas o número de moléculas que têm energia suficiente para escapar da interação com as moléculas vizinhas é maior (ATKINS e DE PAULA, 2006).

Duas fases (líquido e vapor) estão em equilíbrio físico se as seguintes condições são satisfeitas:

- i* A temperatura da fase vapor T^V é igual à temperatura da fase líquida T^L (Equilíbrio térmico);
- ii* A pressão da fase vapor P^V é igual à pressão da fase líquida P^L (Equilíbrio mecânico);

iii A tendência de cada componente para escapar da fase líquida para fase vapor é igual à tendência dos mesmos escaparem da fase vapor para a fase líquida (Equilíbrio termodinâmico químico).

Assim, o estado de equilíbrio se resume em $T^V = T^L = T^{eq}$, $P^V = P^L = P^{eq}$ e $\mu_i^L = \mu_i^V = \mu_i^{eq}(T^{eq}, P^{eq})$. Em casos especiais, a terceira condição pode ser representada pela lei de Raoult, Equação (4.1):

$$P^{eq}y_i = P_i^{sat}(T^{eq})x_i \quad (4.1)$$

em que x_i e y_i são as frações molares do componente i na fase líquida e na fase vapor, respectivamente; $P_i^{sat}(T^{eq})$ é a pressão de vapor (ou pressão de saturação) do componente i na temperatura de equilíbrio T^{eq} do sistema; e P^{eq} é a pressão de equilíbrio do sistema. μ_i^L e μ_i^V são os potenciais químicos do componente i na fase líquida e vapor, respectivamente (HOLLAND, 1981). O termo $.^{eq}$ será ocultado para não sobrecarregar as notações, entendendo-se que T e P estão em equilíbrio.

Para sair do equilíbrio termodinâmico químico, é necessário sair do equilíbrio térmico ou mecânico, adicionando-se energia. A energia adicionada ao sistema causa o aumento do movimento caótico entre as moléculas; ou seja, energia cinética das moléculas. Com isso, as moléculas da fase líquida adquirem energia térmica suficiente para superar a barreira da interação molecular, a barreira mecânica imposta pela pressão da fase vapor e, finalmente, escapar para a fase vapor. Quando a fase líquida chega a uma temperatura em que a pressão de vapor da fase líquida é igual à pressão externa (fase vapor), atinge-se a **temperatura de ebulição** T^{eb} nessa pressão (ATKINS e DE PAULA, 2006).

A palavra destilação se refere à separação física de uma mistura com dois ou mais componentes que tenham diferentes pontos de ebulição. Assim, se uma mistura líquida de dois ou mais componentes é aquecida, uma fase vapor (mistura gasosa) é criada, sendo mais rica no componente com o ponto de ebulição menor dentre aqueles presente na fase líquida. Analogamente, se o vapor (mistura gasosa com dois ou mais componentes) aquecido for resfriado, o líquido condensado tenderá a ser mais rico no componente com o ponto de ebulição maior do que os dos demais componentes (LEI *et al.*, 2005).

Uma coluna de destilação fracionada simples é composta, basicamente, por (PETLYUK, 2004):

- Externos à coluna:

- Uma corrente de entrada (*feed* – F), em que F representa a vazão global de alimentação e \underline{z}_F é o vetor de concentrações global (vapor + líquido) normalizadas da alimentação;
 - A vazão global de alimentação F pode ser parcialmente vaporizada: uma parte na fase vapor (V_F , - vazão de vapor da alimentação, \underline{y}_F - vetor de concentrações normalizadas da fase vapor da alimentação) e outra parte em fase líquida (L_F - vazão de líquido da alimentação, \underline{x}_F - vetor de concentrações normalizadas da fase vapor da alimentação).
 - Um condensador, que recebe a corrente de vapor V_1 que sai no topo da coluna e é contado como o estágio de equilíbrio $N = 1$;
 - Um refeedor (*reboiler*), que recebe a corrente de líquido L_n que sai no fundo da coluna e é contado como o estágio $N = 1 + nt$;
 - Retiradas do condensador: Destilado (*distillate* – D), em que D representa a vazão global do destilado (líquido e/ou vapor); refluxo de topo L_1 , em que parte do líquido do condensador retorna para o topo da coluna (primeiro prato).
 - Para a operação com condensador total, o destilado só contém a fase líquida, enquanto para a operação em condensador parcial, o destilado contém duas fases: Fase vapor (V_D , - vazão de vapor do destilado, \underline{y}_D - vetor de concentrações normalizadas da fase vapor do destilado) e Fase líquida (L_D - vazão de líquido do destilado, \underline{x}_D - vetor de concentrações normalizadas da fase vapor do destilado);
 - Retiradas do refeedor: Fundo (*bottom* – B), em que B representa a vazão global do produto de fundo e \underline{x}_B é o vetor de concentrações normalizadas do fundo; e o refluxo de fundo V_{nt+1} , em que o vapor do refeedor retorna para o fundo da coluna (último prato).
- Interno à coluna:
 - nt pratos, dispositivos que permitem o contato entre as fases vapor e líquida, sendo que a coluna apresenta $1 + nt$ estágios de equilíbrio (pratos + condensador), e dependendo do tipo de refeedor ele pode ser considerado um estágio também ($2 + nt$ estágios de equilíbrio);
 - O estágio (prato) que recebe a carga de alimentação é chamado de estágio f e divide a coluna em duas seções: retificação e esgotamento;
 - A seção superior é chamada de seção de retificação (*rectifying section*), onde se observa maior concentração dos componentes mais voláteis. O líquido (mais frio que os estágios da seção) que retorna do condensador

para a coluna (refluxo de topo), retira energia Q_{cond} e condensa os componentes menos voláteis desta seção, realizando a separação por diferença de volatilidade, aumentando-se a concentração dos componentes mais voláteis em direção ao topo;

- o A seção inferior é chamada de seção de esgotamento (*stripping section*) onde se observa maior concentração dos componentes menos voláteis. O vapor (mais quente que os estágios da seção) que retorna do refeedor para a coluna (refluxo de fundo), injeta energia Q_{reb} e vaporiza os componentes mais voláteis desta seção, realizando a separação por diferença de volatilidade, aumentando-se a concentração dos componentes menos voláteis em direção ao fundo.

A Figura 4.1 mostra o arranjo típico de uma coluna de destilação. Os estágios (pratos, condensador e refeedor) são as etapas que promovem a transferência de calor e de massa entre as fases. Para isso, as colunas de destilação fracionadas são equipadas internamente com dispositivos que promovem o contato entre o líquido que desce do topo ao fundo e o vapor que sobe do fundo ao topo. A fase líquida é canalizada entre um prato e outro por tubos de escoamento descendente, ou chicanas posicionadas junto à parede da coluna, chamados de *downcomer*. A fase vapor ascende de um prato para outro por orifícios encontrados no meio dos pratos, chamados de área ativa. Esses orifícios podem ser válvulas, válvulas com borbulhadores ou apenas furos.

Os pratos são classificados de acordo com o tipo de orifício, podendo ser prato valvulado (*valve tray*), prato com borbulhador (*bubble-cap tray*) ou prato perfurado (*sieve tray*). Além dos pratos, outros dispositivos que promovem o contato entre as fases são os recheios, que caracterizam as colunas empacotadas (*packing*) (CALDAS *et al.*, 2007). A Figura 4.2 ilustra o funcionamento de um dispositivo de contato tipo prato perfurado.

Sabe-se que o equilíbrio termodinâmico de fato não é alcançado, pois nos pratos o tempo de residência para troca térmica e de massa e a área superficial de contato entre o líquido e o vapor não são suficientemente grandes para atingir o equilíbrio teórico. Entretanto, pode-se saber quão próximo um estágio de “equilíbrio” está do equilíbrio teórico, sendo possível estimar a eficiência real do prato ou da coluna.

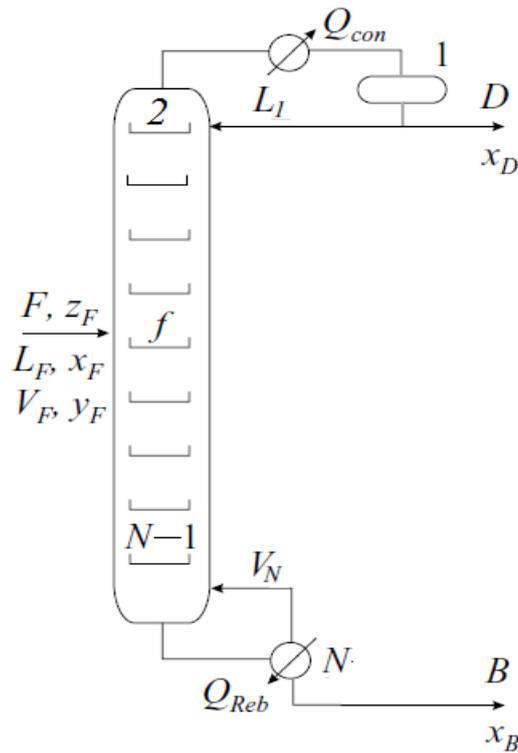


Figura 4.1: Esquema típico de uma coluna de destilação fracionada (PETLYUK, 2004).

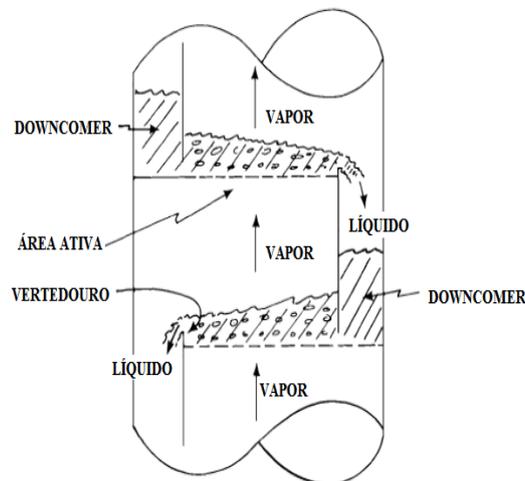


Figura 4.2: Esquema típico de um prato (BUCKLEY *et al.*, 1985).

4.2.1 Fundamentos termodinâmicos

O equilíbrio de fases pode ser classificado como o equilíbrio estabelecido entre duas ou mais fases diferentes ou heterogêneas, como no caso do Equilíbrio Líquido-Vapor (ELV). O ELV é o mais importante para os processos de destilação, particularmente em respeito aos cálculos de energia requerida, para simulação do processo e os projetos de equipamentos de destilação.

A relação de equilíbrio

Conforme descrito por LEI *et al.* (2005), dentro de qualquer sistema fechado em que o equilíbrio de fases existe, a energia livre de Gibbs para todas as fases é mínima, não havendo transferência de calor e massa dentro do sistema. Assim, a diferenciação total da energia livre de Gibbs G das $j = 1, 2, \dots, p$ fases e $i = 1, 2, \dots, c$ componentes do sistema em relação a temperatura T , pressão P e número de moles n é apresentada pela Equação (4.2), em que S^j , V^j e μ_i^j representam a entropia, o volume e o potencial químico do componente i da fase j , respectivamente.

$$dG = \sum_{j=1}^p \left(-S^{(j)}dT + V^{(j)}dP + \sum_{i=1}^c \mu_i^{(j)}dn_i^{(j)} \right) \quad (4.2)$$

A Equação (4.2) também é conhecida como a **equação fundamental da termodinâmica química**. No equilíbrio, com a temperatura e pressão constantes, a Equação (4.2) pode ser reduzida para:

$$dG = \sum_{j=1}^p \left(\sum_{i=1}^c \mu_i^{(j)}dn_i^{(j)} \right) = 0 \quad (4.3)$$

Para cada componente i , o número global de moles é constante e igual ao número de moles inicial antes do equilíbrio (sistema fechado). No equilíbrio, a variação de moles do componente i que sai da fase j é igual à variação de moles de i que entra na fase j ; por isso, a variação de moles em cada fase é nula no equilíbrio:

$$\sum_{j=1}^p \mu_i^{(j)}dn_i^{(j)} = 0 \quad (4.4)$$

Combinando as Equações (4.3) e (4.4), é possível escrever (HOLLAND, 1981):

$$\mu_i^{(1)} = \mu_i^{(2)} = \mu_i^{(3)} = \dots = \mu_i^{(p)} \quad (4.5)$$

$$T^{(1)} = T^{(2)} = T^{(3)} = \dots = T^{(p)} \quad (4.6)$$

$$P^{(1)} = P^{(2)} = P^{(3)} = \dots = P^{(p)} \quad (4.7)$$

Desvios em relação ao comportamento ideal são analisados pela Termodinâmica de forma a preservar as expressões que foram deduzidas para sistemas ideais. Assim, a dependência da energia livre de Gibbs molar com a

pressão e temperatura, em um sistema de fase aberta, é dada como (LEI *et al.*, 2005):

$$d\mu_i^{(j)} = RTd\ln(\hat{f}_i^{(j)}) \quad (4.8)$$

em que R é a constante universal dos gases, $\hat{f}_i^{(j)}$ é a fugacidade parcial do componente i na fase líquida j . Para um componente i e com referência a um estado padrão de baixa pressão, a notação da fugacidade é dada por $f_i^{*(j)}$.

A fugacidade é a “pressão efetiva” que um componente faz de uma fase para a outra. A fugacidade é função da pressão e da temperatura e tem mesma unidade que a pressão. Quando o sistema sai da idealidade, a fugacidade deve ser usada para corrigir a pressão. O nome “fugacidade” vem do latim e indica “tendência de escapar”, sendo algumas vezes chamada de “pressão química”, porque está relacionada às interações intermoleculares. Essa grandeza se origina do trabalho de Gilbert Newton Lewis (1875-1946), físico-químico americano, que também desenvolveu os conceitos de propriedade parcial e de solução ideal (SMITH *et al.*, 2005). Em pressões moderadas, as forças atrativas são maiores, de maneira que a fugacidade pode ser menor que a pressão e a tendência de escape é menor. Em altas pressões, as forças repulsivas são maiores, de forma que a fugacidade é maior que a pressão e as moléculas têm maior tendência de escape (ATKINS e DE PAULA, 2006).

Para a ELV, substituindo a Equação (4.8) na Equação (4.5), obtém-se:

$$\hat{f}_i^{(L)} = \hat{f}_i^{(V)} \quad (4.9)$$

Muitas equações de ELV derivam a partir da Equação (4.9) como, por exemplo, a lei de Raoult da Equação (4.1), que é aplicada para sistemas ideais (gás ideal e solução líquida ideal) e pressões baixas a moderadas. A forma mais simples de calcular a fugacidade de uma mistura gasosa é pela Regra de Lewis-Randall, que é válida quando a compressibilidade Z da mistura varia linearmente com a fração molar dos componentes da mistura. Neste caso a fugacidade do componente i na mistura é proporcional à fração do componente i na mistura e a fugacidade da mistura é calculada com base na fugacidade do componente i puro no seu estado padrão (FERNANDES *et al.*, 2006).

De forma mais abrangente a Equação (4.9) pode ser escrita em função da fugacidade parcial, pressão, coeficiente de fugacidade e coeficiente de atividade, como mostra a equação a seguir:

$$\hat{f}_i^{(L)} = Px_i\hat{\phi}_i^{(L)} = \gamma_i^{(L)}x_i f_i^{*(L)} \quad (4.10)$$

$$\hat{f}_i^{(V)} = Py_i\hat{\phi}_i^{(V)} = \gamma_i^{(V)}y_i f_i^{*(V)} \quad (4.11)$$

em que $\hat{\phi}_i$ e γ_i são, respectivamente, o coeficiente de fugacidade da espécie i e o coeficiente de atividade da espécie i em suas respectivas fases. O sobrescrito $*$ significa o estado padrão para o componente i puro. Igualando as fugacidades da fase vapor e líquida, a equação fica:

$$Py_i\hat{\phi}_i^{(V)} = \gamma_i^{(L)}x_i f_i^{*(L)} \quad (4.12)$$

O conceito de constante de equilíbrio k_i pode ser introduzido e definido como a razão entre as frações molares do componente i da fase vapor e fase líquida (LEI *et al.*, 2005):

$$k_i = \frac{y_i}{x_i} = \frac{\hat{\phi}_i^{(L)}}{\hat{\phi}_i^{(V)}} = \frac{\gamma_i^{(L)} f_i^{*(L)}}{P\hat{\phi}_i^{(V)}} \quad (4.13)$$

Coeficiente de fugacidade da fase vapor para multicomponentes

À medida que se torna necessário modelar sistemas maiores e não ideais, incluem-se novas funções, correções de não idealidade e informações adicionais de nível molecular. Para modelar as propriedades físico-químicas das misturas não ideais, a Equação (4.13) pode ser complexa devido aos termos $\hat{\phi}_i^{(V)}(T, P, \underline{y})$, $\gamma_i^{(L)}(T, P, \underline{x})$ e $f_i^{*(L)}(T, P)$.

O coeficiente de fugacidade da espécie i em solução gasosa $\hat{\phi}_i^{(V)}$ é definido com auxílio da *energia livre de Gibbs parcial residual* \overline{G}_i^R , que representa uma correção da energia livre de Gibbs para gases não ideais.

Devido à maior aproximação das moléculas em um líquido, a interação entre as diferentes moléculas de uma mistura é maior, o que faz com que o comportamento de muitas misturas líquidas não seja descrito corretamente pelas equações de estado; portanto, as regras de mistura não podem ser utilizadas com sucesso. Para calcular corretamente a interação entre as moléculas de uma mistura líquida, pode-se calcular a fugacidade de um composto numa mistura líquida com base no coeficiente de atividade (FERNANDES *et al.*, 2006).

Na abordagem $(\gamma-\phi)$, o coeficiente de fugacidade é utilizado para a fase

vapor, enquanto para a fase líquida utiliza-se o *coeficiente de atividade da espécie i na solução líquida* $\gamma_i^{(L)}$, que é definido por meio da *energia livre de Gibbs parcial em excesso* \overline{G}_i^E , pois as propriedades em excesso inferem o afastamento do comportamento de soluções ideais, análogo às propriedades residuais.

O termo $f_i^{*(L)}$, *fugacidade da espécie i pura na fase líquida*, só depende da temperatura e pressão do sistema, podendo ser descrito com auxílio do fator ou correção de Poynting, Equação (4.14) (FERNANDES *et al.*, 2006; LEI *et al.*, 2005; SMITH *et al.*, 2005).

$$f_i^{*(L)} = f_i^{sat(L)} \left[\exp \left(\int_{P_i^{sat}}^P \frac{V_i^L}{RT} dP \right) \right] = f_i^{sat(L)} (PF)_i \quad (4.14)$$

em que o termo $(PF)_i$ é o fator de Poynting para a espécie i , $f_i^{sat(L)}$ é a fugacidade da espécie i pura (saturada) na fase líquida. Assim, substituindo a Equação (4.14) na Equação (4.13) obtém-se:

$$k_i = \frac{y_i}{x_i} = \frac{\gamma_i^{(L)} f_i^{sat(L)} (PF)_i}{P \hat{\phi}_i^{(V)}} = \frac{\gamma_i^{(L)} P_i^{sat} \phi_i^{sat} (PF)_i}{P \hat{\phi}_i^{(V)}} \quad (4.15)$$

em que P_i^{sat} é a pressão de saturação da espécie i e ϕ_i^{sat} é o coeficiente de fugacidade da espécie i saturada. A pressão de vapor (ou pressão de saturação) pode ser calculada, por exemplo, utilizando as equações de Antoine, Riedel ou Harlecher-Braum, Equações (4.16), (4.17) e (4.18), respectivamente (FERNANDES *et al.*, 2006).

$$\ln P_i^{sat} = A - \frac{B}{T + C} \quad (4.16)$$

$$\ln P_i^{sat} = A + \frac{B}{T} + C \ln(T) + DT^6 \quad (4.17)$$

$$\ln P_i^{sat} = A + \frac{B}{T} + C \ln(T) + \frac{DP_i^{sat}}{T^2} \quad (4.18)$$

em que A , B , C e D são parâmetros estimados com base em cada equação e em dados experimentais de pressão de vapor disponíveis para cada componente. Portanto, os parâmetros geralmente não são iguais para cada equação.

Para sistemas ideais (gás ideal e solução ideal) e baixas pressões, a Equação (4.15) se resume à Equação (4.1), a Lei de Raoult. Porém, para sistemas com

gás ideal para fase vapor, fase líquida com comportamento não ideal e pressões moderadas, a Equação (4.15) se resume à Lei de Raoult modificada, como mostra a Equação (4.19).

$$k_i = \frac{y_i}{x_i} = \frac{\gamma_i^{(L)} P_i^{sat}}{P} \quad (4.19)$$

Os coeficientes de atividade $\gamma_i^{(L)}$ podem ser interpretados de acordo com a natureza do desvio em relação à Lei de Raoult. O comportamento não ideal das soluções está também relacionado com o calor de mistura dos componentes. Por exemplo, se c componentes isolados precisam ganhar calor para se misturar e formar uma solução, entende-se que, os componentes “preferem estar isolados” do que estar em solução. Dessa forma, o desvio em relação à Lei de Raoult seria **positivo**, o que implica dizer que o coeficiente de atividade seria $\gamma_i^{(L)} > 1$. Conseqüentemente, a constante de equilíbrio aumenta, $k_i \uparrow$, fazendo com que os componentes tendem a escaparem com mais facilidades para fase vapor. Por outro lado, se c componentes isolados precisam perder calor para se misturarem e formarem uma solução, o inverso se aplica. Assim, neste caso os componentes “preferem estar em mistura” do que isolados. Portanto, o desvio em relação à Lei de Raoult seria **negativo**, o que implica dizer que o coeficiente de atividade seria $\gamma_i^{(L)} < 1$. Conseqüentemente, diminuindo a constante de equilíbrio, $k_i \downarrow$, fazendo com que os componentes tendem a permanecerem com mais facilidades na fase líquida (KISTER, 1992).

Para sistemas não ideais, especialmente para gases polares, os termos de coeficientes de fugacidades $\hat{\phi}_i^{(V)}$ não podem ser negligenciados. Portanto, outras equações mais representativas para os gases e, conseqüentemente, mais complexas devem ser utilizadas. Entretanto, para deduzir o coeficiente de fugacidade da espécie i parcial $\hat{\phi}_i^{(V)}$, Equação (4.20), e o coeficiente de fugacidade da espécie i pura ϕ_i , Equação (4.21), para o ϕ_i^{sat} por exemplo, é necessário utilizar uma equação de estado *PVT* (SEADER *et al.*, 2010; SMITH *et al.*, 2005).

$$\hat{\phi}_i^{(V)} = \exp \left\{ \frac{1}{RT} \int_V^\infty \left[\left(\frac{\partial P}{\partial n_i} \right)_{T,V,n_{i \neq j}} - \frac{RT}{V} \right] dV - \ln Z \right\} \quad (4.20)$$

em que $V = \bar{V} \sum_{i=1}^c n_i$

$$\phi_i = \exp \left[\frac{1}{RT} \int_{\bar{V}}^\infty \left(P - \frac{RT}{\bar{V}} \right) d\bar{V} - \ln Z + (Z - 1) \right] \quad (4.21)$$

Há muitas Equações de Estados (EOS) que descrevem o comportamento PVT de muitas substâncias e de fácil acesso na literatura, como a equação do virial, a equação de van der Waals, de RK (Redlich-Kwong), de SRK (Soave-Redlich-Kwong), de PR (Peng-Robinson), entre outros (LEI *et al.*, 2005). Por exemplo, para um componente puro, a EOS de SRK é dada por (PRAUSNITZ *et al.*, 1998):

$$P = \frac{RT}{(V - b)} - \frac{a}{V(V + b)} \quad (4.22)$$

em que

$$b = 0,08664 \frac{RT_c}{P_c} \quad (4.23)$$

$$a = 0,42748 \frac{(RT_c)^2}{P_c} \left(1 + f_w(1 - \sqrt{T_r})\right)^2 \quad (4.24)$$

$$T_r = \frac{T}{T_c} \quad (4.25)$$

$$f_w = 0,480 + 1,574w - 0,176w^2 \quad (4.26)$$

T_c e P_c são a temperatura e pressão crítica, respectivamente. w é o fator acêntrico da substância. A EOS de SRK faz parte de uma família de equações de estados cúbicas, e a equação pode ser colocada na forma do fator de compressibilidade $Z = \frac{PV}{RT}$. Para um sistema multicomponentes, as equações de estado precisam das regras de misturas e combinações para calcular os parâmetros do sistema. Colocando a equação cúbica em termos do fator de compressibilidade da mistura \hat{Z} (COKER, 2010):

$$\hat{Z}^3 - \hat{Z}^2 + \hat{Z}(A - B - B^2) - AB = 0 \quad (4.27)$$

em que

$$A = \sum_{i=1}^c \sum_{j=1}^c y_i y_j A_{ij} \text{ ou } \sum_{i=1}^c \sum_{j=1}^c x_i x_j A_{ij} \quad (4.28)$$

$$A_{ij} = (A_i A_j)^{0,5} (1 - k_{ij}) \quad (4.29)$$

$$B = \sum_{i=1}^c y_i B_i \text{ ou } \sum_{i=1}^c x_i B_i \quad (4.30)$$

$$A_i = 0,42747 a_i \frac{P_{ri}^2}{T_{ri}} \text{ e } B_i = 0,08664 \frac{P_{ri}}{T_{ri}} \quad (4.31)$$

Para cada componente, deve ser definida a temperatura $T_{ri} = \frac{T}{T_{ci}}$ e a pressão $P_{ri} = \frac{P}{P_{ci}}$ reduzida, sendo a_i calculado pela Equação (4.24) para cada componente. O parâmetro de iterações binárias k_{ij} , quando não é disponível, é considerado igual a zero (PRAUSNITZ *et al.*, 1998). Pela abordagem $(\phi-\phi)$, o coeficiente de fugacidade é utilizado para as duas fases. Para a EOS de SRK, o coeficiente de fugacidade em função do fator de compressibilidade é dado pela Equação (4.32), tanto para a fase líquida (menor raiz da equação cúbica para \hat{Z}) quanto para a fase vapor (maior raiz da equação cúbica para \hat{Z}).

$$\hat{\phi}_i = \exp \left[(\hat{Z} - 1) \frac{B_i}{B} - \ln(\hat{Z} - B) - \frac{A}{B} \left(\frac{2A_i^{0,5}}{A^{0,5}} - \frac{B_i}{B} \right) \ln \left(\frac{\hat{Z} + B}{\hat{Z}} \right) \right] \quad (4.32)$$

Coeficiente de atividade da fase líquida para multicomponentes

Para sistemas em que a solução não é ideal, o coeficiente de atividade da espécie i na fase líquida não pode ser negligenciado; portanto, $\gamma_i^{(L)}$ deve ser determinado. As equações de Margules e van Laar são exemplos para determinar o coeficiente de atividade. A prática usualmente empregada para o cálculo dos coeficientes de atividade da fase líquida faz uso de modelos derivados de expressões dadas para a energia livre de Gibbs em excesso \overline{G}_i^E , que se relaciona com a composição, pressão e temperatura através da Equação (4.33) (FERNANDES *et al.*, 2006; HOLLAND, 1981).

$$\ln \gamma_i^{(L)} = \frac{1}{RT} \left(\frac{\partial \overline{G}_i^E}{\partial n_i} \right)_{T,P,n_{j \neq i}} \quad (4.33)$$

Em baixas e moderadas pressões, é aceitável desprezar a dependência da energia de Gibbs em relação à pressão, já que o efeito é suficientemente pequeno.

Dessa forma, os modelos adotados para a representação do coeficiente de atividade da fase líquida também ficam como função da temperatura e composição. Estes modelos levam em consideração a energia de interação entre as moléculas, expressa na forma de parâmetros de interação binários. Alguns modelos consideram que a composição local dos componentes é diferente da composição global, considerando também o tamanho e formato da molécula na mistura, com base na Estatística Mecânica e em modelos semi-empíricos. Exemplos de modelos usados para coeficientes de atividade são encontrados na literatura, como Margules, van Laar, Wilson, UNIQUAC (UNIversal QUAsi-Chemical), UNIFAC (UNIQUAC Functional-group Activity Coefficients) e NRTL (Non-Random-Two-Liquid) (FERNANDES *et al.*, 2006; LEI *et al.*, 2005; SMITH *et al.*, 2005).

Entalpia de misturas

Uma das formas de estimar entalpias de mistura de um gás real é pela definição de **propriedade residual**, que mede o afastamento do comportamento de um gás em relação ao comportamento de um gás ideal. Esta propriedade relaciona dados experimentais, equação de estado PVT ou correlações PVT generalizadas. Outra maneira de estimar entalpias, usualmente mais para líquidos, é pelo afastamento do comportamento em relação à solução ideal, ganhando o nome de propriedades **em excesso**. A definição matemática é análoga a das propriedades residuais (POLING *et al.*, 2001; SEADER *et al.*, 2010; SMITH *et al.*, 2005).

Para uma mistura de gases reais, a definição de entalpia mais usual é pela propriedade residual, Equação (4.34), utilizando dados de equação de estado PVT e regras de mistura.

$$\left(\bar{H}^{(V)} - \bar{H}^{(ig)}\right) \equiv \bar{H}^{(V,R)} = P\bar{V} - RT - \int_{\infty}^{\bar{V}} \left[P - T \left(\frac{\partial P}{\partial T} \right)_{\bar{V}} \right] d\bar{V} \quad (4.34)$$

Para uma solução não-ideal, a definição de entalpia mais usual é pela entalpia parcial molar em excesso $\bar{h}_{iL}^{E_p}$, que utiliza modelos derivados de expressões de energia livre de Gibbs em excesso \bar{G}_i^E .

$$\left(\bar{h}_{iL}^p - \bar{h}_{iL}^{is}\right) \equiv \bar{h}_{iL}^{E_p} = RT \left(\frac{\partial \ln \gamma_{iL}}{\partial P} \right)_{T,x} \quad (4.35)$$

Na abordagem $(\phi-\phi)$, a entalpia da fase líquida também é calculada utilizando

a EOS *PVT* pela propriedade residual. Por exemplo, utilizando a EOS de SRK, Equação (4.27), na Equação (4.34) obtém-se a Equação (4.36).

$$\bar{H}^{(VouL,R)} = RT(\hat{Z} - 1) + \frac{P}{B(RT)} \ln \left(\frac{\hat{Z} + B}{\hat{Z}} \right) \left[T \frac{d}{dT} \left(A \frac{(RT)^2}{P} \right) - A \frac{(RT)^2}{P} \right] \quad (4.36)$$

Para um líquido que não esteja em estado saturado, a Equação (4.36) precisa de alguns ajustes, como pode ser encontrados em FERNANDES *et al.* (2006); PRAUSNITZ *et al.* (1998); SEADER *et al.* (2010); SMITH *et al.* (2005).

4.3 O modelo MESH

Na literatura há vasta bibliografia sobre estudos e modelos de destilação, desde a época em que os cálculos eram manuais até os dias atuais. Entretanto, os modelos mais utilizados atualmente começaram a ser estudados um pouco antes do desenvolvimentos dos computadores. Grande parte da literatura é voltada para soluções de modelos estacionários, pois modelos de coluna de destilação são não-lineares, com grandes dimensões e, dependendo da idealidade termodinâmica da composição, constituídos por sistemas não lineares com dificuldades de convergência numérica. Para contornar estes problemas, pesquisadores estudaram diversas estratégias algébricas para atenderem a problemas específicos de colunas, com métodos globais para atenderem aos mais diversos tipos de coluna de destilação. Portanto, uma breve revisão dos modelos estacionários é importante para compreender o desenvolvimento dos modelos dinâmicos, que derivam dos modelos estacionários, tendo como base as equações MESH:

- **Material:** equações de balanço de massa global e por componentes;
- **Equilibrium:** equações de equilíbrio líquido-vapor, incluindo equações de ponto de bolha e orvalho;
- **Summation:** restrições de soma das frações molares normalizadas;
- **Heat or entHalpy:** equações de balanço de energia.

4.3.1 Equações MESH em estado estacionário

THIELE e GEDDES (1933) e LEWIS e MATHESON (1932) foram os pioneiros

nos métodos rigorosos de destilação, sendo considerados como a base dos atuais modelos rigorosos de colunas. O modelo de THIELE e GEDDES (1933) tem como ponto chave a utilização de fatores de esgotamento e absorção para calcular as composições do líquido e do vapor, e as cargas térmicas do condensador e refeedor. O cálculo é feito estágio por estágio, com substituição sucessiva, uma parte do topo a alimentação e outra do fundo a alimentação. Com a composição calculada, a temperatura dos estágios é atualizada com o método de Ponto de Bolha. Tendo a nova temperatura, as vazões globais são calculadas pelas equações de balanço de energia. O modelo requer que sejam especificadas todas as condições da alimentação, o estágio da alimentação, o número de estágios, a razão de refluxo e a vazão total de produto (destilado ou fundo). O método de LEWIS e MATHESON (1932) é bem parecido, sendo que a diferença diz respeito à especificação da distribuição de um componente chave-leve e de um componente chave-pesado. O número de estágios pode ser alterado até que os cálculos convirjam no ponto de encontro, no estágio de alimentação.

FRIDAY e SMITH (1964) apresentaram um método mais adequado para modelar uma coluna de absorção e de *stripping*, que ganhou o nome de método da soma das taxas (SR). O método SR calcula a temperatura solucionando o balanço de energia com a técnica de Newton-Raphson. Em certos casos, a temperatura exerce mais influência nos cálculos termodinâmicos que as composições, como nos cálculos de calor de vaporização. A vazão por componente pode ser calculada pelo método de matriz tridiagonal.

WANG e HENKE (1966) foram os primeiros a citar as equações MESH e também os primeiros a utilizar o método de Pontos de Bolha (*Bubble-point method*). O método de *Bubble-point* utiliza as equações de equilíbrio e as restrições de soma das composições para calcular as temperaturas dos estágios. Neste método, as variáveis calculadas são composições e as vazões totais. O tratamento das equações do modelo feito por WANG e HENKE (1966) gera uma matriz tridiagonal. A matriz tridiagonal pode ser resolvida pelo algoritmo de Thomas (HOLLAND, 1981), tornando o método rápido e preciso. Além disso, a forma matricial tridiagonal permite o armazenamento dos elementos não-nulos em forma de três vetores, em vez de uma matriz ($N \times N$), economizando memória de cálculo. Esse método é a base de muitos métodos até hoje.

Tanto no método de *Bubble-point* quanto no SR as temperaturas e as vazões (líquido e vapor) são calculadas separadamente. Uma solução alternativa foi descrita por HOLLAND (1981), calculando as temperaturas e vazões simultaneamente e utilizando a técnica de Newton-Raphson. A estratégia ganhou o nome de Newton- $2N$, pois significa que são $(N(2c+1))$ equações a serem resolvidas, sendo N o número

de estágios e c o número de componentes (SEADER *et al.*, 2010).

NAPHTALI e SANDHOLM (1971) e HOLLAND (1981) desenvolveram o método Newton Global, que resolve todas as equações MESH simultaneamente com a técnica de Newton-Raphson. As técnicas que utilizam matriz tridiagonal calculam as composições de cada estágio independentemente. Isso não constitui uma desvantagem, especialmente em sistemas altamente não ideais, em que a constante de equilíbrio e a entalpia são altamente dependentes da composição. Basicamente, o método Newton Global inclui o balanço de massa por componente no método Newton-2N. A desvantagem é que o método requer de uma estimativa inicial muito boa para convergir.

Estes métodos descritos formam a base de modelos das equações MESH e estratégias de soluções dos atuais modelos rigorosos. A Figura 4.3 ilustra um estágio genérico j , que envolve as vazões molares de líquido e vapor provenientes dos estágios situados imediatamente acima e abaixo, a entrada de carga neste estágio, bem como adições/retiradas de calor e retiradas laterais de vapor e líquido existentes no mesmo (KISTER, 1992; SEADER *et al.*, 2010). A Figura 4.4 ilustra uma coluna de destilação com múltiplas alimentações, retiradas de líquido e vapor laterais, trocas térmicas, condensador parcial e refeedor. Também é possível com as equações MESH fazer o cálculo do número de graus de liberdade do sistema e determinar quantas e quais variáveis devem ser especificadas e quais devem ser calculadas (SEADER *et al.*, 2010).

As equações que representam o balanço material, escritas na forma vetorial ($i : 1, \dots, c$ componentes e $j : 1, \dots, N$ estágios), podem ser definidas como a Equação (4.37) (HOLLAND, 1981).

$$F_j \left[q \underline{y}_j^f + (q - 1) \underline{x}_j^f \right] + L_{j-1} \underline{x}_{j-1} + V_{j+1} \underline{y}_{j+1} - (U_j + L_j) \underline{x}_j - (W_j + V_j) \underline{y}_j = \underline{0} \quad (4.37)$$

em que q é a fração de vapor na corrente de alimentação j . A relação de equilíbrio entre a composição da fase líquida com a fase vapor é descrita pela Equação (4.38):

$$\underline{y}_j^{eq} = \underline{K}_j \underline{x}_j \quad (4.38)$$

em que \underline{y}_j^{eq} é a fração molar da composição do vapor em equilíbrio teórico com a composição da fase líquida. A constante de equilíbrio, $K_i(T, P, \underline{y}, \underline{x})$, utilizada na Equação (4.38), está descrita na Equação (4.15). A matriz de constantes de equilíbrio $\underline{K}_j(T_j, P_j, \underline{y}_j, \underline{x}_j)$ pode ser escrita na forma diagonal:

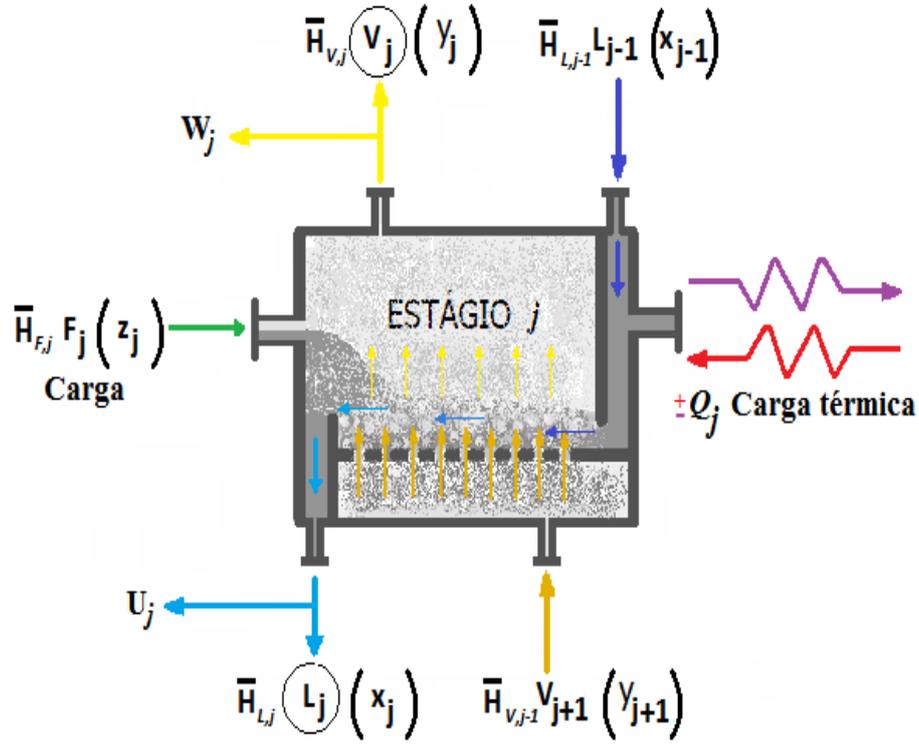


Figura 4.3: Estágio de equilíbrio genérico (KISTER, 1992; SEADER *et al.*, 2010).

$$\underline{\underline{K}}_j = \begin{bmatrix} K_1 & 0 & \cdots & 0 \\ 0 & K_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_n \end{bmatrix}_j \quad (4.39)$$

As restrições de soma das frações molares normalizadas são expressas pela Equação (4.40) ou Equação (4.41).

$$\underline{1}^T \underline{x}_j - 1 = 0 \quad (4.40)$$

ou

$$\underline{1}^T \underline{y}_j - 1 = 0 \quad (4.41)$$

em que $\underline{1}^T$ representa um vetor de dimensão adequada à operação, cujas componentes são todas unitárias.

A equação do balanço de energia é escrita na forma da Equação (4.42).

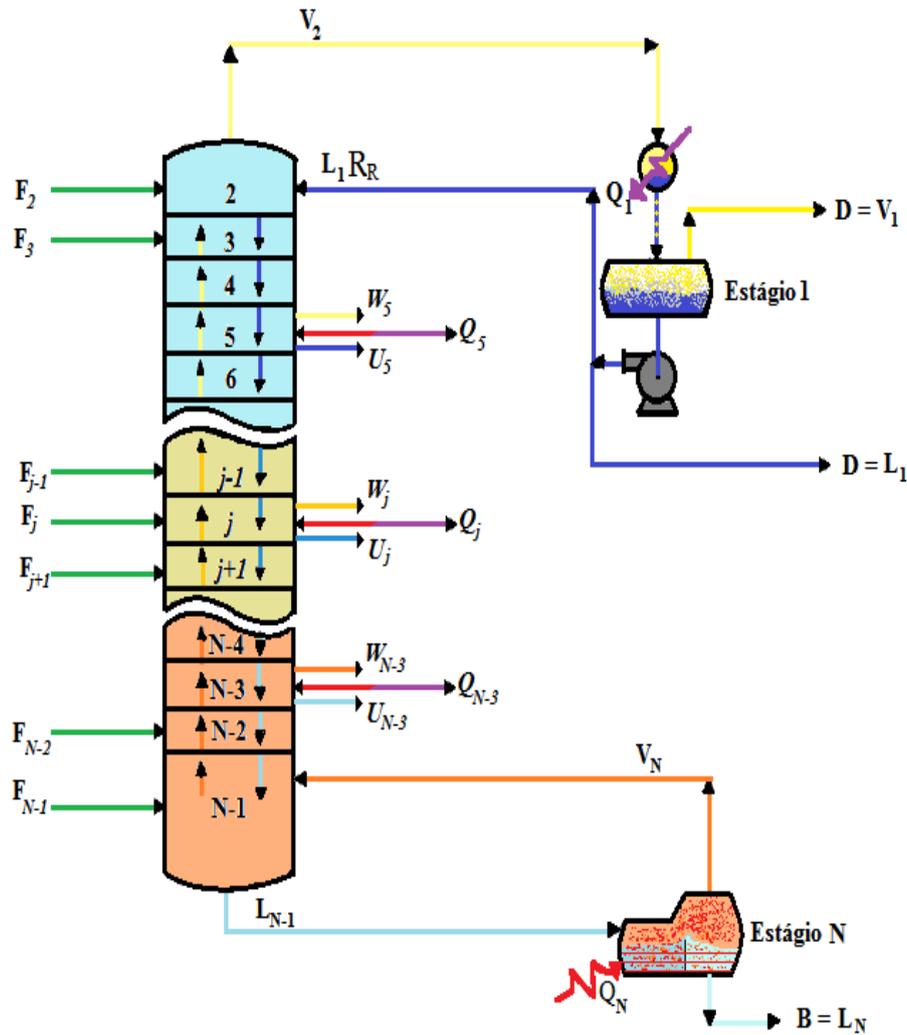


Figura 4.4: Representação de uma coluna de destilação com múltiplas alimentações, retiradas de líquido e vapor laterais, trocas térmicas, condensador parcial e refeedor (KISTER, 1992).

$$F_j [q\bar{H}_{f,j} + (q - 1)\bar{h}_{f,j}] + L_{j-1}\bar{h}_{j-1} + V_{j+1}\bar{H}_{j+1} - (U_j + L_j)\bar{h}_j - (W_j + V_j)\bar{H}_j + Q_j = 0 \quad (4.42)$$

em que \bar{H} se refere à propriedade entalpia molar da fase vapor, \bar{h} é a entalpia molar da fase líquida, Q_j é o calor cedido/retirado do estágio j , U_j é a retirada lateral de líquido em equilíbrio do estágio j e W_j é a retirada lateral de vapor em equilíbrio do estágio j .

Além das equações MESH, a correção dos equilíbrios teóricos termodinâmicos do pratos e estágios são frequentemente utilizados, pois o vapor que atravessa a lâmina de líquido do prato não tem tempo suficiente para atingir o equilíbrio teórico com o líquido, que desce para outro prato pelo *downcomer*. Assim, o uso de Eficiência

de Pratos é empregado, visando a reproduzir o desempenho de um prato real em não-equilíbrio. A eficiência também é utilizada para projetos, a fim de estimar o número de pratos reais em uma coluna de destilação. A eficiência mais empregada é a eficiência de Murphree \underline{E}^M , que é definida para cada componente e por estágio. A eficiência de Murphree se aplica a um estágio genérico j , no qual as vazões são as mesmas e as concentrações reais são comparadas com as de equilíbrio. Para a fase gasosa, a eficiência de Murphree definida para cada componente i ganha a forma da Equação (4.43) (CALDAS *et al.*, 2007).

$$\underline{E}_j^{MG} = \frac{y_j - y_{j+1}}{y_j^{eq} - y_{j+1}} \quad (4.43)$$

4.3.2 Modelo dinâmico de coluna de destilação multicomponente

O estudo do comportamento dinâmico de colunas de destilação tem sido uma das mais interessantes e desafiadoras áreas de simulação de processo químicos e petroquímicos. Desde o pioneiro trabalho de MARSHALL e PIGFORD (1947), muitos trabalhos têm sido publicados a respeito. Comparações com experimentos também foram feitas para validação dos modelos dinâmicos, como nos trabalhos de LUYBEN *et al.* (1964), HUCKABA *et al.* (1965) e KISAKÜREK (1983), que confirmam a aplicação de modelos dinâmicos de coluna para fins de predição.

PEISER e GROVER (1962) foram os primeiros a apresentar os modelos matemáticos detalhados de coluna de destilação multicomponentes. HOLLAND e LIAPIS (1983) apresentaram estratégias de solução para diferentes problemas de separação de processo, tal como o método de Gear para solucionar sistemas de equações diferenciais. Os autores apresentaram também soluções com métodos de Runge-Kutta semi-implícito.

KOOIJMAN e TAYLOR (1995) utilizaram a abordagem de não-equilíbrio entre as fases juntamente com um modelo dinâmico de uma coluna de pratos. A abordagem se baseou na transferência de massa entre as fases, ao invés do equilíbrio termodinâmico de fases. Um dos trabalhos pioneiros nessa abordagem de não-equilíbrio para modelo de coluna foi KRISHNAMURTHY e TAYLOR (1985). Os métodos de não-equilíbrio formam conjuntos de equações parecidos com as equações MESH, as equações MERQ (balanço de massa, balanço de energia, equações das taxas de difusão e relação de equilíbrio de interface).

ALEJSKI e DUPRAT (1996) apresentaram um trabalho em que analisaram um modelo dinâmico de destilação multicomponente reativa, comparando com dados de uma experiência em uma planta piloto real. Os resultados tiveram precisão satisfatória. Pequenas diferenças foram relacionadas às: hipóteses simplificadores do modelo matemático, relacionada a forte não idealidade dos azeótropos presentes e na precisão das medidas experimentais.

NASRI e BINOUS (2009) utilizaram o *software* Mathematica[®] para simular um modelo dinâmico rigoroso de coluna com quatro componentes, 27 estágios e compararam os resultados com o simulador comercial HYSYS, mostrando a eficácia e velocidade do *software* Mathematica[®]. Os autores utilizaram a abordagem de equilíbrio (ϕ - ϕ), a equação de estado SRK, balanços de massa, de entalpia e fórmula do vertedouro de Francis para representar a hidráulica dos pratos.

NICULESCU *et al.* (2017) apresentam um estudo da dinâmica de uma destilação criogênica multicomponente utilizada para separar os isótopos hidrogênio, deutério e trítio ($H_2, H - D, H - T, D_2, D - T, T_2$). A simulação mostrou o comportamento da coluna quando composição e vazão de alimentação são perturbadas, o que impacta diretamente o projeto de controle da coluna e o tempo requerido para alcançar o novo estado estacionário. O modelo dinâmico utilizado tinha como base as equações MESH, resolvidas com o método de LEWIS e MATHESON (1932) e a matriz tridiagonal.

As equações que modelam a coluna de destilação têm como base as equações MESH, sendo necessário adicionar termos de acúmulo temporal que dão as características dinâmicas da coluna. Os termos de acúmulo, são de massa e energia, descritos para cada estágio da coluna. Além disso, o balanço de momentos é negligenciado nos estágios; para compensar, normalmente adicionam-se equações para representar a hidráulica dos pratos. Existem muitas equações para representar a hidráulica dos pratos; entretanto, as mais utilizadas se baseiam na fórmula do vertedouro de Francis FRANCIS (1909) apud WIJN (1999).

Inicialmente, algumas hipóteses devem ser feitas para simplificar e otimizar a descrição matemática do que ocorre dentro da coluna, como suposições sobre as taxas de transferência de massa, de transferência de calor e a eficiência de mistura nos pratos, por exemplo. Os modelos dinâmicos de colunas em geral utilizam as seguintes hipóteses (STATHAKI *et al.*, 1985):

H1- Mistura perfeita da fase líquida em cada prato;

H2- Mistura perfeita da fase vapor em cada prato;

- H3-** Acúmulos molares de vapor (*vapor holdup*) e de entalpia da fase vapor são negligenciados, $\frac{dN_j}{dt} = 0$ e $\frac{d\bar{H}_j}{dt} = 0$, respectivamente;
- H4-** Acúmulos molares de líquido (*liquid holdup*) e de entalpia da fase líquida em cada prato podem variar no tempo, $\frac{dM_j}{dt} \neq 0$ e $\frac{d\bar{h}_j}{dt} \neq 0$, respectivamente;
- H5-** A hidráulica dos pratos pode ser representada por equação extra, como a fórmula do vertedouro de Francis;
- H6-** O perfil de pressão dentro da coluna é linear;
- H7-** Mudanças do calor sensível em cada prato não são negligenciadas;
- H8-** Perda de energia em cada prato é negligenciada;
- H9-** Perda de calor da coluna para ambiente externo é negligenciada;
- H10-** A eficiência dos pratos são parâmetros pré-ajustados pela equação da eficiência de Murphree;
- H11-** Constantes de equilíbrio são estimadas pela equação que melhor convém, de acordo com a composição;
- H12-** Entalpias de mistura são negligenciadas, de forma que as misturas de vapor e misturas de líquido são consideradas ideais;
- H13-** O condensador é parcial, sendo o primeiro ($j = 1$) estágio de equilíbrio; o refeedor é o último ($j = N$);
- H14-** Os líquidos acumulados nos vasos do condensador e refeedor estão bem misturados;
- H15-** A fração vaporizada da alimentação F_j^V entra perfeitamente misturada com o vapor do prato j ;
- H16-** A fração de líquido da alimentação F_j^L entra perfeitamente misturada com o líquido do prato j ;
- H17-** A influência da variação da pressão para a estimativa da constante de equilíbrio e entalpia é negligenciada, $\left(\frac{\partial \bar{h}_{i,j}^P}{\partial P_j}\right)_{T_j, x_{i,j}} = 0$ e $\left(\frac{\partial K_{i,j}}{\partial P_j}\right)_{T_j, x_{i,j}} = 0$, respectivamente.

As equações que governam a dinâmica de uma coluna devem conter: (I) balanços de massa globais; (II) balanços de massa por componente; (III) balanços de entalpia, para todos os estágios (pratos, condensador e refeedor); (IV) constantes

de equilíbrio L-V, temperatura de ponto de bolha e somas das frações molares; (V) hidráulica dos pratos. O conjunto das equações a seguir descreve a dinâmica da coluna conforme as hipóteses admitidas (YU *et al.*, 1975).

(I) Balanços de massa globais ($j : 1, 2, \dots, N$):

$$\frac{d(M_j + N_j)}{dt} \stackrel{\mathbf{H3}}{=} L_{j-1} + V_{j+1} - (U_j + L_j) - (W_j + V_j) + F_j^V + F_j^L \quad (4.44)$$

$$\text{onde: } M_j = Aa_j h_j^t \rho_j^L + M_j^{dc} \Rightarrow \frac{d(M_j)}{dt} = Aa_j \rho_j^L \frac{d(h_j^t)}{dt} \quad (4.45)$$

$$\frac{d(h_j^t)}{dt} = \frac{1}{Aa_j \rho_j^L} [L_{j-1} + V_{j+1} - (U_j + L_j) - (W_j + V_j) + F_j^V + F_j^L] \quad (4.46)$$

em que Aa_j é área ativa do prato j , h_j^t é a altura total do líquido no prato, ρ_j^L a densidade do líquido no prato j e M_j^{dc} é o acúmulo molar de líquido no *downcomer* (considerado constante).

(II) Balanços de massa por componente ($j : 1, 2, \dots, N$ e $i : 1, 2, \dots, c$):

$$\frac{d(M_j x_{i,j} + N_j y_{i,j})}{dt} \stackrel{\mathbf{H3}}{=} \frac{d(M_j x_{i,j})}{dt} = \frac{d(M_j)}{dt} x_{i,j} + \frac{d(x_{i,j})}{dt} M_j \quad (4.47)$$

$$\frac{d(M_j x_{i,j})}{dt} = Aa_j \rho_j^L \frac{d(h_j^t)}{dt} x_{i,j} + \frac{d(x_{i,j})}{dt} M_j \quad (4.48)$$

$$\begin{aligned} \frac{d(M_j x_{i,j})}{dt} = L_{j-1} x_{i,j-1} + V_{j+1} y_{i,j+1} - (U_j + L_j) x_{i,j} - (W_j + V_j) y_{i,j} + \\ + F_j^V y_{i,j}^f + F_j^L x_{i,j}^f \end{aligned} \quad (4.49)$$

Substituindo a Equação (4.46) na Equação (4.48) e depois igualando com a expressão do lado direito da Equação (4.49), obtém-se:

$$\frac{d(x_{i,j})}{dt} = \frac{1}{M_j} \left[L_{j-1}(x_{i,j-1} - x_{i,j}) + V_{j+1}(y_{i,j+1} - x_{i,j}) - \overbrace{(U_j + L_j)(x_{i,j} - x_{i,j})}^{\mathbf{0}} + \right. \\ \left. - (W_j + V_j)(y_{i,j} - x_{i,j}) + F_j^V(y_{i,j}^f - x_{i,j}) + F_j^L(x_{i,j}^f - x_{i,j}) \right] \quad (4.50)$$

$$\frac{d(x_{i,j})}{dt} = \frac{1}{M_j} \left[L_{j-1}(x_{i,j-1} - x_{i,j}) + V_{j+1}(y_{i,j+1} - x_{i,j}) - (W_j + V_j)(y_{i,j} - x_{i,j}) + \right. \\ \left. + F_j^V(y_{i,j}^f - x_{i,j}) + F_j^L(x_{i,j}^f - x_{i,j}) \right] \quad (4.51)$$

(III) Balanços de entalpia ($j : 1, 2, \dots, N$):

$$\frac{d(M_j \bar{h}_j + \cancel{N_j \bar{H}_j})}{dt} \stackrel{\mathbf{H3}}{=} \frac{d(M_j \bar{h}_j)}{dt} = \frac{d(M_j)}{dt} \bar{h}_j + \frac{d(\bar{h}_j)}{dt} M_j \quad (4.52)$$

$$\frac{d(M_j \bar{h}_j)}{dt} = Aa_j \rho_j^L \frac{d(h_j^t)}{dt} \bar{h}_j + \frac{d(\bar{h}_j)}{dt} M_j \quad (4.53)$$

$$\frac{d(M_j \bar{h}_j)}{dt} = L_{j-1} \bar{h}_{j-1} + V_{j+1} \bar{H}_{j+1} - (U_j + L_j) \bar{h}_j - (W_j + V_j) \bar{H}_j + \\ + F_j^V \bar{H}_{f,j} + F_j^L \bar{h}_{f,j} + Q_j \quad (4.54)$$

As substituições feitas no balanço de entalpia são análogas às substituições dos balanços de massa por componente. Assim, o acúmulo de entalpia nos estágios pode ser descrito pela Equação (4.55).

$$\frac{d(\bar{h}_j)}{dt} = \frac{1}{M_j} \left[L_{j-1}(\bar{h}_{j-1} - \bar{h}_j) + V_{j+1}(\bar{H}_{j+1} - \bar{h}_j) - (W_j + V_j)(\bar{H}_j - \bar{h}_j) + \right. \\ \left. + F_j^V(\bar{H}_{f,j} - \bar{h}_j) + F_j^L(\bar{h}_{f,j} - \bar{h}_j) + Q_j \right] \quad (4.55)$$

A estimativa da entalpia pode ser uma função mais complexa, como a Equação (4.36), ou pode usar as hipóteses **H12** e **H17**, que simplificam bastante os cálculos, dependendo da composição, sem perder a acurácia. Então, a estimativa da entalpia pode ser feita para a fase vapor como a de uma mistura de gás ideal, Equação (4.56), enquanto para a fase líquida estima-se a entalpia pelas propriedades parciais molares, Equação (4.57).

$$\bar{H}_j(T_j, y_{i,j}) = \sum_{i=1}^c y_{i,j} \bar{H}_{i,j}^{ig}(T_j, c_{p,i}^V) \quad (4.56)$$

$$\bar{h}_j(T_j, P_j, x_{i,j}) = \sum_{i=1}^c x_{i,j} \bar{h}_{i,j}^p(T_j, P_j, x_{i,j}) \quad (4.57)$$

Utilizando as hipóteses **H3** e **H4** para derivar as entalpias em respeito ao tempo e aplicando a regra da cadeia de diferenciação, obtém-se a Equação (4.58). Como o acúmulo de entalpia molar da fase vapor foi negligenciado, utiliza-se a Equação (4.57).

$$\begin{aligned} \frac{d(\bar{h}_j)}{dt} &= \sum_{i=1}^c \bar{h}_{i,j}^p \frac{d(x_{i,j})}{dt} + \sum_{i=1}^c x_{i,j} \left\{ \left(\frac{\partial \bar{h}_{i,j}^p}{\partial T_j} \right)_{P_j, x_{i,j}} \frac{dT_j}{dt} + \right. \\ &\quad \left. + \left(\frac{\partial \bar{h}_{i,j}^p}{\partial P_j} \right)_{T_j, x_{i,j}} \frac{dP_j}{dt} + \sum_{k=1}^c \left[\left(\frac{\partial \bar{h}_{i,j}^p}{\partial x_{k,j}} \right)_{T_j, P_j, x_{i,j}, \forall i \neq k} \frac{d(x_{k,j})}{dt} \right] \right\} \\ &\xrightarrow{\text{H17}} \\ \frac{d(\bar{h}_j)}{dt} &= \sum_{i=1}^c \bar{h}_{i,j}^p \frac{d(x_{i,j})}{dt} + \sum_{i=1}^c x_{i,j} \left\{ \left(\frac{\partial \bar{h}_{i,j}^p}{\partial T_j} \right)_{P_j, x_{i,j}} \frac{dT_j}{dt} + \right. \\ &\quad \left. + \sum_{k=1}^c \left[\left(\frac{\partial \bar{h}_{i,j}^p}{\partial x_{k,j}} \right)_{T_j, P_j, x_{i,j}, \forall i \neq k} \frac{d(x_{k,j})}{dt} \right] \right\} \quad (4.58) \end{aligned}$$

(IV) Constantes de equilíbrio, pontos de bolha e somas das frações molares ($j : 1, 2, \dots, N$ e $i : 1, 2, \dots, c$):

A temperatura de ponto de bolha requer a restrição de soma das frações molares do vapor em cada estágio, Equação (4.41). Adicionando a eficiência de Murphree a Equação (4.38) e substituindo o resultado na Equação (4.41), obtém-se:

$$\sum_{i=1}^c E_{i,j}^M K_{i,j} x_{i,j} - 1 = 0 \quad (4.59)$$

Observar que a constante de equilíbrio é função de $K_{i,j}(T_j, P_j, x_{i,j})$. Com isso, toda a Equação (4.59) deve ser diferenciada em relação ao tempo, seguindo a regra da cadeia.

$$\begin{aligned} & \sum_{i=1}^c E_{i,j}^M K_{i,j} \frac{d(x_{i,j})}{dt} + \sum_{i=1}^c E_{i,j}^M x_{i,j} \left\{ \left(\frac{\partial K_{i,j}}{\partial T_j} \right)_{P_j, x_{i,j}} \frac{dT_j}{dt} + \right. \\ & \left. + \left(\frac{\partial K_{i,j}}{\partial P_j} \right)_{T_j, x_{i,j}} \frac{dP_j}{dt} + \sum_{k=1}^{H1e7} \left[\left(\frac{\partial K_{i,j}}{\partial x_{k,j}} \right)_{T_j, P_j, x_{i,j}, \forall i \neq k} \frac{d(x_{k,j})}{dt} \right] \right\} = 0 \\ & \sum_{i=1}^c E_{i,j}^M K_{i,j} \frac{d(x_{i,j})}{dt} + \sum_{i=1}^c E_{i,j}^M x_{i,j} \left\{ \left(\frac{\partial K_{i,j}}{\partial T_j} \right)_{P_j, x_{i,j}} \frac{dT_j}{dt} + \right. \\ & \left. + \sum_{k=1}^c \left[\left(\frac{\partial K_{i,j}}{\partial x_{k,j}} \right)_{T_j, P_j, x_{i,j}, \forall i \neq k} \frac{d(x_{k,j})}{dt} \right] \right\} = 0 \quad (4.60) \end{aligned}$$

Substituindo a Equação (4.51) na Equação (4.60), é possível explicitar o termo $\frac{dT_j}{dt}$ para formar mais um conjunto de equações diferenciais. Com o termo $\frac{dT_j}{dt}$ explicitado, pode-se então substituir na Equação (4.58) juntamente com o termo $\frac{d(x_{k,j})}{dt}$, Equação (4.51), para explicitar $\frac{d(\bar{h}_j)}{dt}$ novamente. Com a nova equação de $\frac{d(\bar{h}_j)}{dt}$ explicitada, iguala-se com a Equação (4.55) e obtém-se a expressão para calcular as vazões de vapor em cada estágio V_j .

(V) Hidráulica dos pratos ($j : 2, 3, \dots, N - 1$):

Para modelar as vazões de líquidos L_j nos pratos, a fórmula do vertedouro de Francis é comumente utilizada, como mostrado na Equação (4.61).

$$L_j = \rho_j^L l_j^w \sqrt{g} k \left(h_j^t - \frac{h_j^w}{1,41} \right)^{3/2} \quad (4.61)$$

em que l_j^w é o comprimento do vertedouro do prato j , g é a constante gravitacional, k é uma constante para dimensionar as unidades e h_j^w é a altura do vertedouro do prato j .

Portanto, dadas as condições iniciais, como dados da alimentação, razão de refluxo, energia fornecida ao refeedor e condensador, o problema apresentará $N(c+2)$ equações diferenciais ($\frac{d(x_{k,j})}{dt}$, $\frac{dT_j}{dt}$ e $\frac{d(h_j^t)}{dt}$), $2N-2$ equações algébricas (V_j e L_j exceto

L_1 e L_N) para solucionar $N(c+4)$ incógnitas ($x_{i,j}$, h_j^t , T_j , V_j e L_j) a cada incremento do tempo de integração.

Para completar o problema de $N(c+2) + 2N - 2 = N(c+4) - 2$ equações algébrica-diferenciais e $N(c+4)$ incógnitas, normalmente, adicionam-se mais 2 equações diferenciais ou algébrica de controle de nível do condensador e refeedor.

Importante deixar claro que devido às diversas hipóteses e simplificações há vasta variedade de modelos de coluna de destilação disponíveis na literatura, cada um com suas simplificações, hipóteses, estratégias algébricas, métodos numéricos de integração e abordagem termodinâmica particulares. O modelo apresentado nesta tese constitui apenas um dos exemplos, tendo como base algumas suposições e simplificações. Entretanto, há muitas características em comum entre os diversos modelos, como as equações MESH. Portanto, o modelo apresentado tem como propósito estabelecer as principais características de um modelo dinâmico de coluna de destilação.

4.4 Modelos de ordem reduzida

Rigorosos modelos matemáticos dinâmicos de sistemas de separação por estágio constituído por balanços de massa e de energia apresentam elevado número de equações algébrico-diferenciais. Portanto, o custo computacional para resolver este tipo de sistema é elevado, o que causa grandes dificuldades, sendo muitas vezes impraticáveis (ou inviáveis) as aplicações em tempo real. O desafio de reduzir o custo computacional, sem perder a acurácia e representatividade matemática, motivou diversos estudos de redução de ordem, para adequar o esforço computacional às aplicações.

ESPANA e LANDAU (1978) consideraram que sistemas de separação por estágios podem ser representados por sistemas compartimentais, ao invés de pratos teóricos, divididos em três seções: condensador, seção de retificação e seção de esgotamento. Modelos bilineares obtidos podem ser descritos por 3 variáveis de estado e 8 parâmetros, requerendo o modelo completo para identificar os parâmetros, o que resulta em pouca flexibilidade de simulação.

WONG e LUUS (1980) foram os primeiros a aplicar a colocação ortogonal para redução de ordem de sistemas de separação de estágios. Os autores utilizaram como exemplo uma coluna de absorção de gás de 20 pratos, controlada pela vazão de entrada. O modelo foi reduzido para cinco pontos de colocação.

CHO e JOSEPH (1983a,b) aplicaram método de colocação ortogonal diretamente no domínio discreto, utilizando a família de polinômios de Jacobi. Os autores explicaram que a escolha dos parâmetros α e β do polinômio de Jacobi está associada à posição dos pontos de colocação. Os autores analisaram dois casos: o primeiro foi uma coluna de absorção com 20 pratos, reduzido para 3 pontos de colocação, com $\alpha = 1$ e $\beta = 3$. O segundo caso analisou colunas de destilação com 20 pratos, 3 componentes, reduzido para 4 pontos de colocação.

STEWART *et al.* (1985) mostraram que a família dos polinômios de Hahn apresentava resultados melhores em relação à família dos polinômios de Jacobi, usado por CHO e JOSEPH (1983a,b). Os autores solucionaram 4 problemas de coluna de destilação em que compararam, além dos polinômios Hahn e Jacobi, a seleção dos parâmetros α e β , mostrando que o polinômio Hahn($\alpha = 0, \beta = 0$), ponderado uniformemente, resultava em menores erros absolutos com a utilização de menos pontos de colocação. BENALLOU *et al.* (1986) mostraram que modelos de ordem reduzida baseados na colocação ortogonal eram mais precisos para predição do comportamento dinâmico que os modelos compartimentais.

PINTO e BISCAIA (1988) apresentaram estratégias diferentes de redução de ordem para tratar descontinuidades que ocorrem entre as seções da coluna. Os autores propuseram uma família de polinômios análoga à de Jacobi; entretanto, no domínio discreto, como o de Hahn. Também propuseram a aplicação da técnica de colocação ortogonal por cada seção da coluna, sem extrapolação, sendo no ponto da alimentação os polinômios assumiram o mesmo valor. SECCHI (1988) realizou a extensão do estudo desenvolvido por PINTO e BISCAIA (1988), adotando a aproximação polinomial por seção sem extrapolação. O modelo foi utilizado na estratégia de controle adaptativo *off-line*.

SEFERLIS e HRYMAK (1994) utilizaram a família de polinômios de Hahn e a técnica de colocação ortogonal em elementos finitos, tratando as variáveis como contínuas dentro de cada seção e depois aplicando a colocação ortogonal no domínio discreto.

KIENLE (2000) propôs um método de redução de ordem baseado na Teoria de Propagação de Onda, fenômeno conhecido com formação de padrões espaço-temporais. Os perfis de composição e de temperatura foram descritos por padrões de ondas, em que os parâmetros das ondas foram estimados de acordo com a condição de operação. A técnica se apresentou satisfatória apenas para sistemas ideais e moderadamente não-ideais.

Processos de separação por estágios com elevada razão de refluxo, para alta purificação, apresentam maiores complicações em controle que em condições de razão

de refluxo moderadas, devido às não linearidades de comportamento. KUMAR e DAOUTIDIS (2003) estudaram a influência da redução de ordem em sistemas com elevada razão de refluxo e observaram que o modelo reduzido respondia com dinâmica diferente. Os autores propuseram uma redução de ordem com modelos de dinâmica mais lenta e estudaram aplicações em controle *feedback*.

RIBEIRO *et al.* (2010) apresentaram uma técnica de redução de ordem aplicada a cada seção da coluna, usando a técnica de resíduos ponderados pelos momentos e a forma discreta da quadratura de Gauss-Lobato. O rescalonamento da variável independente discreta foi utilizada para calcular as raízes do polinômio de Hahn com maior exatidão. Resultados satisfatórios mostraram que a nova técnica era mais robusta, precisa e reduzia o custo computacional para soluções de sistemas de equações diferenciais de diferenças.

NASCIMENTO (2013a) apresentou um estudo de otimização para aplicação em tempo real, utilizando modelos de redução de ordem para colunas de destilação utilizado por RIBEIRO *et al.* (2010). O autor mostrou resultados satisfatórios do modelo de redução de ordem constituído com a técnica dos resíduos ponderados pelos momentos em comparação com o modelo de coluna completo. A redução de ordem utilizando resíduos ponderados pelos momentos apresentou resultados superiores aos métodos tradicionais de colocação ortogonal e de outros métodos usados no trabalho.

RIBEIRO *et al.* (2010) e NASCIMENTO (2013a) sugeriram para trabalhos futuros aplicações destes modelos reduzidos em problemas de RTO e controle preditivo. Entretanto, não se observou estudo de RDD utilizando técnicas de redução de ordem nas variáveis de estado em coluna de destilação, o que até o presente momento parece constituir uma oportunidade de estudo com caráter inovador.

4.4.1 Fundamentos teóricos e formulação matemática da técnica de resíduos ponderados pelos momentos

Polinômios ortogonais

Duas funções $p_n(x)$ e $p_m(x)$ são ditas ortogonais em relação a uma função peso $w(x)$ se obedecerem às seguintes condições no intervalo $[a, b]$ para qualquer n num domínio contínuo em x :

$$\int_a^b w(x)p_n(x)p_m(x)dx = 0, \quad (n \neq m) \quad (4.62)$$

$$\int_a^b w(x)[p_n(x)]^2 dx > 0, \quad (n = m) \quad (4.63)$$

em que $\{p_k(x)\}$ pertence a um conjunto infinito e ortogonal de funções polinomiais no domínio contínuo.

Toda função real contínua $f(x)$ que pertence ao intervalo $[a, b]$ pode ser descrita com auxílio de uma soma de funções pertencentes a este conjunto $\{p_k(x)\}$:

$$f(x) = c_0p_0(x) + c_1p_1(x) + \dots + c_np_n(x) + \dots \quad (4.64)$$

em que

$$c_n = \frac{\int_a^b w(x)p_n(x)f(x)dx}{\int_a^b w(x)[p_n(x)]^2 dx} \quad (4.65)$$

Utilizar a aproximação de uma função por um polinômio pode ser vantajoso porque as derivadas e integrais são fáceis de determinar e também são polinômios.

As famílias de polinômios ortogonais podem ser divididas em duas classes:

- **Domínio Contínuo:** Ortogonais em relação à integração;
- **Domínio Discreto:** Ortogonais em relação à soma.

Analogamente, duas funções $\ddot{p}_n(s)$ e $\ddot{p}_m(s)$ são ditas ortogonais em relação a uma função peso $\ddot{w}(s)$ se obedecerem as seguintes condições no intervalo $[a, b]$ para qualquer n num domínio discreto em s :

$$\sum_{s=a}^b \ddot{w}(s)\ddot{p}_n(s)\ddot{p}_m(s) = 0, \quad (n \neq m) \quad (4.66)$$

$$\sum_{s=a}^b \ddot{w}(s)[\ddot{p}_n(s)]^2 > 0, \quad (n = m) \quad (4.67)$$

em que $\{\ddot{p}_k(s)\}$ pertence a um conjunto finito (ou infinito) e ortogonal de funções polinomiais no domínio discreto. Assim, famílias de polinômios ortogonais são construídas em relação ao peso $\ddot{w}(s)$. Funções discretas no intervalo $[a, b]$ são aptas para solucionar equações de diferenças (entre pratos e estágios), domínio discreto, quando são submetidas ao somatório, para obedecerem as condições de ortogonalidade.

Existem na literatura alguns polinômios ortogonais definidos para o domínio discreto; entretanto, para esta tese apenas o polinômio de Hahn será utilizado para técnica de redução de ordem, visto que apresentou bons resultados, como já relatado nos trabalhos de NASCIMENTO (2013a); RIBEIRO *et al.* (2015). O polinômio normalmente é definido por uma representação de função hipergeométrica, e funções hipergeométricas são dadas da seguinte forma:

$${}_nF_m \left(\begin{matrix} c_1, c_2, \dots, c_n \\ d_1, d_2, \dots, d_m \end{matrix} \middle| x \right) = \sum_{k=0}^{\infty} \frac{(c_1)_k (c_2)_k \dots (c_n)_k}{(d_1)_k (d_2)_k \dots (d_m)_k} \frac{x^k}{k!} \quad (4.68)$$

sendo $(c)_k$ o símbolo de *Pochhammer*, que é o produto de k termos do seguinte modo $(c)_k = c(c+1)(c+2) \dots (c+k-1)$, e ${}_nF_m$ é a *Função Hipergeométrica Generalizada*. Utilizando esta notação, o polinômio de Hahn pode ser apresentado na forma da Equação (4.69).

$$Q_n(s; \alpha, \beta, N) = {}_3F_2 \left(\begin{matrix} -n, -s, n + \alpha + \beta + 1 \\ -N, \alpha + 1 \end{matrix} \middle| 1 \right), \quad \alpha, \beta > -1, \quad n \leq N \quad (4.69)$$

em que os zeros do polinômio estão contidos no intervalo descrito pela Inequação (4.70).

$$0 < q_{n,n}(\alpha, \beta, N) < q_{n,n-1}(\alpha, \beta, N) < \dots < q_{n,1}(\alpha, \beta, N) < N \quad (4.70)$$

Interpolação de Lagrange

A interpolação polinomial de Lagrange consiste em aproximar uma função contínua $f(x)$ e definida no intervalo $[0, 1]$, por um polinômio de grau $n-1$, tendo n pontos nodais ou internos $0 < x_1 < x_2 < \dots < x_n < 1$, com $p_{n-1}(x_j) = f(x_j) = f_j$.

Definindo o *polinômio interpolador de Lagrange* de grau $n - 1$:

$$l_j(x) = \prod_{\substack{k=1 \\ k \neq j}}^n \frac{x - x_k}{x_j - x_k} \quad (4.71)$$

O polinômio interpolador de Lagrange satisfaz as seguintes propriedades (função de *Kronecker* $\delta_{k,j}$), onde $x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n$ são as $n - 1$ raízes de $l_j(x)$:

$$l_j(x_i) = \delta_{i,j} = \begin{cases} 1, & \text{para } i = j, \\ 0, & \text{para } i \neq j. \end{cases} \quad (4.72)$$

Outra forma do polinômio interpolador de Lagrange é descrita pela Equação (4.73), em que $P_n(x) = c_n \prod_{k=1}^n (x - x_k)$ é o *polinômio nodal*.

$$l_j(x_i) = \frac{P_n(x)}{(x - x_j)P_n'(x_j)} \quad (4.73)$$

Assim, o polinômio interpolador de Lagrange pode ser definido como:

$$p_{n-1}(x) = \sum_{j=1}^n l_j(x)f(x_j), \quad j = 1, 2, \dots, n. \quad (n \text{ pontos de interpolação}) \quad (4.74)$$

O resíduo (ou erro) da interpolação é dado pela Equação (4.75), sendo o resíduo nos pontos de interpolação $\mathfrak{R}^{(n-1)}(x_i) = 0$, $i = 1, 2, \dots, n$ ou se $f(x)$ for, também, uma função polinomial de grau n , $\frac{d^n f}{dt^n} = 0$ (EVARISTO e ARGIMIRO, 2015).

$$\mathfrak{R}^{(n-1)}(x) = f(x) - p_{n-1}(x) = \prod_{i=1}^n (x - x_i) \frac{1}{n!} \frac{d^n f(t = \xi)}{dt^n}, \quad x_1 < \xi < x_n \quad (4.75)$$

A aplicação da interpolação Lagrangiana no domínio discreto é análoga, sendo necessário conhecer a função ou suas diferenças nos pontos adjacentes, na forma:

$$\ddot{p}(s) = \sum_{j=1}^n \ddot{l}_j(s) f(s^{(j)}) \quad (4.76)$$

$$\ddot{p}(s \pm k) \cong \sum_{j=1}^n \ddot{l}_j(s \pm k) f(s^{(j)}) \quad (4.77)$$

As diferenças podem ser calculadas nos pontos nodais para determinação das matrizes de discretização $\underline{\underline{A}}$ e $\underline{\underline{B}}$. As Equações (4.78) a (4.81) descrevem as matrizes de discretização.

$$A_{i,j}^+ = \ddot{l}_j(s^{(j)} + 1) \quad (4.78)$$

$$A_{i,j}^- = \ddot{l}_j(s^{(j)} - 1) \quad (4.79)$$

$$B_{i,j}^+ = \ddot{l}_j(s^{(j)} + 2) \quad (4.80)$$

$$B_{i,j}^- = \ddot{l}_j(s^{(j)} - 2) \quad (4.81)$$

Com isso, a interpolação Lagrangiana no domínio discreto pode ser descrita pelas Equações (4.82) a (4.85).

$$\ddot{p}(s + 1) \cong \sum_{j=1}^n A_{i,j}^+ f(s^{(j)}) \quad (4.82)$$

$$\ddot{p}(s - 1) \cong \sum_{j=1}^n A_{i,j}^- f(s^{(j)}) \quad (4.83)$$

$$\ddot{p}(s + 2) \cong \sum_{j=1}^n B_{i,j}^+ f(s^{(j)}) \quad (4.84)$$

$$\ddot{p}(s - 2) \cong \sum_{j=1}^n B_{i,j}^- f(s^{(j)}) \quad (4.85)$$

Resíduos ponderados pelos momentos

A técnica dos resíduos ponderados tem como definição aproximar a variável dependente de uma equação por expansões em séries de funções conhecidas, em que os coeficientes são determinados anulando a média ponderada do resíduo, obtido por meio da substituição da variável dependente pela função aproximada no domínio do problema (RIBEIRO, 2011). O método dos mínimos resíduos ponderados consiste

em encontrar os coeficientes do polinômio Lagrangiano que minimizam a seguinte função objetivo no domínio discreto (NASCIMENTO, 2013a):

$$\min_{\underline{c}} \ddot{J}_j(\underline{c}) = \sum_{s=a}^b \ddot{w}(s) \ddot{\mathfrak{R}}(\underline{c}, s), \quad j = 1, 2, \dots, n \quad \text{e} \quad \underline{c} = [c_0 \ c_1 \ \dots \ c_n]^T \quad (4.86)$$

em que \underline{c} são os coeficientes da aproximação polinomial.

Diferentes critérios foram utilizados na literatura para escolher a função peso $\ddot{w}(s)$, levando a diferentes métodos (FINLAYSON e SCRIVEN, 1966): método dos *subdomínios*; método *Galerkin*; método dos *mínimos quadrados*; método dos *momentos*; método da *colocação* e método da *colocação ortogonal*. Para este trabalho o método adotado para ponderar a minimização do resíduo é o método dos **momentos**. O método dos momentos foi desenvolvido por Yamada em 1947 (FINLAYSON e SCRIVEN, 1966). A função peso utilizada para ponderar a minimização do resíduo é expressa por s^{j-1} . Este tipo de ponderação consiste em anular os sucessivos momentos do resíduo, na forma:

$$\ddot{w}_j(s) = s^{j-1}, \quad j = 1, 2, \dots, n \quad (4.87)$$

Redução de ordem

Uma seção de coluna de destilação pode ser representada de forma genérica pela Equação (4.88).

$$\frac{dx_i(t)}{dt} = f[x_{i-1}(t)] + g[x_i(t)] + h[x_{i+1}(t)], \quad i = 1, 2, \dots, N \quad (4.88)$$

com condições de contorno:

$$x_0(t) = p(t) \quad \text{e} \quad x_{N+1}(t) = q(t) \quad (4.89)$$

Observa-se pela Equação (4.88) que $x_i(t) = x(s_i, t)$ é uma função discreta no domínio da variável independente i e no domínio da variável independente t contínua. Os estágios is são discretos no espaço e, também, as espécies são discretamente diferentes. Com isso, a composição admite valores discretos ao longo da coluna e entre as espécies. Entretanto, $x(s_i, t)$ também é uma função contínua no domínio em

t , havendo três possibilidades de aproximações polinomiais, duas no domínio discreto e uma no contínuo. No domínio contínuo, alguns autores aplicaram a aproximação polinomial; para transformar um problema algébrico-diferencial em um sistema só de equações algébricas. No domínio discreto não se observaram relatos de redução de ordem com aproximação polinomial no domínio das espécies químicas, pois não é comum problemas de destilação em que haja mais componentes que estágios. A redução de ordem mais discutida e relatada na literatura é na redução do número de estágios i por aproximação polinomial.

Considerando as variáveis de estado ($x_i(t)$, $y_i(t)$ e $T_i(t)$) que recebem as aproximações polinomiais de grau $n + 1$ com $n + 2$ pontos de interpolação: n pontos internos das raízes do polinômio de Hahn e 2 pontos externos. A notação usada para as raízes do polinômio de Hahn: $s^{(1)}, s^{(2)}, \dots, s^{(n)}$. As Equações (4.90) à (4.92) descrevem a aproximação polinomial dessas variáveis de estado.

$$x(s, t) \cong \ddot{x}_{p_{n+1}}(s, t) = \sum_{j=0}^{n+1} \ddot{l}_j(s) x(s^{(j)}, t) \quad (4.90)$$

$$y(s, t) \cong \ddot{y}_{p_{n+1}}(s, t) = \sum_{j=0}^{n+1} \ddot{l}_j(s) y(s^{(j)}, t) \quad (4.91)$$

$$T(s, t) \cong \ddot{T}_{p_{n+1}}(s, t) = \sum_{j=0}^{n+1} \ddot{l}_j(s) T(s^{(j)}, t) \quad (4.92)$$

em que $\ddot{l}_j(s) = \prod_{\substack{k=0 \\ k \neq j}}^{n+1} \frac{s - s^{(k)}}{s^{(j)} - s^{(k)}}$ são os polinômios interpoladores de Lagrange.

Para simplificar a notação e a demonstração, a formulação segue apenas com $x(s^{(j)}, t) = x^j(t)$. A demonstração e a notação são análogas para $y^j(t)$ e $T^j(t)$, sendo por isso aqui omitidas.

Os pontos internos são $0 < s^{(1)} < s^{(2)} < \dots < s^{(n)} < N + 1$ e os extremos são $s^{(0)} = 0$ e $s^{(n+1)} = N + 1$. RIBEIRO *et al.* (2015) sugeriu a normalização do domínio discreto dos estágios s_i , mantendo normalmente espaçados $[0, 1]$, para maior acurácia no cálculo das raízes do polinômio de Hahn $s^{(j)}$:

$$s_i = \frac{i - 1}{N}, \quad i = 1, 2, \dots, N \quad (4.93)$$

Então, com a normalização dos estágios, os pontos internos ficam $0 < s^{(1)} < s^{(2)} < \dots < s^{(n)} < 1$, e os extremos $s^{(0)} = -\frac{1}{N}$ em $i = 0$, e $s^{(n+1)} = 1$ em $i = N + 1$.

Assim, a discretização não é mais unitária e passa a ser de $\frac{1}{N}$. No caso extremo em que o número de pontos internos de interpolação n se igualou ao número de estágios N , $s^{(j)} = s_i$ para $n = j = i = N$, então o modelo reduzido se equivale ao modelo completo, $x^j(t) = x_i(t)$.

Define-se, então, o resíduo da interpolação, sendo calculado para cada ponto da interpolação $j = 0, 1, \dots, n, n + 1$:

$$\begin{aligned} \ddot{\mathfrak{R}}_{n+1}(s^{(j)}, t) = & \frac{d [\ddot{x}_{p_{n+1}}(s^{(j)}, t)]}{dt} - f [\ddot{x}_{p_{n+1}}(s^{(j)} - 1/N, t)] - g [\ddot{x}_{p_{n+1}}(s^{(j)}, t)] + \\ & - h [\ddot{x}_{p_{n+1}}(s^{(j)} + 1/N, t)] \quad (4.94) \end{aligned}$$

em que as f e h são funções das diferenças de primeira ordem negativa e positiva, respectivamente. A interpolação é feita pelas matrizes $A_{j,k}^- = \ddot{l}_k(s^{(j)} - 1/N)$ e $A_{j,k}^+ = \ddot{l}_k(s^{(j)} + 1/N)$ quadradas de dimensão $n + 2$.

$$\begin{aligned} f [\ddot{x}_{p_{n+1}}(s^{(j)} - 1/N, t)] = & \sum_{k=0}^{n+1} \ddot{l}_k(s^{(j)} - 1/N) f [x^j(t)] = \sum_{k=0}^{n+1} A_{j,k}^- f [x^j(t)] \\ & \text{para } j = 0, 1, \dots, n, n + 1 \quad (4.95) \end{aligned}$$

$$\begin{aligned} h [\ddot{x}_{p_{n+1}}(s^{(j)} + 1/N, t)] = & \sum_{k=0}^{n+1} \ddot{l}_k(s^{(j)} + 1/N) f [x^j(t)] = \sum_{k=0}^{n+1} A_{j,k}^+ f [x^j(t)] \\ & \text{para } j = 0, 1, \dots, n, n + 1 \quad (4.96) \end{aligned}$$

Então, aplicando o método dos momentos para o caso discreto com a soma dos resíduos ponderados pelos momentos, para os n primeiros momentos, o polinômio $x^j(t) = \ddot{x}_{p_{n+1}}(s^{(j)}, t)$ é obtido:

$$\ddot{\mathfrak{R}}_{n+1}^k(t) = \sum_{j=0}^{n+1} (s^{(j)})^{k-1} \ddot{\mathfrak{R}}_{n+1}(s^{(j)}, t) = 0 \quad \text{para } k = 1, 2, \dots, n \quad (4.97)$$

Os métodos utilizados para aproximar uma integral (domínio contínuo) ou um somatório (domínio discreto) são as quadraturas de Gauss. Para este caso, em particular, a classe de quadratura que inclui os pontos extremos é a quadratura de

Gauss-Lobatto. Assim, a soma dos resíduos ponderados pode ser aproximada pelo uso da forma discreta da quadratura de Gauss-Lobatto:

$$\mathfrak{R}_{n+1}^{\dots k}(t) = \sum_{j=0}^{n+1} w_j [s^{(j)}]^{k-1} \ddot{\mathfrak{R}}_{n+1}(s^{(j)}, t) = \sum_{j=0}^{n+1} M_{k,j} \ddot{\mathfrak{R}}_{n+1}(s^{(j)}, t) = 0$$

para $k = 1, 2, \dots, n$ (4.98)

em que $M_{k,j} = w_j [s^{(j)}]^{k-1}$ são os elementos da matriz $\underline{\underline{M}}$ de dimensão $(n \times n+2)$, w_j e $s^{(j)}$ são os pesos e as abscissas da quadratura de Gauss-Lobatto, respectivamente. Essa quadratura é exata para uma função polinomial de grau até $2n + 1$, o qual ocorre sempre para o caso linear, porque $\ddot{\mathfrak{R}}_{n+1}(s, t)$ é um polinômio de grau $n + 1$. Então, como as n equações descritas em (4.98) são lineares nos coeficientes da matriz $\underline{\underline{M}}$, elas podem ser reescritas conforme a Equação (4.99) (RIBEIRO *et al.*, 2015).

$$\ddot{\mathfrak{R}}_{n+1}(s^{(j)}, t) + V_{k,0}^{(0)} \ddot{\mathfrak{R}}_{n+1}(s^{(0)}, t) + V_{k,1}^{(n+1)} \ddot{\mathfrak{R}}_{n+1}(s^{(n+1)}, t) = 0 \quad \text{para } k = 1, 2, \dots, n$$

(4.99)

em que $\underline{\underline{V}} = \begin{bmatrix} \underline{V}^{(0)} & \underline{V}^{(n+1)} \end{bmatrix} = \tilde{M}^{-1} \begin{bmatrix} \underline{b}^{(0)} & \underline{b}^{(n+1)} \end{bmatrix}$, $b_{k,0}^{(0)} = M_{k,0}$ são os elementos da primeira coluna da matriz $\underline{\underline{M}}$, $b_{k,1}^{(n+1)} = M_{k,n+1}$ são os elementos da última coluna da matriz $\underline{\underline{M}}$, e \tilde{M} é uma matriz quadrada $(n \times n)$ formada pela remoção da primeira e última coluna de $\underline{\underline{M}}$.

Portanto, substituindo a Equação (4.94) na Equação (4.99), obtém-se a Equação (4.100). Além disso, para completar o sistema de $n + 2$, equações é necessária a inclusão das condições de contorno $x_0(t) = p(t) = x^0(t)$ e $x_{N+1}(t) = p(t) = x^{n+1}(t)$.

$$\frac{dX^k(t)}{dt} = \sum_{j=0}^{n+1} \underline{D}_{k,j}^- f [x^j(t)] + G [x^k(t)] + \sum_{j=0}^{n+1} \underline{D}_{k,j}^+ h [x^j(t)]$$

(4.100)

para $k = 1, 2, \dots, n$

em que

$$X^k(t) = x^k(t) + V_{k,0}^{(0)} x^0(t) + V_{k,1}^{(n+1)} x^{n+1}(t) \quad (4.101)$$

$$D_{k,j}^- = A_{k,j}^- + V_{k,0}^{(0)} A_{0,j}^- + V_{k,1}^{(n+1)} A_{n+1,j}^- \quad (4.102)$$

$$D_{k,j}^+ = A_{k,j}^+ + V_{k,0}^{(0)} A_{0,j}^+ + V_{k,1}^{(n+1)} A_{n+1,j}^+ \quad (4.103)$$

$$G[x^k(t)] = g[x^k(t)] + V_{k,0}^{(0)} g[x^0(t)] + V_{k,1}^{(n+1)} g[x^{n+1}](t) \quad (4.104)$$

4.5 Comentários Finais

Com base no que foi discutido observa-se a escassez de estudos nos problemas de RD utilizando modelos rigorosos de destilação, principalmente utilizando os modelos reduzidos. Dessa forma, é oportuno analisar o desempenho e acurácia desses modelos em conjunto com estimadores robustos em problemas de RD e DMEG, como já discutido. Além disso, a abordagem de solução sequencial dos problemas de RDDR em colunas de destilação deve ser avaliada, para o monitoramento em tempo real.

Capítulo 5

Metodologia

“Mais importante que as riquezas naturais são as riquezas artificiais da educação e tecnologia.”

“O mundo não será salvo pelos caridosos, mas pelos eficientes.”

“O respeito ao produtor de riqueza é o começo da solução da pobreza.”

Roberto de Oliveira Campos (1917 – 2001)

Este capítulo tem como objetivo apresentar a metodologia desenvolvida e implementada para monitorar processos de separação baseado RD. Para isso, 2 problemas de RD em processos de separação são apresentados: (a) RDDR em coluna de destilação, com dados simulados; (b) RD para monitoramento em tempo real em processo de separação de CO_2 por membrana, com dados reais. Processo (a): A metodologia foi desenvolvida e implementada observando os seguintes passos: o sistema de aquisição de dados; a modelagem do processo; a escolha da frequência ou tempo de amostragem; a sintonização dos estimadores-M robustos para detecção de erros grosseiros; a estratégia para acelerar o procedimento de otimização; e a implementação de um novo termo na função objetivo, para forçar a similaridade temporal, quando o processo se encontra em regime estacionário dentro do horizonte móvel. Processo (b): A metodologia foi desenvolvida e implementada de acordo com as seguintes etapas: aquisição de dados; pré-tratamento de dados; caracterização de dados; reconciliação de dados; detecção de erros grosseiros; e avaliação crítica dos dados medidos com um sensor virtual. Detalhes metodológicos mais específicos e resultados de cada processo estudado são descritos no Capítulo 6 e no Apêndice A, respectivamente. Portanto, a metodologia de forma geral tem como contribuição: a caracterização da natureza dos dados (dinâmica lenta ou rápida, que depende da frequência de discretização e aquisição de dados); o grau de similaridade temporal

dos dados (independente do processo ter dinâmica lenta ou rápida), que define indiretamente o tamanho da janela móvel; estratégias para acelerar a otimização como restrições de desigualdade adaptativa; e a flexibilidade da metodologia poder ser utilizada tanto com estimadores robustos ou não.

5.1 Processo (a) - Coluna de destilação

5.1.1 Aquisição de dados e frequência de amostragem

O sistema de aquisição de dados tem como finalidade transmitir a informação do Sistema Distribuído de Controle Digital (SDCD) até o programa matemático de reconciliação de dados. Entretanto, para esse presente trabalho os dados são gerados artificialmente e já inseridos dentro do programa de RD. Dessa forma, a metodologia omite, a princípio, etapas como pré-filtragem de dados, pois é comum em um sistema de aquisição de dados real a informação vir do SDCD corrompida, vazia (*not a number - nan*) ou com sinal fora de escala. Detalhes sobre pré-tratamentos de dados encontra-se no Apêndice A. Com isso, a primeira etapa da metodologia baseia-se primeiramente na seleção da frequência de amostragem (PRATA, 2009).

A frequência de amostragem é um parâmetro bastante relevante para a resolução do processo, pois possibilita a correta identificação da dinâmica e frequência natural do processo. Outro parâmetro importante é a resolução da amplitude do sinal; entretanto, esse parâmetro está associado à própria instrumentação, sendo um dos fatores que causam os erros intrínsecos aleatórios. A Figura 5.1 ilustra esses dois tipos de discretização do sinal contínuo de um processo.

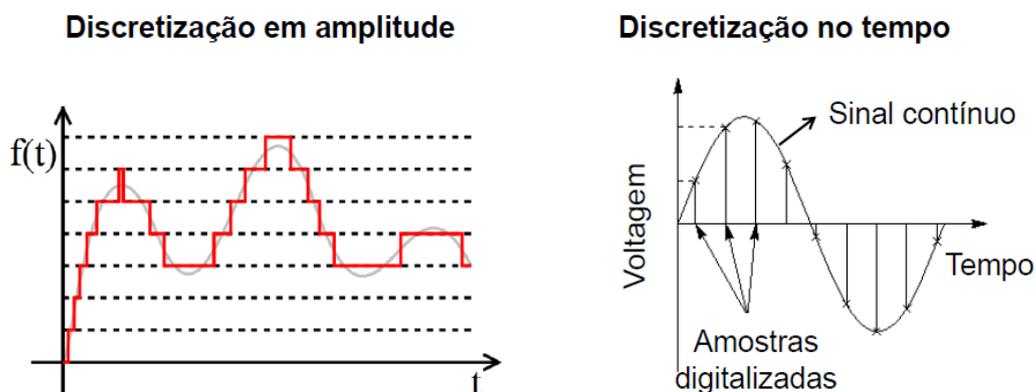


Figura 5.1: Discretização de sinal em amplitude e no tempo.

Um processo químico real tem dinâmica contínua; entretanto, para que se

entenda o processo, a aquisição de dados discretiza a informação obtida no processo contínuo e armazena em um historiador, normalmente após uma pré-filtragem. Dessa forma, a frequência deve ser compreendida como um parâmetro que ajuda a melhorar resolução da informação do processo. Por outro lado, uma frequência alta significa uma enorme quantidade de dados armazenados de forma desnecessária para a compreensão do processo. A Figura 5.2 ilustra a amostragem de um sinal contínuo.

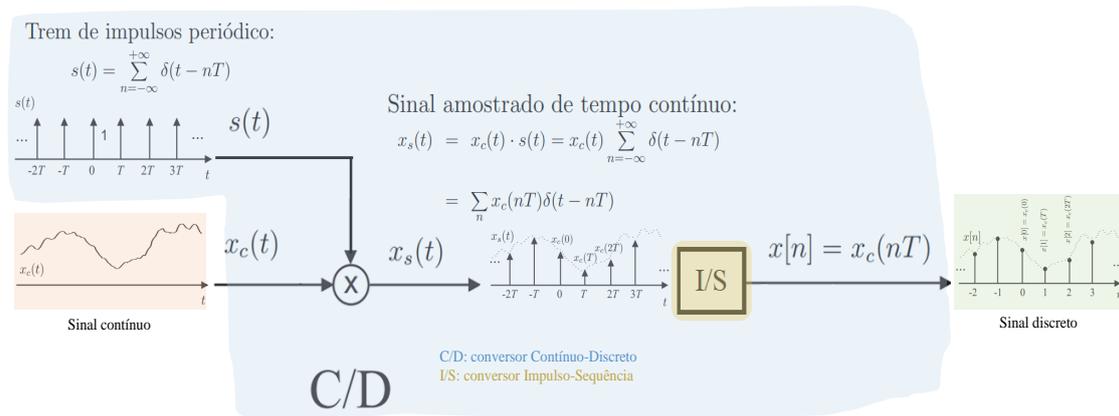


Figura 5.2: Exemplo de amostragem de sinal contínuo.

O Teorema da Amostragem de Nyquist–Shannon (NYQUIST, 1928; SHANNON, 1949) diz que um sinal analógico, limitado em banda, que foi amostrado, pode ser perfeitamente recuperado a partir de uma sequência finita de amostras, se a taxa de amostragem for maior que $2f_m$ amostras por segundo, em que f_m é a maior frequência do sinal original. Porém, se um sinal contiver uma componente exatamente em f_m , e amostras espaçadas de exatamente $1/(2f_m)$ segundos, não se consegue recuperar totalmente o sinal. A prática do teorema leva ao entendimento do *aliasing*, que ocorre quando o sistema amostrador não satisfaz as condições do teorema. A Figura 5.3 ilustra o fenômeno de *aliasing*, quando um sinal contínuo (em vermelho) é discretizado com uma frequência de amostragem menor que sua frequência original.

Observar a frequência contínua natural em processos químicos nem sempre é possível. Com isso, outra forma de selecionar a frequência de amostragem, mais adequada a processos químicos, deve ser utilizada. As Equações (5.1) e (5.2) representam as relações sugeridas em ÅSTRÖM e WITTENMARK (2013); SEBORG *et al.* (2010).

“A seleção do período de amostragem permanece mais como uma arte do que uma ciência” (SEBORG *et al.*, 2010).

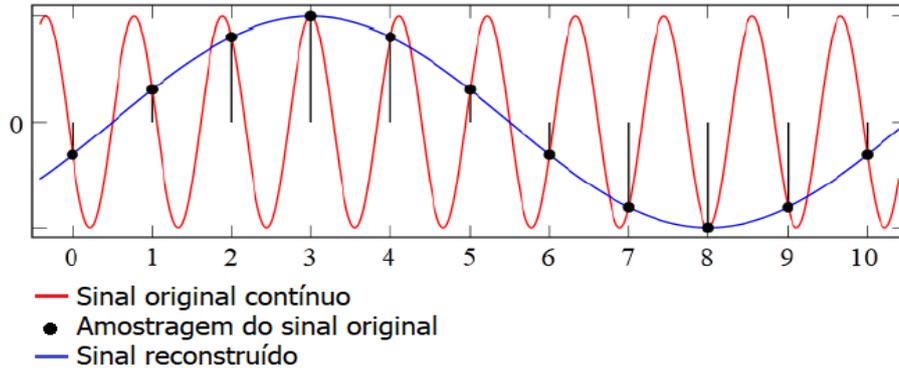


Figura 5.3: Esquema ilustrativo do fenômeno de *aliasing*.

Entretanto, a Equação (5.3) que foi a utilizada para encontrar a melhor frequência de amostragem t_a neste trabalho, sugerida em nota pelos orientadores.

$$0,01\tau_d \leq t_a \leq 0,05\tau_d \quad (5.1)$$

ou

$$\frac{t_{ss}}{15} \leq t_a \leq \frac{t_{ss}}{6} \quad (5.2)$$

ou

$$\frac{t_c}{12} \leq t_a \leq \frac{t_c}{10} \quad (5.3)$$

onde τ_d é o tempo característico dominante; t_{ss} é o tempo necessário para atingir o próximo estado estacionário; t_c é o tempo característico (ou constante de tempo τ). Sendo $t_c = \frac{1}{5}t_{ss}$ para a Equação (5.3). A Figura 5.4 exemplifica a escolha do tempo característico e a frequência de amostragem.

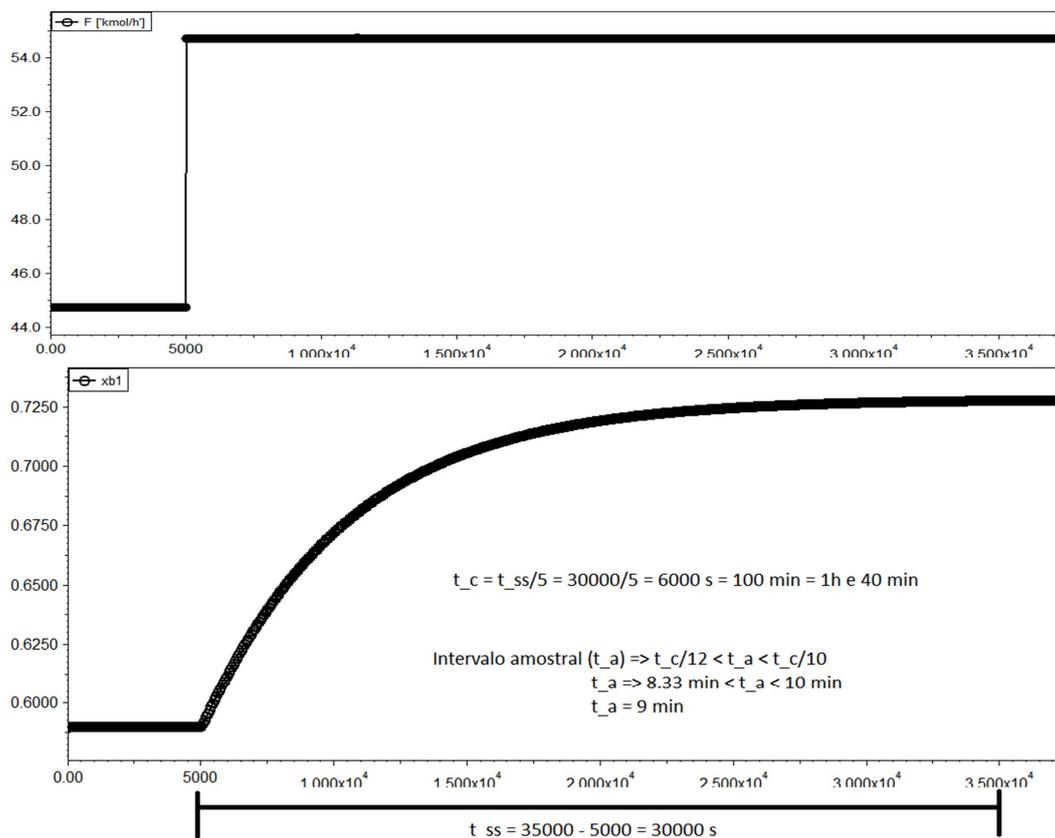


Figura 5.4: Esquema ilustrativo do processo de escolha do melhor período de amostragem.

5.1.2 Análises de séries temporais

Dados de um processo químico real é medido, armazenado e tratado, seguindo os seguintes passos, simplificadamente: medições através de sensores; aquisição desse sinal em tempos discretos; conversão e armazenamento do sinal em dados numéricos; pré-tratamento e tratamento estatístico dos dados. Com isso, as aquisições de dados geram por um certo período de tempo uma série temporal aleatória, ou seja, um processo estocástico. Dessa forma, podemos analisar os dados de um processo químico dinâmico sobre a ótica de uma grande área da estatística, conhecida como Análise de Séries Temporais.

Uma série temporal é uma coleção ordenada de medições feitas em intervalos regulares. Em modelos de regressão linear com dados de variáveis a ordem das observações é irrelevante para a análise. Entretanto, em séries temporais a ordem dos dados é fundamental. Dessa forma, os principais objetivos da Análise de Séries Temporais são: (i) compreender o mecanismo gerador da série; (ii) prever o comportamento futuro da série. Uma característica muito importante de uma série

temporal aleatória é que as observações vizinhas são dependentes, e o interesse é analisar e modelar essa dependência. Boa parte das séries têm características típicas, as principais são: tendência, ciclo e sazonalidade (MIGON, 2007). Com isso, para analisar e modelar esses dados, precisamos entender o processo estocástico e suas propriedades.

Uma variável aleatória $X(t)$ que é uma função do tempo gera um processo estocástico. A Figura 5.5 ilustra um processo estocástico, que podemos comparar, analogamente, com um processo químico. Exemplo: 4 medidores de temperatura, ζ_i com $i = 1 \dots 4$, no mesmo ponto do processo, medindo a mesma variável e de forma sincronizada (YNOGUTI, 2011).

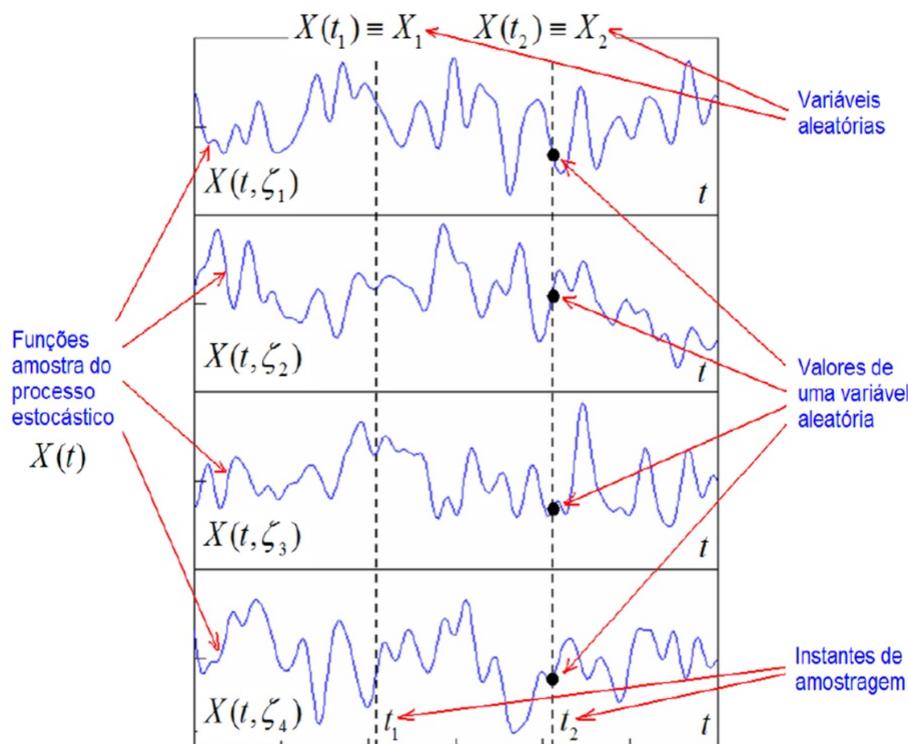


Figura 5.5: Esquema ilustrativo de um processo estocástico $X(t)$.
 Fonte: GUIMARÃES (2007).

Assim, podemos dizer que uma FDP de um processo estocástico também varia no tempo. De fato, $X(t)$ é completamente especificada se a FDP de X é especificada para cada valor de t . Pois, a FDP de X pode ser diferente para diferentes valores de t em geral. Dessa forma, a FDP de X no instante t é expressa por $f_X(X; t)$. A Figura 5.6 ilustra como uma FDP varia no tempo. A FDP $f_X(X; t)$ é conhecida com FDP de 1ª ordem. Entretanto, apenas o conhecimento de uma FDP de 1ª ordem é insuficiente para especificar um processo estocástico (YNOGUTI, 2011). Assim, para ilustrar o porquê da insuficiência da FDP de 1ª ordem, a Figura 5.7 exemplifica duas séries temporais com mesma média e variância, porém com a frequência do

processo estocástico bem distintas. Assim, duas séries temporais bem distintas apresentam a mesma FDP de 1ª ordem.

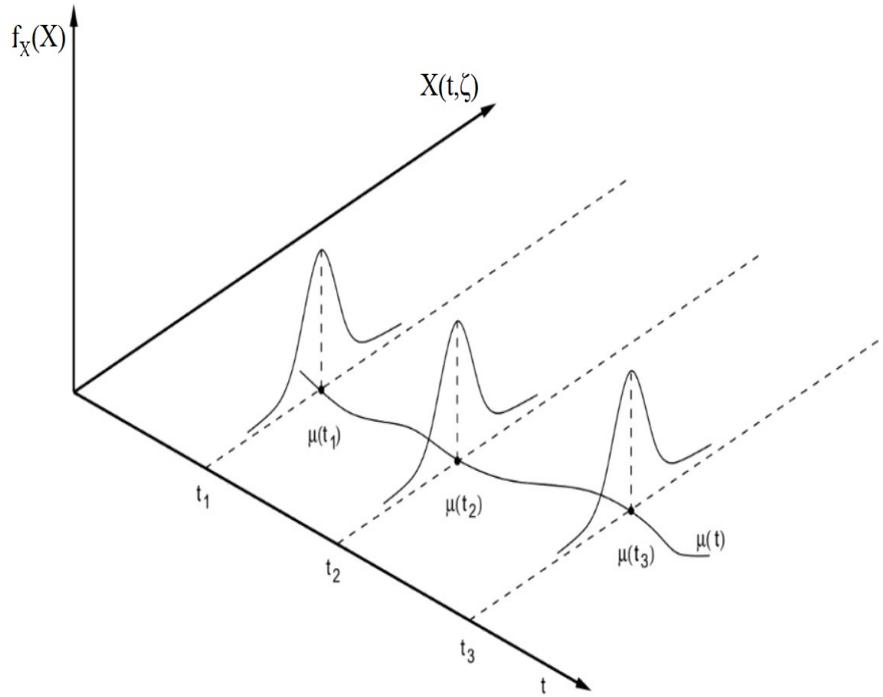


Figura 5.6: Processo estocástico interpretado como uma família de FDPs. Fonte: MORETTIN e TOLOI (2018).

Portanto, para diferenciar duas séries temporais de um processo estocástico devemos observar a correlação entre as amplitudes, em instantes temporais consecutivos, como por exemplo em t_1 e $t_2 = t_1 + \tau$ (GUIMARÃES, 2007). Dessa forma, essa medida de correlação nos dará a informação de similaridade das amplitudes em t_1 e t_2 . Conforme POLO (2017), podemos definir a função de autocorrelação, $R_X(t_1, t_2)$, como:

$$R_X(t_1, t_2) = \mathbf{E}[X(t_1), X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1), X(t_2)}(x_1, x_2) dx_1 dx_2 \quad (5.4)$$

De fato, $R_X(t_1, t_2)$ fornece informações importantes sobre o comportamento da frequência do processo estocástico, sendo conhecida como a FDP de 2ª ordem, e pode ser derivada da FDP conjunta de $X(t_1)$ e $X(t_2)$. Em resumo, para especificar um processo estocástico, precisamos não só da FDP de 1ª ordem mas também da FDP de 2ª ordem (YNOGUTI, 2011). A Figura 5.8 ilustra a diferenciação de duas séries temporais com mesmas médias e variâncias, porém distintas em suas FDPs de 2ª ordem.

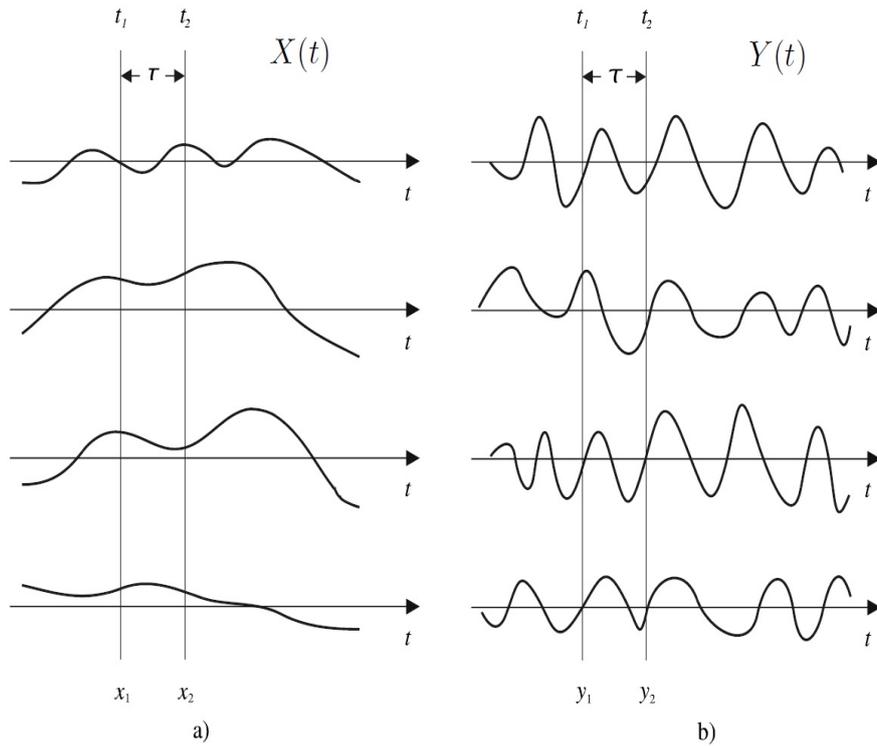


Figura 5.7: Processos estocásticos com mesma FDP de 1ª ordem.
 Fonte: YNOGUTI (2011).

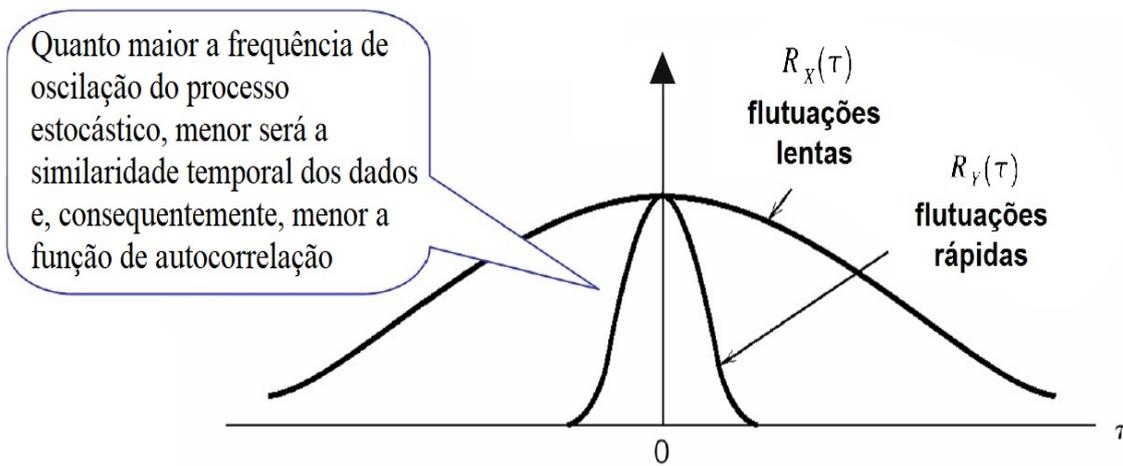


Figura 5.8: Função de autocorrelação para os processos $X(t)$ e $Y(t)$.
 Fonte: GUIMARÃES (2007).

A função de autocorrelação também pode ser estimada via média amostral se colhermos um número suficientemente grande de amostras do processo estocástico analisado, ou seja:

$$R_X(t_1, t_2) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N X_j(t_1) X_j(t_2) \quad (5.5)$$

Um processo estocástico cujas características estatísticas não variam com o tempo é classificado como um **processo estocástico estacionário**. Para um processo estocástico estacionário, podemos dizer que uma mudança da origem de tempo será impossível de detectar; o processo irá parecer o mesmo, estatisticamente (YNOGUTI, 2011). Para um processo estocástico estacionário a função de autocorrelação independe do momento em que as amostras são colhidas, dependendo somente do espaçamento temporal entre elas (GUIMARÃES, 2007; POLO, 2017). Assim teremos:

$$R_X(t_1, t_2) = R_X(t_2 - t_1) = R_X(\tau) \quad \text{para qualquer valor de } t_1 \text{ e de } t_2 \quad (5.6)$$

A função de autocovariância de um processo estocástico $X(t)$ é a covariância das variáveis aleatórias $X(t_1)$ e $X(t_2)$. Assim, a função de autocovariância pode ser interpretada como uma autocorrelação do erro da variável aleatória (descontando a média dos dados aleatórios), sendo definida pela Equação (5.7). Para um processo estocástico estacionário a função de autocovariância é definida pela Equação (5.8).

$$K_X(t_1, t_2) = \mathbf{E} \{ [X(t_1) - \mu_X(t_1)] [X(t_2) - \mu_X(t_2)] \} = R_X(t_1, t_2) - \mu_X(t_1) \mu_X(t_2) \quad (5.7)$$

$$K_X(t_1, t_2) = K_X(\tau) = R_X(\tau) - \mu_X^2 \quad (5.8)$$

Um tipo de processo estocástico estacionário importante é o processo ergódico, devido sua utilidade em sistemas de comunicações e processamentos de sinais, além de suas propriedades estatísticas especiais. Assim, para compreender e aplicar suas propriedade adequadamente aos dados de processos químicos, precisamos definir o processo estocástico ergódico.

As médias de um processo estocástico são, por definição, médias estatísticas aferidas “através” do processo, ou seja, aplicando o operador $\mathbf{E}[\]$ no conjunto de funções amostra. Para os processos ergódicos, as médias estatísticas podem ser obtidas por meio de médias temporais realizadas a partir de uma única função amostra, ou seja, “ao longo” do processo GUIMARÃES (2007). A Figura 5.9 ilustra a diferença de aferir médias estatísticas em um processo ergódico (“ao longo”) de um

processo estocástico não estacionário (“através”).

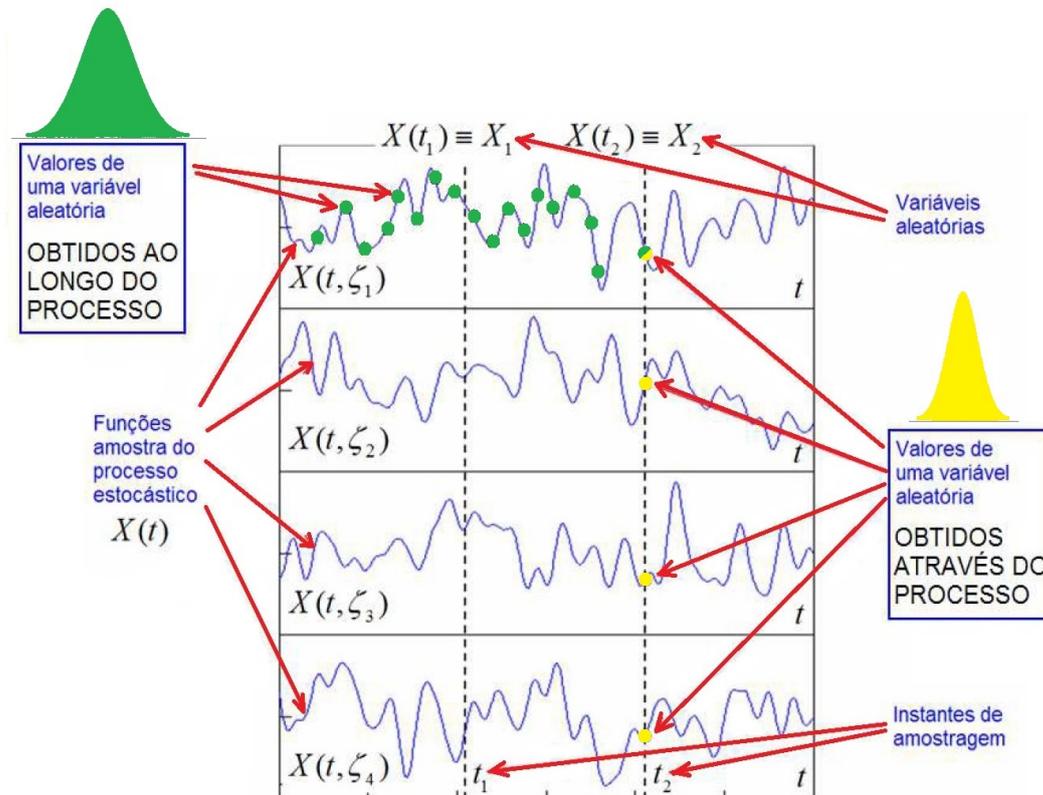


Figura 5.9: Diferença na aferição de amostras “através” e “ao longo” de um processo estocástico.

Fonte: GUIMARÃES (2007).

De fato, na maior parte de um processo químico é desejável o controle e estabilidade, como forma de garantir a qualidade e evitar acidentes. Com isso, podemos concluir que em boa parte das séries temporais oriundas dos dados de processo, tenham um comportamento de um processo estocástico estacionário, ou até, processo estocástico ergódico.

Felizmente, para a maioria dos casos as estatísticas de 1ª e 2ª ordem são suficientes para determinar o processo (GUIMARÃES, 2007). Contudo, na prática muitos dos processos estacionários são usualmente ergódicos com relação pelo menos às estatísticas de 2ª ordem, tais como a média e a função de autocorrelação (YNOGUTI, 2011).

Uma ferramenta bastante útil para estudar uma série temporal estacionária e modelar o seu comportamento são os diagramas de autocorrelações (correlogramas). Existem dois tipos de correlogramas importantes na análise de séries temporais: Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP) (MORAES, 2017).

- **Função de Autocorrelação (FAC):**

- A FAC é o correlograma da autocorrelação contra a defasagem (lag - k).

- **Função de Autocorrelação Parcial (FACP):**

- Filtra os efeitos lineares de correlações implícitas entre as defasagens, mantendo-se apenas a correlação pura (parcial) entre duas defasagens.

A definição para estimar uma FAC é descrita pela Equação (5.9) e ilustrada pela Figura 5.10, onde as diagonais da matriz de autocorrelação revelam como a correlação entre quaisquer dois valores da série muda conforme sua defasagem muda (BOX *et al.*, 2015). As Figuras 5.11 e 5.12 ilustram exemplos de correlograma FAC e FACP.

$$R_k = \frac{\mathit{cov}(X_t, X_{t-k})}{\sqrt{\mathit{var}(X_t, X_{t-k})}} = \frac{\mathit{cov}(X_t, X_{t-k})}{\mathit{var}(X_t)}$$

onde:

$$\mathit{var}(X_{t-k}) = \mathit{var}(X_t);$$

$$R_0 = 1;$$

$$R_k = R_{-k};$$

$$-1 \leq R_k \leq 1$$

assim:

$$\hat{r}_k = \frac{\sum_{t=k+1}^N (X_t - \mu_X)(X_{t-k} - \mu_X)}{\sum_{t=1}^N (X_t - \mu_X)^2} \quad (5.9)$$

A análise desses correlogramas na prática auxilia a escolha dos modelos preditivos de séries temporais. Existem vários modelos de séries temporais desenvolvidos e relatados na literatura de Análise de Séries Temporais. Entretanto, para este trabalho apenas 3 modelos clássicos de Box-Jenkins são exemplificados: Autoregressivo AR(p), Média Móvel - MA(q) e Autoregressivo com Média Móvel - ARMA(p,q) (BEZERRA, 2006; BOX *et al.*, 2015).

A análise inicial das séries temporais através das FAC e FACP nos fornecem informações para a escolha adequada do modelo (BOX *et al.*, 2015; MORAES, 2017). Em resumo, ao analisar os correlogramas devemos observar tais comportamentos:

- **Um processo AR(p) tem:**

- Um decaimento geométrico da FAC.

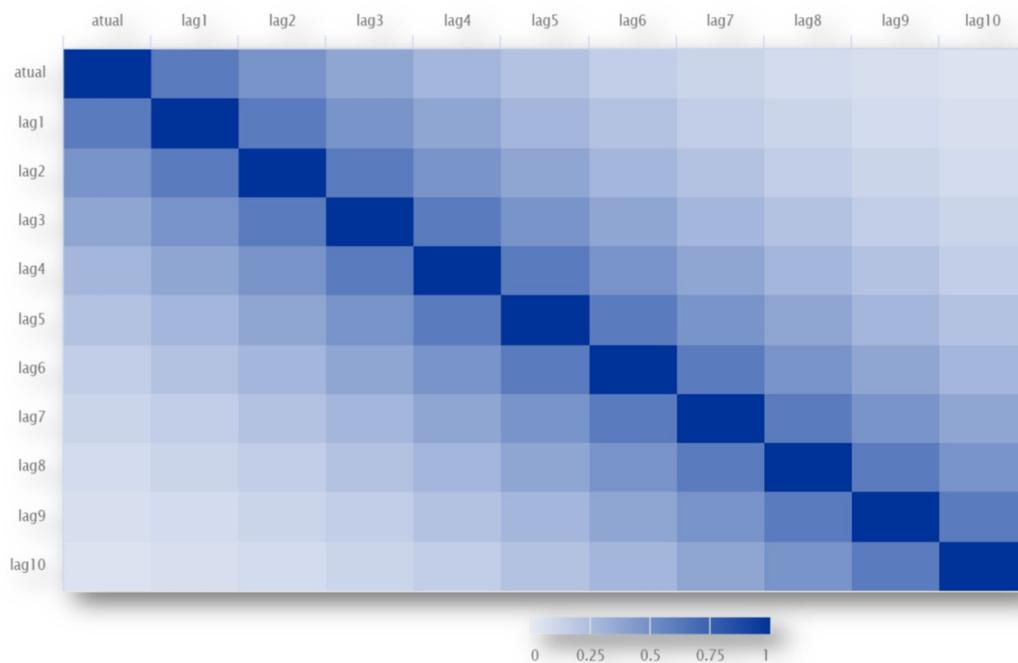


Figura 5.10: Matriz de autocorrelação correspondente de um processo estacionário. Fonte: BOX *et al.* (2015).

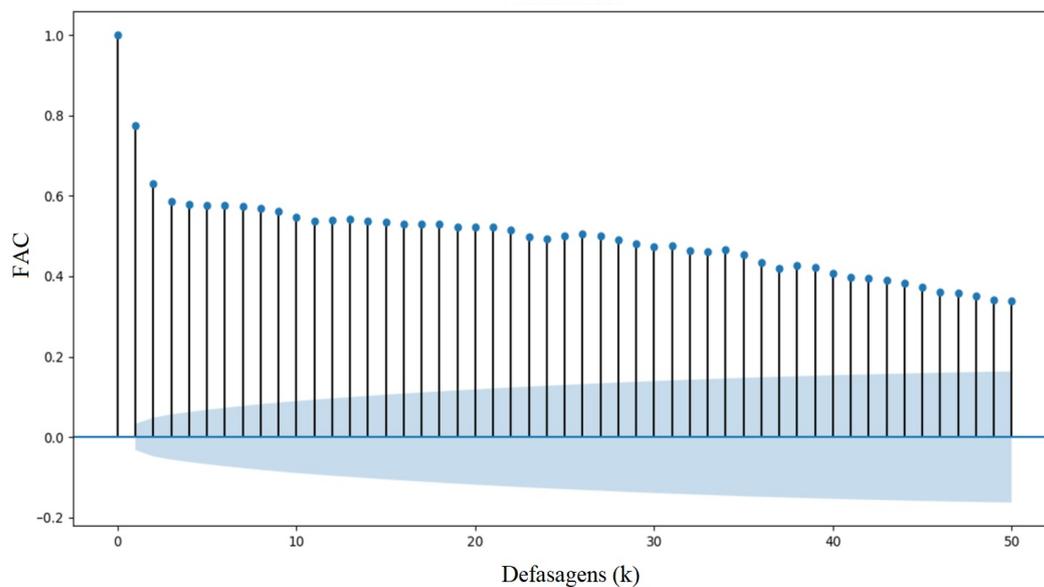


Figura 5.11: Exemplo de FAC - gráfico gerado no Python com a biblioteca *StatsModels* (lag=50).

- Um número de defasagens ($k > 0$) acima do corte (teste de significância) da $FACP = p$ (ordem da AR).
- **Um processo MA(q) tem:**
 - Um número de defasagens ($k > 0$) acima do corte (teste de significância) da $FAC = q$ (ordem da MA).

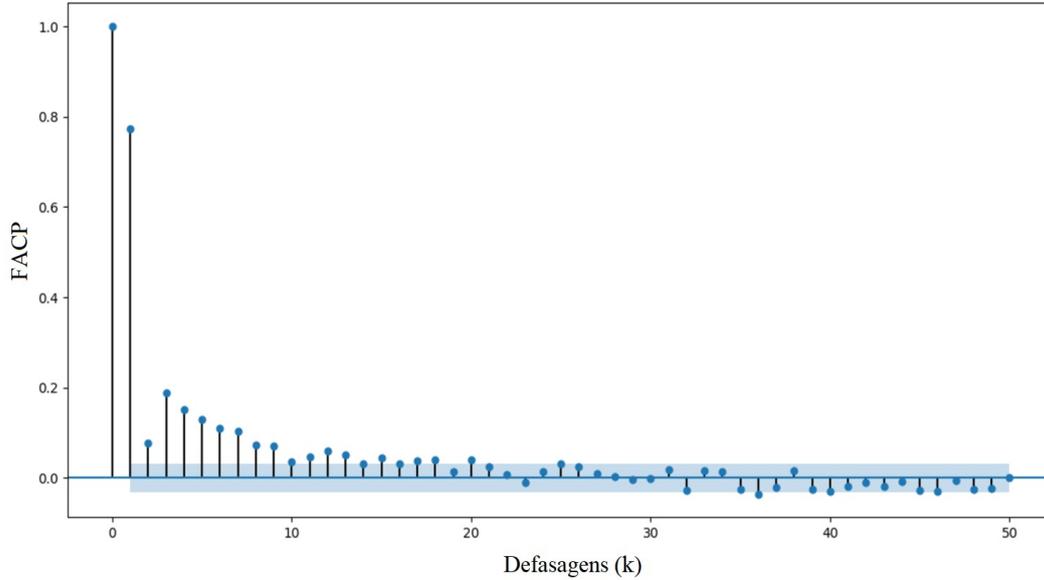


Figura 5.12: Exemplo de FACP - gráfico gerado no Python com a biblioteca *StatsModels* ($k=50$).

- Um decaimento geométrico da FACP.
- **Um processo ARMA(p,q) tem:**
 - Um decaimento geométrico da FAC.
 - Um decaimento geométrico da FACP.

Observe que há uma região em tom de azul claro nas Figuras 5.11 e 5.12, que representam os limites do teste de significância, sendo que valores acima ou abaixo da região são estatisticamente significantes. Os cálculos dos testes de significância podem ser consultados em BOX *et al.* (2015) e COSTA (2020).

O modelo Autoregressivo representa séries temporais em que o dado presente depende apenas do próprio valor da variável em defasagens anteriores e de um termo de erro. A notação AR(p) refere a ordem p do modelo Autoregressivo. O processo estocástico AR estabelece uma dependência linear sobre as defasagens temporais da variável. Assim, uma série temporal segue um processo AR de ordem p é representado pela Equação (5.10) (BEZERRA, 2006; BOX *et al.*, 2015; MORAES, 2017).

$$\hat{x}_t^* = \delta_t + \sum_{i=1}^p \phi_i x_{t-i} \quad (5.10)$$

onde ϕ_i são os parâmetros do modelo AR, p é a ordem do modelo AR, x_{t-i} são as medidas defasadas no tempo e δ_t é termos de erro de um ruído branco.

O modelo Média Móvel representa séries temporais em que o dado presente depende apenas dos erros passados, que é simplesmente uma combinação linear de termos de erro de um ruído branco. A notação MA(q) refere a ordem q do modelo “*Moving Average*”. Além disso, os processos de MA são estacionários porque suas médias, variâncias e covariâncias são finitas e invariantes no tempo. O modelo MA de ordem q é representado pela Equação (5.11) (BEZERRA, 2006; BOX *et al.*, 2015; MORAES, 2017).

$$\hat{x}_t^* = \mu_x + \delta_t - \sum_{i=1}^q \theta_i \delta_{t-i} \quad (5.11)$$

onde θ_i são os parâmetros do modelo MA, q é a ordem do modelo MA, μ_x é o valor esperado de \hat{x}_t^* , δ_t e δ_{t-i} são termos de erro de um ruído branco.

Com isso, o modelo Autoregressivo com Média Móvel representa séries temporais com o comportamento de processo AR e MA, simultaneamente. A notação ARMA(p, q) refere as ordens dos processos AR(p) e MA(q) do modelo combinado. Com isso, um processo ARMA(p, q) é representado pela Equação (5.12) (BEZERRA, 2006; BOX *et al.*, 2015; MORAES, 2017).

$$\hat{x}_t^* = \delta_t + \sum_{i=1}^p \phi_i x_{t-i} + \mu_x - \sum_{i=1}^q \theta_i \delta_{t-i} \quad (5.12)$$

Os parâmetros do modelo AR(p) podem ser estimados utilizando as equações de Yule-Walker. Os parâmetros do modelo AR(p) são em função da autocorrelação teórica, mas colocando em função da autocorrelação estimada temos as estimativas de Yule-Walker (BOX *et al.*, 2015):

$$\underline{\phi}_p = \underline{\underline{P}}_p^{-1} \hat{\underline{r}}_p \quad (5.13)$$

onde:

$$\underline{\underline{P}}_p = \begin{bmatrix} 1 & \hat{r}_1 & \cdots & \hat{r}_{p-1} \\ \hat{r}_1 & 1 & \cdots & \hat{r}_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_{p-1} & \hat{r}_{p-2} & \cdots & 1 \end{bmatrix} \quad \underline{\phi}_p = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} \quad \hat{\underline{r}}_p = \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_p \end{bmatrix} \quad (5.14)$$

Portanto, modelos de séries temporais podem ser úteis para representar

a similaridade da variável de processo. De fato, quando o regime permanece estacionário, os dados podem ser tratados como processo estocástico estacionário e, com isso, modelos de séries temporais de processo estacionários são perfeitamente aplicáveis. Com isso, para esta tese, apenas o modelo Autoregressivo foi aplicado no caso de RDD em coluna de destilação. Além disso, o processo autoregressivo aplicado foi o de primeira ordem ($p = 1$). Assim, a estratégia de similaridade temporal implementada para esta tese tem como fundamento na Análise de Séries Temporais. Além disso, o caso do monitoramento através de RD em processo de membrana, descrito no Apêndice A, aplicou os conceitos das Funções de Autocorrelações e Autocorrelação Parcial na análise dos dados reais do processo. A análise faz parte da metodologia para caracterizar e analisar a estacionaridade do processo, gerando informação importante para selecionar o modelo de restrição e a estratégia de RD mais adequados.

5.1.3 Modelagem matemática da coluna e geração dos erros de medida

O modelo utilizado para descrever a coluna baseia-se nas equações MESH, descritas detalhadamente no Capítulo 4. Para o presente trabalho, o modelo descrito no *software* EMSO utilizou apenas as equações de balanço de Massa, constantes de Equilíbrio de fases e Somatório das frações com as equações de ponto de bolha, ou seja, equações MES. As equações de balanço de energia foram negligenciadas a fim de constituir um modelo simplificado e, conseqüentemente, rápido na simulação, mas o suficientemente acurado para estudar e entender a metodologia desenvolvida sem gastar tempo excessivo computacional. Maiores detalhes sobre as colunas de cada caso estão descritas no Capítulo 6.

O modelo descrito também foi considerado perfeito; ou seja, não utilizou dados reais para estimar os parâmetros de eficiência dos pratos. Outras equações importantes que simulam um sistema de controle regulatório também não foram utilizadas, como as equações de controle de nível ou temperatura. O modelo reduzido de coluna também está descrito no *software* EMSO, que utiliza o *plugin* “Discrete” desenvolvido para calcular as raízes do polinômio de Hahn, pesos das quadraturas e as matrizes de interpolação das diferenças de primeira ordem. Assim, o modelo pode ser descrito conforme já detalhado no Capítulo 4. Tanto para o modelo completo quanto para o modelo reduzido, as equações algébrico-diferenciais foram resolvidas pelo integrador *dasslc* já pré-programado no EMSO.

Para os casos 1 e 2, o modelo de coluna foi dividido em duas etapas, em

que o processo encontra-se em regime permanente. Devido as dificuldades para minimizar o erro grosseiro na variável de entrada, uma estratégia de reconciliação fosse desenvolvida e implementada. Quando um erro grosseiro encontra-se na variável de entrada, a reconciliação pela estratégia sequencial parece ter dificuldade de corrigir esse erro. Para solucionar esse problema, a reconciliação de dados foi dividida em duas etapas: na primeira etapa as equações de balanço de massa global foram solucionadas separadamente do resto das equações do modelo; na segunda etapa, com as vazões globais já reconciliadas, todas as outras variáveis são reconciliadas utilizando o modelo completo. Dessa forma, foi possível eliminar o efeito negativo do erro grosseiro na vazão de alimentação da coluna, por exemplo.

Os erros aleatórios pequenos foram gerados dentro um intervalo de desvio padrão ($\pm 1,96\sigma$) para simular variáveis medidas, baseados nos respectivos desvios padrões relatados na literatura:

- Vazão: 2% do valor medido (SWARTZ, 1989);
- Composição: 1% do valor medido (RAO e NARASIMHAN, 1996);
- Temperatura: 0,75°C (SWARTZ, 1989).

5.1.4 Sintonia dos estimadores-M robustos

A sintonia dos estimadores é uma etapa fundamental para a detecção e eliminação dos erros grosseiros. Essa etapa pode ser feita de forma simultânea com a RD, utilizando a minimização do critério Akaike, ou de forma sequencial (*a priori*). Entretanto, para esse trabalho o procedimento utilizado foi a sintonia *a priori* com o critério de Eficiência Relativa, descrita detalhadamente no Capítulo 3 e no Apêndice B. A rotina para sintonizar os parâmetros dos estimadores-M robustos foi desenvolvida no *software* Maple. Esses parâmetros foram sintonizados para uma Eficiência Relativa de 95%. Entretanto, o trabalho descrito no Apêndice B apresenta sintonias para Eficiência Relativa de 90, 95, 98 e 99%, de 45 estimadores-M robustos. Os parâmetros de sintonia de alguns estimadores podem ser encontrados na Tabela 3.4 do Capítulo 3. No Apêndice B encontra-se o trabalho com todas as sintonias dos 45 estimadores-M robustos. A Figura 5.13 ilustra, como exemplo, a rotina desenvolvida no *software* Maple para sintonizar os parâmetros do estimador de Hampel.

```

▼ Hampel
▼ Tuning
> restart
> Tuning := proc(aH, bH, cH)
  local  $\psi$ ,  $V$ ,  $f$ ,  $f2$ ,  $Vf$ ,  $\xi$ ,  $Eff$ ;

 $\psi$  :=  $\xi$  :  $f$  :=  $\frac{e^{-\frac{1}{2}\xi^2}}{\sqrt{2\cdot\pi}}$  :  $f2$  :=  $\frac{d}{d\xi} f$  :  $Vf$  :=  $\frac{2\cdot\int_0^{+\infty} \psi \xi^2 \cdot f d\xi}{\left(2\cdot\int_0^{+\infty} \psi \xi \cdot f d\xi\right)^2}$  :

 $\psi$  := piecewise  $\left( |\xi| \leq aH, \xi, aH < |\xi| \text{ and } |\xi| \leq bH, aH \cdot \text{signum}(\xi), bH < |\xi| \text{ and } |\xi| \leq cH, aH \left(\frac{cH - |\xi|}{cH - bH}\right) \cdot \text{signum}(\xi), cH < |\xi|, 0 \right)$  :

 $V$  := evalf  $\left( \frac{2\cdot\left(\int_0^{+\infty} \psi^2 \cdot f d\xi\right)}{\left(-2\cdot\left(\int_0^{+\infty} \psi \cdot f2 d\xi\right)\right)^2} \right)$  :

 $Eff$  :=  $\frac{Vf}{V}$  :

  return  $Eff$ ; end proc :
> #Digits:=30:
>  $f$  := simplify(evalf( $\Re$ (Tuning(aH, 2·aH, 4·aH))), 'assume = real') :
> aH := fsolve( $f$  - 0.90, aH, 1..3)
                                     aH := 1.105233041 (1.1.1)
> Tuning(aH, 2·aH, 4·aH)
                                     0.9000000009 (1.1.2)

```

Figura 5.13: Sintonização do estimador de Hampel utilizando o *software* Maple.

5.1.5 Comunicação EMSO-Matlab

As equações do modelo foram resolvidas no software EMSO a cada iteração e os estimadores robustos foram minimizados pelos otimizadores (pré-programados) no *software* Matlab a cada iteração. A interface criada com o EMSO-Matlab e o Simulink permite resolver o problema utilizando uma técnica conhecida como *feasible path* (caminho viável), em que o modelo do simulador EMSO recebe as variáveis de decisão (entrada, variáveis independentes) e os parâmetros de processo, resolvendo o modelo e gerando as variáveis de saída (variáveis dependentes). Assim, o modelo (restrição do problema de RD) é sempre satisfeito. As variáveis de saída são lidas pelo *software* Matlab por meio da interface Simulink e, por fim, todas as variáveis medidas são computadas pelos otimizadores. O ciclo se repete até atingir um critério de parada (tolerância).

A tolerância utilizada é normalmente definida pela diferença entre uma variável da iteração i e a mesma variável na iteração $i + 1$, com valor de parada sendo uma constante desejada $k \geq |x^{(i+1)} - x^{(i)}|$. As Figuras 5.14 e 5.15 ilustram o ambiente criado para fazer a comunicação e a interface entre os softwares. A Figura 5.16 ilustra

o fluxograma de iterações da abordagem sequencial de otimização entre os *softwares* EMSO e Matlab, com as interfaces sendo feitas pelos ambientes de comunicação EMSO-Matlab e Simulink.

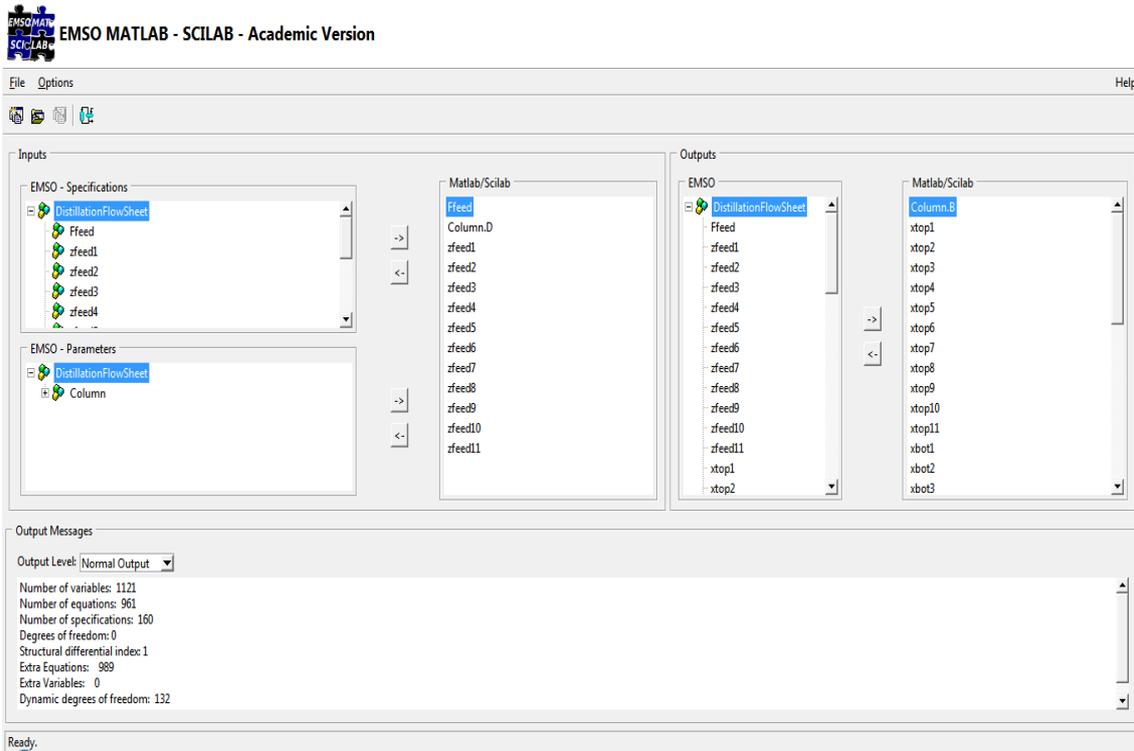


Figura 5.14: Interface EMSO-Matlab.

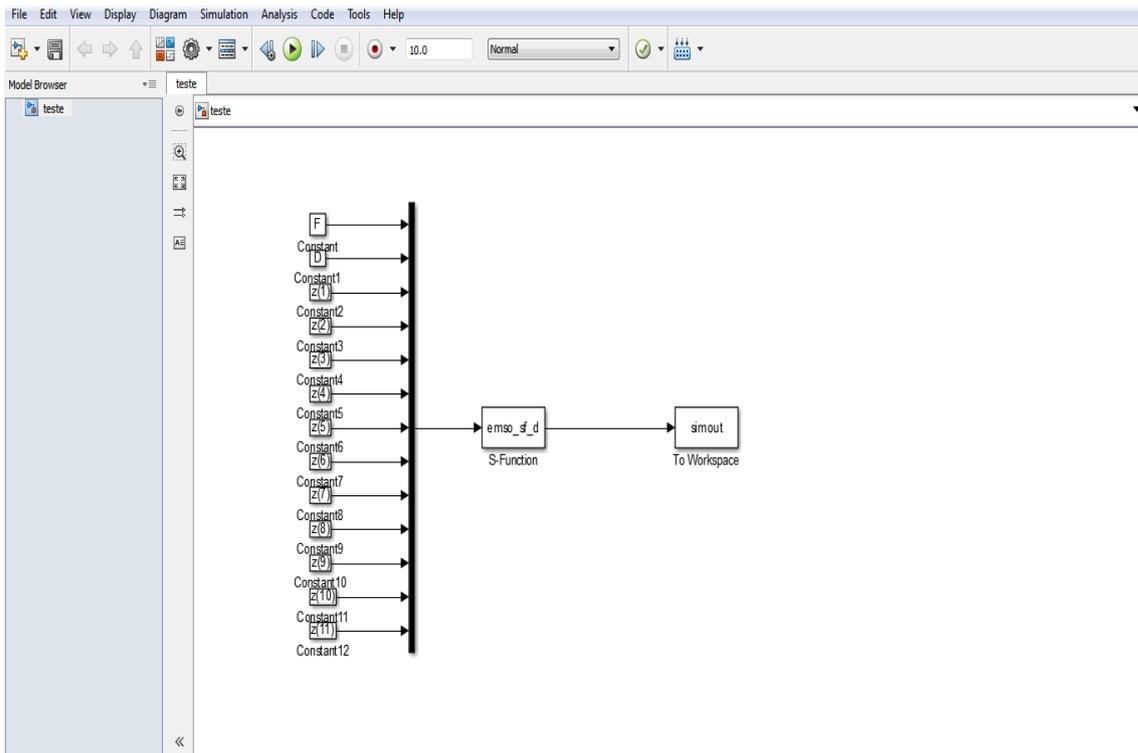


Figura 5.15: Comunicação com o Simulink (EMSO-Matlab).

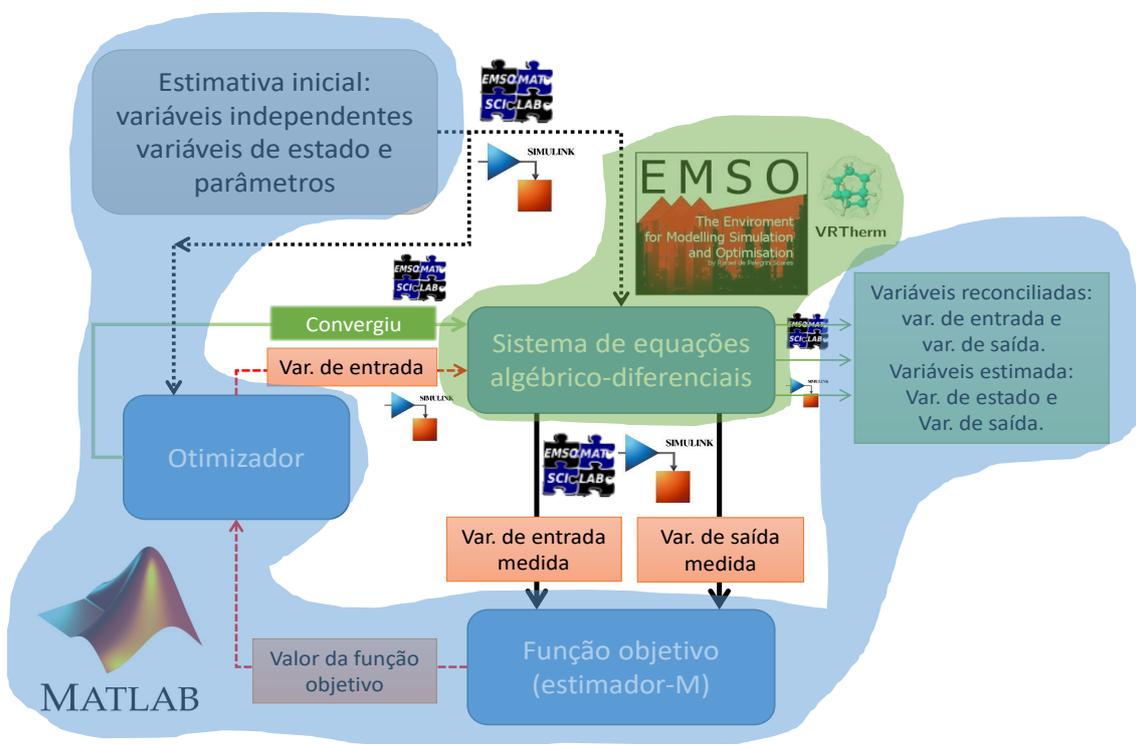


Figura 5.16: *Softwares* EMSO e Matlab aplicados ao esquema de solução sequencial para o problema de RDD.

5.1.6 Estratégias na RDDR

As variáveis são constantemente atualizadas no *software* EMSO pelo Matlab, conforme ilustrado pela Figura 5.16; entretanto, outro conjunto de variáveis (ou parâmetros iniciais) importantes são as condições iniciais do modelo. Então, para que a janela móvel se mova para a próxima região temporal, deve-se sempre atualizar as condições iniciais, tornando o problema mais fluido no tempo, pois o tempo de integração é limitado pelo tamanho da janela. Assim, o procedimento adotado foi utilizar como condições iniciais da nova janela o vetor das variáveis de estado da coluna, da janela anterior. Além disso, esse vetor estimado (ótimo) deve constituir o primeiro ponto da janela (ou o mais antigo no tempo). Com isso, a cada passo da janela móvel as condições iniciais são atualizadas.

De fato, foi considerado apenas o ponto corrente $\hat{z}(i, t_c)$ como variável de otimização dentro da janela. Isso foi feito para reduzir o custo computacional e, além disso, a redundância e similaridade temporal é condicionada ao modelo preditivo Autoregressivo. Outros pontos da janela (variáveis de entrada e condições iniciais) foram considerados constantes e utilizados apenas na integração do modelo.

Outro parâmetro importante é a estimativa inicial dada ao otimizador para a variável de entrada (ou variável de otimização). Essa estimativa é feita em duas etapas: a primeira tem um critério específico para a primeira janela móvel da RDDR; a segunda utiliza um critério que vale para as demais janelas. O critério adotado para a primeira janela utiliza com estimativa da variável 75% do valor medido no tempo corrente mais 25% da mediana dos valores dentro da janela, ou seja, $z_0(i) = 0,75z(i, t_c) + 0,25med(\underline{z}(i))$. Para a segunda etapa, a estimativa dada ao otimizador é apenas o valor presente reconciliado da janela anterior; ou seja, $z_0(i) = \hat{z}(i, t_c - 1)$. Nos desenvolvimentos iniciais utilizando o próprio valor medido como estimativa inicial, observou alguns problemas de convergência, como demora para convergir e resultados piores que os valores medidos, principalmente para as variáveis de entrada. Com a implementação dessas estratégias estes problemas foram resolvidos.

Uma estratégia desenvolvida e implementada para reduzir custo computacional foi denominada de Restrições de Desigualdade Adaptativa. Essa estratégia tem como objetivo reduzir o espaço dimensional de cada variável de otimização, sem prejudicar a região de confiança do valor medido. Outro objetivo é diminuir a possibilidade de regiões com mínimos locais, evitando problemas comuns em métodos determinísticos de otimização. Portanto, clausuras foram desenvolvidas para que as fronteiras das variáveis de otimização fossem adaptativas ao ponto corrente. Essas clausuras seguem a seguinte lógica:

- Se o valor medido no tempo corrente for **maior** que a mediana dos pontos da janela então:
 - Limite inferior: $\hat{z}_{lb} = 0,4med(\underline{z}(i)) + 0,6\hat{z}(i, t_c - 1) - 1,96\sigma_i$
 - Limite superior: $\hat{z}_{ub} = z(i, t_c) + 1,96\sigma_i$
- Se o valor medido no tempo corrente for **menor** que a mediana dos pontos da janela então:
 - Limite inferior: $\hat{z}_{ub} = z(i, t_c) - 1,96\sigma_i$
 - Limite superior: $\hat{z}_{lb} = 0,4med(\underline{z}(i)) + 0,6\hat{z}(i, t_c - 1) + 1,96\sigma_i$

Por fim, uma outra estratégia inovadora foi desenvolvida com o objetivo de captar a similaridade temporal das informações, quando o processo se encontra em regime pseudo-estacionário. Primeiramente, é necessária uma estratégia para identificar o período de regime pseudo-estacionário e, em seguida, habilitar um termo extra na função objetivo que força a similaridade temporal (ou penaliza a discrepância temporal), que representa o termo Autoregressivo de ordem 1 - AR(1). A estratégia de identificação é feita utilizando uma regressão linear simples, utilizando os pontos da janela móvel e calculado para cada variável i de entrada, conforme Equação (5.15). O termo AR(1) na função objetivo é feito utilizando o estimador MQP. Portanto, clausuras foram desenvolvidas para os termos de similaridade temporal. Essas clausuras seguem a seguinte lógica:

$$z(i, t) = \beta t + \alpha \quad (5.15)$$

- Lógica que “liga/desliga” a similaridade temporal (termo AR) da função objetivo:
 - Se $|\beta| > 0,035 \implies P_1(i) = 0$.
 - Se $|\beta| < 0,035 \implies P_1(i) = 1$.
- Lógica que “aumenta/diminui” a variância temporalmente:
 - Se $|\beta| > 0,015 \implies P_2(i) = 1$.
 - Se $|\beta| < 0,015 \implies P_2(i) = 0$.

A Fobj e o termo extra AR(1) formulado para o caso da RDDR em coluna é expresso pela Equação (5.16).

$$\min_{\hat{z}(i,t)} \sum_{t=t_c-t_H}^{t_c} \sum_{i=1}^{NV} \rho_r [\xi(i,t)] + \sum_{i=1}^{NV} \frac{P_1(i) [\hat{z}(i,t_c-1) - \hat{z}(i,t_c)]^2}{[\sigma(i) + P_2(i)]} \quad (5.16)$$

Entretanto, esta nova metodologia pode ser formulada com base no algoritmo de BAI *et al.* (2006) (descrito na subseção 2.4.2 a partir da Equação (2.102)) e estendido para modelos preditivos de séries temporais, mais especificamente, modelo AR(p).

BAI *et al.* (2006) formulam um problema de RDD baseado em modelos preditivos lineares, como uma alternativa ao filtro de Kalman, chamado de “Algoritmo de RDD baseado em predição-correção”. Além disso, o problema apresenta solução analítica, ou seja, não necessita de estratégia de solução sequencial ou simultânea.

O problema proposto neste trabalho apresenta, além do modelo de predição AR(p), restrições não lineares e dinâmicas. Sendo as restrições baseadas em modelos fenomenológicos e de primeiros princípios, admitido como modelo perfeito (sem erros). Com isso, a estratégia de otimização sequencial foi adotada para solucionar a nova metodologia, chamada de “Algoritmo sequencial de RDDR baseado em modelos fenomenológicos e autorregressivos”. A metodologia proposta representa um problema de Reconciliação de Dados Dinâmica e Robusta, suportada por um “filtro de dados” baseado em modelos lineares de séries temporais.

Dessa forma, a formulação mais abrangente para a metodologia desenvolvida e representada pelas Equações (5.17) a (5.20):

$$\min_{\hat{\underline{z}}_t} J(\hat{\underline{z}}_t) = \rho(\hat{\underline{z}}_t, \underline{z}_t, \underline{\underline{\Sigma}}) + [\underline{P}_1(\hat{\underline{x}}_t^* - \hat{\underline{z}}_t)]^T \underline{\underline{\Omega}}^{-1} [\underline{P}_1(\hat{\underline{x}}_t^* - \hat{\underline{z}}_t)] \quad (5.17)$$

sujeito a:

$$\underline{f} \left[\frac{d\hat{\underline{z}}(t, \hat{\underline{\theta}})}{dt}, \hat{\underline{z}}(t, \hat{\underline{\theta}}), \underline{u}(t, \hat{\underline{\theta}}) \right] = \underline{0} \quad (5.18)$$

$$\underline{h} \left[\hat{\underline{z}}(t, \hat{\underline{\theta}}), \underline{u}(t, \hat{\underline{\theta}}) \right] = \underline{0} \quad (5.19)$$

$$\underline{g} \left[\hat{\underline{z}}(t, \hat{\underline{\theta}}), \underline{u}(t, \hat{\underline{\theta}}) \right] \geq \underline{0} \quad (5.20)$$

onde:

$$\begin{aligned} \hat{\underline{x}}_t^* &= \underline{\underline{\Phi}}_p \hat{\underline{z}}_{t-p} + \underline{\delta}_t \\ \underline{\underline{\Phi}}_p &= \begin{bmatrix} \underline{\phi}_{-p1}^T & \underline{0}^T & \cdots & \underline{0}^T \\ \underline{0}^T & \underline{\phi}_{-p2}^T & \cdots & \underline{0}^T \\ \vdots & \vdots & \ddots & \vdots \\ \underline{0}^T & \underline{0}^T & \cdots & \underline{\phi}_{-pNV}^T \end{bmatrix}_{(NV \times \sum_{i=1}^{NV} pi)} \\ \hat{\underline{x}}_t^* &= \begin{bmatrix} \hat{x}_{t,1}^* \\ \hat{x}_{t,2}^* \\ \vdots \\ \hat{x}_{t,NV}^* \end{bmatrix}_{(NV \times 1)} \quad \hat{\underline{z}}_{t-p} = \begin{bmatrix} \hat{z}_{t-p1} \\ \hat{z}_{t-p2} \\ \vdots \\ \hat{z}_{t-pNV} \end{bmatrix}_{(\sum_{i=1}^{NV} pi \times 1)} \quad \underline{\delta}_t = \begin{bmatrix} \delta_{t,p1} \\ \delta_{t,p2} \\ \vdots \\ \delta_{t,pNV} \end{bmatrix}_{(NV \times 1)} \end{aligned} \quad (5.21)$$

onde $\rho(\cdot)$ é o estimador-M robusto; \underline{z}_t é o vetor das NV s variáveis medidas no tempo t ; $\hat{\underline{z}}_t$ é o vetor das NV s variáveis reconciliadas no tempo t ; $\underline{\underline{\Sigma}}$ é a matriz de covariâncias dos erros de medição; $\underline{\underline{\Omega}} = \mathbf{E}[\underline{\delta}_t \underline{\delta}_t^T] + \text{diag}(\underline{P}_2)$ é a matriz de covariâncias dos erros dos modelos de predição AR, juntamente com o termo que “aumenta/diminui” a variância temporalmente; $\underline{f}(\cdot)$ é vetor das equações diferenciais do modelo fenomenológico e de primeiros princípios; $\underline{h}(\cdot)$ é vetor das equações algébricas; $\underline{g}(\cdot)$ é vetor das inequações algébricas; $\underline{\phi}_{-pi}$ são os parâmetros do modelo AR para a variável i e é calculado conforme Equação (5.13); $\delta_{t,pi}$ é o termo de erro de um ruído branco do modelo AR da variável i no tempo t , definido na Equação (5.10); $\hat{\underline{z}}_{t-pi}$ é o vetor correspondente a variável reconciliada i nos tempos $t-1, t-2, \dots, t-p(i)$, com dimensão e ordem do modelo AR igual a $p(i)$, representando as $p(i)$ s medidas reconciliadas temporais defasadas de $\hat{z}_{t,i}$; $\hat{x}_{t,i}^*$ é o valor predito pelo modelo AR(pi) da variável i no tempo t .

As Figuras 5.17 a 5.21 ilustram o procedimento utilizado para a atualização das condições iniciais do modelo, as estimativas iniciais do otimizador, as restrições de desigualdade adaptativa das variáveis de otimização e os termos e parâmetros de similaridade temporal.

```

% Estimativa para otimização da próxima janela-> últimos valores reconciliados
for i=1:2+nc-1,
    x0(i) = 0.6.*p.xm(i,t+MH) + 0.2.*median(p.xm(i,1+MH:t+MH)) + 0.2.*Xrec_t(i,t+MH-1);
end

for i=1:2+nc-1,
    x0(i) = 0.75.*p.xm(i,t+MH) + 0.25.*median(p.xm(i,1+MH:t+MH));
end

x1guess = x1ot(:,1)';

% 95% de confiança (+- 1.96dp):
% Lower boundaries for independent variable: 60% do ponto atual da janela + 20% da mediana dos valores medic
% Upper boundaries for independent variable: 60% do ponto atual da janela + 20% da mediana dos valores medic
for i=1:2+nc-1,
    xlb(i)=(0.6.*p.xm(i,t+MH) + 0.2.*median(p.xm(i,1+MH:t+MH)) + 0.2.*Xrec_t(i,t+MH-1)) - 1.96.*p.dp(i);
    xub(i)=(0.6.*p.xm(i,t+MH) + 0.2.*median(p.xm(i,1+MH:t+MH)) + 0.2.*Xrec_t(i,t+MH-1)) + 1.96.*p.dp(i);
end

```

Figura 5.17: Vetor das estimativas iniciais do otimizador, das condições iniciais do modelo e os limites das variáveis de otimização no Matlab.

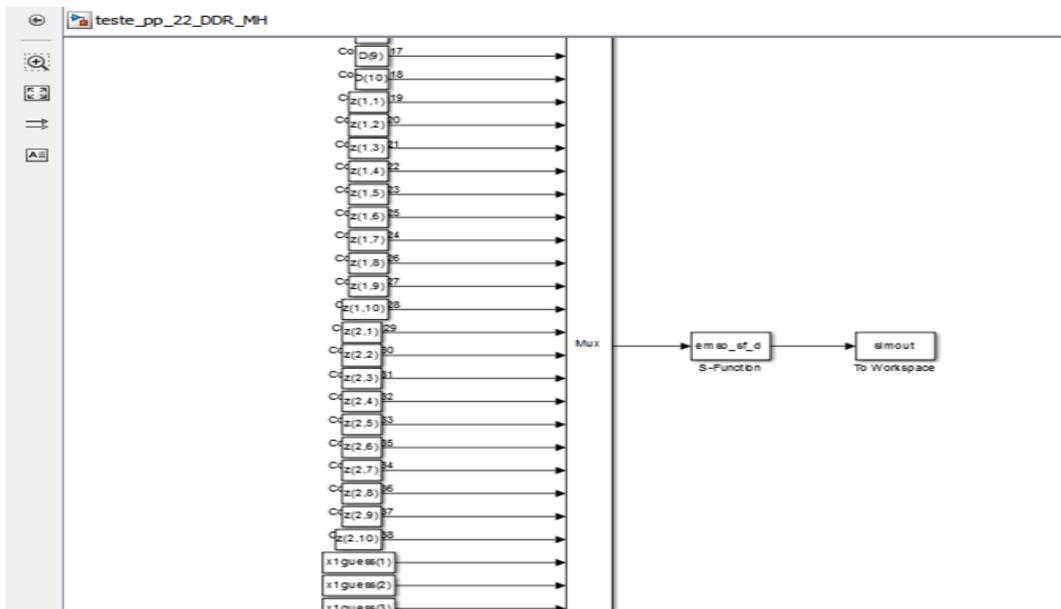


Figura 5.18: Vetor das condições iniciais e de entrada no Simulink.

```

2228
2229
2230
2231 # GUESS
2232 # col.L = 19.7 * 965 * 'kmol/d';
2233 # col.V = (19.7+1) * 965 * 'kmol/d';
2234 # col.T = [320:(330-320)/col.N1:320] * 'K';
2235 # col.y = 0.5;
2236
2237 # INITIAL
2238 # col.T = [320:(330-320)/col.N1:320] * 'K';
2239 # col.x(1,:) = [0.8:(0.2-0.8)/col.N1:0.2];
2240 # col.x(2,:) = 1-col.x(1,:);
2241
2242 #
2243
2244 col.x(1,:) = [col.x1guess1, col.x1guess2, col.x1guess3, col.x1guess4,
2245             col.x1guess6, col.x1guess7, col.x1guess8, col.x1guess9,
2246             col.x1guess11, col.x1guess12, col.x1guess13, col.x1gues
2247             col.x1guess16, col.x1guess17, col.x1guess18, col.x1gues
2248             col.x1guess21, col.x1guess22];
2249 col.x(2,:) = 1-col.x(1,:);
2250
2251 #
2252
2253 **

```

Figura 5.19: Vetor das condições iniciais no EMSO.

EMSO-MATLAB - SCILAB - Academic Version

File Options

Inputs

- EMSO - Specifications
 - Test_Coluna_completa_peq
 - kv
 - F1
 - F2
 - F3
 - F4
 - F5
- EMSO - Parameters
 - col
 - M
 - x1guess1
 - x1guess2
 - x1guess3
 - x1guess4
 - x1guess5

Matlab/Scilab

- z2_9
- z2_10
- col.x1guess1
- col.x1guess2
- col.x1guess3
- col.x1guess4
- col.x1guess5
- col.x1guess6
- col.x1guess7
- col.x1guess8
- col.x1guess9
- col.x1guess10
- col.x1guess11
- col.x1guess12
- col.x1guess13
- col.x1guess14

Outputs

- EMSO
 - Test_Coluna_completa_peq
 - kv
 - F1
 - F2
 - F3
 - F4
 - F5
 - F6
 - F7
 - F8
 - F9
 - F10
 - D1
 - D2
 - D3
- Matlab/Scilab
 - T12
 - T13
 - T14
 - T15
 - T16
 - T17
 - T18
 - T19
 - T20
 - T21
 - T22
 - col.F
 - x1
 - x2
 - x3
 - x4

Output Messages

Output Level: Normal Output

Number of equations: 316
 Number of specifications: 67
 Degrees of freedom: 0
 Structural differential index: 1
 Extra Equations: 339
 Extra Variables: 0
 Dynamic degrees of freedom: 44
 Number of initial Conditions: 44

Figura 5.20: Vetor das condições iniciais e de entrada no *software* de conexão EMSO-Matlab.

```

if MH>0,
% Para melhorar a similaridade temporal dos valores reconciliados de entrada
Psimtemp=ones(2,2+nc-1);
    for i=1:2+nc-1,
        if abs(P(1,i))>0.035,
            Psimtemp(1,i)=0;
        end
        if abs(P(1,i))<0.015,
            Psimtemp(2,i)=0;
        else
            Psimtemp(2,i)=abs(P(1,i));
        end
    end
    f = f + (Psimtemp(1,1).*(Xrec_t(1,MH+t-1)-F(t)).^2./(p.var(1)+Psimtemp(2,1)) ...
        + Psimtemp(1,2).*(Xrec_t(2,MH+t-1)-D(t)).^2./(p.var(2)+Psimtemp(2,2)) ...
        + Psimtemp(1,3).*(Xrec_t(3,MH+t-1)-z(1,t)).^2./(p.var(3)+Psimtemp(2,3)));
end
% Fobj com o estimador NC:
for j=1:t,

```

Figura 5.21: Termo extra da função objetivo que força a similaridade temporal.

5.2 Processo (b) - Separação de CO_2 por membrana

5.2.1 Introdução

Um problema comum na produção de petróleo é o excesso de gás CO_2 presente nas correntes de gás natural. O primeiro e mais notório problema está relacionado à emissão desse gás no ambiente. Porém, além dos possíveis problemas ambientais, o excesso de CO_2 nessas correntes pode causar problemas na planta de processo, como o congelamento devido à queda de pressão nas seções de compressões e resfriamento da planta e corrosão de dutos metálicos (MOKHATAB *et al.*, 2018). De acordo com a resolução da ANP (Agência Nacional de Petróleo, Gás Natural e Biocombustíveis), o gás natural comercial deve conter no máximo 3% (mol) de CO_2 (ANP, 2008). Portanto, uma solução possível para lidar com o CO_2 produzido é a reinjeção do CO_2 no poço de petróleo, o que também pode permitir o aumento da produtividade do poço. Isso certamente pode minimizar os impactos e problemas ambientais nas plantas de processamento de gás natural. Portanto, a separação do CO_2 constituiu uma etapa fundamental durante o tratamento do gás natural em campos de produção de petróleo.

Diferentes processos físicos/químicos podem ser usados para separar o CO_2 do gás natural, como destilação criogênica, absorção ou processos de osmose reversa de membrana (MOKHATAB *et al.*, 2018; SPEIGHT, 2018). Particularmente, a remoção de CO_2 do gás natural com a ajuda de processos de separação por membrana tem sido usada desde 1981 (HENIS e TRIPODI, 1981). No entanto, as aplicações foram inicialmente limitadas devido aos riscos econômicos intrínsecos associados à atividade de produção de petróleo e às restrições operacionais relacionadas às separações por membrana. No entanto, o cenário vem mudando devido às vantagens relacionadas ao menor consumo de energia, baixo investimento de capital, baixos custos operacionais e natureza mais compacta desses equipamentos (AL-OBAIDI *et al.*, 2018; KOVVALI *et al.*, 1994; SINGH *et al.*, 2013).

Diante do aumento da importância industrial dos processos de separação por membranas, as demandas por desenvolvimento de modelagem matemática, simulação, otimização, controle, tratamento de dados estatísticos e procedimentos de monitoramento online também aumentaram, visto que essas técnicas são fundamentais para o projeto e monitoramento de processos químicos. Como consequência, o desempenho do processo analisado pode ser avaliado de forma mais precisa e monitorada, permitindo a detecção de falhas em linha e em tempo real. Com base nessas tecnologias, os riscos e o tempo necessários para a tomada de decisão podem ser minimizados (SOARES, 2003).

Portanto, DE MENEZES *et al.* (2020) desenvolveram e implementaram uma metodologia pioneira, para uma aplicação web que possibilite o monitoramento online e em tempo real, em processos de separação de CO_2 por membrana em escala industrial, com base em RD, DMEG e métricas estatísticas. O aplicativo também pode ser usado para fornecer informações sobre variáveis não medidas (sensor virtual) e para diagnosticar a ocorrência de medições de erros grosseiros e mau funcionamento do instrumento. A metodologia proposta compreende as seguintes etapas: (i) pré-tratamento e caracterização dos dados do processo; (ii) reconciliação dos dados do processo para minimizar as incertezas da medição, com o auxílio de equações de balanço de massa; (iii) detecção de desvios sistemáticos (*bias*) para identificação de falhas e mau funcionamento do processo; e (iv) observação de variáveis não medidas (sensor virtual ou gêmeo digital).

5.2.2 Processo de Separação por Membrana

As membranas constituem excelentes alternativas para a separação de gases devido aos baixos custos de instalação e manutenção. No ambiente industrial e para esta aplicação, as membranas são geralmente do tipo: densa, material polimérico e morfologia assimétrica. Elas são organizadas em módulos com geometria cilíndrica de folhas enrolada em espiral ou de fibras ocas. As unidades de separação com fibras ocas são normalmente aplicadas a fluxos relativamente menores quando comparados aos módulos com folhas enroladas em espiral. Por outro lado, os módulos com folhas enroladas em espiral são mais baratos, capazes de lidar com pressões operacionais mais elevadas e são mais resistentes à incrustação, pois as partículas presentes no fluxo de gás de alimentação podem bloquear as fibras da membrana (DORTMUNDT e DOSHI, 1999; RACKLEY, 2017). Por isso, as unidades com folhas enroladas em espiral são amplamente utilizadas em processos industriais de separação de gases.

DIAS *et al.* (2020) desenvolveram um modelo matemático baseado em uma abordagem fenomenológica para uma folha de membrana em espiral. O modelo foi validado em quatro estudos de caso de separações de gases comuns, com muito bom desempenho e robustez. O modelo também permitiu a previsão de vazões e concentrações ao longo da folha da membrana, características importantes para o entendimento dos processos de operação da membrana. Além disso, um método de discretização foi proposto para resolver o modelo, que se mostrou mais rápido e mais eficiente, comparado ao método de *shooting*.

Importante também enfatizar que as unidades industriais de separação por membrana espiralada para aplicações de CO_2 são formadas por várias folhas, as quais são enroladas em um tubo coletor perfurado central, formando um elemento de

separação modular. Esses elementos podem ser dispostos em série para construir um tubo de separação por membrana. Em seguida, os tubos podem ser organizados em paralelo para formar um banco. Por fim, os bancos podem ser alinhados em paralelo para compor um trem de separação por membrana, enquanto os trens podem ser dispostos em paralelo para formar um estágio (DORTMUNDT e DOSHI, 1999). A Figura 5.22 mostra um exemplo desse tipo de unidade.

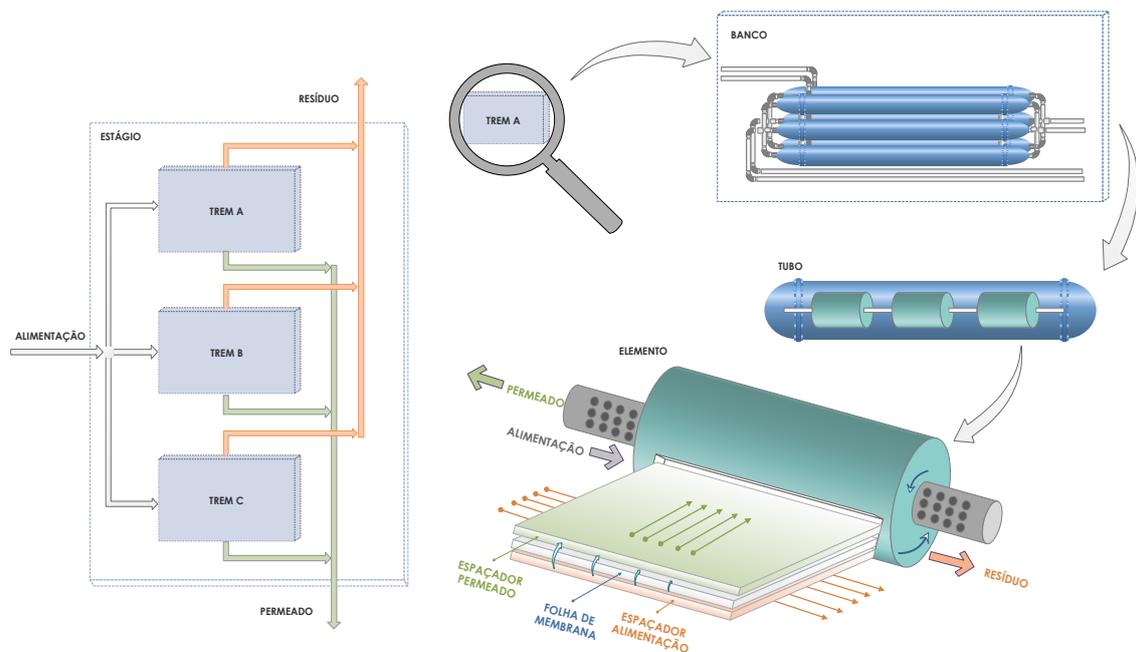


Figura 5.22: Representação esquemática de uma unidade com folhas de membrana em espirais.

Fonte: DE MENEZES *et al.* (2020).

Atualmente, poucos artigos estão de alguma forma envolvidos com a reconciliação de dados em processos de separação por membrana. No entanto, o trabalho desenvolvido por LASHKARI e KRUCZEK (2012) mostrou que é possível “reconciliar” as propriedades do processo de separação por membrana, utilizando dados relativos ao tempo morto do processo, que é afetado pela resistência que a membrana oferece ao fluxo de gás. Na verdade, este trabalho desenvolveu procedimentos para estimar permeabilidades efetivas e coeficientes de difusão, permitindo a modelagem de efeitos de resistência dependentes da posição, usando uma unidade em escala de laboratório como exemplo.

CRIVELLARI (2016) propôs um modelo para simular a separação do CO_2 do gás natural através de uma membrana polimérica espiral contra-corrente. O modelo foi utilizado para analisar a influência de diferentes variáveis nas condições de operação do processo e foi baseado em equações de equilíbrio fenomenológico.

O modelo foi validado com dados coletados na literatura e em algumas unidades industriais. Além disso, um procedimento de DR foi usado para tratar os dados disponíveis e estimar os parâmetros do modelo; entretanto, o estudo foi implementado offline e não permitiu nenhum tipo de monitoramento em tempo real da operação do processo.

5.2.3 Metodologia

A metodologia implementada no presente trabalho compreende seis etapas: aquisição de dados; pré-tratamento de dados; caracterização de dados; reconciliação de dados; detecção de erros grosseiros e monitoramento de processo (sensor virtual ou gêmeo digital). As primeiras etapas do procedimento implementado envolveram o pré-tratamento e a caracterização dos dados. De fato, o correto entendimento de algumas características dos dados é fundamental para a implementação adequada da etapa de reconciliação de dados (FEITAL *et al.*, 2014). A caracterização inicial dos dados foi realizada offline e utilizando dados históricos disponíveis no sistema de aquisição de dados da unidade industrial. Os dados disponíveis foram usados para determinar os períodos de amostragem apropriados (com base nos tempos de resposta do processo) e calcular as variâncias de medição (usadas para formular o problema de estimativa) e correlações de variáveis (para caracterizar a independência dos dispositivos de medição). A classificação das variáveis também foi realizada para determinar os conjuntos de variáveis observáveis e não observáveis (com o auxílio do modelo proposto, conforme descrito a seguir) (STANLEY e MAH, 1981b).

Usando os dados disponíveis e as equações do modelo, o procedimento de RD foi resolvido *offline* para validar o procedimento proposto e determinar alguns índices de desempenho. Particularmente, uma métrica estatística foi usada para descrever as magnitudes dos desvios entre as variáveis medidas e reconciliadas. Em seguida, o modelo foi usado *offline* para cálculo de variáveis não medidas, fornecendo a resposta do sensor virtual (ou gêmeo digital). Por fim, os procedimentos propostos foram implementados *online* e em tempo real. Os procedimentos e códigos numéricos foram desenvolvidos e implementados em Python 3.7.6 e os detalhes da metodologia desenvolvida são explicados no Apêndice A.

5.3 Comentários Finais

Com base no que foi discutido, a metodologia desenvolvida apresentou propostas que podem solucionar com um bom desempenho o problema de RDDR. As

etapas da metodologia abordaram: o sistema de aquisição de dados, admitindo que as etapas de pré-filtragem foram tratadas *a priori*, trabalhando apenas com os dados gerados no próprio simulador de RDDR; a escolha da frequência de amostragem discutiu e apresentou uma abordagem para este tratamento; a modelagem do processo, que admitiu, para este presente trabalho, apenas equações de balanços de massas, equilíbrio químico, ponto de bolha e soma das concentrações; a sintonização dos estimadores-M robustos, que foi feita *a priori* utilizando o método de Eficiência Relativa; as estratégias para acelerar o procedimento de otimização, que incluiu cláusulas para seleção das estimativas iniciais das variáveis de otimização, dos limites adaptativos das variáveis de otimização, das condições iniciais do modelo dinâmico; a similaridade temporal, que incluiu um novo termo na Fobj relacionado ao modelo preditivo Autoregressivo, aumentando a robustez e redundância das variáveis reconciliadas de entrada, sendo o termo ativado por cláusulas que detectam um estado estacionário dos dados dentro da janela móvel. Assim, a metodologia apresentada com características inovadoras satisfaz um dos objetivos da tese, que é o desenvolvimento de uma metodologia de RD para monitoramento de processos de separação (dinâmica ou estacionária) em tempo real. Outro fato, foi a dificuldades para tratar erros grosseiros nas variáveis de entrada em modelos dinâmicos. Essas dificuldades foram contornadas com a aplicação do termo de similaridade, provendo robustez para eliminar erros grosseiros, aumentando a redundância e grau de liberdade estatístico nas variáveis de entrada. Dessa forma, o novo termo busca intensificar a similaridade temporal dos dados, filtrando ruídos e erros grosseiros das variáveis de entrada do processo. Além disso, o monitoramento em tempo real de sistema de separação de CO_2 por membrana foi implementada pela primeira vez. Com isso, a metodologia de caracterização dos dados, RD e métricas de análises de erros grosseiros atenderam um dos objetivos da tese, que é o desenvolvimento de um aplicativo-*web* para monitoramento em tempo real do processo de separação por membrana, baseado em RD e métricas para DMEG.

Capítulo 6

Resultados

“Não há problema em não saber todas as respostas. É melhor admitir nossa ignorância do que acreditar em respostas que podem estar erradas. Fingir que sabe tudo fecha a porta para a descoberta do que realmente estava lá.”

Neil deGrasse Tyson (1958 –)

Este capítulo tem como objetivo apresentar os estudos de casos selecionados da literatura e os respectivos resultados. Os problemas foram implementados nos pacotes computacionais EMSO e Matlab. Particularmente, serão apresentados os modelos matemáticos que constituem os estudos de casos, suas variáveis (medidas e não medidas), suas variâncias, os parâmetros do processo, estimadores (Fobj), métodos de otimização utilizados e os parâmetros de convergência. Serão apresentadas também as variáveis simuladas (exatas), medidas (erros aleatórios) e também aquelas corrompidas por erros grosseiros, bem como as respectivas magnitudes. Por fim, para o estudo de caso, serão apresentados os objetivos específicos, caracterizando a metodologia aplicada nesta tese.

6.1 Apresentação dos estudos de caso

6.1.1 Caso 1: Avaliação dos métodos de otimização

KRISHNAMURTHY e TAYLOR (1985) compararam dados medidos em uma coluna de destilação com recheio real com os dados previstos por modelo

de não equilíbrio. O modelo de não equilíbrio soluciona, simultaneamente, os balanços materiais e os balanços energéticos para cada estágio, com as equações de transferência de massa e transferência de calor, e também as equações de equilíbrio de interfaces, conhecido como equações MERQ (*component Mass balance, Energy balance, Rate equations and interface eQ* *Equilibrium relations*). Os autores utilizaram dados medidos nas variáveis independentes; ou seja, simulando erros para iniciar a simulação. Assim, os valores simulados para as variáveis dependentes (estados) propaga estes erros. Neste caso, a RD juntamente com modelo rigoroso de coluna foi necessário para prever variáveis não medidas e corrigir qualquer erro de medição das variáveis medidas (tanto de entrada e de saída de simulação).

Os dados da coluna real foram adquiridos por medições experimentais das vazões de alimentação F , composição global da alimentação z_F , razão de refluxo RD e vazão de destilado D . A composição da alimentação foi calculada por balanço material com os produtos de topo e fundo e comparado com os valores experimentais. A coluna de recheio consiste de duas seções, com uma alimentação entre as seções, um condensador total e um *reboiler*. A alimentação consiste de uma mistura de 11 hidrocarbonetos, dentre eles alcanos e seus isômeros. A Tabela 6.1 apresenta as configurações da coluna de recheio e as especificações da simulação.

Tabela 6.1: Configuração e especificações da coluna de recheio (KRISHNAMURTHY e TAYLOR, 1985).

Altura de um recheio (seção superior = seção inferior)	5,1816 m (total de 10,3642 m)
Diâmetro do recheio (seção superior)	0,9144 m
Diâmetro do recheio (seção inferior)	1,2192 m
Tipo de recheio	Aleatório com anéis Pall metálicos (2 in)
Pressão da coluna P	1,1235 MPa
Razão de Refluxo RD	1,19
Vazão total da alimentação F	0,05113 kmol/s
Temperatura da alimentação T_F	455,22 K
Temperatura de sub-resfriamento do líquido de refluxo ΔT_{sub}	15,60 K
Vazão total do fundo B	0,031734 kmol/s

A Tabela 6.2 apresenta a composição da alimentação, utilizada pelos autores para a simulação.

Tabela 6.2: Composição da alimentação F (KRISHNAMURTHY e TAYLOR, 1985).

Componente	Vazão [kmol/s $\times 10^4$]	Componente	Vazão [kmol/s $\times 10^4$]
etano (C_2)	1,0505	hexano (C_6)	8,7775
propano ($i-C_3$)	98,9570	heptano (C_7)	41,3300
isobutano ($i-C_4$)	39,2580	octano (C_8)	114,1480
n -butano ($n-C_4$)	30,8690	nonano (C_9)	86,6360
isopentano ($i-C_5$)	10,2710	decano (C_{10})	74,0520
n -pentano ($n-C_5$)	5,9489		

A Tabela 6.3 compara os dados experimentais com os resultados da simulação. Segundo os autores, os dados experimentais de 4 componentes foram imprecisos; portanto, não foram relatadas as composições das vazões de destilado e de fundo para os componentes etano, propano, nonano e decano.

Tabela 6.3: Resultado da simulação (KRISHNAMURTHY e TAYLOR, 1985).

Componente	Vazão por componente do destilado $D \cdot x_{Di}$ [kmol/s $\times 10^4$]		Vazão por componente do fundo $B \cdot x_{Bi}$ [kmol/s $\times 10^5$]	
	Exp.	Sim.	Exp.	Sim.
etano (C_2)	–	1,0501	–	0,0041
propano ($i-C_3$)	–	98,7530	–	2,0562
isobutano ($i-C_4$)	39,1780	38,9280	0,7903	3,2944
n -butano ($n-C_4$)	30,7800	30,5455	0,7300	3,2478
isopentano ($i-C_5$)	9,9031	9,6769	3,6781	5,9381
n -pentano ($n-C_5$)	5,5578	5,5150	3,9125	4,3400
hexano (C_6)	4,8984	5,0658	38,7910	37,117
heptano (C_7)	3,3464	3,8067	379,8300	375,2200
octano (C_8)	0,3153	0,5902	1138,3300	1135,5800
nonano (C_9)	–	0,0236	–	866,1400
decano (C_{10})	–	0,0062	–	740,4400

Exp.=dados experimentais; **Sim.**=dados simulados;

Assim, com base nas Tabelas 6.1, 6.2 e 6.3, as restrições do problema são

descritas com modelo dinâmico, admitindo que o volume acumulado de líquido em cada estágio seja constante, conhecido como *constant liquid holdup*. O modelo foi inserido no pacote computacional EMSO e resolvido pelo integrador *dasslc* já pré-programado no EMSO. O resultado simulado com o modelo no EMSO é comparado com o simulado pelo trabalho de KRISHNAMURTHY e TAYLOR (1985). Além disso, utilizam-se as variáveis simuladas (consideradas exatas) para gerar erros aleatórios e erros grosseiros sobre elas, para avaliação do problema de RD e o desempenho dos otimizadores.

O modelo completo descrito no EMSO é constituído de:

- Números de variáveis: 1121;
- Número de equações: 961;
- Número de especificações: 160;
- Graus de liberdade: 0;
- Índice da estrutura diferencial: 1;
- Graus de liberdade dinâmicos: 132.

Os autores relataram também dados experimentais das temperaturas ao longo da coluna. Entretanto, os dados estão ilustrados em figura de difícil visualização, podendo causar erros maiores ao obter os valores com observações ao olho nu. Assim, fez-se necessário a utilização de um *software* “pega ponto”, ou seja, um programa que adquire os dados de um gráfico ou uma figura, o McLabEn (Matemática Computacional: Laboratório em Engenharia e Ciências) versão 1.33. A Figura 6.1 ilustra o gráfico com o perfil de temperatura experimental do trabalho de KRISHNAMURTHY e TAYLOR (1985) e o perfil de temperatura simulado com o *software* EMSO, incluindo as temperaturas do refeedor e condensador, indicando que o modelo apresenta boa acurácia no prato de alimentação, no condensador e no refeedor. A discrepância das temperaturas nos demais pratos poderia ser corrigida por uma estimativa adequada da constante de eficiência dos pratos; porém, o parâmetro utilizado para a eficiência foi de 90%, um valor muito alto de eficiência. Esse foi o menor valor com que se conseguiu simular, pois a simulação não convergia com valores menores que 90%. Outro fator, que pode ter impactado na convergência, pode ter sido no número de estágios. O artigo relatou apenas 10 pontos de temperatura ao longo da coluna, por isso adotou o número de 10 estágios.

Um dado importante para RD de destilação com modelos rigorosos é a temperatura. Assim, faz-se necessária a medida de temperatura em cada estágio de

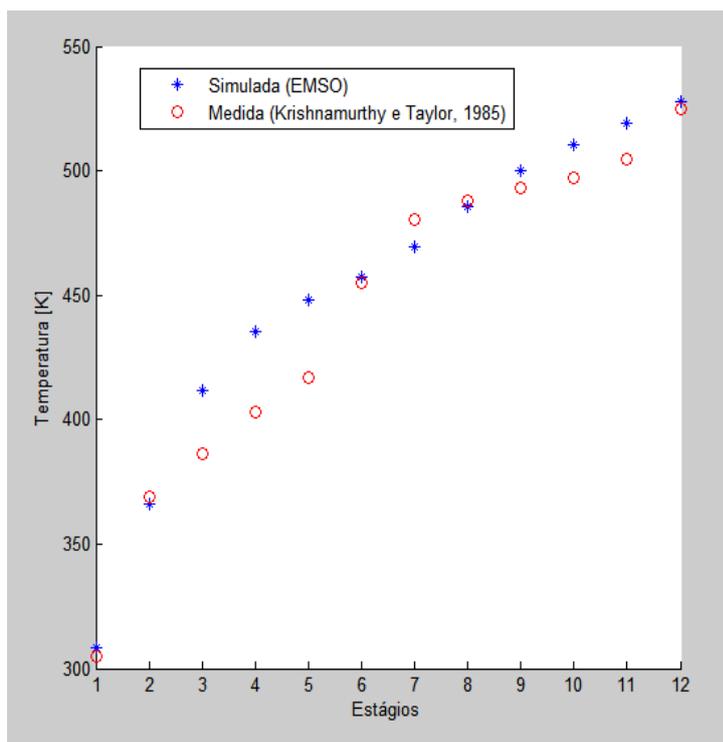


Figura 6.1: Comparação dos perfis de temperatura dos dados experimentais e do simulado (KRISHNAMURTHY e TAYLOR, 1985).

separação. Outro fator importante é que inferir temperatura é mais fácil, mais rápido e mais econômico que inferir composição do líquido ou do vapor. Além disso, apenas inferindo a temperatura em cada estágio é suficiente para resolver um problema de RD em colunas com modelo rigoroso, admitindo-se a pressão constante na coluna e conhecendo a relação de equilíbrio líquido-vapor (FARZI *et al.*, 2008). A Tabela 6.4 apresenta os dados das temperaturas adquiridas com o software McLabEn.

Assim, o problema de RD passou a conter mais 12 variáveis medidas, formando um total de 48 variáveis medidas (simuladas e consideradas como exatas):

- 3 vazões (F , D e B);
- 33 componentes (11 em F , 11 em D e 11 em B);
- 12 temperaturas (12 estágios de equilíbrio).

O problema pode ser interpretado por colunas de pratos, de forma que o modelo de equações MES (MESH sem balanço de energia) pode ser usado. Além disso, o modelo MERQ utilizado no trabalho de KRISHNAMURTHY e TAYLOR (1985) não apresentou significativa acurácia em relação ao MES, embora o modelo MES seja computacionalmente mais rápido. A alimentação em colunas de pratos

Tabela 6.4: Perfil de temperatura da coluna (KRISHNAMURTHY e TAYLOR, 1985).

Estágios	Temperatura [K]
0	305,000
1	368,850
2	386,230
3	402,941
4	416,979
5	455,220
6	480,481
7	487,834
8	492,848
9	497,193
10	504,546
11	525,000

é realizada no quinto estágio de cima para baixo, conforme ilustra a Figura 6.1. Com isso, o problema de RD em colunas pode ser resolvido com modelo MES e pacotes termodinâmicos adequados utilizando o *software* EMSO + VRTherm (pacote termodinâmico externo do EMSO) (SOARES e SECCHI, 2003) em conjunto com o software Matlab. Assim, tem-se por objetivo comparar modelo MES com diferentes Funções Objetivos (estimadores-M) no problema de RD com os dados experimentais e simulados por KRISHNAMURTHY e TAYLOR (1985).

Uma avaliação importante também é analisar como o modelo MES se comporta com erros grosseiros gerados propositalmente nas variáveis de saída, entrada da coluna e na temperatura, e com isso avaliar o melhor estimador. Foram usados os estimadores MQP, Normal Contaminada e Welsch. Nesse caso, foram avaliados também os métodos de otimização Nelder-Mead (rotina *fminsearch*), Enxame de Partículas (PSO, rotina *swarm*), quasi-Newton (BFGS, rotina *fminunc*), SQP e Ponto Interior (rotina *fmincon*). Portanto, uma análise ampla dos métodos é apresentada, levando-se os aspectos: determinístico/não determinístico; restrita/irrestrita e busca local/busca global. O caso foi resolvido nos *softwares* computacionais EMSO e Matlab. Os parâmetros de convergência (tolerâncias absolutas) e de desempenho utilizados nos métodos de otimização foram:

- **Neader-Mead:** 'Tolerância em X' = 1e-4, 'Tolerância no valor da Fobj' = 1e-2, 'Máximo de Iterações' = 500;
- **PSO:** 'Número de gerações' = 8, 'Número de partículas' = 100, 'Parâmetro de pesquisa local' = 1, 'Parâmetro de pesquisa global' = 1;
- **BFGS:** 'Tolerância em X' = 1e-5, 'Tolerância no valor da Fobj' = 1e-4, 'Máximo de Iterações' = 500;
- **SQP:** 'Tolerância em X' = 1e-5, 'Tolerância no valor da Fobj' = 1e-4, 'Máximo de Iterações' = 500;
- **Ponto Interior:** 'Tolerância em X' = 1e-5, 'Tolerância no valor da Fobj' = 1e-4, 'Máximo de Iterações' = 500.

Foram gerado erros grosseiros em 3 variáveis, conforme apresenta a Tabela 6.5. As variáveis exatas (entrada e saída) são apresentadas na Tabela 6.6. Neste caso não havia variável não medida, considerando-se que todas as variáveis foram medidas. As variáveis corrompidas (entrada e saída) são apresentadas na Tabela 6.7. A Tabela 6.1 apresenta os parâmetros utilizados para gerar as variáveis exatas, e que foram mantidos na RD. Consideraram-se como variáveis fixas a razão de refluxo ($RD = 1,19$) e a temperatura de sub-resfriamento do líquido de refluxo ($\Delta T_{sub} = 15,60$ K) (KRISHNAMURTHY e TAYLOR, 1985). Utilizou-se o simulador EMSO para calcular os valores exatos do problema proposto e geraram-se erros aleatórios pequenos ($\pm 1,96\sigma$) para simular variáveis medidas com base nos respectivos desvios padrões assumidos na literatura, conforme apresentado na Tabela 6.8, que resume a metodologia.

Tabela 6.5: Variáveis corrompidas para simular erros grosseiros no estudo de caso 1.

Variável	Sem Erro Grosseiro	Com Erro Grosseiro	Magnitude
Alimentação F	184,0680 kmol/h	220,8816 kmol/h	$10 \cdot \sigma$
Isobutano em D (x_{D3})	$39,1217 \cdot 10^{-4}$ kmol/s	$46,9461 \cdot 10^{-4}$ kmol/s	$20 \cdot \sigma$
Temperatura do Estágio 2	411,89 K	396,89 K	$-20 \cdot \sigma$

Tabela 6.6: Valores exatos para o estudo de caso 1.

	F [kmol/h]	D [kmol/h]	B [kmol/h]	Temp. [K]	Estágio
	184,0680	69,8256	114,2424	T_i	i
	$\cdot 10^{-4}$ [kmol/s]	$\cdot 10^{-4}$ [kmol/s]	$\cdot 10^{-5}$ [kmol/s]	308,44	0
C_2	1,0737	1,0474	0,0000	366,34	1
C_3	98,9366	98,9584	0,0000	411,89	2
$i-C_4$	39,2678	39,1217	1,2694	435,09	3
$n-C_4$	30,8825	30,6263	2,5387	447,56	4
$i-C_5$	10,2771	9,4071	8,5682	457,00	5
$n-C_5$	5,9311	5,1787	7,6162	469,70	6
C_6	8,7944	4,2089	45,6970	485,93	7
C_7	41,3130	4,4029	369,3838	500,42	8
C_8	114,1733	0,9504	1131,9518	511,03	9
C_9	86,6142	0,0388	865,7035	519,35	10
C_{10}	74,0362	0,0000	740,3542	527,98	11

Tabela 6.7: Valores corrompidos para o estudo de caso 1.

	F [kmol/h]	D [kmol/h]	B [kmol/h]	Temp. [K]	Estágio
	220,8816*	70,9685	112,8579	T_i	i
	$\cdot 10^{-4}$ [kmol/s]	$\cdot 10^{-4}$ [kmol/s]	$\cdot 10^{-5}$ [kmol/s]	307,47	0
C_2	1,0878	1,0607	0,0000	367,52	1
C_3	100,8490	100,5740	0,0000	396,89*	2
$i-C_4$	39,7448	46,9461*	1,2547	434,09	3
$n-C_4$	30,4771	31,2138	2,5818	446,44	4
$i-C_5$	10,1372	9,3089	8,7260	458,13	5
$n-C_5$	5,8544	5,1020	7,5091	468,62	6
C_6	8,9321	4,1557	45,0529	486,88	7
C_7	40,7005	4,3406	376,3223	499,35	8
C_8	112,6523	0,9626	1119,3437	512,16	9
C_9	87,9944	0,0384	855,3165	520,79	10
C_{10}	75,1479	0,0000	751,9176	529,29	11

*erro grosseiro

A metodologia utilizada neste trabalho consistiu, basicamente, na implementação do problema de RD e RD robusta (RRD) no pacote computacional EMSO e no *softwares* Matlab (conjuntamente), para avaliação do problema de RD, do modelo e do desempenho dos métodos de otimização. A Tabela 6.8 resume a metodologia aplicada ao caso em maiores detalhes.

Tabela 6.8: Resumo da metodologia para o caso 1 KRISHNAMURTHY e TAYLOR (1985).

Característica/Item	Descrição
Processo:	Destilação de hidrocarbonetos leves.
Característica do processo:	Descrito pelas Tabelas 3.2, 3.3 e 3.4.
Operação:	Modelo dinâmico operando em regime permanente.
Restrições:	Não lineares.
Número de variáveis de otimização (entrada):	12 $[2 + nc - 1]$
Número de variáveis medidas:	48
Limites nas variáveis:	Sim (todas as variáveis devem ser não negativas).
Modelo matemático:	Equações MES dinâmico.
Valores exatos (livre de erros):	Foram gerados para o caso.
Dados com erros grosseiros:	Foram gerados para o caso.
	<ul style="list-style-type: none"> • Vazão: 2% do valor medido (SWARTZ, 1989);
Desvios padrão:	<ul style="list-style-type: none"> • Composição: 1% do valor medido (RAO e NARASIMHAN, 1996); • Temperatura: 0,75°C (SWARTZ, 1989).
<i>Softwares:</i>	EMSO + Matlab.
Otimizadores:	Nelder-Mead, Enxame de Partículas (PSO), quasi-Newton (BFGS), SQP e Ponto Interior
Objetivo:	É avaliado o desempenho do modelo MES, o desempenho dos estimadores MQP, Normal Contaminada e Welsch para DMEG com base em dados simulados e avaliar o desempenho dos otimizadores (Método de Nelder-Mead, Enxame de Partículas (PSO), quasi-Newton (BFGS), SQP e Ponto Interior).
Critério de avaliação dos estimadores:	Critério TER (Total Error Reduction).
Critério de avaliação dos otimizadores:	Menor Fobj e menor tempo.

6.1.2 Caso 2: Avaliação da redução de ordem do modelo de coluna

O segundo estudo de caso para este trabalho apresenta um sistema binário propeno-propano (*C₃-splitter*) com uma variação no número de estágios (177, 101 e 51), que conseqüentemente, mude o tamanho do sistema de equações. O caso foi descrito por SEFERLIS e HRYMAK (1994) como uma coluna de 177 estágios. SIMÕES (2000) utilizou como referência o estudo de SEFERLIS e HRYMAK (1994) para aplicar uma estratégia de redução de ordem de modelos de separação por estágio baseada em colocação ortogonal. Um estudo seguindo esta mesma linha e com o mesmo caso foi feito por RIBEIRO *et al.* (2010), visando a comparar a redução do custo computacional e acurácia entre os métodos de colocação ortogonal e dos resíduos ponderados pelos momentos. A Tabela 6.9 ilustra os parâmetros específicos para o sistema de separação propeno-propano.

Tabela 6.9: Configuração e especificações da coluna de destilação propeno-propano RIBEIRO (2011).

Número de estágios	22 - 51 - 101 - 177
Componentes	Propeno - Propano
Pressão da coluna P	18,3627 atm
Razão de Refluxo RD	19,7
Vazão total da alimentação F	44,725 kmol/h
Temperatura da alimentação T_F	319,26 K
Composição da alimentação \underline{z}	$[0, 89730, 1027]^T$
Vazão total do destilado D	40,2083 kmol/h
Modelo termodinâmico (EoS)	Peng-Robinson

A fim de criar um cenário fictício de dados medidos (corrompidos por erros aleatórios e erros grosseiros), o problema foi resolvido inicialmente pelo modelo completo. Dentre as possibilidades de números de estágios de separação, escolheu-se o modelo com 51 estágios, pois o modelo com 22 estágios não apresentou relevante diferença nos tempos de integração. Já para os modelos de 101 e 177 estágios, suas contrapartes reduzidas não resultaram em um bom tempo de integração.

A Tabela 6.10 ilustra esta decisão de acordo com os tempos de simulação de cada modelo para uma iteração utilizando a interface EMSO-Matlab, o que gera uma perda de tempo computacional, e o tempo de uma iteração somente utilizando o *software* EMSO. A Tabela 6.10 também apresenta os tempos de simulação de uma

integração de 10800 segundos no simulador EMSO, sem a interface de integração dos *softwares* EMSO-Matlab. Os valores gerados pelo modelo completo e pelo modelo reduzido com 51 estágios estão apresentados na Tabela 6.11.

Tabela 6.10: Tempo gasto em segundos para uma iteração.

Estágios (pontos de colocação)	Modelo completo	Método dos momentos
Com interface Matlab-EMSO		
22 (3)	9,188 s	6,426 s
51 (3)	14,207 s	6,815 s
101 (7)	23,524 s	48,287 s
177 (8)	–	–
Sem interface Matlab-EMSO		
22 (3)	2,539 s	1,379 s
51 (3)	7,962 s	1,880 s
101 (7)	14,874 s	4,319 s
177 (8)	23,543 s	8,457 s

Utilizando os parâmetros iniciais para a simulação com base na Tabela 6.9 com modelo MES dinâmico (completo), considerou-se que o volume acumulado de líquido em cada estágio fosse constante, conhecido como *constant liquid holdup*. O modelo foi inserido no pacote computacional EMSO e resolvido pelo integrador *dasslc* já pré-programado no EMSO. Todos os modelos foram simulados com o tempo total de integração de 10800 segundos. O resultado simulado com o modelo completo foi comparado com o simulado pelo modelo reduzido pelo método dos momentos. Além disso, utilizaram-se as variáveis simuladas do modelo completo (consideradas exatas) para gerar erros aleatórios e erros grosseiros sobre elas, para avaliação do problema de RD e o desempenho dos estimadores e dos modelos, como ilustra a Tabela 6.11.

Tabela 6.11: Comparação entre os modelos completo e reduzido pelos momentos com 51 estágios no estudo de caso 2.

Variável	Modelo completo	Modelo momentos	Erro (%)
F [kmol/h]	44,7250	44,7250	0,00
D [kmol/h]	40,2083	40,2083	0,00
z_1	0,8973	0,8973	0,00
z_2	0,1027	0,1027	0,00
B [kmol/h]	4,5167	4,5167	0,00
$x_{D,1}$	0,9687	0,9504	1,89
$x_{D,2}$	0,0313	0,0496	58,35
$x_{B,1}$	0,2427	0,2493	2,75
$x_{B,2}$	0,7573	0,7507	0,88
T_1 [K]	318,7849	318,9270	0,04
T_5 [K]	318,9153	319,1208	0,06
T_{10} [K]	319,1273	319,3988	0,09
T_{15} [K]	319,4024	319,7165	0,10
T_{20} [K]	319,7448	320,0738	0,10
T_{25} [K]	320,1498	320,4708	0,10
T_{30} [K]	320,6011	320,9074	0,10
T_{35} [K]	321,2577	321,4914	0,07
T_{40} [K]	322,3292	322,5068	0,06
T_{45} [K]	323,4756	323,4971	0,01
T_{51} [K]	324,7094	324,6524	0,02

Tabela 6.12: Valores medidos (erro grosseiro em negrito), valores exatos e a magnitude dos erros no estudo de caso 2.

Variável	Valor exato	Valor medido	Magnitude
F [kmol/h]	44,7250	46,2328	-1,690
D [kmol/h]	40,2083	41,1145	-1,130
z_1	0,8973	0,8854	-2,000
z_2	0,1027	0,0981	1,270
B [kmol/h]	4,5167	4,6829	-1,840
$x_{D,1}$	0,9687	0,8523*	10,030
$x_{D,2}$	0,0313	0,0310	-1,310
$x_{B,1}$	0,2427	0,2377	-1,550
$x_{B,2}$	0,7573	0,7196	1,490
T_1 [K]	318,7849	320,0077	-1,630
T_5 [K]	318,9153	319,8464	-1,241
T_{10} [K]	319,1273	318,1802	1,263
T_{15} [K]	319,4024	318,4596	1,257
T_{20} [K]	319,7448	335,2229*	-20,637
T_{25} [K]	320,1498	321,5953	-1,927
T_{30} [K]	320,6011	319,6236	1,303
T_{35} [K]	321,2577	322,6052	-1,797
T_{40} [K]	322,3292	320,8906	1,918
T_{45} [K]	323,4756	324,3804	-1,206
T_{51} [K]	324,7094	323,6858	1,365

O modelo completo descrito no EMSO foi constituído de:

- Números de variáveis: 690;
- Número de equações: 628;
- Número de especificações: 62;
- Graus de liberdade: 0;
- Índice da estrutura diferencial: 1;
- Graus de liberdade dinâmicos: 102.

A Figura 6.2 ilustra os perfis (completo e reduzido) de temperatura ao longo da coluna, incluindo a temperatura no refeedor ($i = 51$) e condensador ($i = 1$). Foram utilizadas apenas 11 temperaturas igualmente espaçadas dentre os 51 estágios para simular as temperaturas “medidas” no problema de RD. Assim, o problema de RD passa a obter mais 11 variáveis medidas, formando um total de 20 variáveis medidas (simuladas e consideradas como exatas pelo modelo completo):

- 3 vazões (F , D e B);
- 6 componentes (2 na corrente F , 2 em D e 2 em B);
- 11 temperaturas ao longo da coluna.

O problema de RD em colunas pode ser resolvido com modelo MES (completo) e com o modelo reduzido pelo método dos momentos, utilizando o pacote termodinâmico adequado (VRTherm - pacote termodinâmico externo do EMSO) em conjunto com o *software* Matlab, em que foram criadas as rotinas de otimização. Assim, tem-se por objetivo comparar modelo MES completo com o modelo reduzido pela técnica dos resíduos ponderados pelos momentos. Além disso, foi comparado os desempenhos dos diferentes estimadores-M no problema de RD e DMEG: Mínimos Quadrados Ponderados, Normal Contaminada e o estimador de Welsch. O método de otimização utilizado foi o Ponto Interior, incluído nos pacotes de otimização do *software* Matlab. Os erros aleatórios pequenos foram gerados conforme descrito no Capítulo 5.

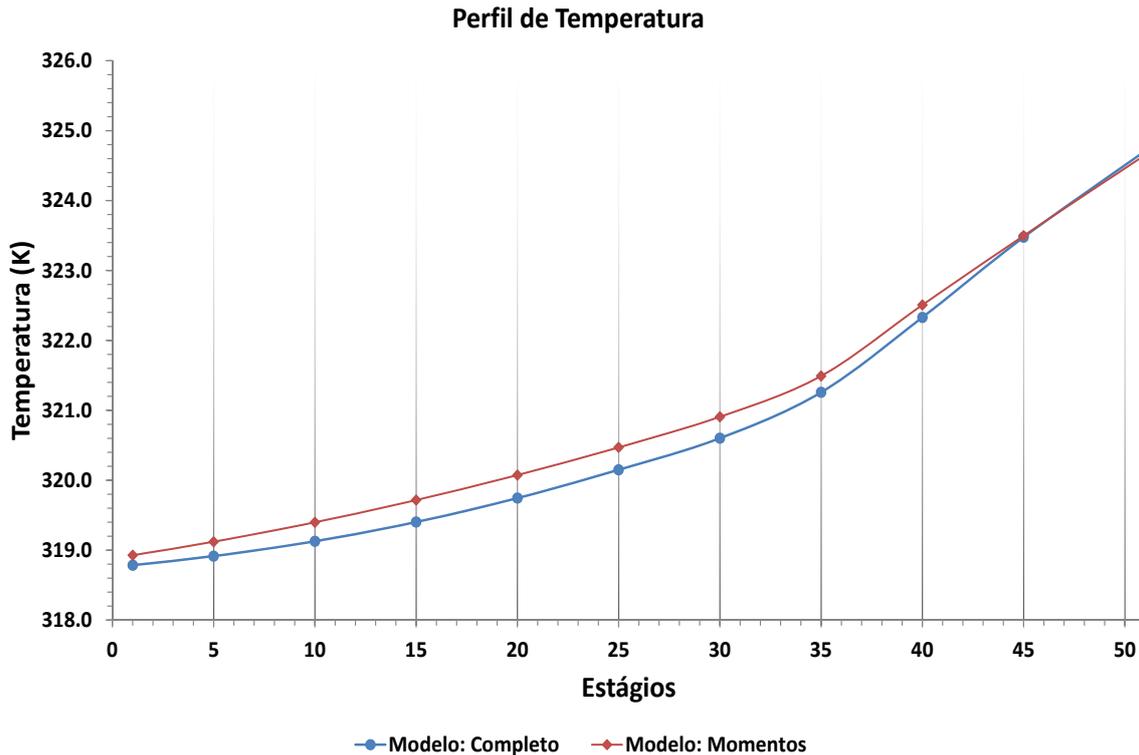


Figura 6.2: Perfis de temperatura dos modelos: completo e resíduo ponderado pelos momentos no estudo de caso 2.

6.1.3 Caso 3: Avaliação do termo de similaridade com modelo autoregressivo

O terceiro estudo de caso trata o problema da Reconciliação de Dados Dinâmica Robusta com modelo Autoregressivo de ordem 1. O objetivo é avaliar o comportamento da janela móvel quando a simulação apresenta perturbação na variável de entrada F . Para esta parte, os dados apresentaram erros grosseiros de baixa magnitude ($3\sigma \leq \xi \leq 6\sigma$), todas as variáveis foram medidas e o modelo de coluna utilizado foi o completo com 22 estágios, para poupar tempo de otimização.

Os dados exatos foram rodados primeiramente no *software* EMSO e armazenados no Matlab. Em seguida, os erros aleatórios foram gerados seguindo a distribuição Normal, com as respectivas variâncias e médias de cada variável. Um estudo de condicionamento matemático do problema foi feito *a priori*, com a finalidade de adequar os erros dentro das flutuações dinâmicas de cada variável. Uma nuvem de erros muito maiores que as flutuações não capta a dinâmica do modelo, causando mau-condicionamento matemático no problema (por exemplo: o erro de medição é maior que a oscilação e pequenas mudanças do processo). Então, as variâncias para este caso não seguiram um valor de referência na literatura.

O desvio-padrão, $\sigma = dp$, foi considerado constante dentro da janela e seguiu constante durante todo o problema:

- dp Vazão: 0,1 [kmol/h];
- dp Composição: 0,002;
- dp Temperatura: 0,025 [K].

A coluna usou os mesmos parâmetros descritos na Tabela 6.9. Inicialmente, utilizou-se como base para outros parâmetros da RDD o trabalho de NICHOLSON *et al.* (2014), como tamanho da janela H , número de pontos totais, tempo de amostragem e tempo de integração. Entretanto, para uma boa reconciliação, um estudo da frequência de amostragem ideal deve ser feito para cada modelo de coluna, pois a dinâmica afeta a escolha da frequência de amostragem. Para este caso, o problema de destilação se diferencia dos demais parâmetros da coluna do trabalho de NICHOLSON *et al.* (2014). A relação utilizada para encontrar a frequência de amostragem ideal está descrita detalhadamente no Capítulo 5.

Um problema que também foi observado é que cada variável responde com uma dinâmica diferente à perturbação de entrada. Durante o processo de escolha do tempo de amostragem, o resultado que pareceu mais adequado para a RDD foi com o $t_a = 15$ minutos, utilizando como referência a variável com a resposta mais lenta. Os resultados com os tempos $t_a = 1$ minuto e $t_a = 9$ minutos não apresentaram bom comportamento e foram descartados deste trabalho. Assim, os parâmetros utilizados e as principais descrições do problema de RDD são apresentados na Tabela 6.13. A Equação (6.1) foi utilizada para fazer o degrau suavizado na variável de entrada F :

$$y_{final} = y_{inicial} \pm \frac{\Delta y}{2} \left\{ 1 + \tanh \left[k_v \left(\frac{t}{t_d} - 1 \right) \right] \right\} \quad (6.1)$$

em que y_{final} e $y_{inicial}$ são os valores final e inicial da variável manipulada, Δy é a amplitude do degrau, k_v é a constante de velocidade do degrau (para valores muito altos de k_v , a equação equivale ao degrau brusco tradicional) e t_d é o instante de tempo da aplicação do degrau na variável. Os valores exatos e corrompidos por erros aleatórios são ilustrados nas Figuras 6.3 a 6.14.

Tabela 6.13: Descrição e parâmetros utilizados no problema de RDD.

Degrau em F	Suave utilizando tanh: $44,7250 \rightarrow 47,5583$ [kmol/h] e $t_d = 150$ minutos
Tamanho da janela H	10 pontos
Tempo de amostragem t_a	15 - 9 - 1 minutos
Tempo de integração $H \times t_a$	150 - 90 - 10 minutos
Número total de pontos	300
Passo da janela	De 1 em 1 ponto
Estimador-M	Normal Contaminada, $E_{ff} = 95\%$
Variáveis medidas (12)	$F, D, B, z, \underline{x}_D, \underline{x}_B, T_c, T_f$ e T_r
Método de otimização	Ponto-Interior

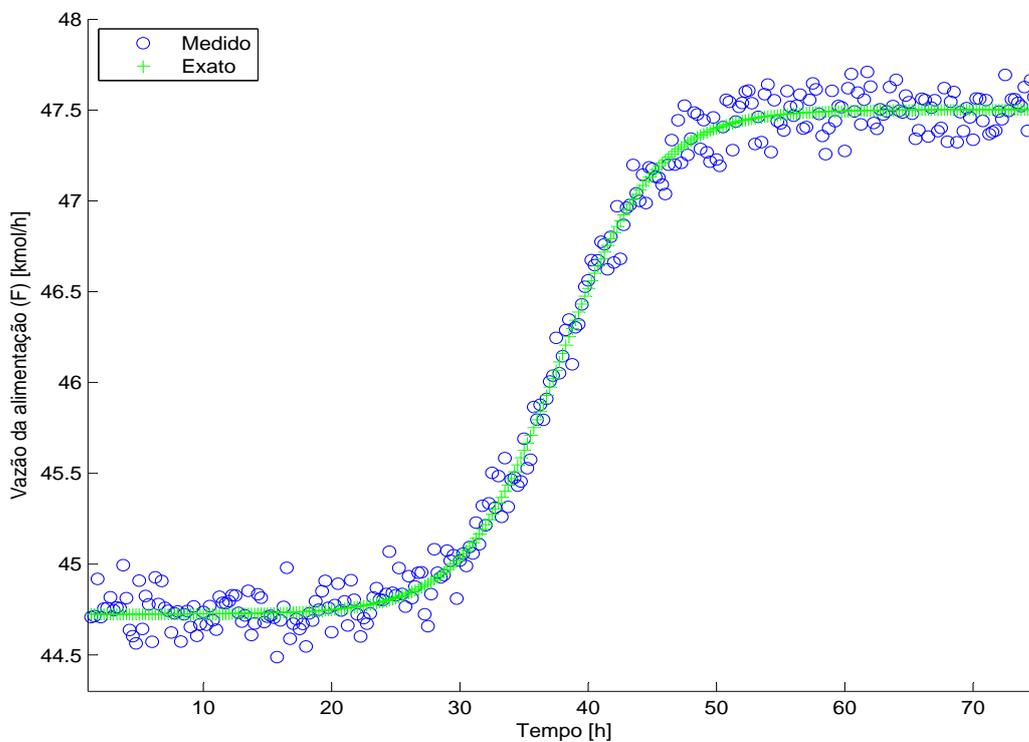


Figura 6.3: Valores exatos e medidos (corrompidos) - F no estudo de caso 3.

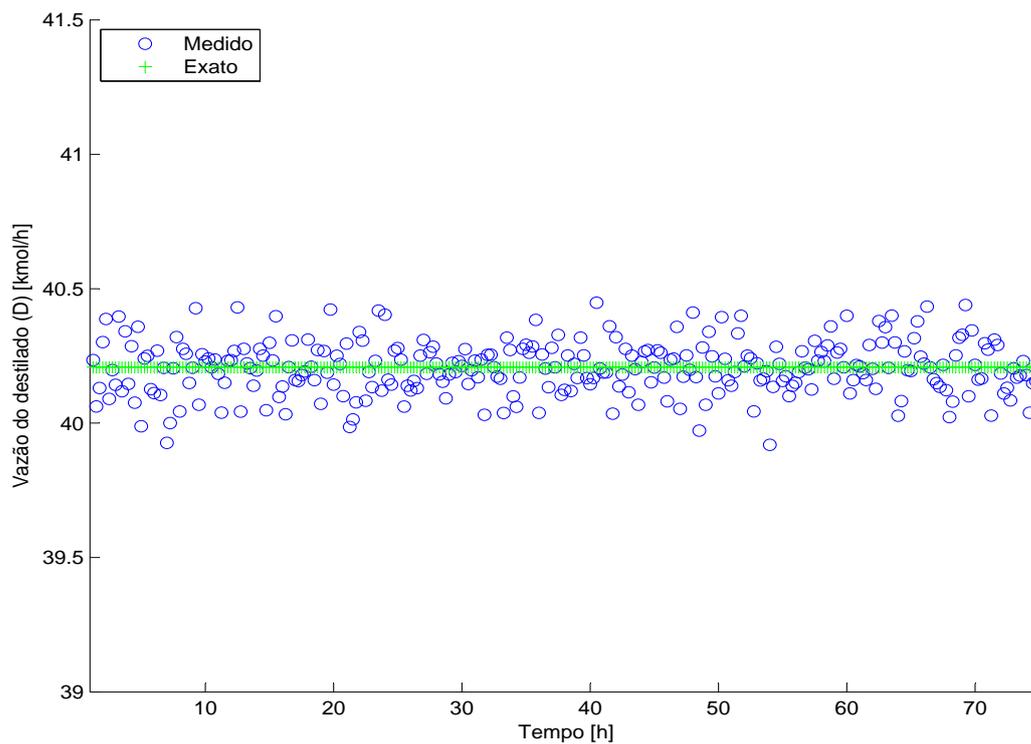


Figura 6.4: Valor exatos e medidos (corrompidos) - D no estudo de caso 3.

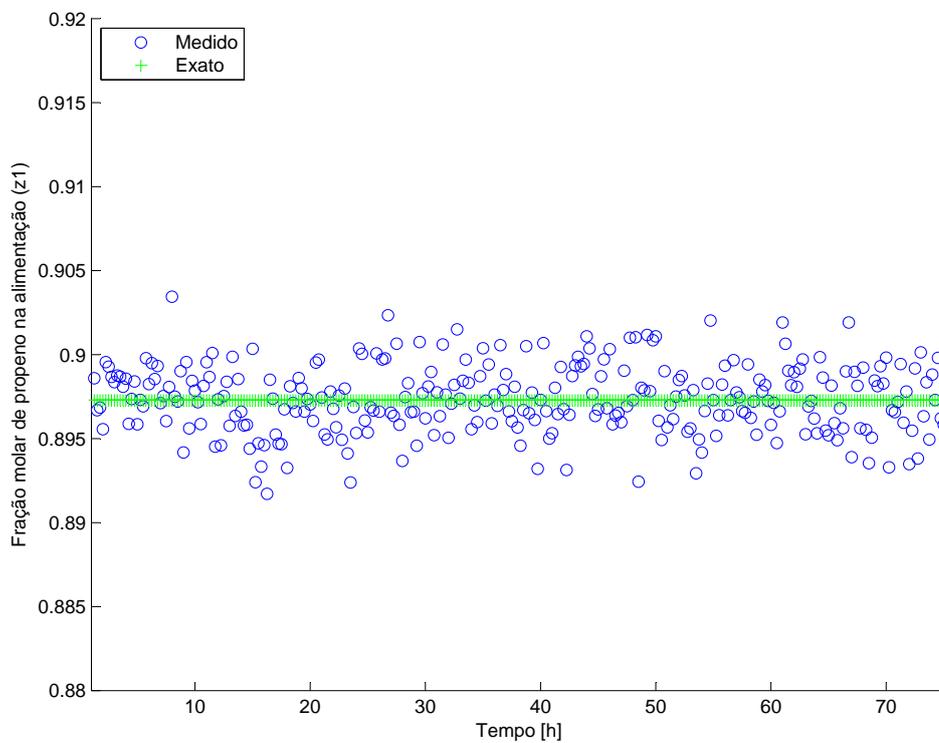


Figura 6.5: Valores exatos e medidos (corrompidos) - z_1 no estudo de caso 3.

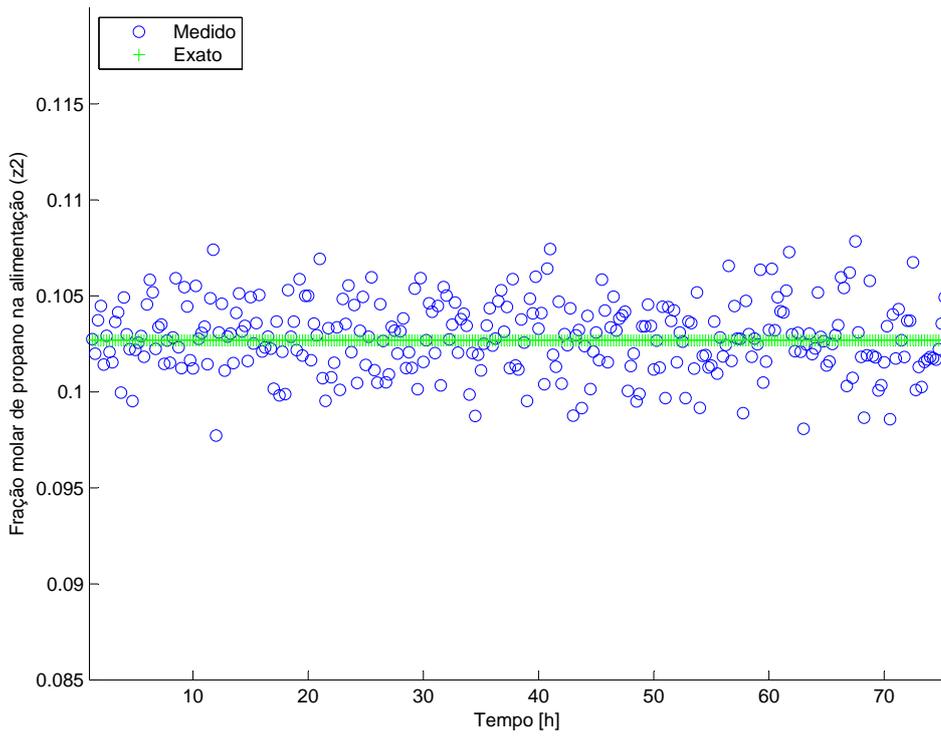


Figura 6.6: Valores exatos e medidos (corrompidos) - z_2 no estudo de caso 3.

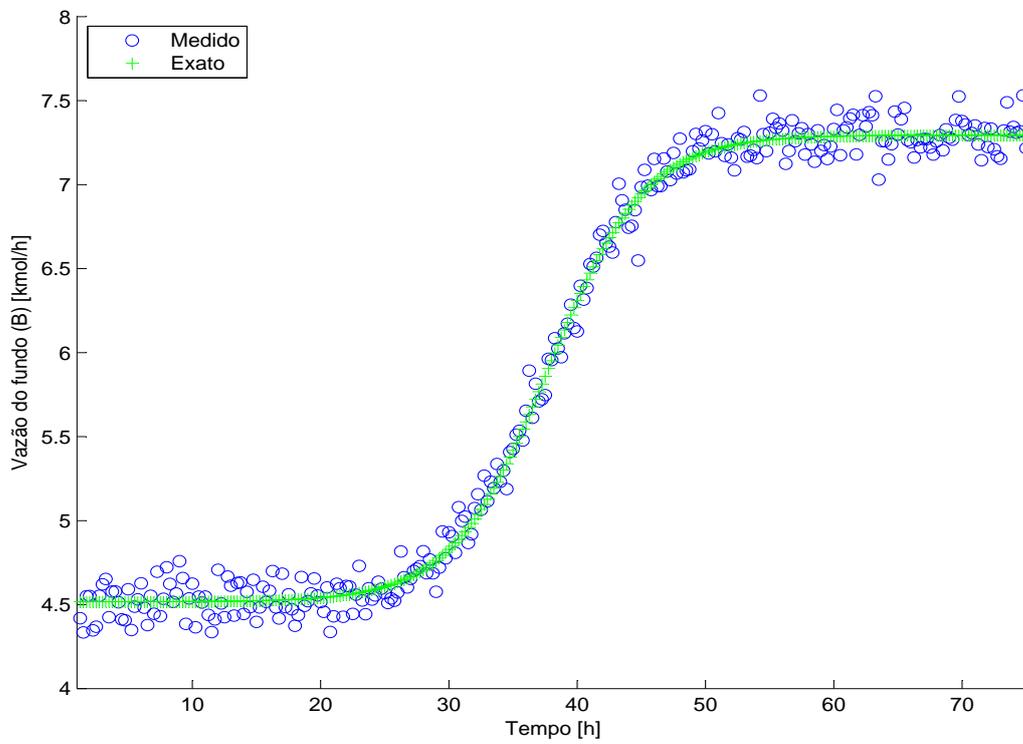


Figura 6.7: Valores exatos e medidos (corrompidos) - B no estudo de caso 3.

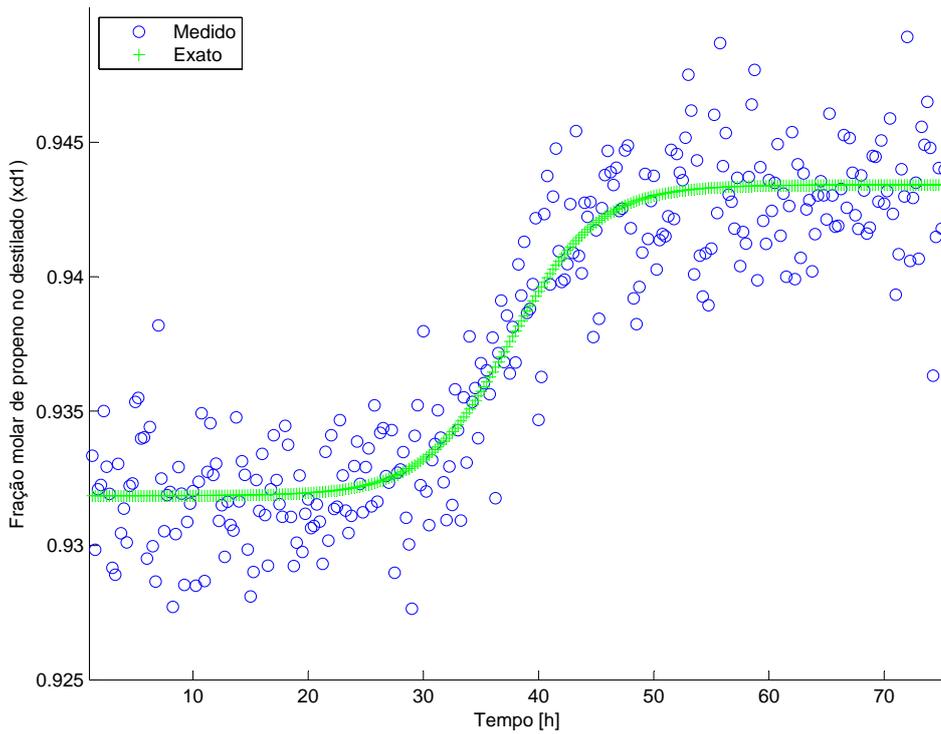


Figura 6.8: Valores exatos e medidos (corrompidos) - $x_{D,1}$ no estudo de caso 3.

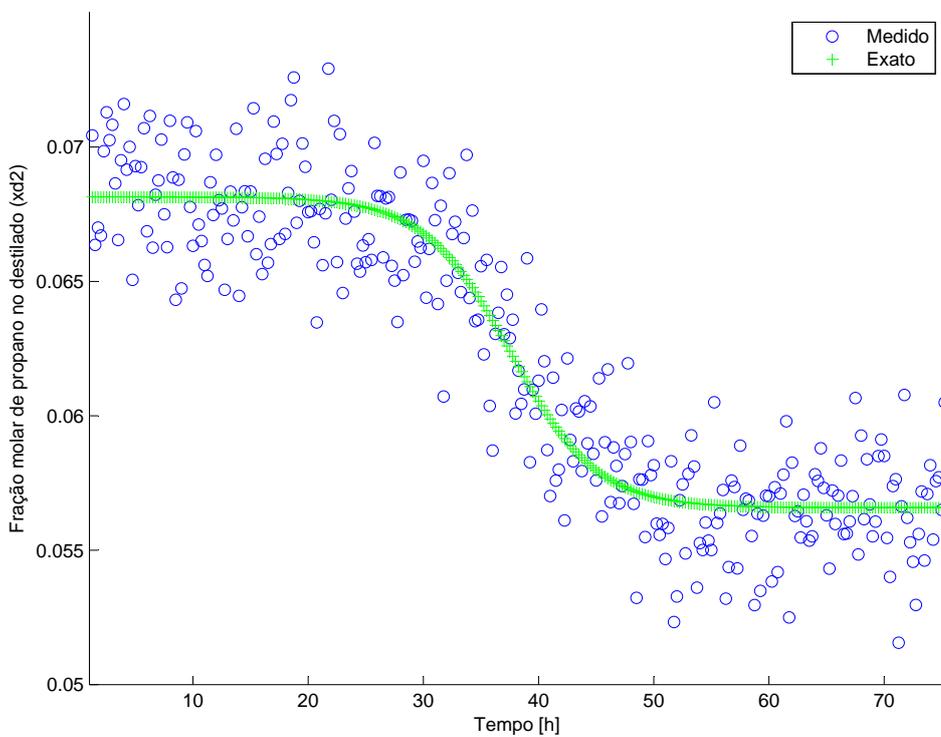


Figura 6.9: Valores exatos e medidos (corrompidos) - $x_{D,2}$ no estudo de caso 3.

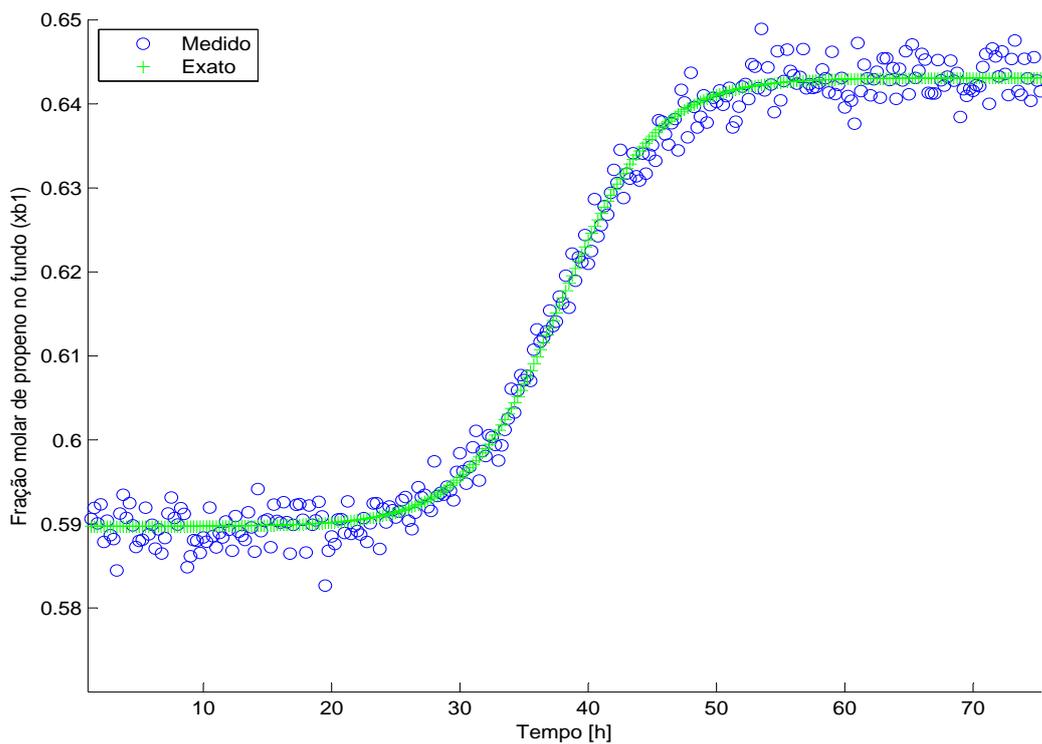


Figura 6.10: Valores exatos e medidos (corrompidos) - $x_{B,1}$ no estudo de caso 3.

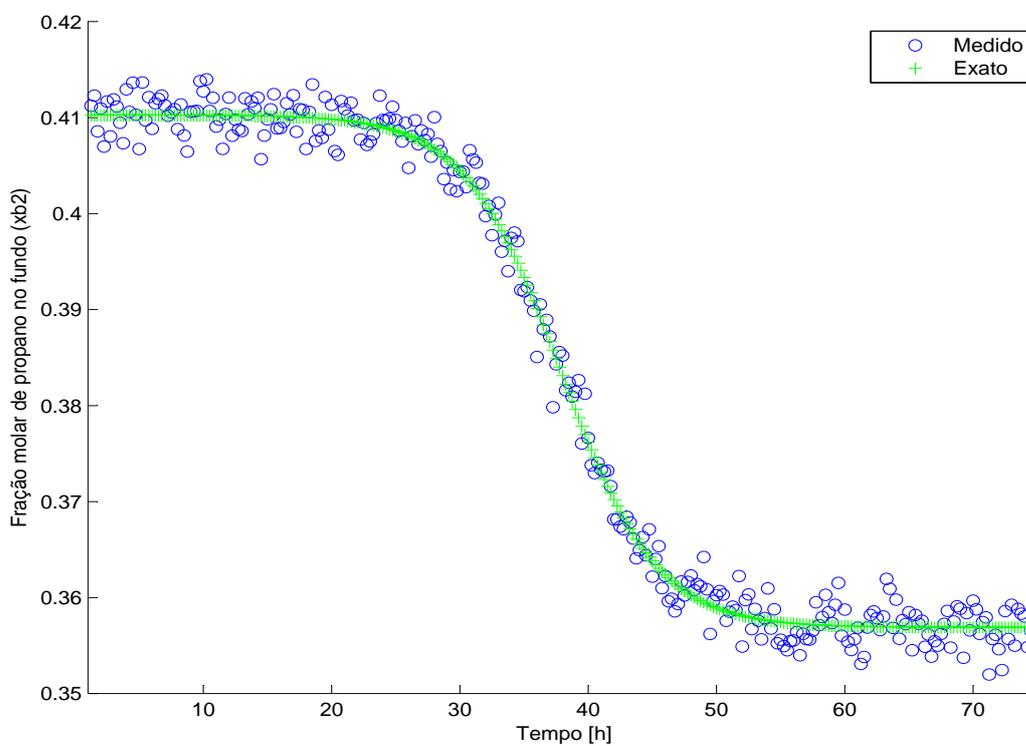


Figura 6.11: Valores exatos e medidos (corrompidos) - $x_{B,2}$ no estudo de caso 3.

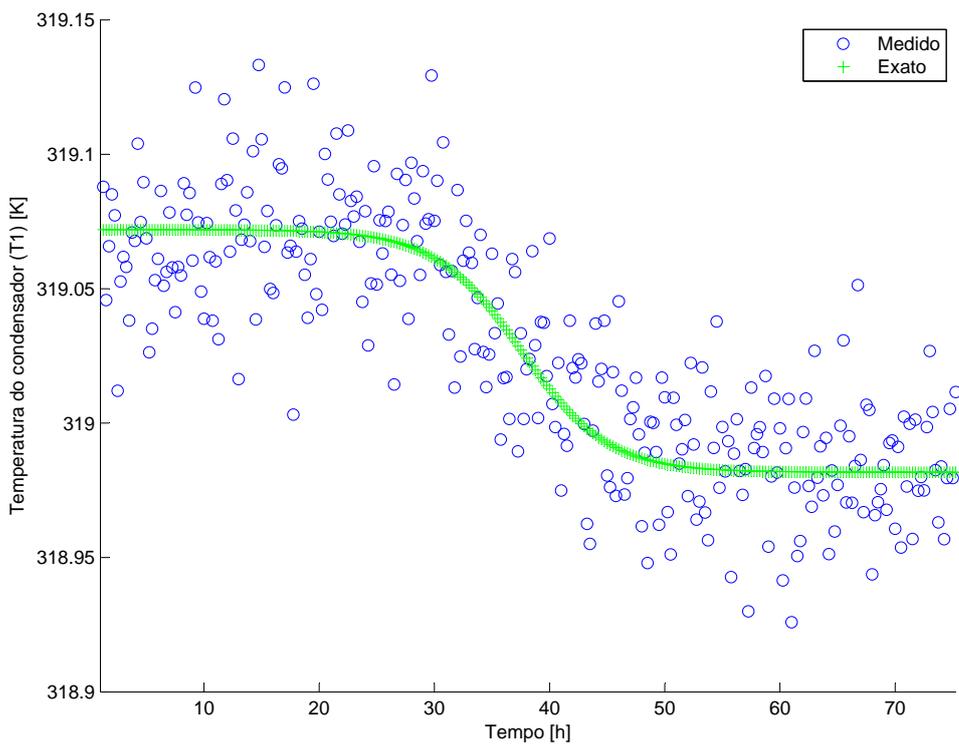


Figura 6.12: Valores exatos e medidos (corrompidos) - T_c no estudo de caso 3.

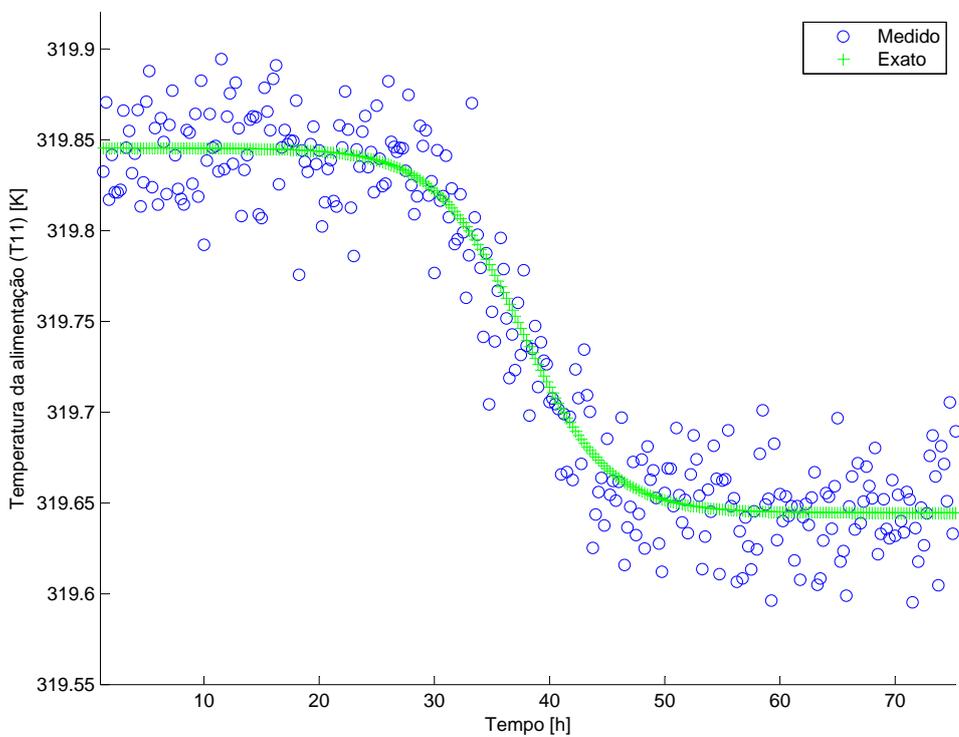


Figura 6.13: Valores exatos e medidos (corrompidos) - T_f no estudo de caso 3.

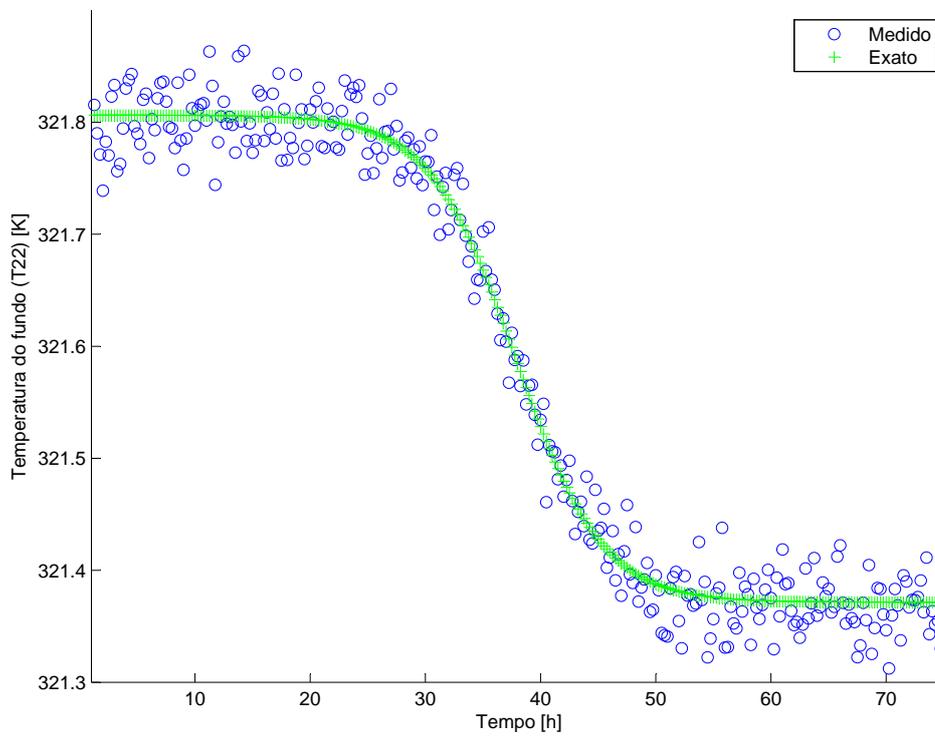


Figura 6.14: Valores exatos e medidos (corrompidos) - T_r no estudo de caso 3.

6.1.4 Caso 4: Monitoramento de um processo de separação por membrana

O quarto caso estuda o monitoramento de um processo de separação de CO_2 por membrana *on-line* e em tempo real. O monitoramento é realizado através de aplicativo-*web* desenvolvido em linguagem Python na plataforma Dash, onde os dados podem ser monitorados em *dashboard* iterativos e visualizados em qualquer navegador-*web*. Este caso é pioneiro no uso de procedimentos de RD para monitorar processos de separação por membrana de forma confiável, online e em tempo real. Para tanto, foi utilizada durante este caso uma planta de membranas localizada em uma das Unidades *Offshore* da Petrobras. Por fim, vale ressaltar que, por motivos de sigilo industrial, os resultados numéricos são apresentados de forma normalizada. Para mais detalhes desse estudo de caso consultar o Apêndice A.

6.2 Resultados dos estudos de caso

6.2.1 Caso 1: Avaliação dos métodos de otimização

Avaliação do modelo

Nas Tabelas 6.14, 6.15 e 6.16 são apresentados os resultados obtidos utilizando o software EMSO com o modelo MES, e os resultados simulado e experimental relatados por KRISHNAMURTHY e TAYLOR (1985), para as correntes de alimentação (F), destilado (D) e fundo (B), respectivamente.

As Tabelas 6.14, 6.15, 6.16 e 6.17, apresentam os resultados obtidos utilizando-se o software EMSO, pacote termodinâmico VRTherm (Líquido: Peng-Robinson; Gás: Peng-Robinson), com o modelo MES dinâmico, supondo *liquid holdup* constante, para as correntes de alimentação, destilado e fundo, descritas por suas respectivas vazões e concentrações. Na Tabela 6.14 pode-se observar que os valores usados no trabalho de KRISHNAMURTHY e TAYLOR (1985) foram mantidos praticamente os mesmos, pois usaram-se os mesmos valores para variáveis de entrada. A Tabela 6.15 apresenta os resultados da corrente do destilado, sendo observadas pequenas alterações em relação ao valor experimental e simulado, exceto a vazão do decano no destilado, que foi computada como zero pelo modelo MES no EMSO.

A Tabela 6.16 apresenta os resultados da corrente de fundo, e podem ser observados grandes desvios em relação ao valor experimental e simulado.

Tabela 6.14: Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte I.

		(KRISHNAMURTHY e TAYLOR, 1985)		EMSO		
		Variável [kmol/s]	Experimental	Simulado	MES	Erro relativo [%]
		$F \Rightarrow$ [kmol/s]	0,051130	0,051130	0,051130	0,000
Componentes [kmol/s] $\times 10^4$	C_2	$x_{F,1}$	1,0505	1,0505	1,0737	2,161
	C_3	$x_{F,2}$	98,9570	98,9570	98,9366	0,021
	$i-C_4$	$x_{F,3}$	39,2580	39,2580	39,2678	0,025
	$n-C_4$	$x_{F,4}$	30,8690	30,8690	30,8825	0,044
	$i-C_5$	$x_{F,5}$	10,2710	10,2710	10,2771	0,059
	$n-C_5$	$x_{F,6}$	5,9489	5,9489	5,9311	0,299
	C_6	$x_{F,7}$	8,7775	8,7775	8,7944	0,193
	C_7	$x_{F,8}$	41,3300	41,3300	41,3130	0,041
	C_8	$x_{F,9}$	114,1480	114,1480	114,1733	0,022
	C_9	$x_{F,10}$	86,6360	86,6360	86,6142	0,025
C_{10}	$x_{F,11}$	74,0520	74,0520	74,0362	0,021	

Primeiramente, deve-se considerar o valor simulado pelo trabalho de KRISHNAMURTHY e TAYLOR (1985) mais correto, pois se aproxima mais do valor experimental. Além disso, o modelo utilizado para a simulação foi mais rigoroso, devido à utilização das equações de fenômeno de transporte, parâmetros do recheio e hidráulica, que descreve as perdas de carga, área de troca térmica e de massa, e *liquid holdup* dentro da coluna recheada.

Portanto, é notório que o modelo utilizado no simulador EMSO tem suas limitações quando se trata de uma corrente que se encontra em elevados patamares de pressão e temperatura, para esta composição; Assim, essas limitações podem estar relacionadas às das Equações de Estados e uma baixa eficiência de Murphree utilizada no modelo MES ($E^M = 0,9$). Além disso, o número de estágio admitido pode ter sido o fator mais relevante para a diferença entre os valores preditos e os medidos relatados em KRISHNAMURTHY e TAYLOR (1985).

Além disso, a abordagem termodinâmica utilizada foi $\phi - \phi$, ou seja, tanto

Tabela 6.15: Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte II.

		(KRISHNAMURTHY e TAYLOR, 1985)		EMSO		
		Variável [kmol/s]	Experimental	Simulado	MES	Erro relativo [%]
		$D \Rightarrow$ [kmol/s]	–	0,019396	0,019396	0,000
Componentes [kmol/s] × 10 ⁴	C_2	$x_{D,1}$	–	1,050100	1,0474	0,257
	C_3	$x_{D,2}$	–	98,753000	98,9584	0,208
	$i-C_4$	$x_{D,3}$	39,17800	38,928000	39,1217	0,498
	$n-C_4$	$x_{D,4}$	30,78000	30,544000	30,6263	0,269
	$i-C_5$	$x_{D,5}$	9,90310	9,676900	9,4071	2,788
	$n-C_5$	$x_{D,6}$	5,55780	5,515000	5,1787	6,098
	C_6	$x_{D,7}$	4,89840	5,065800	4,2089	16,915
	C_7	$x_{D,8}$	3,34640	3,806700	4,4029	15,662
	C_8	$x_{D,9}$	0,31530	0,590200	0,9504	61,030
	C_9	$x_{D,10}$	–	0,023600	0,0388	64,407
C_{10}	$x_{D,11}$	–	0,006200	0,0000	100,000	

para fase líquida e gasosa foram utilizadas as equações de estados para o cálculo da fugacidade. O que poderia melhorar as aproximações dos cálculos seria um modelo energia livre de Gibbs em excesso para o coeficiente de fugacidade da fase líquida e uma eficiência de Murphree mais realista, pois um típico valor de eficiência de pratos varia entre 0,5 e 0,7 (DOUGLAS, 1988). Entretanto, valores de eficiência de Murphree abaixo de 0,9 não convergiram e o melhor pacote termodinâmico disponível foi o Peng-Robinson.

A Tabela 6.17 e a Figura 6.1 apresentam os resultados do perfil de temperatura ao longo da coluna simulado no EMSO e o perfil de temperatura relatado através de um gráfico no trabalho de KRISHNAMURTHY e TAYLOR (1985). Primeiramente, deve-se ter em mente que, ao extrair a informação de um gráfico ou figura utilizando software de aquisição ou interpolação de pontos, inevitavelmente erros de imprecisão do software são adicionados aos dados. Os erros gerados nesta etapa naturalmente foram agregados aos erros experimentais descritos no trabalho de

Tabela 6.16: Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte III.

		(KRISHNAMURTHY e TAYLOR, 1985)		EMSO		
Variável		Experimental	Simulado	MES	Erro relativo [%]	
$B \Rightarrow$ [kmol/s]		0,031734	0,031734	0.031734	0,000	
Componentes [kmol/s] $\times 10^5$	C_2	$x_{B,1}$	–	0,0041	0,0000	100,000
	C_3	$x_{B,2}$	–	2,0562	0,0000	100,000
	$i-C_4$	$x_{B,3}$	0,7903	3,2944	1,2694	61,468
	$n-C_4$	$x_{B,4}$	0,7300	3,2478	2,5387	21,833
	$i-C_5$	$x_{B,5}$	3,6781	5,9381	8,5682	44,292
	$n-C_5$	$x_{B,6}$	3,9125	4,3400	7,6162	75,488
	C_6	$x_{B,7}$	38,7910	37,1170	45,6970	23,116
	C_7	$x_{B,8}$	379,8300	375,2200	369,3838	1,555
	C_8	$x_{B,9}$	1.138,3300	1.135,5800	1.131,9518	0,320
	C_9	$x_{B,10}$	–	866,1400	865,7035	0,050
C_{10}	$x_{B,11}$	–	740,4400	740,3542	0,012	

KRISHNAMURTHY e TAYLOR (1985). Os valores com maiores desvios foram colocados em negrito, a fim de permitir uma melhor visualização. Portanto, pode-se dizer que as mesmas limitações que geraram valores com grandes desvios na corrente de fundo, que foram atribuídas às equações de estados, também podem ser atribuídas aos desvios em relação ao perfil de temperatura.

A Figura 6.15 mostra que no fundo da coluna a temperatura reduzida de cada componente ultrapassa muito a temperatura crítica, para os componentes mais leves, aproximando-se da temperatura crítica para os componentes mais pesados. Entretanto, a pressão reduzida permanece abaixo de 1, o que pode mostrar que o fluido não está acima de seu ponto crítico e, portanto, dentro dos limites de utilização de equação de estado.

Tabela 6.17: Resultados do modelo MES no caso KRISHNAMURTHY e TAYLOR (1985) - parte IV.

Estágios (i)	(KRISHNAMURTHY e TAYLOR, 1985)		EMSO
	Temperatura Medida [K]	Temperatura MES [K]	Erro relativo [%]
Condensador: T_0	305,00	308,44	1,128
Prato 1: T_1	368,85	366,34	0,680
Prato 2: T_2	386,23	411,89	6,644
Prato 3: T_3	402,94	435,09	7,979
Prato 4: T_4	416,98	447,56	7,334
Prato 5: T_5	455,22	457,00	0,391
Prato 6: T_6	480,48	469,70	2,244
Prato 7: T_7	487,83	485,93	0,389
Prato 8: T_8	492,85	500,42	1,536
Prato 9: T_9	497,19	511,03	2,784
Prato 10: T_{11}	504,55	519,35	2,933
Refervedor: T_{11}	525,00	527,98	0,568

	C_2	C_3	$i-C_4$	$n-C_4$	$i-C_5$	$n-C_5$	$n-C_6$	$n-C_7$	$n-C_8$	$n-C_9$	$n-C_{10}$	
Tc (K)	305.300	369.900	407.800	425.000	461.000	469.800	507.600	540.000	568.900	595.000	617.800	
Pc (bar)	49.000	42.500	36.500	38.000	33.800	33.600	30.200	27.400	24.900	23.000	21.100	
T (K)	Pr	0.229	0.264	0.308	0.296	0.332	0.334	0.372	0.410	0.451	0.488	0.532
305.000	Tr	0.999	0.825	0.748	0.718	0.662	0.649	0.601	0.565	0.536	0.513	0.494
368.850		1.208	0.997	0.904	0.868	0.800	0.785	0.727	0.683	0.648	0.620	0.597
386.230		1.265	1.044	0.947	0.909	0.838	0.822	0.761	0.715	0.679	0.649	0.625
402.941		1.320	1.089	0.988	0.948	0.874	0.858	0.794	0.746	0.708	0.677	0.652
416.979		1.366	1.127	1.023	0.981	0.905	0.888	0.821	0.772	0.733	0.701	0.675
455.220		1.491	1.231	1.116	1.071	0.987	0.969	0.897	0.843	0.800	0.765	0.737
480.481		1.574	1.299	1.178	1.131	1.042	1.023	0.947	0.890	0.845	0.808	0.778
487.834		1.598	1.319	1.196	1.148	1.058	1.038	0.961	0.903	0.858	0.820	0.790
492.848		1.614	1.332	1.209	1.160	1.069	1.049	0.971	0.913	0.866	0.828	0.798
497.193		1.629	1.344	1.219	1.170	1.079	1.058	0.979	0.921	0.874	0.836	0.805
504.546		1.653	1.364	1.237	1.187	1.094	1.074	0.994	0.934	0.887	0.848	0.817
525.000		1.720	1.419	1.287	1.235	1.139	1.117	1.034	0.972	0.923	0.882	0.850

Figura 6.15: Perfil de temperatura e pressão reduzida dos componentes na coluna no estudo de caso 1.

Avaliação dos estimadores

Na etapa de reconciliação de dados, utilizaram-se os estimadores MQP, quasi-robusto Normal Contaminada (NC) e o estimador robusto *redescending* Welsch com DMEG, juntamente com o modelo MES através do software EMSO, utilizando os métodos de otimização Nelder-Mead, Enxame de Partículas (PSO), quasi-Newton (BFGS), SQP e Ponto Interior através do *software* Matlab.

- Otimizador Nelder-Mead

Nas Tabelas de 6.18 à 6.21 são apresentados os resultados obtidos para o caso de reconciliação de dados com o otimizador Nelder-Mead, em que se avaliou o modelo MES na presença de variáveis com erros grosseiros. Propositamente, as variáveis vazão da alimentação F , vazão de isobutano no destilado $x_{D,3}$ e temperatura do segundo prato (estágio 2) foram corrompidas para simular dados com valores espúrios, conforme a Tabela 6.5, já apresentada no caso anterior. Todas as demais variáveis foram corrompidas com erros aleatórios randomicamente distribuídos com base nos valores simulados no EMSO.

As Tabelas 6.18, 6.19, 6.20 e 6.21, apresentam os resultados da RD e RRD obtidos utilizando-se o *software* EMSO e o *software* Matlab, para as correntes de alimentação, destilado e fundo, descritas por suas respectivas vazões e concentrações. Utilizando os estimadores MQP, Normal Contaminada (NC) e Welsch.

Tabela 6.18: Resultados com o otimizador Nelder-Mead - F no estudo de caso 1.

		Dados Reconciliados - EMSO/Matlab				
		Exato	Medido	MQP	NC	Welsch
F [kmol/h] \Rightarrow		184,0680	220,8816*	<u>185,1120</u>	<u>183,5640</u>	<u>183,5640</u>
Componentes [kmol/s] $\times 10^4$	$x_{F,1}$	1,0737	1,0878	1,0798	1,0708	1,0708
	$x_{F,2}$	98,9366	100,8490	100,9375	100,0934	100,0934
	$x_{F,3}$	39,2678	39,7448	39,7477	39,4153	39,4153
	$x_{F,4}$	30,8825	30,4771	30,4921	30,2371	30,2371
	$x_{F,5}$	10,2771	10,1372	10,1297	10,0450	10,0450
	$x_{F,6}$	5,9311	5,8544	5,8105	5,7619	5,7619
	$x_{F,7}$	8,7944	8,9321	8,8957	8,8213	8,8213
	$x_{F,8}$	41,3130	40,7005	40,7246	40,3841	40,3841
	$x_{F,9}$	114,1733	112,6523	112,7641	111,8211	111,8211
	$x_{F,10}$	86,6142	87,9944	88,0825	87,3459	87,3459
	$x_{F,11}$	74,0362	75,1479	75,2275	74,5984	74,5984

* erro grosseiro.

Tabela 6.19: Resultados com o otimizador Nelder-Mead - D no estudo de caso 1.

		Dados Reconciliados - EMSO/Matlab				
		Exato	Medido	MQP	NC	Welsch
D [kmol/h] \Rightarrow		69,8256	70,9685	71,7120	70,6680	70,6680
Componentes [kmol/s] $\times 10^4$	$x_{D,1}$	1,0474	1,0607	1,0757	1,060000	1,0600
	$x_{D,2}$	98,9584	100,574	101,6119	100,1326	100,1326
	$x_{D,3}$	39,1217	46,9461*	<u>40,15870</u>	<u>39,574100</u>	<u>39,5741</u>
	$x_{D,4}$	30,6263	31,2138	31,41380	30,956500	30,9565
	$x_{D,5}$	9,4071	9,3089	9,64130	9,500900	9,5009
	$x_{D,6}$	5,1787	5,102	5,29870	5,221600	5,2216
	$x_{D,7}$	4,2089	4,1557	4,30270	4,240100	4,2401
	$x_{D,8}$	4,4029	4,3406	4,56170	4,495300	4,4953
	$x_{D,9}$	0,9504	0,9626	0,996	0,9815	0,9815
	$x_{D,10}$	0,0388	0,0384	0,0398	0,0393	0,0393
	$x_{D,11}$	0,0000	0,0000	0,0000	0,0000	0,0000

* erro grosseiro.

Tabela 6.20: Resultados com o otimizador Nelder-Mead - B no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
B [kmol/h] \Rightarrow	114,2424	112,8579	113,3640	112,8960	112,8960
$x_{B,1}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,2}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,3}$	1,2694	1,2547	1,25960	1,254400	1,2540
$x_{B,4}$	2,5387	2,5818	2,51920	2,508800	2,5080
$x_{B,5}$	8,5682	8,726	8,50230	8,467200	8,4645
$x_{B,6}$	7,6162	7,5091	7,55760	7,526400	7,5240
$x_{B,7}$	45,697	45,0529	45,34560	45,158400	45,1440
$x_{B,8}$	369,3838	376,3223	365,91380	364,403200	364,2870
$x_{B,9}$	1131,9518	1119,3437	1122,6185	1117,984	1117,6275
$x_{B,10}$	865,7035	855,3165	859,0472	855,5008	855,228
$x_{B,11}$	740,3542	751,9176	734,6617	731,6288	731,3955

Tabela 6.21: Resultados com o otimizador Nelder-Mead - Temperaturas no estudo de caso 1.

Estágio i	Exato	Medido	Dados Reconciliados - EMSO/Matlab		
			MQP	NC	Welsch
0	308,44	307,47	308,45	308,45	308,45
1	366,34	367,52	366,31	366,31	366,31
2	411,89	396,89*	<u>411,89</u>	<u>411,89</u>	<u>411,89</u>
3	435,09	434,09	435,13	435,13	435,13
4	447,56	446,44	447,61	447,61	447,61
5	457,00	458,13	457,02	457,02	457,02
6	469,70	468,62	469,68	469,68	469,68
7	485,93	486,88	485,84	485,84	485,84
8	500,42	499,35	500,30	500,30	500,30
9	511,03	512,16	510,93	510,93	510,93
10	519,35	520,79	519,29	519,29	519,29
11	527,98	529,29	527,92	527,92	527,92

* erro grosseiro.

Para todos os casos apresentados a seguir (estimadores e otimizadores) o problema de reconciliação de dados não detectou o erro grosseiro na variável de entrada, com isso uma estratégia foi montada visando eliminar o erro grosseiro da variável F . A estratégia utilizada foi reconciliar as vazões globais em um primeiro passo, em que o modelo (restrição do problema) foi inserido na função objetivo como uma penalidade, com isso o otimizador pôde manipular a terceira variável e assim detectar o erro grosseiro. Assim, o valor reconciliado da primeira etapa tornou-se o novo valor medido para a segunda etapa, onde utilizaram todas as variáveis medidas, inclusive as novas variáveis medidas. Utilizou-se apenas um tipo de otimizador na primeira etapa, o otimizador *fmincon* com algoritmo do Ponto Interior, pois foi o que apresentou o menor tempo durante as rodadas de simulação. Para a segunda etapa foram mantidos os outros cinco otimizadores.

As possíveis explicações para a não detecção de erros grosseiros nas variáveis de entrada são:

- A princípio a técnica conhecida como otimização por *feasible path* dificulta a reconciliação de dados de detectar o erro na variável de entrada, pois um dos motivos é que das três variáveis do balanço de massa global (F , D e B) duas são variáveis de decisão, ou seja, manipulada pelo otimizador, onde a terceira é função dessas duas primeiras e, sendo assim, o otimizador só manipula duas variáveis e a terceira é recebida na função objetivo com um parâmetro em função do modelo e das variáveis de entrada. Na outra estratégia as restrições de igualdade também devem ser satisfeitas, e ela reduz os graus de liberdade em mesmo número;
- Outro problema que pode ocorrer é quando duas variáveis são medidas e uma está com erro grosseiro, aproximando assim do *breack point* dos estimadores robusto, sendo que o limite máximo de *breack point* de qualquer estimador robusto é de 50%. Quando a porcentagem de erros grosseiros nos dados ultrapassa o *breack point* do estimador as estimativas são corrompidas, invertendo a discriminação entre os dados com erros aleatórios e erros grosseiros. Com isso o estimador “quebra”. O mesmo problema pode ser inevitável se uma das vazões não for medida em coluna, ou seja, estimada pelo modelo. Assim, para atingir a redundância espacial no balanço de massa global no problema de RD em colunas é necessário que todas as vazões sejam medidas, e se possível que todas elas sejam manipuladas pelo otimizador;
- Outra explicação possível para este problema está na natureza dinâmica do modelo de coluna. Modelos dinâmicos apresentam termos de acúmulos das composições e massas, que por sua vez reflete o comportamento dinâmico das

variáveis de saída, como atraso no fechamento de balanço de massa global. Com isso, variações bruscas nas variáveis de entrada não vão se traduzir em respostas com mesmas intensidades e instantaneamente. Isso pode explicar o porquê a RD robusta não eliminou o erro grosseiro na variável F , pois durante os passos iterativos das otimizações as alterações nas variáveis F e D (manipuladas pelo otimizador) não se refletiram em mudanças significativas na variável de saída B . Dessa forma, o erro grosseiro presente na variável F não influenciou na variável de saída, mas os desvios do valor medido (com erro grosseiro) durante os passos iterativos foram significativos na Fobj. Portanto, o modelo dinâmico pode ser entendido como um “filtro” de grandes oscilações e perturbações na operação, pois as respostas são suavizadas e atrasadas. De fato, quando manipulações na operação de uma coluna ocorre como mudança brusca de *set-point*, parada ou inicialização o processo, o procedimento de RD pode entender que essas mudanças podem ser um tipo de erro grosseiro. Com isso, esses dados podem corromper as métricas de estimativas de variâncias e alteram os parâmetros e variáveis de estado do modelo. Por isso, monitoramento de dados baseado em RD necessita de etapas de “proteções” numéricas e pré-tratamentos, após a aquisição de dados e anterior a RD.

Na Tabela 6.18 comparando-se os resultados entre os estimadores e os valores exatos e medidos observa-se que houve detecção do erro grosseiro da variável corrompida F . Entretanto, os estimadores que apresentaram melhor reconciliação do erro grosseiro foram o estimador Welsch e o estimador NC, o estimador MQP apresentou a pior correção em relação ao valor exato. Apesar da pequena diferença entre as variáveis reconciliadas, o estimador MQP mostrou um pouco de sua não robustez, gerando alterações em outras variáveis não corrompidas com erros grosseiros, sujando (*smearing*) outras variáveis, pois as presenças de erros grosseiros geram um peso na função objetivo de um estimador MQP muito maior que esses mesmos erros gerariam a um estimador robusto. Portanto, para ponderar esse peso e respeitar as restrições, outras variáveis são penalizadas (manchadas), a fim de encontrar um mínimo satisfatório. Isso não acontece com o estimador robusto Welsch, por exemplo, pois o erro grosseiro não gera um peso grande na função objetivo. Isso explica por que outras variáveis não sofreram alterações significantes com o estimador Welsch e NC.

Nas Tabelas 6.19, 6.20 e 6.21 comparam-se os resultados da variável de saída (corrente de destilado, corrente de fundo e o perfil de temperaturas), podendo-se notar que todos os estimadores detectaram e reconciliaram bem as variáveis com erro grosseiro (vazão de isobutano no destilado - $x_{D,3}$ e temperatura do estágio 2

T_2). A temperatura do estágio 2 não foi influenciada pelo erro grosseiro na vazão de entrada e, também, não foi influenciada pelo erro grosseiro da variável concentração de isobutano no destilado $x_{D,3}$, pois no procedimento de solução do modelo MES a temperatura do estágio 2 é função do vetor de concentrações da fase vapor, do vetor de concentrações da fase líquida e da pressão do estágio (na verdade, utilizam-se os pontos de bolha e uma estimativa da temperatura para achar o ponto de orvalho, depois corrigindo-se a temperatura). Como o vetor de concentrações da alimentação e a temperatura de alimentação não apresentaram erros grosseiros, apenas erros aleatórios se propagaram para o cálculo da temperatura no estágio 2. Portanto, o mérito maior da correção do erro grosseiro na temperatura é do próprio modelo MES. Para a concentração de isobutano no destilado $x_{D,3}$, observa-se uma correção um pouco mais próxima do valor exato com os estimadores NC e Welsh.

A Tabela 6.22 apresenta o critério de avaliação de desempenho utilizado neste caso com o otimizador Nelder-Mead. O critério TER foi selecionado para a avaliação, pois não depende do estimador utilizado e deve ser utilizado quando as variáveis exatas são conhecidas.

Tabela 6.22: Avaliação de desempenho com o otimizador Nelder-Mead (TER).

TER	MQP	NC	Welsh
Vazões (Ponto Interior)	0,8194	0,9063	0,9044
TOTAL	0,8603	0,8630	0,8630

Observando-se a Tabela 6.22 e a Figura 6.16 conclui-se que os estimadores NC e Welsh apresentam um desempenho superior em relação ao estimador MQP no primeiro passo (reconciliando apenas as vazões), o que contribui, mais uma vez, à utilização do procedimento de RRD em colunas de destilação. Entretanto, no total os estimadores NC e Welsh apresentaram um TER levemente superior, o que torna a estratégia de eliminação do erro grosseiro utilizando parte do modelo em uma etapa inicial uma técnica eficiente, visto que o estimador MQP obteve bons resultados e eliminando o efeito *smearing* na primeira etapa. Uma possível explicação para o bom resultado do estimador não robusto MQP pode ser atribuída ao fato que na primeira etapa, a variável medida de decisão (F) com erro grosseiro foi tratada, tornando assim todo o novo conjunto de variáveis medidas de decisão com apenas erros aleatórios. Assim, a nova distribuição dos erros das variáveis medidas de entrada se aproximou mais de uma distribuição Normal, tornando o estimador MQP mais adequado para este caso.

A Figura 6.17 ilustra (como exemplo a função Normal Contaminada com o

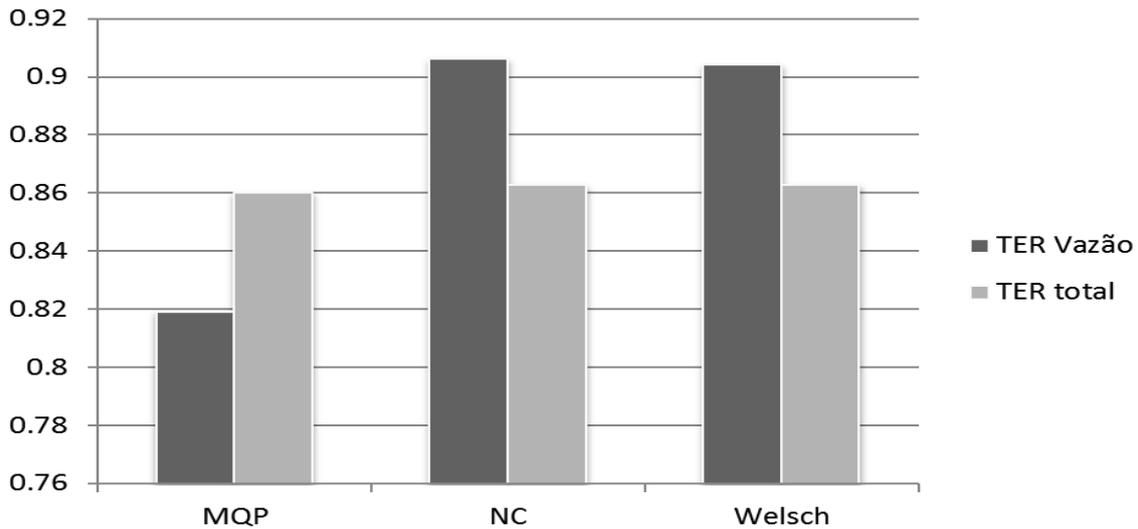


Figura 6.16: Comparação entre os critérios TER de avaliação (Nelder-Mead).

otimizador Nelder-Mead) o procedimento de otimização ao longo de suas iterações. Como se observa, foram necessárias mais de 100 iterações para reduzir o tamanho do polítopo formado por $n + 1$ vértices no espaço de n dimensões. Uma alternativa para contornar esse problema seria alterar parte do código, a fim de adequar ao problema de RD. O algoritmo original do método faz as distribuições dos pontos (vértices do polítopo) utilizando uma distribuição uniforme. Pois, se a geração os pontos dos vértices for normalmente distribuída, o procedimento de otimização pelo método de Nelder-Mead pode ser acelerado, devido a redução do espaço de busca. Entretanto, esta estratégia não foi concluída.

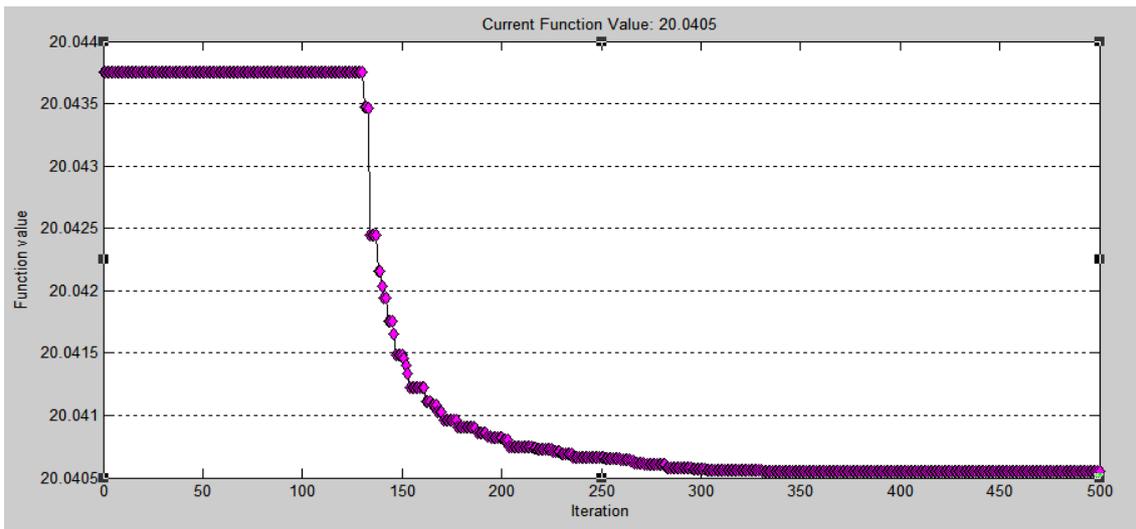


Figura 6.17: Progressão da otimização com o método Nelder-Mead utilizando o estimador NC.

- Otimizador PSO (*Particle Swarm Optimization*)

Nas Tabelas de 6.23 a 6.26 são apresentados os resultados obtidos para o caso de reconciliação de dados com o otimizador PSO, em que se avalia o modelo MES na presença de variáveis com erros grosseiros. Analogamente ao caso com o otimizador Nelder-Mead, utilizaram-se as mesmas variáveis e com os mesmos erros, conforme a Tabela 6.5.

Tabela 6.23: Resultados com o otimizador PSO - F no estudo de caso 1.

	Exato	Medido	Dados Reconciliados - EMSO/Matlab		
			MQP	NC	Welsch
F [kmol/h] \Rightarrow	184,0680	220,8816*	<u>187,9560</u>	<u>187,5240</u>	<u>183,9600</u>
$x_{F,1}$	1,0737	1,0878	1,0964	1,0418	1,0731
$x_{F,2}$	98,9366	100,8490	102,9059	99,8565	102,7110
$x_{F,3}$	39,2678	39,7448	40,3583	38,8591	40,0624
$x_{F,4}$	30,8825	30,4771	31,0650	30,9415	30,8644
$x_{F,5}$	10,2771	10,1372	10,2854	10,7305	10,2711
$x_{F,6}$	5,9311	5,8544	5,9519	5,9904	5,7232
$x_{F,7}$	8,7944	8,9321	9,0323	9,5325	9,0447
$x_{F,8}$	41,3130	40,7005	41,4547	40,8907	40,5734
$x_{F,9}$	114,1733	112,6523	114,9142	110,4829	112,6244
$x_{F,10}$	86,6142	87,9944	89,6968	90,5845	86,1546
$x_{F,11}$	74,0362	75,1479	76,8009	79,4893	76,1390

* erro grosseiro.

Tabela 6.24: Resultados com o otimizador PSO - D no estudo de caso 1.

		Dados Reconciliados - EMSO/Matlab				
		Exato	Medido	MQP	NC	Welsch
D [kmol/h] \Rightarrow		69,8256	70,9685	71,4240	72,3600	70,8480
Componentes [kmol/s] $\times 10^4$	$x_{D,1}$	10,474	10,607	1,0714	1,0854	1,0627
	$x_{D,2}$	98,9584	100,574	101,2038	102,5301	100,3877
	$x_{D,3}$	39,1217	46,9461*	<u>39,9974</u>	<u>40,5216</u>	<u>39,6749</u>
	$x_{D,4}$	30,6263	31,2138	31,2877	31,6977	31,0354
	$x_{D,5}$	94,071	93,089	96,026	97,284	95,251
	$x_{D,6}$	51,787	51,02	52,774	53,466	52,349
	$x_{D,7}$	42,089	41,557	42,854	43,416	42,509
	$x_{D,8}$	44,029	43,406	45,434	46,029	45,067
	$x_{D,9}$	0,9504	0,9626	0,9920	1,0050	0,9840
	$x_{D,10}$	0,0388	0,0384	0,0397	0,0402	0,0394
	$x_{D,11}$	0,0000	0,0000	0,0000	0,0000	0,0000

* erro grosseiro.

Tabela 6.25: Resultados com o otimizador PSO - B no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
B [kmol/h] \Rightarrow	114,2424	112,8579	116,5320	115,1640	113,0760
$x_{B,1}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,2}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,3}$	12,694	12,547	12,948	12,796	12,564
$x_{B,4}$	25,387	25,818	25,896	25,592	25,128
$x_{B,5}$	85,682	87,26	87,399	86,373	84,807
$x_{B,6}$	76,162	75,091	77,688	76,776	75,384
$x_{B,7}$	45,697	45,0529	46,6128	46,0656	45,2304
$x_{B,8}$	369,3838	376,3223	376,1394	371,7238	364,9842
$x_{B,9}$	1131,9518	1119,3437	1153,9905	1140,4435	1119,7665
$x_{B,10}$	865,7035	855,3165	883,0536	872,6872	856,8648
$x_{B,11}$	740,3542	751,9176	755,1921	746,3267	732,7953

Tabela 6.26: Resultados com o otimizador PSO - Temperaturas no estudo de caso 1.

Estágio i	Exato	Medido	Dados Reconciliados - EMSO/Matlab		
			MQP	NC	Welsch
0	308,44	307,47	308,45	308,45	308,45
1	366,34	367,52	366,31	366,31	366,31
2	411,89	396,89*	<u>411,89</u>	<u>411,89</u>	<u>411,89</u>
3	435,09	434,09	435,13	435,13	435,13
4	447,56	446,44	447,61	447,61	447,61
5	457,00	458,13	457,02	457,02	457,02
6	469,70	468,62	469,68	469,68	469,68
7	485,93	486,88	485,84	485,84	485,84
8	500,42	499,35	500,30	500,30	500,30
9	511,03	512,16	510,93	510,93	510,93
10	519,35	520,79	519,29	519,29	519,29
11	527,98	529,29	527,92	527,92	527,92

* erro grosseiro.

Nas Tabelas 6.23, 6.24 e 6.26 observa-se que todos os estimadores detectaram o erro grosseiro nas variáveis corrompidas (F , $x_{D,3}$ e T_2). Assim, como no caso do otimizador Nelder-Mead, as variáveis de saída são reconciliadas mais por mérito do modelo, pois o otimizador não manipula as variáveis dependentes diretamente. Assim, analogamente, é de se esperar que os estimadores robustos tenham pouca sensibilidade, já que apenas as variáveis de entrada podem ser manipuladas. Além disso, como o modelo é dinâmico nas concentrações, pequenas perturbações nas variáveis de concentração da alimentação gera uma resposta lenta e suavizada nas saídas (destilado e fundo).

A Tabela 6.27 apresenta o critério de avaliação de desempenho utilizado neste caso com o otimizador PSO. Observando-se a Tabela 6.27 e a Figura 6.18 conclui-se que o estimador MQP apresentou um desempenho superior em relação aos estimadores NC e Welsch no TER total, já que o TER das vazões foi sempre o mesmo, pois o método de otimização foi determinístico (Ponto Interior). Entretanto, no TER total os estimadores NC e Welsch apresentaram um TER inferior, o que contribui ainda mais para a explicação do bom resultado do estimador não robusto MQP. Possivelmente a nova distribuição dos erros das variáveis medida de entrada se aproximou mais de uma distribuição Normal do que uma distribuição de um estimador robusto, tornando o estimador MQP mais adequado para este caso. Outro fato particular desse otimizador é que é de natureza estocástica, ou seja, pode ser que outras rodadas de simulação o otimizador ache um ponto ótimo melhor ou pior que este, tornando a avaliação de desempenho TER em função do método de otimização, também. Uma alternativa é rodar a otimização várias vezes e tirar a média do ponto ótimo.

Tabela 6.27: Avaliação de desempenho com o otimizador PSO (TER) no estudo de caso 1.

TER	MQP	NC	Welsch
Vazões (Ponto Interior)	0,8194	0,9063	0,9044
TOTAL	0,8642	0,7021	0,8170

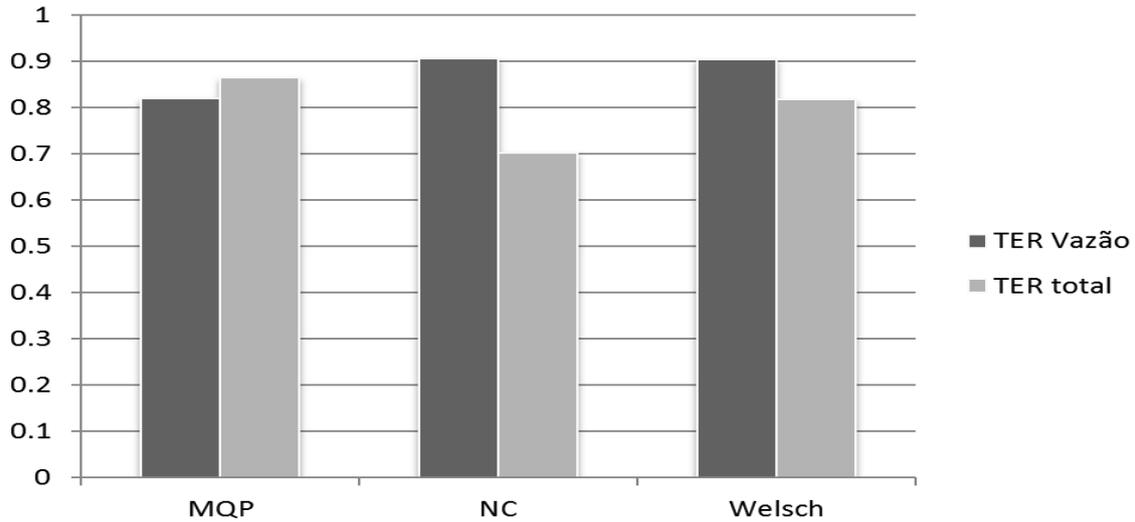


Figura 6.18: Comparação entre os critérios TER de avaliação (PSO) no estudo de caso 1.

- Otimizador quasi-Newton (BFGS)

Nas Tabelas de 6.28 a 6.31 são apresentados os resultados obtidos para o caso de reconciliação de dados com o otimizador *fminunc* do Matlab, que utiliza o método quasi-Newton (BFGS), em que se avaliou o modelo MES na presença de variáveis com erros grosseiros. Analogamente ao caso com o otimizador Nelder-Mead, utilizaram-se as mesmas variáveis e com os mesmos erros, conforme a Tabela 6.5.

Nas Tabelas 6.28, 6.29 e 6.31 observa-se, novamente, que todos os estimadores detectaram o erro grosseiro nas variáveis corrompidas (F , $x_{D,3}$ e T_2). Assim, como nos outros casos, as variáveis de saída foram reconciliadas, pois o modelo atua como se fosse um “filtro” de ruído das variáveis de entrada. Assim, é de se esperar que as Fobj's tenham pouca sensibilidade em relação às variáveis de saída. De fato, como o modelo é dinâmico nas concentrações, pequenas perturbações nas variáveis de concentração da alimentação geram uma resposta lenta e suavizada nas saídas (destilado e fundo).

A Tabela 6.32 apresenta o critério de avaliação de desempenho utilizado neste caso com o otimizador BFGS. Observando a Tabela 6.32 e a Figura 6.19, conclui-se que o estimador MQP apresentou um desempenho levemente inferior em relação aos estimadores NC e Welsch no TER total, já que o TER das vazões foi sempre o mesmo, pois o método de otimização da primeira etapa é determinístico (Ponto Interior).

Tabela 6.28: Resultados com o otimizador BFGS - F no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab					
	Exato	Medido	MQP	NC	Welsch	
F [kmol/h]⇒	184,0680	220,8816*	<u>187,1280</u>	<u>183,7440</u>	<u>183,6720</u>	
Componentes [kmol/s]×10 ⁴	$x_{F,1}$	1,0737	1,0878	1,0916	1,0718	1,0714
	$x_{F,2}$	98,9366	100,8490	102,0367	100,1915	100,1523
	$x_{F,3}$	39,2678	39,7448	40,1805	39,4539	39,4385
	$x_{F,4}$	30,8825	30,4771	30,8241	30,2667	30,2549
	$x_{F,5}$	10,2771	10,1372	10,2401	10,0549	10,0509
	$x_{F,6}$	5,9311	5,8544	5,8737	5,7675	5,7653
	$x_{F,7}$	8,7944	8,9321	8,9925	8,8299	8,8265
	$x_{F,8}$	41,3130	40,7005	41,1682	40,4237	40,4078
	$x_{F,9}$	114,1733	112,6523	113,9921	111,9307	111,8869
	$x_{F,10}$	86,6142	87,9944	89,0417	87,4315	87,3973
	$x_{F,11}$	74,0362	75,1479	76,0467	74,6715	74,6423

* erro grosseiro.

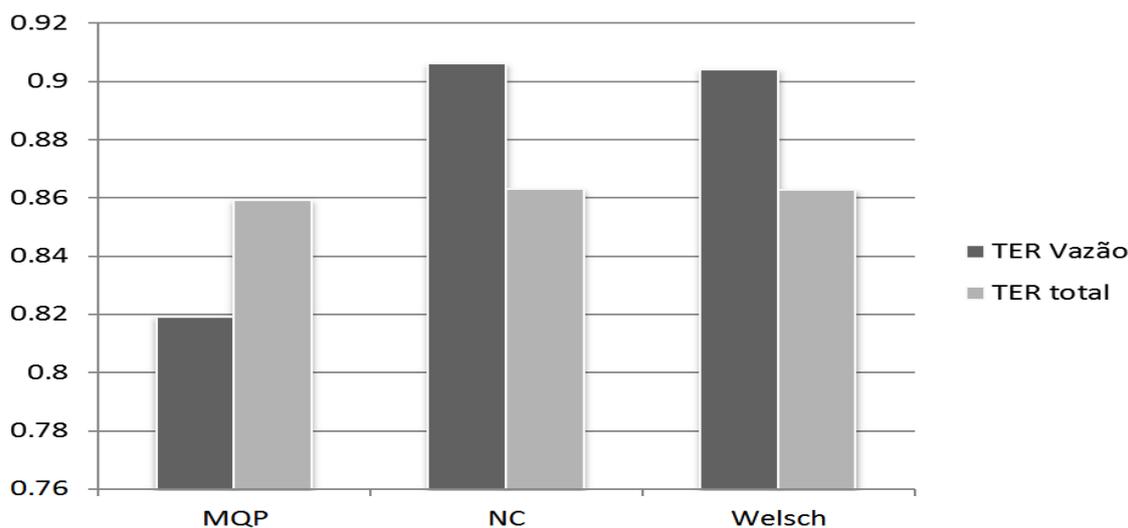


Figura 6.19: Comparação entre os critérios TER de avaliação (BFGS) no estudo de caso 1.

Tabela 6.29: Resultados com o otimizador BFGS - D no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
D [kmol/h] \Rightarrow	69,8256	70,9685	71,4960	70,6320	70,6320
$x_{D,1}$	10,4740	1,0607	1,0724	1,0595	1,0595
$x_{D,2}$	98,9584	10,0574	101,3059	100,0816	100,0816
$x_{D,3}$	39,1217	46,9461*	<u>40,0378</u>	<u>39,5539</u>	<u>39,5539</u>
$x_{D,4}$	30,6263	31,2138	31,3192	30,9407	30,9407
$x_{D,5}$	9,4071	9,3089	9,6122	9,4961	9,4961
$x_{D,6}$	5,1787	5,1020	5,2828	5,2189	5,2189
$x_{D,7}$	4,2089	4,1557	4,2898	4,2379	4,2379
$x_{D,8}$	4,4029	4,3406	4,5479	4,4930	4,4930
$x_{D,9}$	0,9504	0,9626	0,9930	0,9810	0,9810
$x_{D,10}$	0,0388	0,0384	0,0397	0,0392	0,0392
$x_{D,11}$	0,0000	0,0000	0,0000	0,0000	0,0000

* erro grosseiro.

Tabela 6.30: Resultados com o otimizador BFGS - B no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
B [kmol/h] \Rightarrow	114,2424	112,8579	115,5960	113,0760	113,0040
$x_{B,1}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,2}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,3}$	12,6940	12,547	12,844	12,564	12,556
$x_{B,4}$	25,3870	25,818	25,688	25,128	25,112
$x_{B,5}$	85,6820	87,260	86,697	84,807	84,753
$x_{B,6}$	76,1620	75,091	77,064	75,384	75,336
$x_{B,7}$	45,6970	45,0529	46,2384	45,2304	45,2016
$x_{B,8}$	369,3838	376,3223	373,1182	364,9842	364,7518
$x_{B,9}$	1131,9518	1119,3437	1144,7215	1119,7665	1119,0535
$x_{B,10}$	865,7035	855,3165	875,9608	856,8648	856,3192
$x_{B,11}$	740,3542	751,9176	749,1263	732,7953	732,3287

Tabela 6.31: Resultados com o otimizador BFGS - Temperaturas no estudo de caso 1.

Estágio i	Exato	Medido	Dados Reconciliados - EMSO/Matlab		
			MQP	NC	Welsch
0	308,44	307,47	308,45	308,45	308,45
1	366,34	367,52	366,31	366,31	366,31
2	411,89	396,89*	<u>411,89</u>	<u>411,89</u>	<u>411,89</u>
3	435,09	434,09	435,13	435,13	435,13
4	447,56	446,44	447,61	447,61	447,61
5	457,00	458,13	457,02	457,02	457,02
6	469,70	468,62	469,68	469,68	469,68
7	485,93	486,88	485,84	485,84	485,84
8	500,42	499,35	500,3	500,3	500,3
9	511,03	512,16	510,93	510,93	510,93
10	519,35	520,79	519,29	519,29	519,29
11	527,98	529,29	527,92	527,92	527,92

* erro grosseiro.

Tabela 6.32: Avaliação de desempenho com o otimizador BFGS (TER) no estudo de caso 1.

TER	MQP	NC	Welsch
Vazões (Ponto Interior)	0,8194	0,9063	0,9044
TOTAL	0,8595	0,8632	0,8631

- Otimizador *fmincon* (SQP)

Nas Tabelas de 6.33 a 6.36 são apresentados os resultados obtidos para o caso de reconciliação de dados com o otimizador *fmincon* do Matlab que utiliza o método SQP (*Successive Quadratic Programming*), em que se avaliou o modelo MES na presença de variáveis com erros grosseiros. Analogamente ao caso com o otimizador Nelder-Mead, utilizaram-se as mesmas variáveis e com os mesmos erros, conforme mostra a Tabela 6.5.

Tabela 6.33: Resultados com o otimizador SQP - F no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
F [kmol/h] \Rightarrow	184,0680	220,8816*	<u>185,2920</u>	<u>183,7440</u>	<u>183,6720</u>
$x_{F,1}$	1,0737	1,0878	1,0809	1,0718	1,0714
$x_{F,2}$	98,9366	100,8490	101,0356	100,1915	100,1523
$x_{F,3}$	39,2678	39,7448	39,7863	39,4539	39,4385
$x_{F,4}$	30,8825	30,4771	30,5217	30,2667	30,2549
$x_{F,5}$	10,2771	10,1372	10,1396	10,0549	10,0509
$x_{F,6}$	5,9311	5,8544	5,8161	5,7675	5,7653
$x_{F,7}$	8,7944	8,9321	8,9043	8,8299	8,8265
$x_{F,8}$	41,3130	40,7005	40,7642	40,4237	40,4078
$x_{F,9}$	114,1733	112,6523	112,8737	111,9307	111,8869
$x_{F,10}$	86,6142	87,9944	88,1681	87,4315	87,3973
$x_{F,11}$	74,0362	75,1479	75,3006	74,6715	74,6423

* erro grosseiro.

Nas Tabelas 6.33, 6.34 e 6.36 observa-se, novamente, que todos os estimadores detectaram o erro grosseiro nas variáveis corrompidas (F , $x_{D,3}$ e T_2). A Tabela 6.37 apresenta o critério de avaliação de desempenho utilizado neste caso com o otimizador SQP. Observando-se a Tabela 6.37 e a Figura 6.20, conclui-se que o estimador MQP apresentou um desempenho levemente inferior em relação aos estimadores NC e Welsch no TER total, resultado muito parecido com o resultado gerado pelo método de otimização BFGS.

Tabela 6.34: Resultados com o otimizador SQP - D no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
D [kmol/h]⇒	69,8256	70,9685	71,8200	70,6320	70,6320
$x_{D,1}$	10,4740	1,0607	1,0773	1,0595	1,0595
$x_{D,2}$	98,9584	10,0574	101,7650	100,0816	100,0816
$x_{D,3}$	39,1217	46,9461*	<u>40,2192</u>	<u>39,5539</u>	<u>39,5539</u>
$x_{D,4}$	30,6263	31,2138	31,4612	30,9407	30,9407
$x_{D,5}$	9,4071	9,3089	9,6558	9,4961	9,4961
$x_{D,6}$	5,1787	5,1020	5,3067	5,2189	5,2189
$x_{D,7}$	4,2089	4,1557	4,3092	4,2379	4,2379
$x_{D,8}$	4,4029	4,3406	4,5686	4,4930	4,4930
$x_{D,9}$	0,9504	0,9626	0,9975	0,9810	0,9810
$x_{D,10}$	0,0388	0,0384	0,0399	0,0392	0,0392
$x_{D,11}$	0,0000	0,0000	0,0000	0,0000	0,0000

* erro grosseiro.

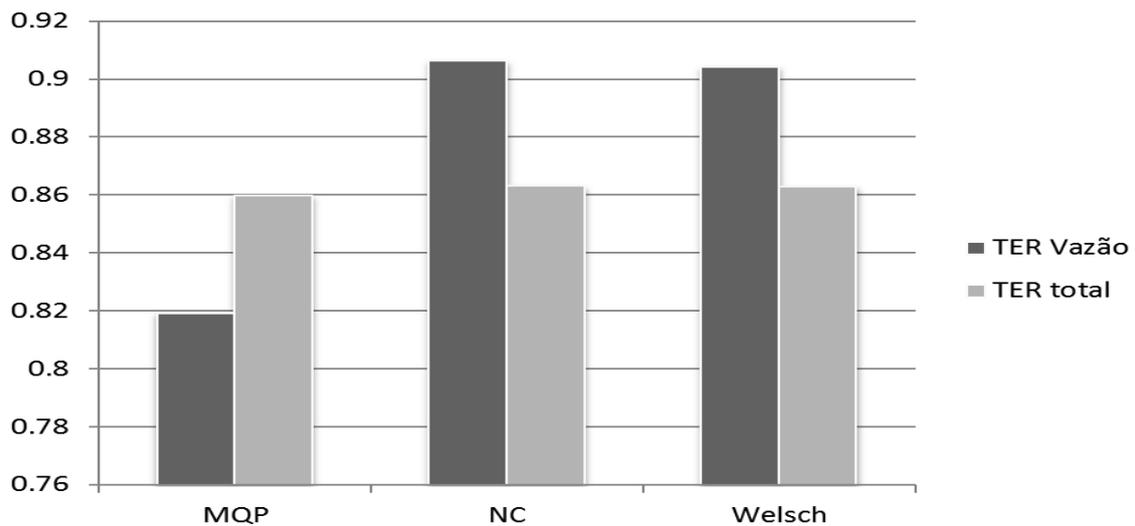


Figura 6.20: Comparação entre os critérios TER de avaliação (SQP) no estudo de caso 1.

Tabela 6.35: Resultados com o otimizador SQP - B no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
B [kmol/h] \Rightarrow	114,2424	112,8579	113,4720	113,0760	113,0040
$x_{B,1}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,2}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,3}$	12,6940	12,547	12,608	12,564	12,556
$x_{B,4}$	25,3870	25,818	25,216	25,128	25,112
$x_{B,5}$	85,6820	87,260	85,104	84,807	84,753
$x_{B,6}$	76,1620	75,091	75,648	75,384	75,336
$x_{B,7}$	45,6970	45,0529	45,3888	45,2304	45,2016
$x_{B,8}$	369,3838	376,3223	366,2624	364,9842	364,7518
$x_{B,9}$	1131,9518	1119,3437	1123,6880	1119,7665	1119,0535
$x_{B,10}$	865,7035	855,3165	859,8656	856,8648	856,3192
$x_{B,11}$	740,3542	751,9176	735,3616	732,7953	732,3287

Tabela 6.36: Resultados com o otimizador SQP - Temperaturas no estudo de caso 1.

Estágio i	Exato	Medido	Dados Reconciliados - EMSO/Matlab		
			MQP	NC	Welsch
0	308,44	307,47	308,45	308,45	308,45
1	366,34	367,52	366,31	366,31	366,31
2	411,89	396,89*	<u>411,89</u>	<u>411,89</u>	<u>411,89</u>
3	435,09	434,09	435,13	435,13	435,13
4	447,56	446,44	447,61	447,61	447,61
5	457,00	458,13	457,02	457,02	457,02
6	469,70	468,62	469,68	469,68	469,68
7	485,93	486,88	485,84	485,84	485,84
8	500,42	499,35	500,30	500,30	500,30
9	511,03	512,16	510,93	510,93	510,93
10	519,35	520,79	519,29	519,29	519,29
11	527,98	529,29	527,92	527,92	527,92

* erro grosseiro.

Tabela 6.37: Avaliação de desempenho com o otimizador SQP (TER) no estudo de caso 1.

TER	MQP	NC	Welsch
Vazões (Ponto Interior)	0,8194	0,9063	0,9044
TOTAL	0,8599	0,8632	0,8631

- Otimizador *fmincon* (*Interior-Point*)

Nas Tabelas de 6.38 a 6.41 são apresentados os resultados obtidos para o caso de reconciliação de dados com o otimizador *fmincon* do Matlab, que utiliza o método Ponto Interior, em que se avaliou o modelo MES na presença de variáveis com erros grosseiros. Analogamente ao caso com o otimizador Nelder-Mead, utilizaram-se as mesmas variáveis e com os mesmos erros, conforme a Tabela 6.5.

Tabela 6.38: Resultados com o otimizador Ponto Interior - F no estudo de caso 1.

		Dados Reconciliados - EMSO/Matlab				
		Exato	Medido	MQP	NC	Welsch
F [kmol/h]⇒		184,0680	220,8816*	<u>186,2280</u>	<u>183,7080</u>	<u>183,6720</u>
Componentes [kmol/s]×10 ⁴	$x_{F,1}$	1,0737	1,0878	1,0863	1,0716	1,0714
	$x_{F,2}$	98,9366	100,8490	101,5460	100,1719	100,1523
	$x_{F,3}$	39,2678	39,7448	39,9873	39,4462	39,4385
	$x_{F,4}$	30,8825	30,4771	30,6759	30,2608	30,2549
	$x_{F,5}$	10,2771	10,1372	10,1908	10,0529	10,0509
	$x_{F,6}$	5,9311	5,8544	5,8455	5,8174	5,7653
	$x_{F,7}$	8,7944	8,9321	8,9493	8,8282	8,8265
	$x_{F,8}$	41,3130	40,7005	40,9702	40,4158	40,4078
	$x_{F,9}$	114,1733	112,6523	113,4439	111,9088	111,8869
	$x_{F,10}$	86,6142	87,9944	88,6135	87,4144	87,3973
	$x_{F,11}$	74,0362	75,1479	75,6810	74,6569	74,6423

* erro grosseiro.

Nas Tabelas 6.38, 6.39 e 6.41 observa-se que todos os estimadores detectaram o erro grosseiro nas variáveis corrompidas (F , $x_{D,3}$ e T_2). A Tabela 6.42 apresenta o critério de avaliação de desempenho utilizado neste caso com o otimizador Ponto Interior. Observando a Tabela 6.42 e a Figura 6.21, conclui-se que o estimador NC apresentou um desempenho superior em relação aos estimadores MQP e Welsch no TER total. Além disso, o TER para o estimador NC foi o melhor resultado entre todos os métodos de otimização e estimadores.

A Tabela 6.43 apresenta o resumo dos resultados obtidos para todos os casos em relação ao critério TER total.

Tabela 6.39: Resultados com o otimizador Ponto Interior - D no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
D [kmol/h] \Rightarrow	69,8256	70,9685	72,1800	70,6680	70,6680
$x_{D,1}$	10,4740	1,0607	1,0827	1,0600	1,0600
$x_{D,2}$	98,9584	10,0574	102,2751	100,1326	100,1326
$x_{D,3}$	39,1217	46,9461*	<u>40,4208</u>	<u>39,5741</u>	<u>39,5741</u>
$x_{D,4}$	30,6263	31,2138	31,6189	30,9565	30,9565
$x_{D,5}$	9,4071	9,3089	9,7042	9,5009	9,5009
$x_{D,6}$	5,1787	5,1020	5,3333	5,2216	5,2216
$x_{D,7}$	4,2089	4,1557	4,3308	4,2401	4,2401
$x_{D,8}$	4,4029	4,3406	4,5915	4,4953	4,4953
$x_{D,9}$	0,9504	0,9626	1,0025	0,9815	0,9815
$x_{D,10}$	0,0388	0,0384	0,0401	0,0393	0,0393
$x_{D,11}$	0,0000	0,0000	0,0000	0,0000	0,0000

* erro grosseiro.

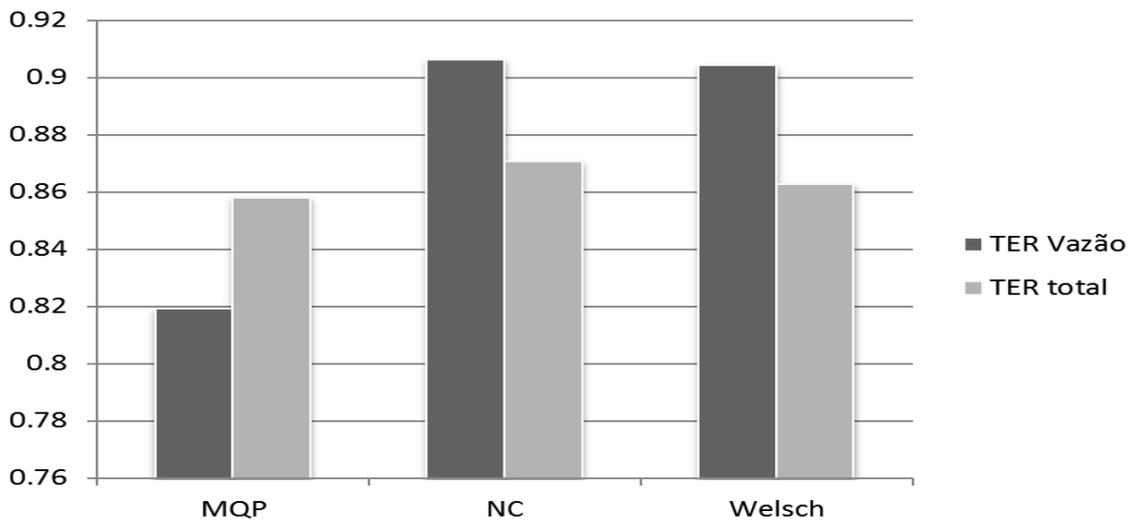


Figura 6.21: Comparação entre os critérios TER de avaliação (Ponto Interior) no estudo de caso 1.

Tabela 6.40: Resultados com o otimizador Ponto Interior - B no estudo de caso 1.

	Dados Reconciliados - EMSO/Matlab				
	Exato	Medido	MQP	NC	Welsch
B [kmol/h] \Rightarrow	114,2424	112,8579	114,0120	113,0040	112,9680
$x_{B,1}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,2}$	0,0000	0,0000	0,0000	0,0000	0,0000
$x_{B,3}$	12,6940	12,547	12,668	12,556	12,552
$x_{B,4}$	25,3870	25,818	25,336	25,112	25,104
$x_{B,5}$	85,6820	87,260	85,509	84,753	84,726
$x_{B,6}$	76,1620	75,091	76,008	75,336	75,312
$x_{B,7}$	45,6970	45,0529	45,6048	45,2016	45,1872
$x_{B,8}$	369,3838	376,3223	368,0054	364,7518	364,6356
$x_{B,9}$	1131,9518	1119,3437	1129,0355	1119,0535	1118,6970
$x_{B,10}$	865,7035	855,3165	863,9576	856,3192	856,0464
$x_{B,11}$	740,3542	751,9176	738,8611	732,3287	732,0954

Tabela 6.41: Resultados com o otimizador Ponto Interior - Temperaturas no estudo de caso 1.

Estágio i	Exato	Medido	Dados Reconciliados - EMSO/Matlab		
			MQP	NC	Welsch
0	308,44	307,47	308,45	308,45	308,45
1	366,34	367,52	366,31	366,31	366,31
2	411,89	396,89*	<u>411,89</u>	<u>411,89</u>	<u>411,89</u>
3	435,09	434,09	435,13	435,13	435,13
4	447,56	446,44	447,61	447,61	447,61
5	457,00	458,13	457,02	457,02	457,02
6	469,70	468,62	469,68	469,68	469,68
7	485,93	486,88	485,84	485,84	485,84
8	500,42	499,35	500,30	500,30	500,30
9	511,03	512,16	510,93	510,93	510,93
10	519,35	520,79	519,29	519,29	519,29
11	527,98	529,29	527,92	527,92	527,92

* erro grosseiro.

Tabela 6.42: Avaliação de desempenho com o otimizador Ponto Interior (TER) no estudo de caso 1.

TER	MQP	NC	Welsch
Vazões (Ponto Interior)	0,8194	0,9063	0,9044
TOTAL	0,8581	0,8708	0,8630

Tabela 6.43: Resumo das avaliações de desempenho (TER total) no estudo de caso 1.

Otimizadores	MQP	NC	Welsch
Simplex	0,8603	0,8630	0,8630
PSO	0,8642	0,7021	0,8170
BFGS	0,8595	0,8632	0,8631
SQP	0,8599	0,8632	0,8631
Ponto Interior	0,8581	0,8708	0,8630

Avaliação dos otimizadores

Nas Figuras de 6.22 a 6.27 são apresentados os resultados obtidos para analisar o desempenho dos métodos de otimização. Utilizou-se como critério para análise de desempenho o valor da função objetivo (Fobj) e o tempo gasto para o término da simulação. Para o critério Fobj, quanto menor o valor melhor o desempenho do método de otimização. Para o critério tempo gasto, quanto menor o tempo, melhor e mais rápido é o método de otimização. Todas as rotinas de otimização convergiram e suas paradas ocorreram por critério de tolerância atingida ou por número de iterações limites. Como os valores das Fobj's ficaram muito grande (entre 3.936.513.788,40 e 3.936.513.789,40) para os casos com o estimador de MQP, subtraiu-se o valor 3.936.513.780,0 das Fobj's, deixando apenas a variação numérica relevante (entre 7,80 e 10,00), conforme ilustra a Figura 6.22.

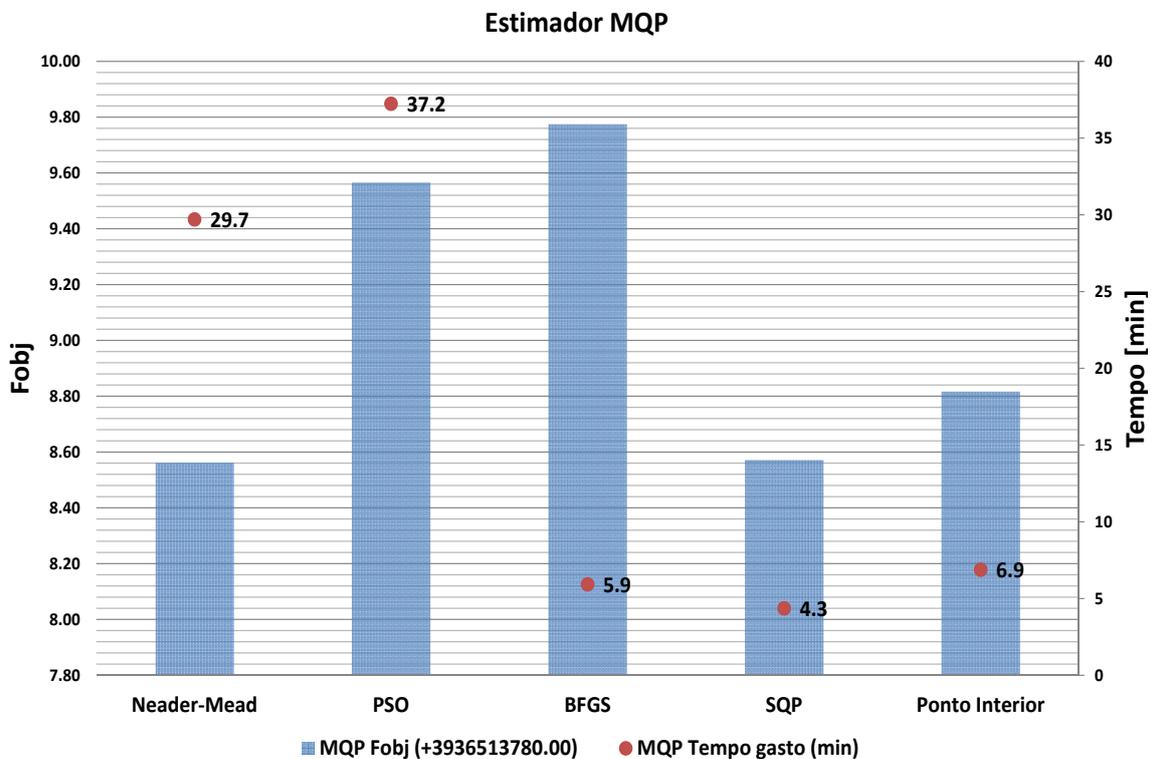


Figura 6.22: Comparação entre os otimizadores com os critérios Fobj e tempo de simulação (MQP).

Observa-se que, em relação ao método de otimização mais rápido, o otimizador *fmincon* utilizando o algoritmo SQP foi o mais rápido que todos os demais estimadores (Função Objetivo). Por outro lado, o otimizador PSO obteve o pior tempo, com um tempo gasto em média de 30 minutos de simulação. Vale ressaltar que o tempo gasto em uma simulação do modelo no EMSO com o integrador *dasslc* é em torno de 9 segundos, assim uma rotina de otimização que analisa intensamente

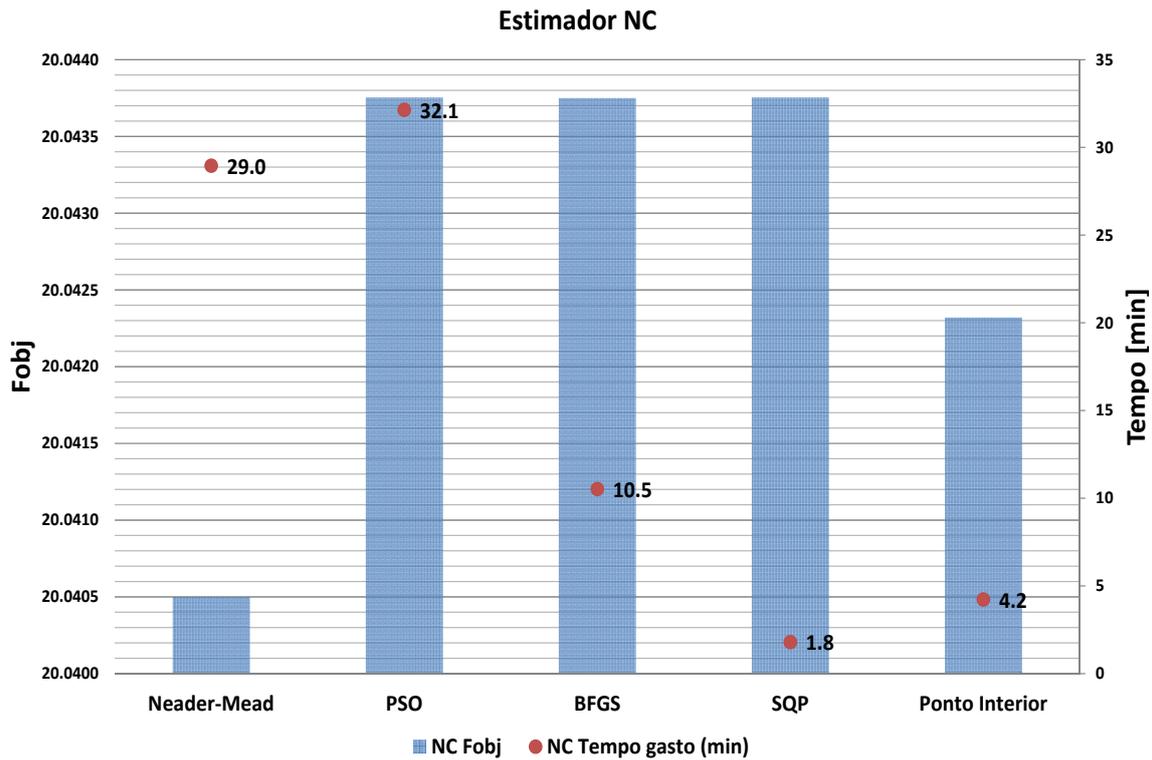


Figura 6.23: Comparação entre os otimizadores com os critérios Fobj e tempo de simulação (NC).

uma função objetivo como o PSO, sendo aplicado na abordagem “*feasible path*”, o modelo será rodado tantas de vezes que o otimizador precisar calcular a função objetivo. Além disso, o número de iterações com a função objetivo depende de 2 parâmetros fundamentais: o número de gerações (8) e número de partículas (100); por isso, o método PSO obteve o pior tempo.

Os métodos que utilizam derivadas foram mais rápidos. Esses métodos determinísticos geram pequenas perturbações nas variáveis de entrada para produzir a Jacobiana numérica. Porém, essas pequenas perturbações são simuladas em um tempo mais rápido que uma perturbação grande em relação ao valor medido. Deve-se notar, que os métodos determinísticos perturbam as variáveis de decisão, que são apenas 12 variáveis de entrada ($F, D, x_{F,1}, \dots, x_{F,10}$). O método Nelder-Mead apresentou o menor valor de função objetivo; entretanto, os valores das Fobj estão muito próximos, o que pode indicar uma região comum de convergência. Outras estimativas iniciais não foram testadas para confirmar se foi um caso de mínimo local. Uma possibilidade de teste seria, por exemplo, usar o valor ótimo do método de Nelder-Mead para confirmar os mínimos locais dos métodos determinísticos.

Nas Figuras 6.25, 6.26 e 6.27, observa-se que a maioria dos métodos de otimização resultaram em um valor do TER próximo, exceto o otimizador PSO,

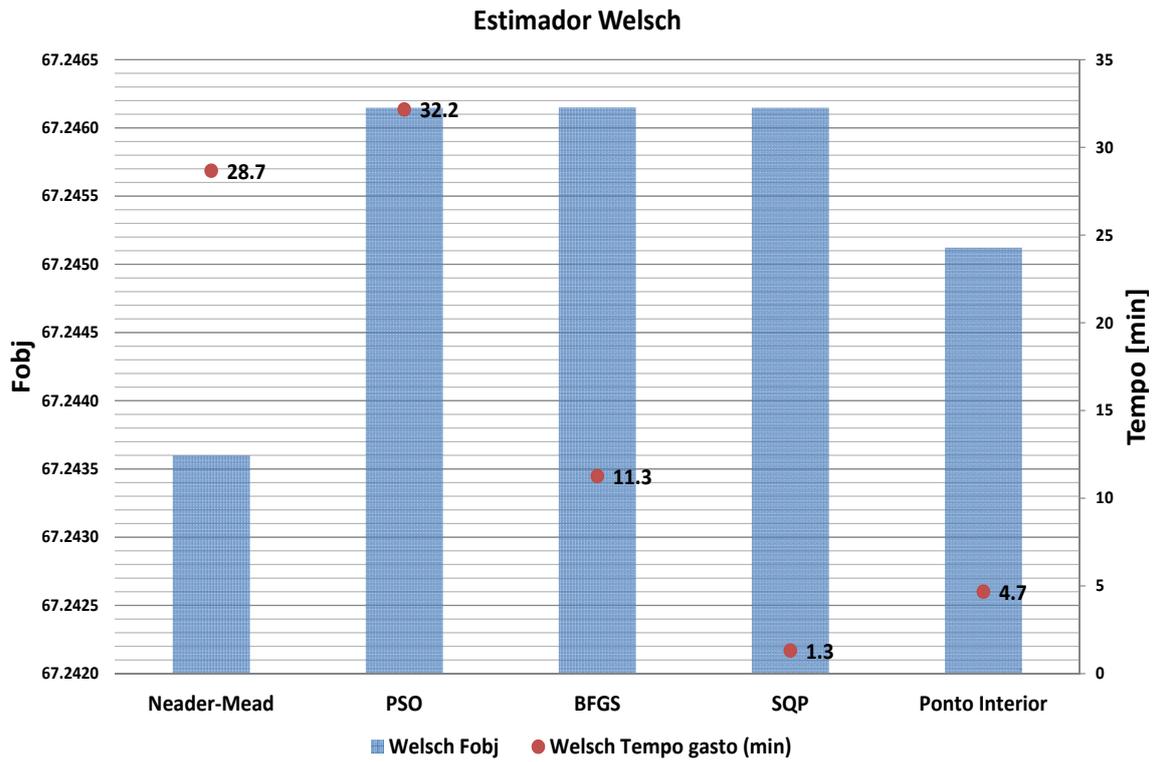


Figura 6.24: Comparação entre os otimizadores com os critérios Fobj e tempo de simulação (Welsch).

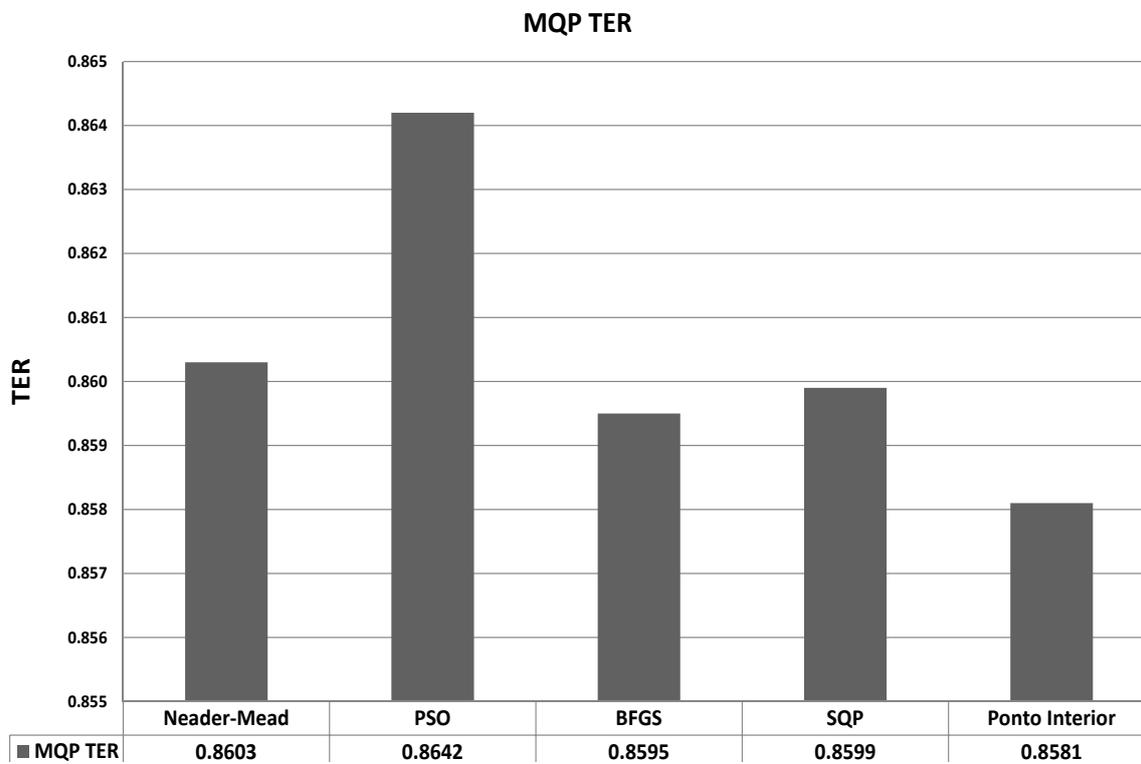


Figura 6.25: Comparação entre os otimizadores com o critério TER (MQP).

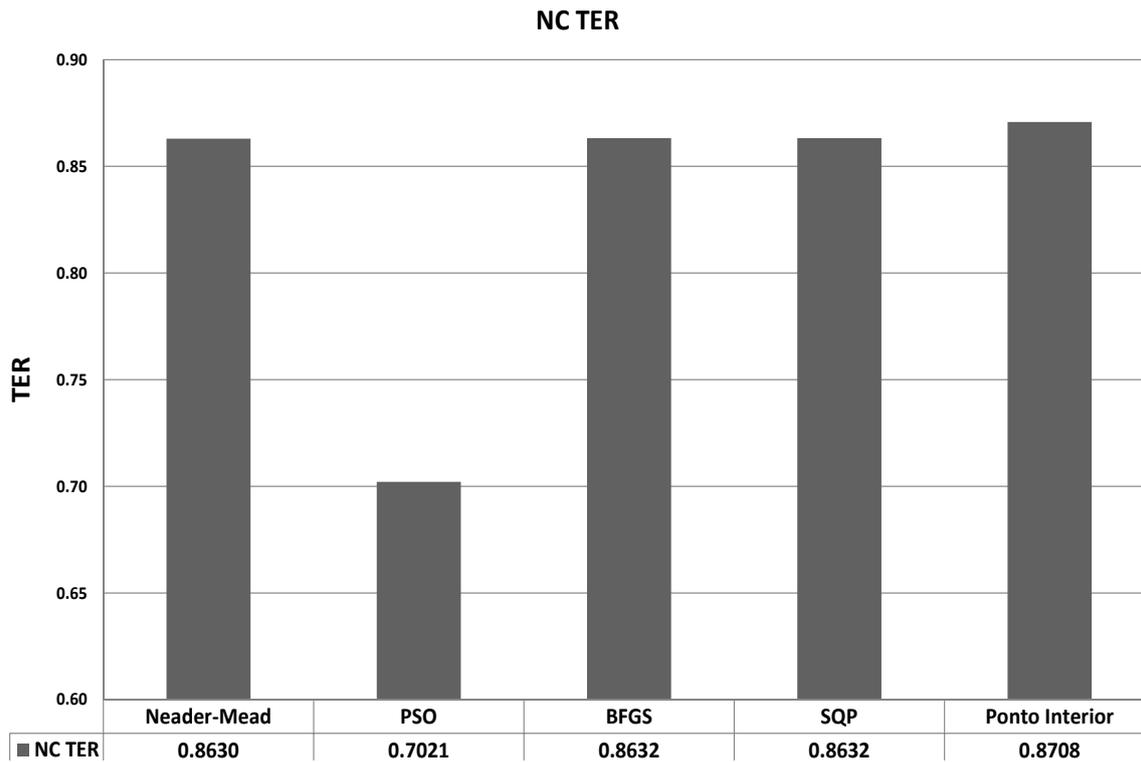


Figura 6.26: Comparação entre os otimizadores com o critério TER (NC).

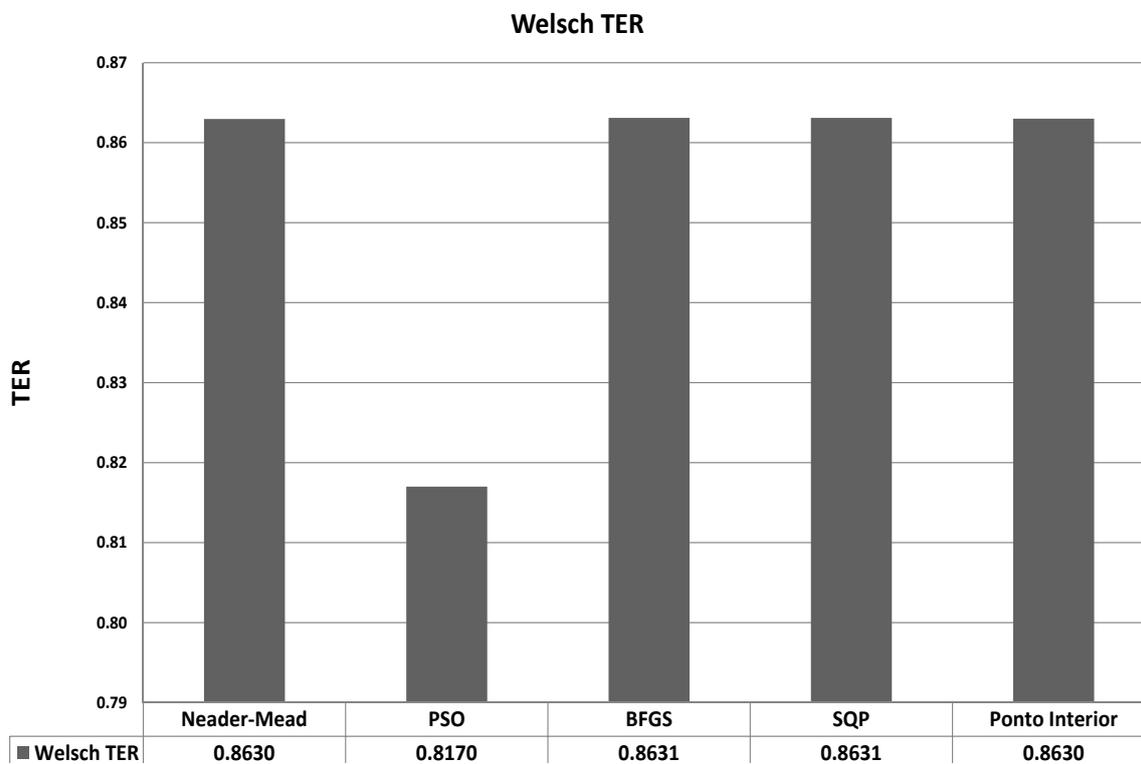


Figura 6.27: Comparação entre os otimizadores com o critério TER (Welsch).

que obteve valores discrepantes. Para o estimador MQP, obteve-se o melhor valor de TER (0,8642); para o estimador NC e Welsch, obtiveram-se os piores valores de TER (0,7021 e 0,8170), respectivamente. Como só foi feita uma rodada de simulação com o otimizador PSO (estocástico), não é possível analisar o caso do TER (0,8642) com propriedade. Seria necessário observar o seu comportamento em média para entender a probabilidade de um valor de TER com os métodos estocásticos.

6.2.2 Caso 2: Avaliação da redução de ordem do modelo de coluna

A estratégia utilizada no procedimento de RD foi reconciliar as vazões globais em um primeiro passo, em que o modelo (restrição do problema) foi inserido na função objetivo como uma penalidade (como no Caso 1). Com isso, o otimizador pôde manipular a terceira variável (a vazão de fundo B). Essa estratégia foi implementada a fim de detectar e eliminar os efeitos de erros grosseiros nas variáveis de entrada. Assim, o valor reconciliado da primeira etapa tornou-se o valor fixo (já reconciliado e ausente da segunda etapa da reconciliação). Utilizou-se o mesmo método de otimização na primeira etapa e na segunda etapa, a rotina *fmincon* (algoritmo Ponto Interior).

Para a segunda etapa, utilizou-se a concentração de propano como a variável de decisão da otimização, já que a concentração de propano é calculada com a diferença da soma das concentrações igual a um. Assim, a restrição das somas das concentrações é sempre respeitada, e com isso a otimização fica com apenas uma variável manipulada (dimensão 1). As variáveis de saída são recebidas do EMSO para o Matlab pela interface e calculadas nas funções objetivo.

Critério TER

A Figura 6.28 apresenta o critério de avaliação de desempenho utilizado neste caso. O critério TER foi selecionado para a avaliação, pois não depende do estimador utilizado e deve ser utilizado quando as variáveis exatas são conhecidas.

Observando a Figura 6.28, conclui-se que o problema de RD resolvido com o modelo completo mostrou-se mais eficiente em reduzir os erros aleatórios e grosseiros. Entretanto, os valores considerados exatos foram gerados com o modelo completo, o que explica os melhores TER utilizando o modelo completo. O modelo reduzido apresentou valores bem próximos aos valores do modelo completo, conforme ilustra a Tabela 6.11; porém há uma variável com um erro bem considerável ($\mathbf{x}_{D,2}$), sinalizada

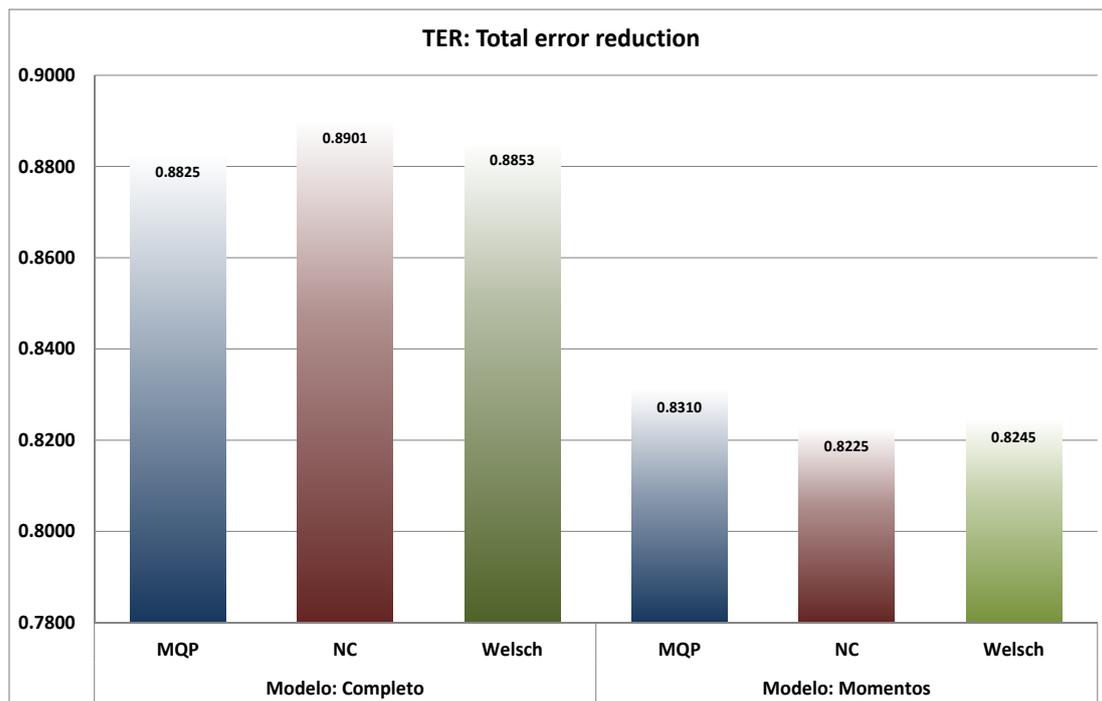


Figura 6.28: Comparação entre os critérios TER no estudo de caso 2.

em negrito na Tabela 6.11. Com isso, a eficiência do método dos momentos ficou prejudicada no procedimento de RD.

Analisando o TER entre os estimadores, observaram-se pequenas diferenças na redução total dos erros e pouca relação entre os modelos. Os melhores resultados da avaliação TER foram obtidos com os estimadores robusto (Welsch $TER = 0,8853$) e quasi-robusto (Normal Contaminada $TER = 0,8901$), utilizando o modelo completo. O estimador MQP resultou no pior valor já que não é robusto, pois na presença de erros grosseiros o estimador MQP apresenta resultados inferiores aos estimadores robustos.

As Figuras 6.29 e 6.30 ilustram o perfil de temperatura com os valores medidos, valores exatos e os valores reconciliados com o estimador Normal Contaminada (NC) utilizando o modelo completo e o modelo de ordem reduzida, respectivamente. Para ilustrar, escolheu-se o valor reconciliado do estimador NC, porque foi o que resultou no melhor TER, quando calculado só com as variáveis de temperatura: ($TER_{temp} = 0,9846$) modelo completo e ($TER_{temp} = 0,9660$) modelo reduzido. As Figuras 6.29 e 6.30 ilustram bem a reconciliação dos erros aleatórios, assim como a detecção e correção do erro grosseiro pelo estimador NC.

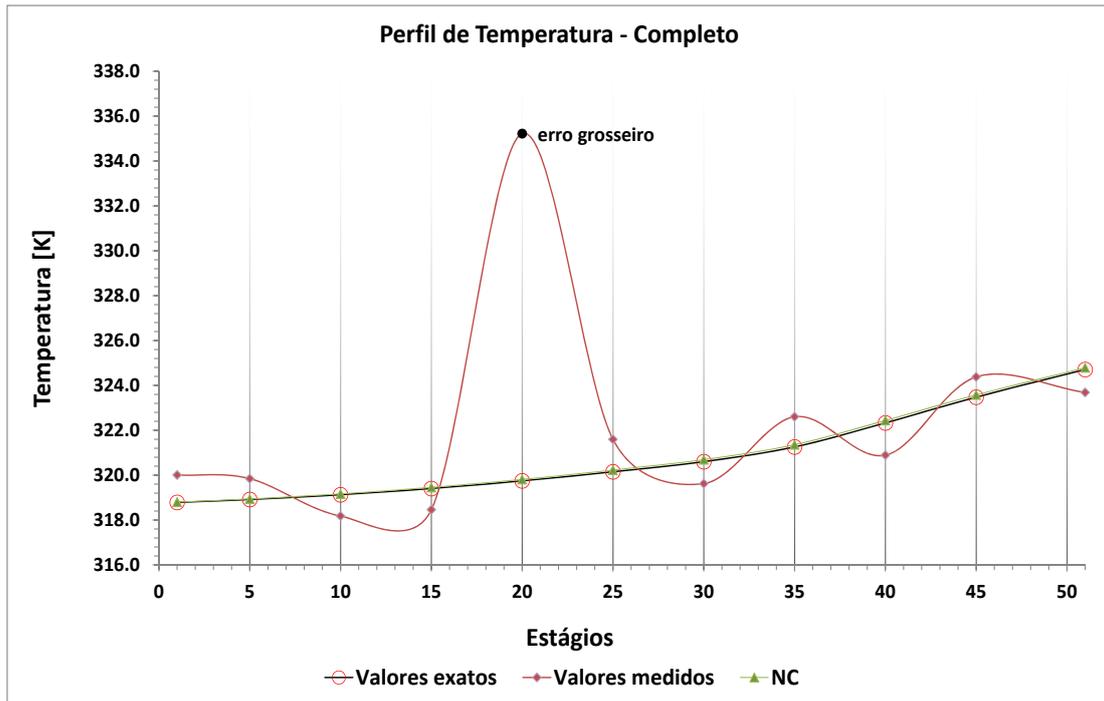


Figura 6.29: Perfil de temperatura com modelo completo: dados exatos, medidos e reconciliados no estudo de caso 2.

Avaliação da função objetivo e número de iterações

A Figura 6.31 ilustra o valor mínimo encontrado da função objetivo e o número de iteração na otimização, com os três estimadores e utilizando os dois modelos. Observam-se pequenas mudanças no valor da função objetivo, o que tudo indica é que a otimização caminha coerentemente para um ponto ótimo tanto para o modelo completo quanto para o modelo reduzido, quando comparado entre as mesma Fobjjs. Por outro lado, observa-se que para um mesmo estimador (no caso MQP) o valor mínimo da função objetivo não significou que os valores reconciliados foram melhores, pois nesse caso o estimador MQP em conjunto com o modelo reduzido encontrou resultados que apresentaram valores da função objetivo menores, e que ao mesmo tempo apresentaram um valor de TER pior. Observando o número de iterações, é notório que em média o procedimento de reconciliação com modelo reduzido teve mais iterações; ou seja, calculou-se mais vezes o modelo e função objetivo até atingir o critério de parada do otimizador.

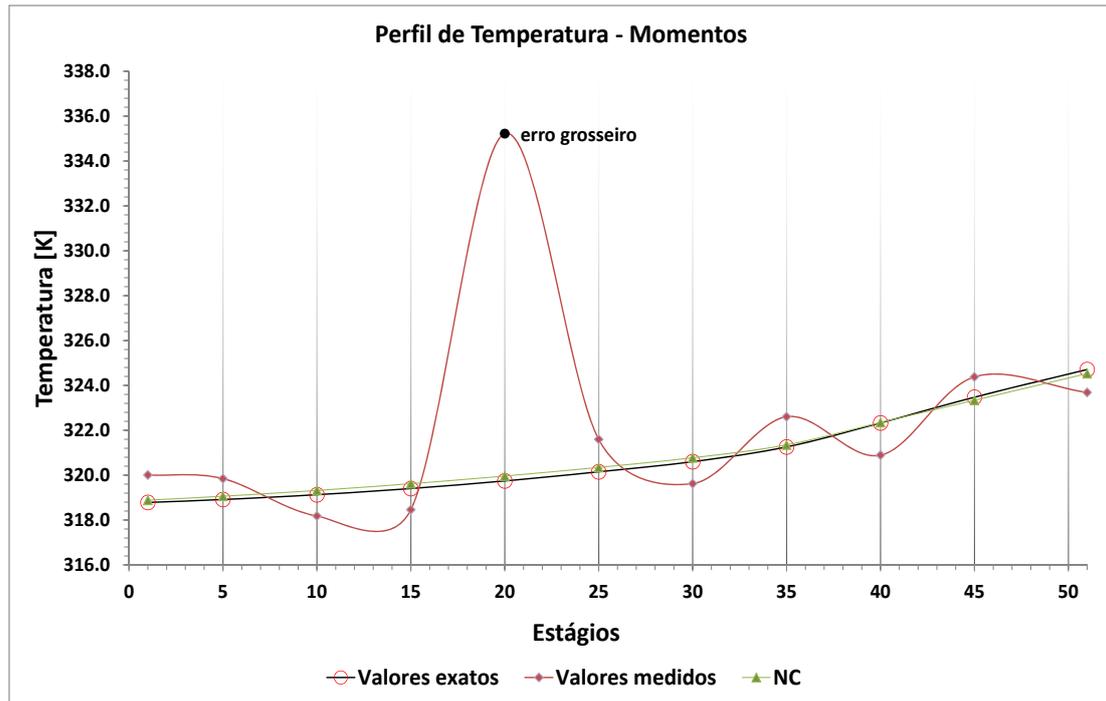


Figura 6.30: Perfil de temperatura com modelo reduzido: dados exatos, medidos e reconciliados no estudo de caso 2.

Avaliação do tempo total de otimização e do tempo gasto em cada iteração

A Figura 6.32 ilustra o valor total gasto pelo procedimento de RD em minutos, assim como o tempo em segundos gasto em média em cada iteração. Observa-se claramente que o procedimento de RD com o modelo reduzido foi o mais rápido, com exceção do estimador Welsch, que resultou em custo computacional inusitado. O procedimento de RD utilizando o modelo reduzido com o estimador de Welsch apresentou uma anomalia durante as iterações, pois durante certa iteração a otimização apresentou uma lentidão incomum e diferente do tempo gasto entre outras iterações no mesmo processo de otimização. Isso acabou gerando um custo grande para o processo como um todo, aumentando a média do tempo gasto por cada iteração. O procedimento de RD utilizando o modelo reduzido com o estimador NC apresentou o melhor tempo e o menor custo computacional.

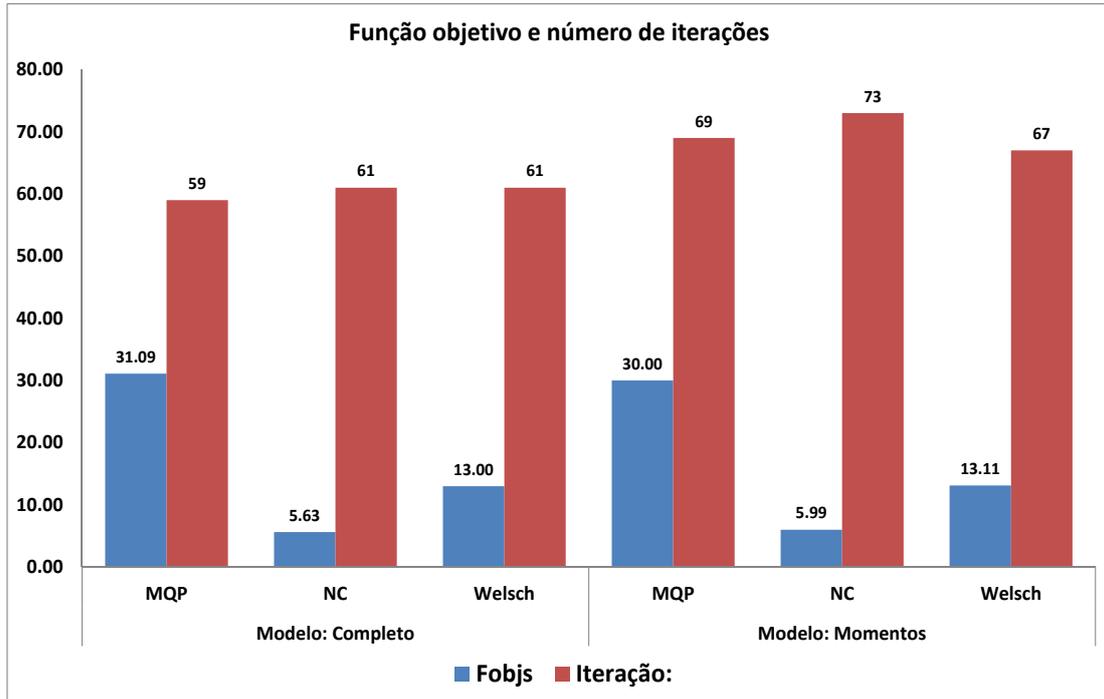


Figura 6.31: Comparação entre os valores de Função objetivo e número de iterações no estudo de caso 2.

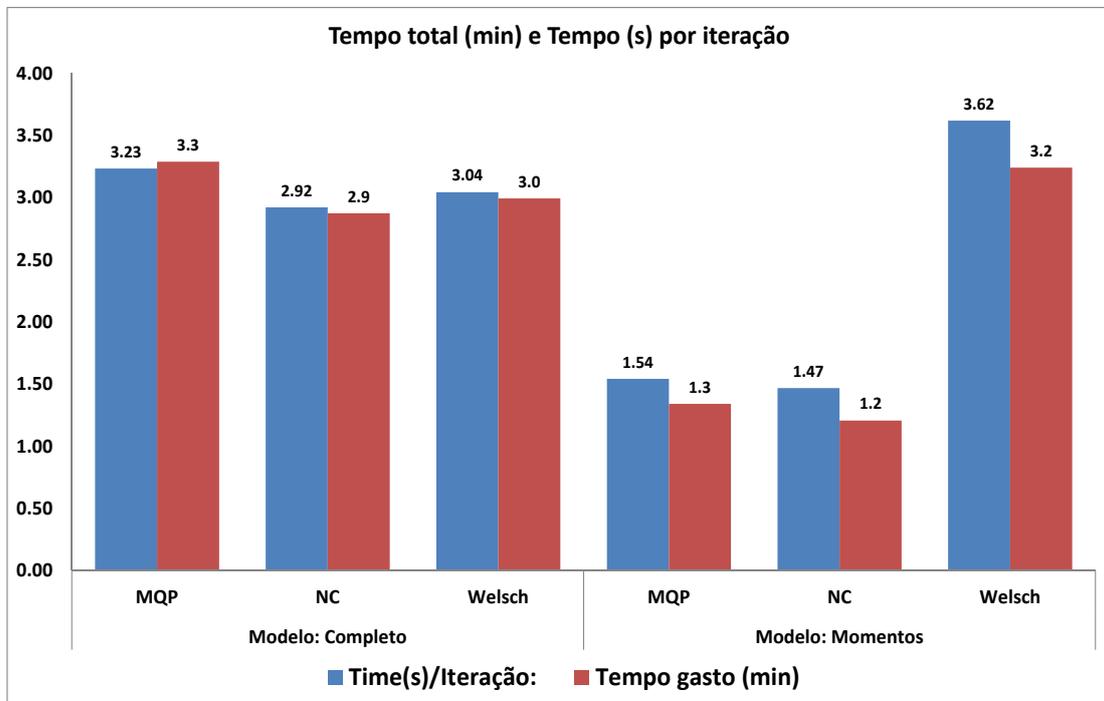


Figura 6.32: Comparação entre os tempos totais de otimização e os tempos médios de cada iteração no estudo de caso 2.

6.2.3 Caso 3: Avaliação do termo de similaridade com modelo autoregressivo

Para este caso, o problema de RDDR foi solucionado utilizando apenas o valor presente (ou atual) da janela como variável de otimização. Os demais pontos da janela (pontos passados) mantiveram-se inalterados, porém não constantes. Dentre as variáveis medidas, apenas as variáveis de entradas foram utilizadas como variável de otimização. Então F , D e z_1 foram selecionadas como variáveis de otimização. A variável z_2 foi calculada pela restrição de soma das concentrações.

As estimativas utilizadas para as variáveis de otimização foram calculadas de acordo com a seguinte regra: 20% do ponto atual medido + 80% do ponto atual reconciliado da janela anterior. Isso faz com que a estimativa não se aproxime tanto do valor medido, dando maior confiança para o último valor ótimo. Dessa forma o tempo de otimização por janela foi mais rápido, pois evita iterações desnecessárias até o ponto ótimo.

Para os limites das variáveis de otimização, foi feita uma regra também. O intervalo das variáveis de otimização segue uma lógica móvel; ou seja, as fronteiras das variáveis de otimização mudam de acordo com a seguinte regra: 60% do valor do último ponto reconciliado + 40% da mediana dos valores medidos na janela $\pm 1.96 \times \sigma_i$. Além disso, a regra apresenta uma lógica *if* para a fronteira expandir por todo o espaço dimensional, no caso do valor medido ser maior ou menor que a mediana da janela, como já detalhado na subseção 5.1.6.

Os resultados são ilustrados nas Figuras 6.33 a 6.44 e na Tabela 6.44. Foram considerados erros grosseiros valores com magnitude de erro acima de 3σ em relação ao valor exato.

Tabela 6.44: Resultados de tempo e número de iterações do caso SEFERLIS e HRYMAK (1994) - no estudo de caso 2 parte II.

Tempo total das otimizações	13,7057 h
Tempo médio gasto na otimização por janela	2,8162 min/janela
Número total de iterações	12.982 iterações
Tempo médio de cada iteração	3,8007 s/iteração
Número de iterações média por janela	44,4589 iterações/janela

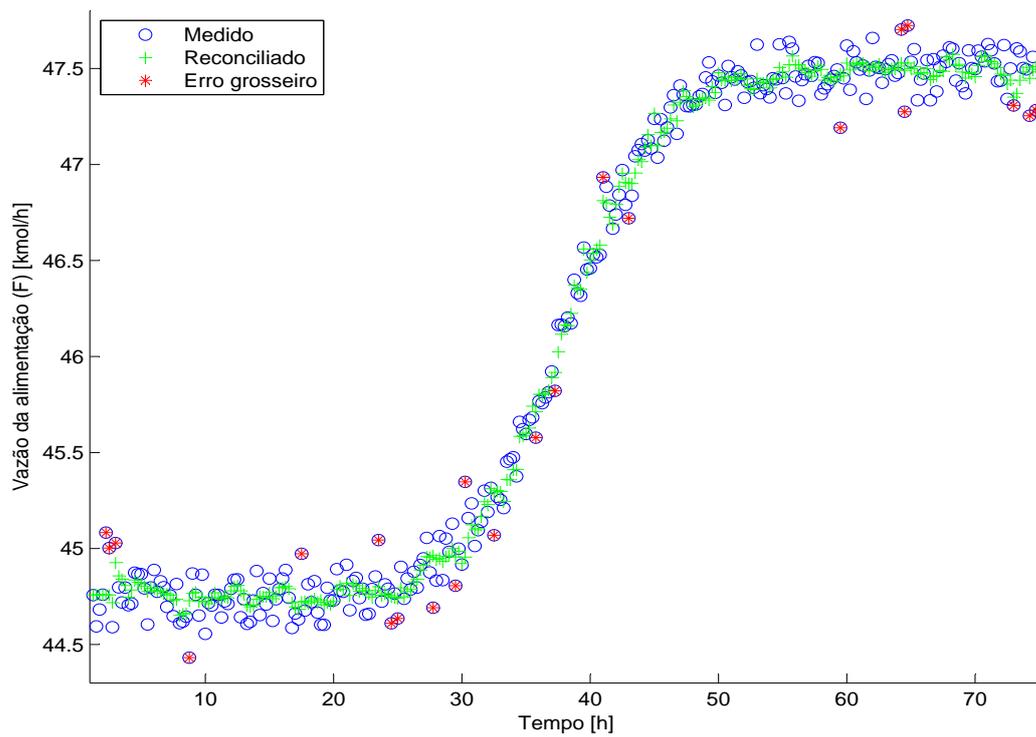


Figura 6.33: Resultados - Vazão de alimentação F no estudo de caso 2 parte II.

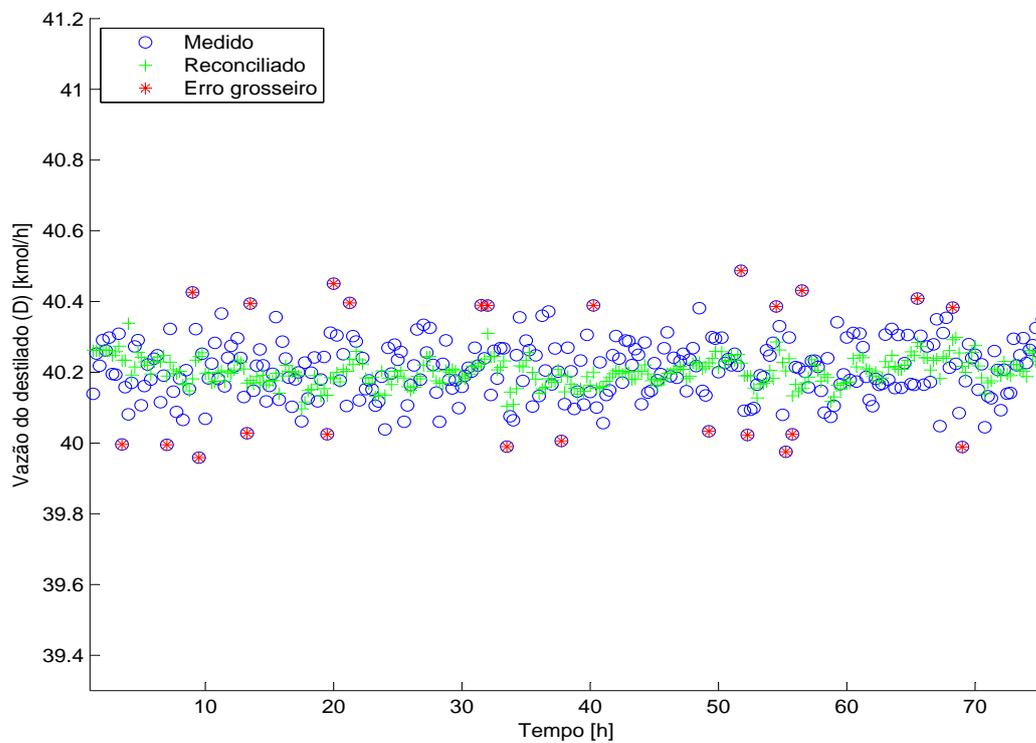


Figura 6.34: Resultados - Vazão de destilado D no estudo de caso 2 parte II.

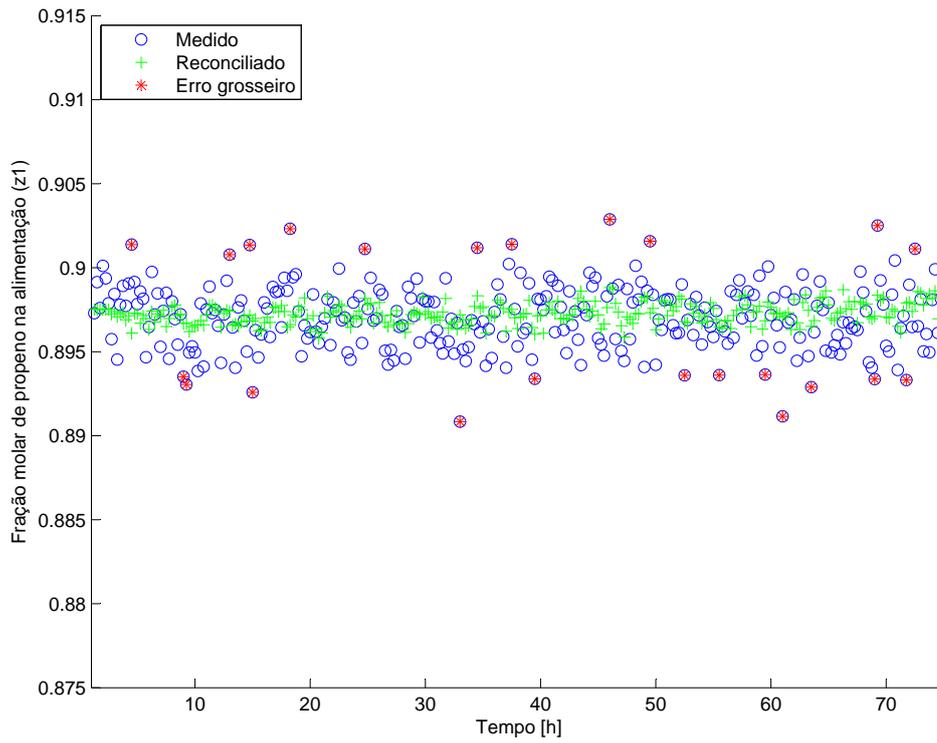


Figura 6.35: Resultados - Fração molar de propeno na alimentação z_1 no estudo de caso 2 parte II.

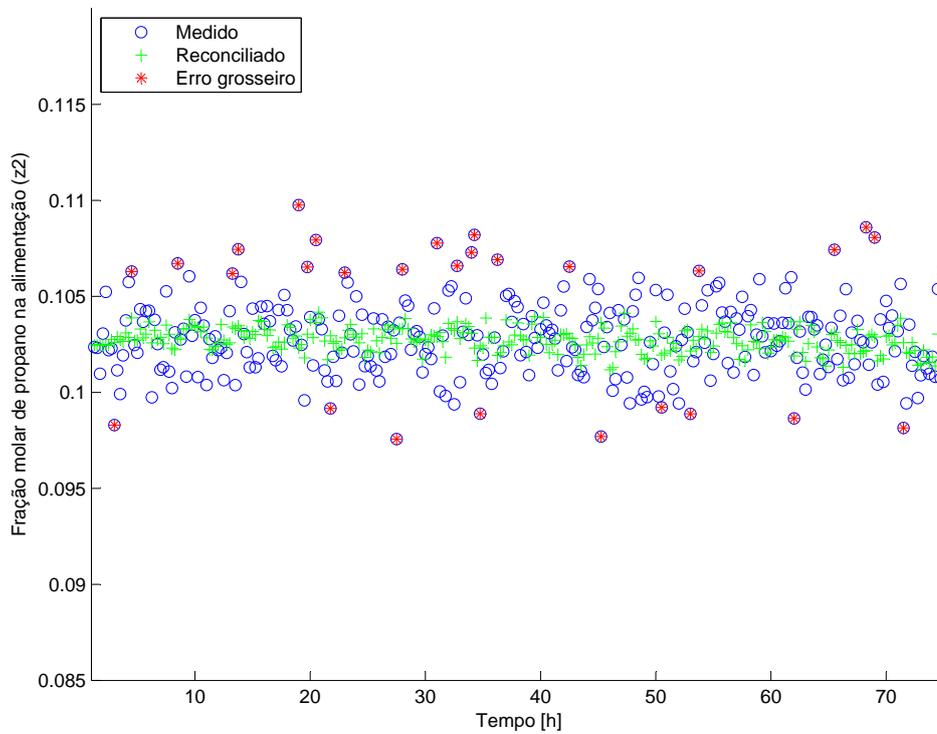


Figura 6.36: Resultados - Fração molar de propano na alimentação z_2 no estudo de caso 2 parte II.

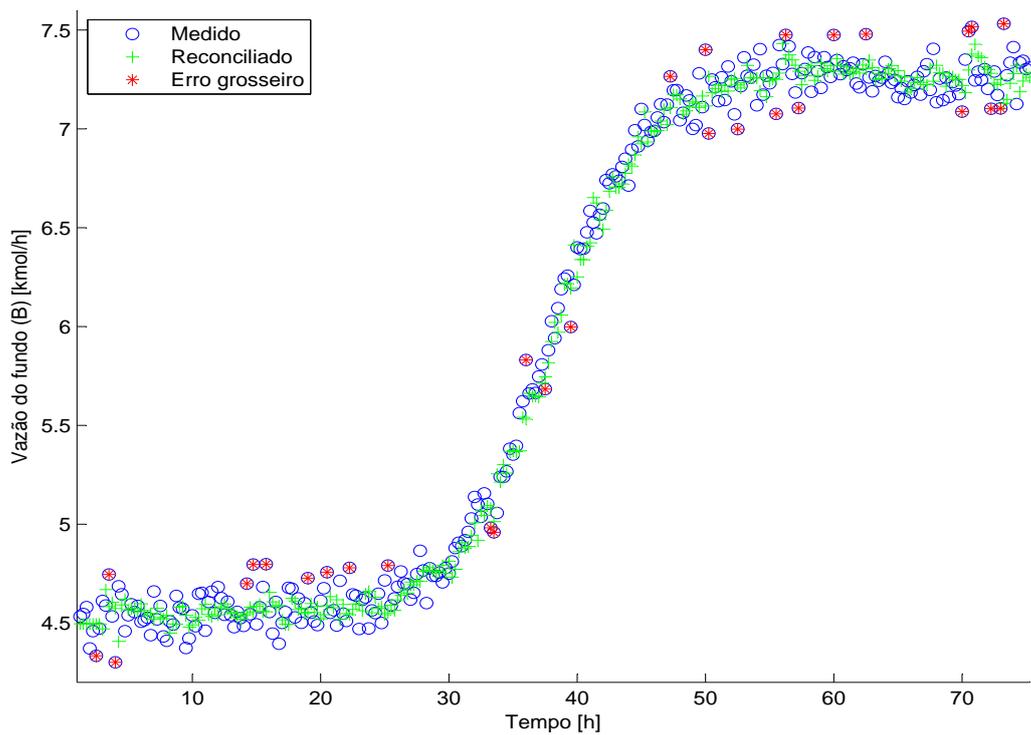


Figura 6.37: Resultados - Vazão do fundo B no estudo de caso 2 parte II.

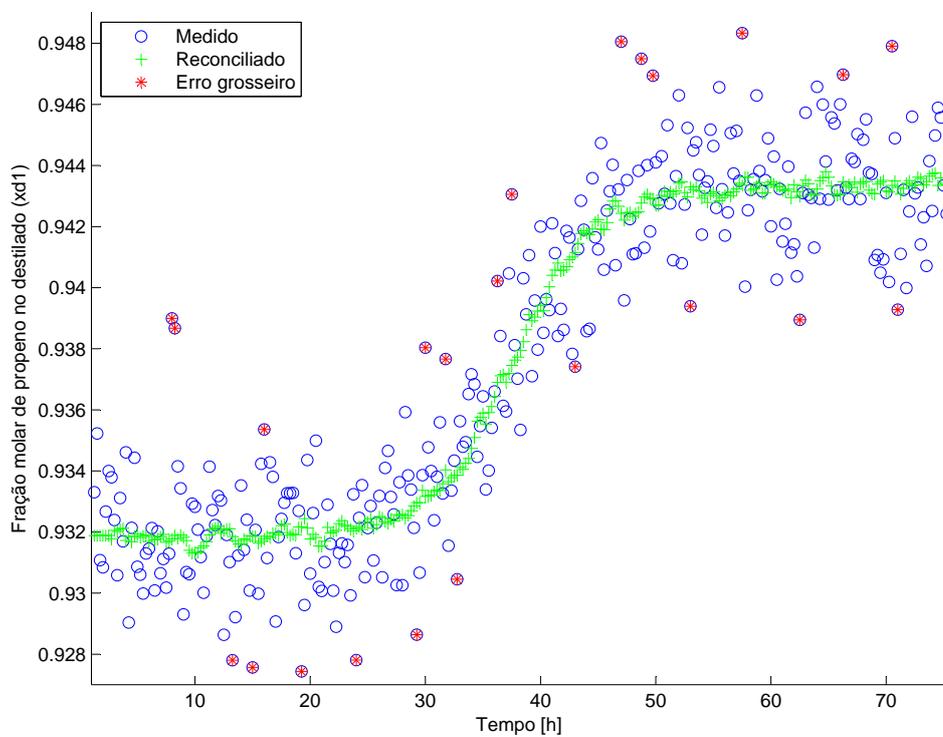


Figura 6.38: Resultados - Fração molar de propeno no destilado $x_{D,1}$ no estudo de caso 2 parte II.

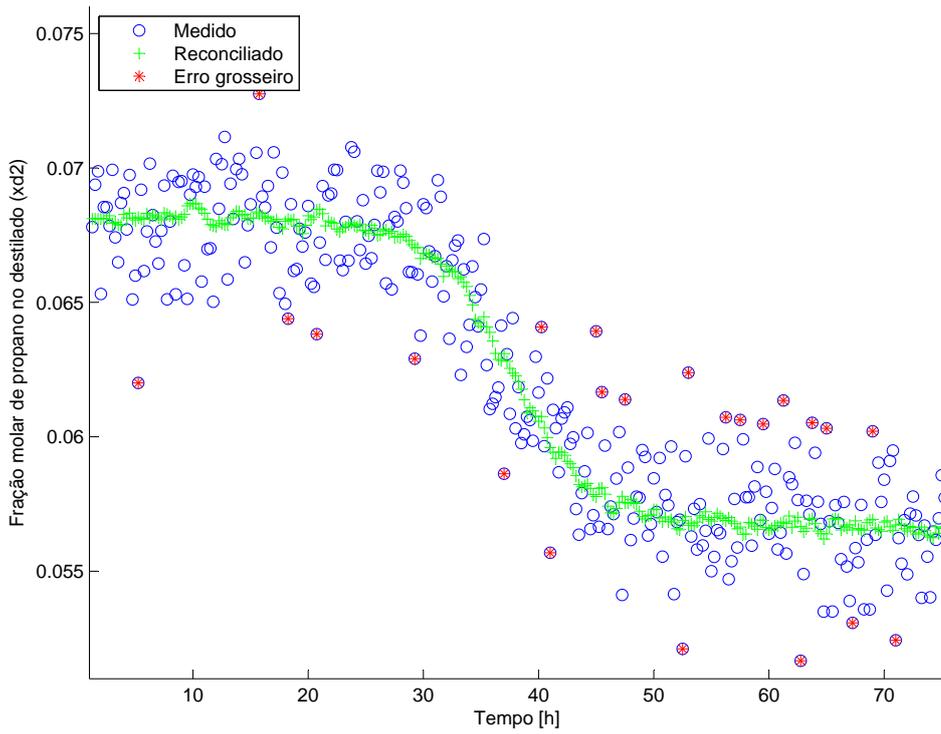


Figura 6.39: Resultados - Fração molar de propano no destilado $x_{D,2}$ no estudo de caso 2 parte II.

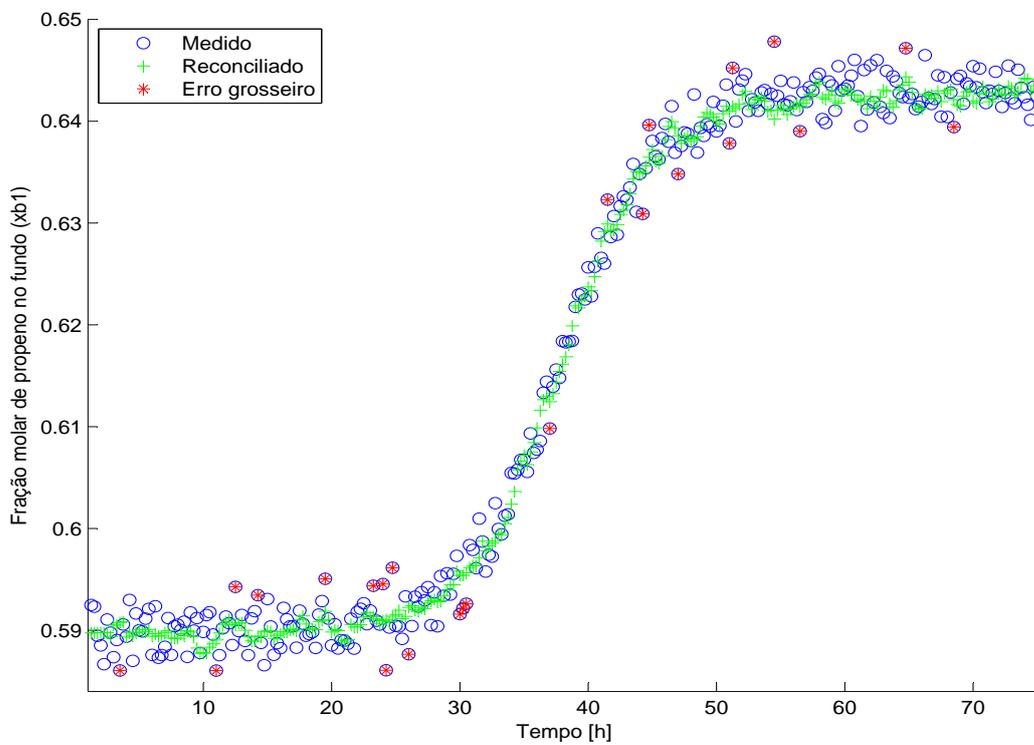


Figura 6.40: Resultados - Fração molar de propano no fundo $x_{B,1}$ no estudo de caso 2 parte II.

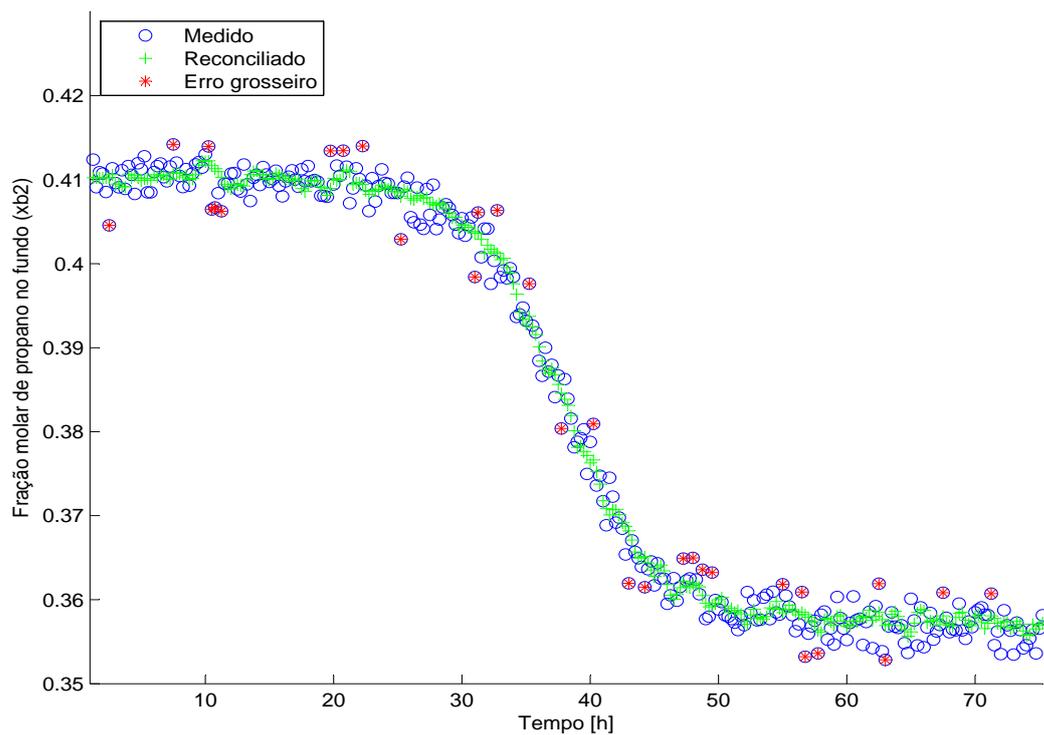


Figura 6.41: Resultados - Fração molar de propano no fundo $x_{B,2}$ no estudo de caso 2 parte II.

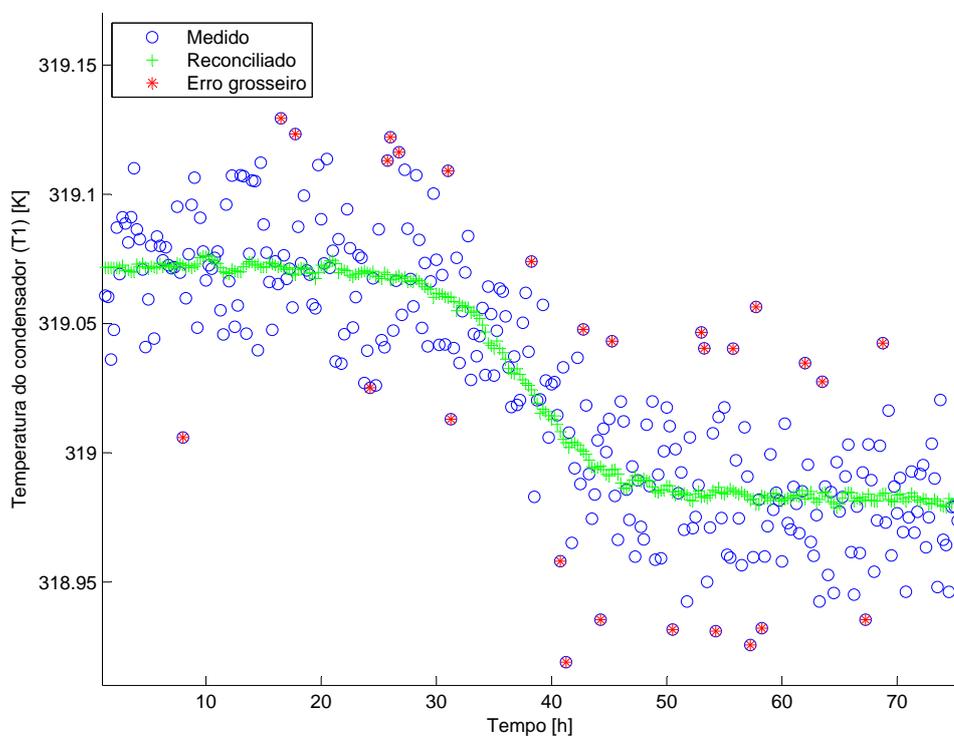


Figura 6.42: Resultados - Temperatura do condensador T_c no estudo de caso 2 parte II.

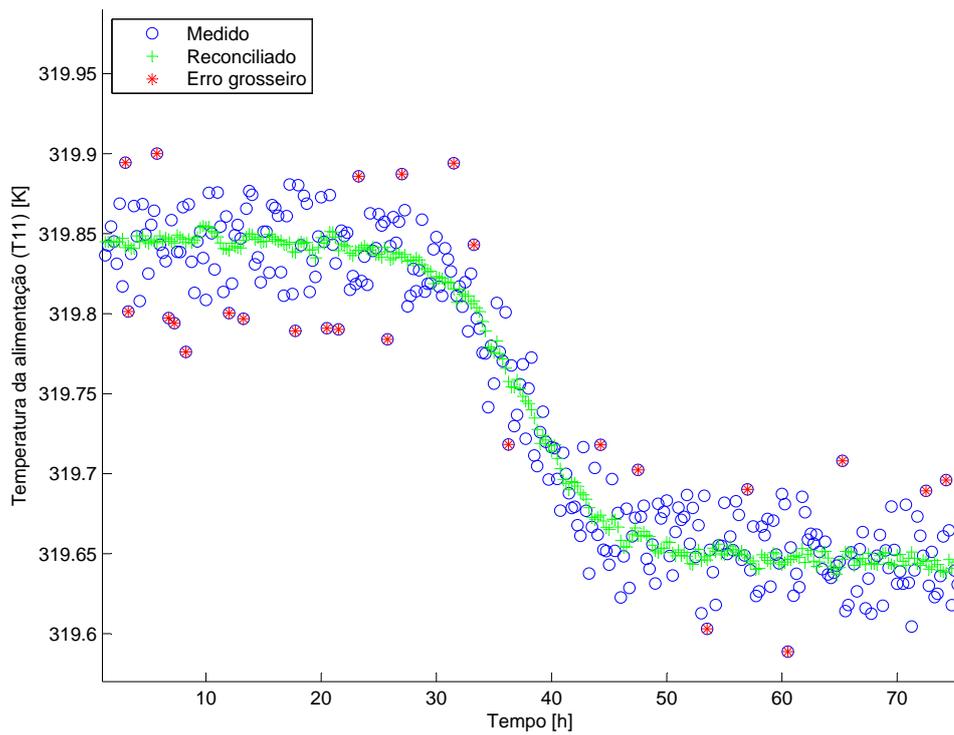


Figura 6.43: Resultados - Temperatura do prato de alimentação T_f no estudo de caso 2 parte II.

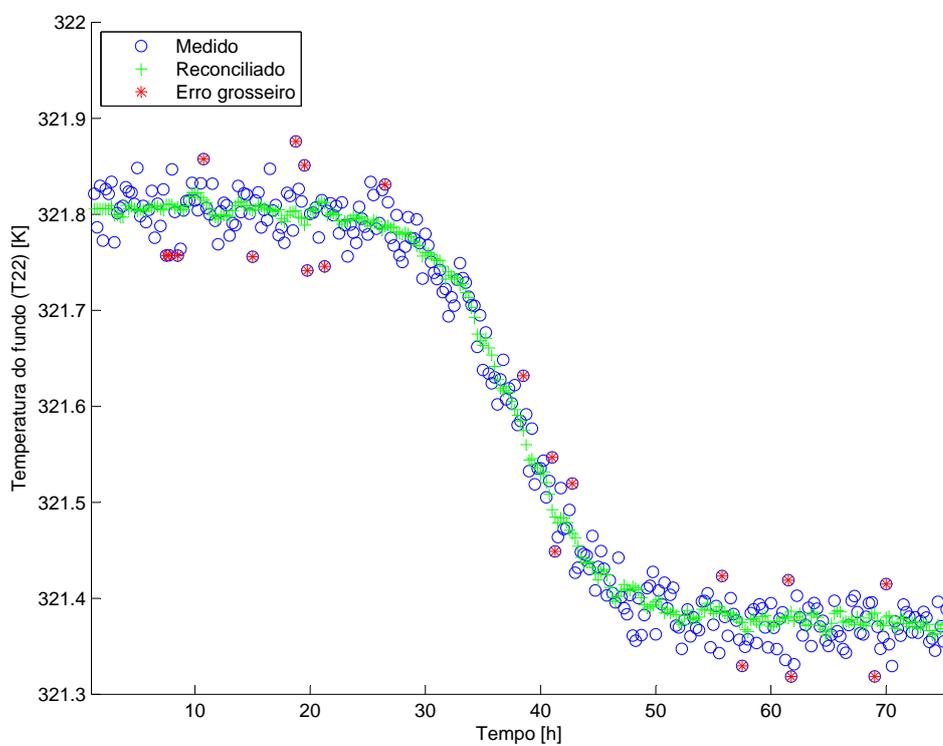


Figura 6.44: Resultados - Temperatura do refrigerador T_r no estudo de caso 2 parte II.

Observa-se que, para as variáveis de saída, a reconciliação é mais “suave”; ou seja, não é tão ruidosa como as variáveis de entrada. Isso é devido à própria natureza matemática do modelo de equações diferenciais, pois pequenas perturbações na variável manipulada provocam muito pouca ou nenhuma oscilação na resposta (variável de saída). O próprio modelo de equações diferenciais atua como um “filtro” para os ruídos das variáveis de entrada. Porém, por outro lado essa mesma natureza matemática atua para dificultar a eliminar erros grosseiros na variável de entrada, pois grandes variações nas variáveis de entrada não correspondem às mesmas variações nas variáveis de respostas, devido ao “efeito filtro” das equações diferenciais. Dessa forma, é possível concluir que essa natureza matemática tem robustez quando os erros grosseiros se encontram nas variáveis de saída (respostas). Por outro lado, perde-se robustez quando os erros grosseiros se encontram nas variáveis de entrada.

Por este motivo, para as variáveis de entrada, adicionou-se um termo a mais na F_{obj} para amenizar o ruído e aumentar a robustez. O termo equivale à similaridade temporal das variáveis na janela, fazendo com que os valores não fiquem tão discrepantes no tempo. De fato, o termo tem como fundamento os modelos lineares da Análise de Séries Temporais, que são modelos preditivos de séries temporais, e para o caso o modelo aplicado se encaixa como um modelo Autoregressivo de ordem 1, AR(1). O que demonstra a sua eficácia como filtro (suavizando os dados de entrada), predição consistente com os dados (autocorrelação) e robustez em presença de erros grosseiros (similaridade e redundância estatística).

Portanto, observando as Figuras 6.33 a 6.44, pode-se afirmar que a RDD obteve sucesso. Entretanto, é previsível que a RDD dê certo utilizando o mesmo modelo que gerou os dados corrompidos (com erros grosseiros de baixa magnitude). Desta forma, novos estudos devem ser feitos para averiguar a RDD em condições de falhas, erros grosseiros, com modelo reduzido e outros estimadores. Além disso, modelo Autoregressivo de outras ordens devem ser estudados, para entender melhor os efeitos desses modelos preditivos de séries temporais.

6.2.4 Caso 4: Monitoramento de um processo de separação por membrana

Os resultados deste caso estão apresentados e discutidos no Apêndice A.

Capítulo 7

Conclusão

“A humildade é a única base sólida de todas as virtudes.”

“Se não sabes, aprende; se já sabes, ensina.”

Confúcio “Kung-fu-tzu” (551 a.C. – 479 a.C.)

7.1 Caso 1: Avaliação dos métodos de otimização

Demostrou-se, juntamente com os *softwares* EMSO e Matlab, a eficiência e o desempenho em problemas de reconciliação de dados em coluna de destilação, assim como a eficiência e robustez dos estimadores (MQP, Normal Contaminada e Welsch) já introduzidos na reconciliação de dados em engenharia química. Comprovou-se também a possibilidade de utilizar modelos rigorosos de destilação MESH em problemas de reconciliação de dados com variáveis medidas e erros grosseiros, reconciliando não somente vazões e concentrações, mas também as temperaturas dos estágios de equilíbrio.

A reconciliação robusta de dados realizada por estimadores-M, como Normal Contaminada e Welsch, mostrou-se para este caso um pouco superior, frente ao estimador MQP (não robusto). O melhor desempenho para o estimador foi comprovado pelo critério de avaliação TER, em que o estimador Normal Contaminada resultou no melhor valor de TER (0,8708). Isto mostra a eficiência do procedimento de reconciliação de dados e detecção de erros grosseiros simultaneamente, utilizando estimadores-M (pseudo robusto e robustos), evitando o uso de estratégias iterativas.

Os métodos de otimização determinísticos que utilizam derivadas, para este caso, mostram-se superiores ao método não determinístico PSO e ao método de busca direta Nelder-Mead, pois o tempo gasto na otimização com os métodos que utilizam derivadas foi quase 30 vezes mais rápidos, e com o valor mínimo da função objetivo satisfatório.

Portanto, os estimadores Normal Contaminada e Welsch com a rotina *fmincon* utilizando o algoritmo do Ponto Interior resultaram, no geral, nos melhores desempenhos da reconciliação de dados com detecção simultânea de erros grosseiros, com o modelo MESH em coluna de destilação.

7.2 Caso 2: Avaliação da redução de ordem do modelo de coluna

Demostrou-se, juntamente com os *softwares* EMSO e Matlab, a eficiência e o desempenho em problemas de reconciliação de dados em coluna de destilação, assim como a eficiência e velocidade da técnica de redução de ordem com o método dos resíduos ponderados pelos momentos em modelos de sistema de separação com estágios. Com isso, comprovou-se a possibilidade de utilizar modelos reduzidos de destilação em problemas de reconciliação de dados com variáveis medidas e erros grosseiros, reconciliando não somente vazões e concentrações, mas também as temperaturas dos estágios de equilíbrio.

A reconciliação robusta de dados realizada por estimadores-M, como Normal Contaminada (quasi-robusto), mostrou-se para este caso um pouco superior, frente ao estimador MQP (não robusto). O melhor desempenho para o estimador foi comprovado pelo critério de avaliação TER, em que o estimador Normal Contaminada resultou no melhor valor de TER (0,8901) e melhor tempo de otimização (1,2 minutos). Isto mostra a eficiência do procedimento de reconciliação de dados e detecção de erros grosseiros simultaneamente, evitando o uso de estratégias iterativas. O desempenho do procedimento dos estimadores em conjunto com o modelo reduzido mostrou-se satisfatório, pois apresentou também bons resultados para o TER, o que não descarta a utilização do método.

A reconciliação de dados em conjunto com o modelo reduzido mostrou-se mais veloz em relação ao modelo completo, o que pode tornar o procedimento de RD praticável em tempo real. Ressalta-se também que, apesar de resultar um valor do TER com o modelo reduzido inferior ao do TER obtido com o modelo completo, os dados de maneira geral foram bem reconciliados. Como num procedimento em

tempo real os dados medidos não seguem uma distribuição fixa e muito menos um modelo de referência, isso reforça ainda mais a hipótese da utilização de modelos reduzidos, pois o mais importante é reconciliar em tempo suficiente do que não ter tempo de reconciliar com um modelo rigoroso.

7.3 Caso 3: Avaliação do termo de similaridade com modelo autoregressivo

A solução do problema de RDDR em coluna de destilação pela técnica *feasible path* resultou em resultados satisfatórios e animador. Com isso, abre oportunidades de novos estudos e aplicações, como em casos de dados reais de colunas.

A solução apresentou dificuldades para casos com tempo de amostragem menor que 15 minutos, o que poderia ser um empecilho para aplicação da RDDR *on-line* voltada para atender a demanda do controle supervisionado e regulatório. O estudo indica que pode ter havido um equívoco na seleção do tempo de amostragem; entretanto, este problema pode ser contornado com a aplicação de modelos autoregressivos. Dessa forma, modelo preditivos de autoregressão evita o uso de dados reconciliados dentro de uma janela, tratando como variável reconciliada apenas o valor medido atual. Além disso, indiretamente a autoregressão estima o tamanho da “janela” do modelo (ordem do modelo), com base nos próprios dados. Porém, a janela de integração deve ser calculada com base no tempo característico do modelo dinâmico, diferente da janela do modelo autoregressivo. De fato, é importante respeitar o tamanho correto da janela de integração do modelo dinâmico, pois uma janela curta demais pode perder informação nas variáveis de estado como a “inércia”, que conseqüentemente reflete em defasagens nas variáveis de saída, fazendo com que o modelo não alcance os dados medidos e induzindo a interpretação de erros do tipo *drifts* ou *bias*.

O termo de similaridade temporal (Autoregressivo de ordem 1), mostrou-se um excelente desempenho. Reduziu os ruídos das variáveis de entrada, aumentando sua robustez e dando redundância estatística ao problema de RDDR. Portanto, como previsto pela demonstração na subseção 2.4.2, os efeitos da RDD baseados em modelos preditor-corretor diminuiu a variabilidade da variável reconciliada em relação a variabilidade dos dados brutos, atuando como um filtro na variável de entrada. As cláusulas de ativação do termo de similaridade e detecção do estado estacionário demonstraram ser capazes de atender mudanças de *set-point* na operação de uma coluna. Pois, em um monitoramento em tempo real e em linha, não se espera que os dados estejam sempre caracterizados como um

processo estocástico estacionário. O monitoramento em tempo real requer métricas, proteções e tratamentos dos dados, para prover o melhor da metodologia de monitoramento baseado em RDDR. Dessa forma, é importante ressaltar que o termo de similaridade deu robustez para as variáveis de entrada, evitando problemas de não detecção/eliminação de erros grosseiros nas variáveis de entrada, devido ao “efeito filtro” do modelo dinâmico de coluna de destilação.

O estimador Normal Contaminada mostrou-se satisfatório na detecção de pequenos erros grosseiros, como já demonstrado em outros estudos. O método de otimização Ponto-Interior também não apresentou problemas, pois foi selecionado devido a sua eficácia e rapidez, baseado nos estudos anteriores.

7.4 Caso 4: Monitoramento de um processo de separação por membrana

Foi desenvolvida e implementada pela primeira vez uma metodologia em forma de aplicação web para permitir a monitorização dos processos de separação por membranas online e em tempo real, recorrendo a técnicas estatísticas para tratamento dos dados do processo. A metodologia proposta compreende as seguintes etapas: (i) pré-tratamento e caracterização dos dados do processo; (ii) reconciliação dos dados do processo para minimizar as incertezas da medição, com o auxílio de equações de balanço de massa; (iii) detecção de desvios sistemáticos para identificação de falhas e mau funcionamento do processo; e (iv) observação de variáveis não medidas (funcionando como um sensor virtual ou gêmeo digital).

As etapas de pré-tratamento e caracterização dos dados foram fundamentais para a compreensão e correta formulação do problema. A etapa de caracterização pode ter ampla aplicação, pois este procedimento pode ser aplicado em qualquer processo químico. Esta etapa é essencial para a seleção apropriada de técnicas de reconciliação de dados e procedimentos de detecção de erros grosseiros. Depois disso, as etapas propostas de reconciliação de dados e detecção de erros brutos mostraram robustez, bom desempenho e velocidade. O esquema proposto foi baseado em equações de balanço de massa por componente em estado estacionário, detalhadas e validadas após a caracterização adequada dos dados reais de operação. Os procedimentos numéricos foram validados *offline* e, em seguida, implementados *online* e em tempo real pela primeira vez, permitindo a identificação bem-sucedida dos *bias* e *outliers* de medição, além de fornecer estimativas para dados não medidos.

Os procedimentos desenvolvidos podem ser usados para detecção *online* e

em tempo real de falhas de processo e diagnóstico de processo. Além disso, o procedimento fornece dados confiáveis para futuras etapas de simulações e estimação de parâmetros, permitindo a implementação de gêmeos digitais, conforme o modelo proposto no trabalho de DIAS *et al.* (2020), parte I deste projeto de pesquisa. As etapas do Sistema de Gerenciamento de Produção e do Planejamento de Recursos Empresariais também podem se beneficiar da disponibilidade de dados mais confiáveis e variáveis inferidas por um sensor virtual. Portanto, as principais vantagens do procedimento são o manuseio confiável dos dados, o diagnóstico de erros/falhas graves e o monitoramento em tempo real do processo.

7.5 Revisão sobre estimadores-M robustos para análises de regressão

A maior parte da revisão descrita no Apêndice B foi dedicada à apresentação e uso de estimadores-M, porque embora algumas aplicações desses estimadores sejam focadas especificamente em estatísticas robustas, engenharia elétrica/eletrônica e processamento de sinais, outras áreas do conhecimento também desenvolveram, estudaram e encontraram uso para os estimadores-M, particularmente para solução de problemas de Reconciliação de Dados e Estimação de Parâmetros no campo da engenharia química. Por este motivo, a revisão apresenta pela primeira vez uma compilação de 50 estimadores-M (48 robustos), com suas respectivas funções características (a função ρ , a função ψ e a função ψ') e parâmetros sintonizados para os níveis de Eficiência Relativa de 90%, 95%, 98% e 99% em relação à distribuição normal. É importante ressaltar que muitos desses estimadores nunca haviam sido apresentados anteriormente na forma gráfica, tornando mais difícil a correta implementação e validação dos estimadores. Além disso, muitos dos estimadores analisados foram usados poucas vezes (às vezes uma única vez), mostrando a importância de ajustá-los e apresentá-los à comunidade científica. Como um todo, este artigo apresenta uma coleção de estimadores-M robustos que estão espalhados por várias áreas da ciência que estudam problemas de análises de regressão. Assim, o presente trabalho preenche uma lacuna e permite a troca de conhecimento entre essas diversas comunidades científicas, apresentando um grande e bem documentado número de estudos de caso, resultados e conclusões sobre o uso de estimadores-M robustos. Finalmente, deve-se notar através da Tabela 2.1, que os estimadores-M de Normal Contaminada (quasi-robusto), Welsch, Hampel, Fair, Lorenziana, Correntropia e Cauchy foram os mais usados para análise de regressão em problemas de engenharia química, mostrando que a preferência desta área representa um

conjunto reduzido no universo dos estimadores-M, como os 50 estimadores-M (MPQ, Normal contaminada e 48 robustos) apresentados nesta revisão.

7.6 Continuação da pesquisa e desenvolvimento - trabalhos futuros

As metodologias até aqui desenvolvidas apresentaram bons resultados, capacidade de aplicação em monitoramento em tempo real e abriram novas possibilidades de investigação, principalmente para o caso de RDDR em coluna de destilação. O caso da RDDR com modelo Autoregressivo precisa ser estudado e analisado com mais detalhe e profundidade. Os efeitos do modelo de predição Autoregressivo devem ser estudados com base na formulação generalizada do problema proposto, como as Equações (5.17) a (5.20). Além disso, o problema da janela de integração do modelo dinâmico pode ser resolvido utilizando o método de integração *dasslc* no *software* Matlab, que utiliza informação das variáveis de estado, tanto como informação de entrada como de saída, carregando toda a informação da “inércia” do modelo dinâmico para a janela seguinte, evitando defasagem nas variáveis de resposta. Com isso, a janela de integração do modelo dinâmico pode ser reduzida a 2 pontos, reduzindo o tempo de integração em relação a um intervalo de integração baseado no tempo característico, que se refletiria em muitos pontos. Entretanto, o modelo dinâmico de coluna deve ser modelado dentro *software* Matlab, o que se tornou uma desvantagem cronológica na pesquisa, pois o Matlab não tem pacote termodinâmico como o EMSO (VRTherm). Portanto, este estudo ficará para os trabalhos futuros.

Outra frente de trabalhos futuros é orientada pelo trabalho apresentado no Apêndice B, onde são apresentados 48 estimadores-M robustos, com sintonias para 90, 95, 98 e 99% de Eficiência Relativa. Este trabalho pode nortear diversas pesquisas envolvendo problemas de Análises de Regressão (Reconciliação de Dados e Estimação de Parâmetros) com dados contaminados por erros grosseiros. Com isso, tanto o estudo com RDDR em coluna de destilação, como o monitoramento baseado RD em separação de CO_2 por membrana ganham novas oportunidades de pesquisa e desenvolvimento. As subseções 2.4.1 e 2.4.1 apresentam desenvolvimentos que poderão ser aplicáveis ao caso do monitoramento de processo de membranas, utilizando estimadores-M robustos e suas funções pesos $w(\xi)$, para a detecção/eliminação simultânea de erros grosseiros em tempo real.

Para concluir, a metodologia proposta chamada de “Algoritmo sequencial de RDDR baseado em modelos fenomenológicos e autorregressivos” tem potencial para

contribuir em novas frentes de pesquisas em problemas de RDDR, não apenas em coluna de destilação, pois a metodologia pode ser aplicada em processos de separação de dinâmica rápida, caracterizado por modelos estacionários, como o processo de separação de CO_2 por membranas. A metodologia de caracterização dos dados e do processo, desenvolvida para monitoramento em tempo real em sistema de separação de CO_2 por membrana, pode ser aplicável a outros processos químicos. Principalmente pela etapa de caracterização, por ter ampla aplicação e, com isso, abrindo novas possibilidades de pesquisa e desenvolvimento. Assim como a promissora frente de pesquisa e desenvolvimento, conduzida pelo trabalho no Apêndice B, também possa vir a contribuir em futuras teses do Programa de Engenharia Química da COPPE, principalmente na área de Análises de Regressão e Monitoramento de processo.

A Tabela 7.1 apresenta um resumo dos objetivos e resultados de cada caso.

Tabela 7.1: Resumo dos objetivos e resultados.

Casos	Objetivos	Resultados
Caso 1	Avaliar os métodos de otimização e estimadores-M.	Os métodos determinísticos apresentaram melhor performance. O estimador-M Normal Contaminada e Welsch obtiveram os melhores desempenhos.
Caso 2	Avaliar a velocidade e acurácia do modelo reduzido de coluna.	O modelo reduzido pelo método dos resíduos ponderados apresentou melhora significativa na velocidade e boa acurácia.
Caso 3	Avaliar estratégias que aceleram a otimização e aumenta a robustez.	As métricas de estimativa e desigualdade adaptativa apresentaram melhorias na velocidade de otimização. O modelo de predição autoregressivo resolveu o problema de robustez nas variáveis de entrada.
Caso 4	Implementar um sistema de monitoramento em tempo real de um sistema de separação.	A implementação do monitoramento em tempo baseado em RD foi bem sucedido. A metodologia de detecção de erros grosseiros mostrou-se uma boa ferramenta para diagnosticar falhas e erros.
<i>Review</i>	Apresentar uma revisão ampla dos estimadores-M robustos, suas sintonias e características	Foi apresentado 50 estimadores-M (48 robustos e 2 não robustos); Todos os estimadores sintonizáveis foram sintonizados para os níveis de Eficiência Relativa de 90, 95, 98 e 99%; As famílias de funções (ρ , ψ e ψ') foram apresentadas na forma gráfica e funcional (ρ , ψ); Descrição detalhada do procedimento de sintonização para estimadores com mais de 1 parâmetro de sintonia; Uma tabela guia para ajudar a selecionar o estimador-M mais adequado ao seu problema de regressão.

Referências Bibliográficas

- AKSHAYA, S., 2016. “Analysis of Energy Consumption in Industrial Distillation Columns”. Disponível em: <<https://www.aiche.org/academy/videos/conference-presentations/analysis-energy-consumption-industrial-distillation-columns>>.
- AL-OBAIDI, M. A., KARA-ZAÏTRI, C., MUJTABA, I. M., 2018, “Simulation and sensitivity analysis of spiral wound reverse osmosis process for the removal of dimethylphenol from wastewater using 2-D dynamic model”, *Journal of Cleaner Production*, v. 193, pp. 140–157.
- ALBUQUERQUE, J. S., BIEGLER, L. T., 1995, “Decomposition algorithms for on-line estimation with nonlinear models”, *Computers & chemical engineering*, v. 19, n. 10, pp. 1031–1039.
- ALBUQUERQUE, J. S., BIEGLER, L. T., 1996, “Data reconciliation and gross-error detection for dynamic systems”, *AIChE journal*, v. 42, n. 10, pp. 2841–2856.
- ALDRICH, J., OTHERS, 1997, “RA Fisher and the making of maximum likelihood 1912-1922”, *Statistical Science*, v. 12, n. 3, pp. 162–176.
- ALEJSKI, K., DUPRAT, F., 1996, “Dynamic simulation of the multicomponent reactive distillation”, *Chemical Engineering Science*, v. 51, n. 18, pp. 4237–4252.
- ALHAJ-DIBO, M., MAQUIN, D., RAGOT, J., 2008, “Data reconciliation: A robust approach using a contaminated distribution”, *Control Engineering Practice*, v. 16, n. 2, pp. 159–170.
- ALMASY, G. A., SZTANO, T., 1975, “Checking and correction of measurements on the basis of linear system model”, *Problems of Control and Information Theory*, v. 4, pp. 57–69.

- ANDRADE, M. D. C. V., 2004, *Aplicação de redes neurais para detecção de erros grosseiros em dados de processo de separação de isótopos de urânio por ultracentrifugação*. Tese de Mestrado, Instituto de Pesquisa Energéticas e Nucleares, São Paulo, SP.
- ANP, ANP RESOLUTION Nº 16, . D. ., 2008, *Agência Nacional do Petróleo, Gás Natural e Biocombustíveis*. Diário Oficial da União, June 18.
- ARORA, N., BIEGLER, L. T., 2001, “Redescending estimators for data reconciliation and parameter estimation”, *Computers & Chemical Engineering*, v. 25, n. 11, pp. 1585–1599.
- ARORA, N., BIEGLER, L. T., 2004, “Parameter estimation for a polymerization reactor model with a composite-step trust-region NLP algorithm”, *Industrial & engineering chemistry research*, v. 43, n. 14, pp. 3616–3631.
- ASPENTECH, 2021. “Aspen Operations Reconciliation and Accounting”. AspenTech. Disponível em: <<https://www.aspentech.com/en/products/msc/aspen-operations-reconciliation-and-accounting>>.
- ÅSTRÖM, K. J., WITTENMARK, B., 2013, *Computer-controlled systems: theory and design*. Courier Corporation.
- ATKINS, P., DE PAULA, J., 2006, *Atkins físico-química*, v. 1. LTC.
- AUREL, 2021. “MASSBAL”. Aurel Systems Inc. Disponível em: <<https://www.aurelsystems.com/pulp-paper/>>.
- AVEVA, 2021. “ROMeO Material Balance”. AVEVA. Disponível em: <https://www.aveva.com/content/dam/aveva/documents/datasheets/Datasheet_AVEVA_ROMeOMaterialBalanceModule_06-19.pdf.coredownload.inline.pdf>.
- BAGAJEWICZ, M. J., ROLLINS, D. K., 2002, “Instrument Engineers’ Handbook: Process Software and Digital Networks”. cap. 5, pp. 687–702, CRC Press.
- BAGAJEWICZ, M. J., 2000a, *Process plant instrumentation: design and upgrade*. CRC Press.
- BAGAJEWICZ, M. J., JIANG, Q., 1997, “Integral approach to plant linear dynamic reconciliation”, *AIChE Journal*, v. 43, n. 10, pp. 2546–2558.

- BAGAJEWICZ, M., 2000b, “A brief review of recent developments in data reconciliation and gross error detection/estimation”, *Latin American Applied Research*, v. 30, n. 4, pp. 335–342.
- BAI, S., THIBAUT, J., 2010, *Dynamic Data Reconciliation: Theory and Practice*. VDM Publishing.
- BAI, S., THIBAUT, J., MCLEAN, D. D., 2005, “Closed-loop data reconciliation for the control of a binary distillation column”, *Chemical Engineering Communications*, v. 192, n. 11, pp. 1444–1467.
- BAI, S., THIBAUT, J., MCLEAN, D. D., 2006, “Dynamic data reconciliation: Alternative to Kalman filter”, *Journal of Process Control*, v. 16, n. 5, pp. 485–498.
- BAI, S., MCLEAN, D. D., THIBAUT, J., 2007a, “Impact of model structure on the performance of dynamic data reconciliation”, *Computers & chemical engineering*, v. 31, n. 3, pp. 127–135.
- BAI, S., MCLEAN, D. D., THIBAUT, J., 2007b, “Simultaneous measurement bias correction and dynamic data reconciliation”, *The Canadian Journal of Chemical Engineering*, v. 85, n. 1, pp. 111–117.
- BARD, Y., 1974, *Nonlinear parameter estimation*. Academic press.
- BASELGA, S., 2007, “Global optimization solution of robust estimation”, *Journal of surveying Engineering*, v. 133, n. 3, pp. 123–128.
- BELSIM, 2021. “VALI”. Belsim. Disponível em: <<https://www.belsim.com/business/software/data-validation/>>.
- BENALLOU, A., SEBORG, D., MELLICHAMP, D., 1986, “Dyanmic compartmental models for separation processes”, *AIChE Journal*, v. 32, n. 7, pp. 1067–1078.
- BENQLILOU, C., 2004, *Data reconciliation as a framework for chemical processes optimization and control*. Tese de Doutorado, Universitat Politècnica de Catalunya.
- BESSA, R. J. G. D. S., 2009, *Treino on line de redes neuronais com critérios de informação aplicado à previsão eólica*. Tese de Mestrado, Faculdade de Economia da Universidade do Porto, Porto, Portugal.

- BEZERRA, M. I. S., 2006. “Apostila de análise de séries temporais”. Curso de Estatística - DMEC/ FCT / UNESP. Disponível em: <<https://www.docsity.com/pt/apostila-de-analise-de-series-temporais/4770523/>>.
- BOUROUIS, M., PIBOULEAU, L., FLOQUET, P., et al., 1998, “Simulation and data validation in multistage flash desalination plants”, *Desalination*, v. 115, n. 1, pp. 1–14.
- BOX, G. E., TIAO, G. C., 2011, *Bayesian inference in statistical analysis*, v. 40. John Wiley & Sons.
- BOX, G. E., JENKINS, G. M., REINSEL, G. C., et al., 2015, *Time series analysis: forecasting and control*. John Wiley & Sons.
- BUCKLEY, P. S., LUYBEN, W. L., SHUNTA, J. P., 1985, *Design of distillation column control systems*. Instrument Society of America,.
- CALDAS, J. N., DE LACERDA, A. I., VELOSO, E., et al., 2007, *Internos de torres: pratos & recheios*. Interciência, 2^a edição.
- CÂMARA, M. M., SOARES, R. M., FEITAL, T., et al., 2017, “Numerical Aspects of Data Reconciliation in Industrial Applications”, *Processes*, v. 5, n. 4, pp. 56.
- CAUCHY, A. L., 1853, “Sur les résultats moyens d’observations de même nature, et sur les résultats les plus probables”, *CR Acad. Sci. Paris*, v. 37, pp. 198–206.
- CHEMPLANT, 2021. “RECON”. ChemPlant Thecnology. Disponível em: <<https://www.chemplant.cz/inpage/recon/>>.
- CHEN, J., ROMAGNOLI, J. A., 1998, “A strategy for simultaneous dynamic data reconciliation and outlier detection”, *Computers & chemical engineering*, v. 22, n. 4-5, pp. 559–562.
- CHEN, J., BANDONI, A., ROMAGNOLI, J. A., 1997, “Robust estimation of measurement error variance/covariance from process sampling data”, *Computers & chemical engineering*, v. 21, n. 6, pp. 593–600.
- CHEN, J., BANDONI, A., ROMAGNOLI, J. A., 1998, “Outlier detection in process plant data”, *Computers & chemical engineering*, v. 22, n. 4-5, pp. 641–646.
- CHEN, J., PENG, Y., MUNOZ, J. C., 2013, “Correntropy estimator for data reconciliation”, *Chemical Engineering Science*, v. 104, pp. 1019–1027.

- CHEN, X., 1998, *The Optimal Implementation Of On-Line Optimization For Chemical And Refinery Processes Vol. I*. Tese de Doutorado, University of Mississippi.
- CHO, Y., JOSEPH, B., 1983a, “Reduced-order steady-state and dynamic models for separation processes. Part I. Development of the model reduction procedure”, *AIChE journal*, v. 29, n. 2, pp. 261–269.
- CHO, Y., JOSEPH, B., 1983b, “Reduced-order steady-state and dynamic models for separation processes. Part II. Application to nonlinear multicomponent systems”, *AIChE Journal*, v. 29, n. 2, pp. 270–276.
- CHUI, C. K., CHEN, G., OTHERS, 1999, “Kalman filtering”, *With real time applications*.
- COIMBRA, J. C., MELO, P. A., PRATA, D. M., et al., 2017, “On-Line Dynamic Data Reconciliation in Batch Suspension Polymerizations of Methyl Methacrylate”, *Processes*, v. 5, n. 3, pp. 51.
- COKER, A. K., 2010, *Ludwig’s Applied process design for chemical and petrochemical plants*, v. 2. 4th edition, Gulf Professional Publishing.
- CONGLI, M., HONGYE, S., JIAN, C., 2006, “An NT-MT Combined Method for Gross Error Detection and Data Reconciliation”, *Chinese Journal of Chemical Engineering*, v. 14, n. 5, pp. 592–596.
- CORDÃO, I. J. G., 2005, *Implantação de sistema de conciliação de dados em uma refinaria de petróleo*. Tese de Mestrado, Universidade Estadual de Campinas, Campinas, SP.
- COSTA, H. C., 2020. “Correlação, Autocorrelação e Autocorrelação Parcial”. RPubS by RStudio. Disponível em: <https://rpubs.com/hudsonchavs/fac_facp>.
- CRIVELLARI, G. P., 2016, *Modelagem matemática e simulação de um permeador de gases para separação de CO₂ de gás natural*. Tese de Mestrado, Universidade de São Paulo, São Paulo, SP.
- CROWE, C. M., CAMPOS, Y. A., HRYMAK, A., 1983, “Reconciliation of process flow rates by matrix projection. Part I: linear case”, *AIChE Journal*, v. 29, n. 6, pp. 881–888.
- CROWE, C., 1986, “Reconciliation of process flow rates by matrix projection. Part II: the nonlinear case”, *AIChE Journal*, v. 32, n. 4, pp. 616–623.

- DA CUNHA, A. S., SANTOS, L. S., PEIXOTO, F. C., et al., 2017, “Robust data reconciliation in a chemical reactor through simulated annealing optimization”, *Latin American Applied Research*, v. 47, pp. 131–136.
- DA CUNHA, A. S., PEIXOTO, F. C., PRATA, D. M., 2020, “Robust Data Reconciliation in Chemical Reactors”, *Computers & Chemical Engineering*, p. 107170.
- DAROUACH, M., RAGOT, J., FAYOLLE, J., et al., 1986, “Validation des mesures par équilibrage de bilans-matière”, *International journal of mineral processing*, v. 17, n. 3-4, pp. 273–285.
- DE ALBUQUERQUE, A. R. L., 2008, “Otimização na destilação de etanol?” *Mecatrônica Atual - Automação Industrial de Processos e Manufatura*, v. 39, n. 6, pp. 20–23.
- DE JESUS, N. J. C., 2011, *Otimização em tempo real em um processo industrial de produção de etileno*. Tese de Doutorado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- DE LAPLACE, P., 1774, “Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards”, *Mém. Acad. Roy. Sci. Paris*, v. 6, pp. 353–371.
- DE MENEZES, D. Q. F., 2015, *Reconciliação robusta de dados em colunas de destilação*. Tese de Mestrado, Universidade Federal Fluminense, Niterói, RJ.
- DE MENEZES, D. Q. F., DE SÁ, M. C. C., FONTOURA, T. B., et al., 2020, “Modeling of Spiral Wound Membranes for Gas Separations—Part II: Data Reconciliation for Online Monitoring”, *Processes*, v. 8, n. 9 (aug), pp. 1035. doi: 10.3390/pr8091035. Disponível em: <<https://doi.org/10.3390/pr8091035>>.
- DE SOUZA, P. N. M., 2015, *Implementação em paralelo de reconciliador robusto de dados com simultânea estimação de parâmetros para uso em tempo real*. Tese de Doutorado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- DIAS, A. C. S., DE SÁ, M. C. C., FONTOURA, T. B., et al., 2020, “Modeling of spiral wound membranes for gas separations. Part I: An iterative 2D permeation model”, *Journal of Membrane Science*, v. 612, pp. 118278.

- DO VALLE, E. C., SOARES, R. P., FINKLER, T. F., et al., 2008, “A New Tool Providing an Integrated Framework for Process Optimization”, *EngOpt - International Conference on Engineering Optimization*.
- DO VALLE, E. C., DE ARAÚJO KALID, R., SECCHI, A. R., et al., 2018, “Collection of Benchmark Test Problems for Data Reconciliation and Gross Error Detection and Identification”, *Computers & Chemical Engineering*.
- DORTMUNDT, D., DOSHI, K., 1999, “Recent developments in CO₂ removal membrane technology”, *UOP LLC*, pp. 1–30.
- DOUGLAS, J. M., 1988, *Conceptual design of chemical processes*, v. 1110. McGraw-Hill New York.
- EDGAR, T. F., HIMMELBLAU, D. M., LASDON, L. S., 2001, *Optimization of chemical processes*.
- ELECTRIC’S, S., 2018. “PRO/II Process Engineering”. Disponível em: <<http://software.schneider-electric.com/products/simsci/design/pro-ii/>>.
- ENERGY.GOV, 2016. “New Software Will Enable Chemical Manufacturers to Optimize Distillation Column Configuration and Save Energy”. U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy - 11/agosto/2016. Disponível em: <<https://energy.gov/eere/amo/articles/new-software-will-enable-chemical-manufacturers-optimize-distillation-column-configuration-and-save-energy>>.
- EPE, E. D. P. E., 2020a. “Balanco Energético Nacional 2020”. Ministério de Minas e Energia, a. Disponível em: <<http://shinyepe.brazilsouth.cloudapp.azure.com:3838/ben/>>.
- EPE, E. D. P. E., 2020b. “Balanco Energético Nacional 2020”. Ministério de Minas e Energia, b. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-479/topico-521/Relato%CC%81rio%20Si%CC%81ntese%20BEN%202020-ab%202019_Final.pdf>.
- ESPANA, M., LANDAU, I. D., 1978, “Reduced order bilinear models for distillation columns”, *Automatica*, v. 14, n. 4, pp. 345–355.
- EVARISTO, ARGIMIRO, 2015, “Apostila de Interpolação Polinomial”. PEQ/COPPE/UFRJ.

- FABER, R., LI, B., LI, P., et al., 2006, “Data reconciliation for real-time optimization of an industrial coke-oven-gas purification process”, *Simulation Modelling Practice and Theory*, v. 14, n. 8, pp. 1121–1134.
- FABER, R., ARELLANO-GARCIA, H., LI, P., et al., 2007, “An optimization framework for parameter estimation of large-scale systems”, *Chemical Engineering and Processing: Process Intensification*, v. 46, n. 11, pp. 1085–1095.
- FAIR, R. C., 1974, “On the robust estimation of econometric models”. In: *Annals of Economic and Social Measurement, Volume 3, number 4*, NBER, pp. 667–677.
- FARIAS, A. C., 2009, *Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros*. Tese de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS.
- FARZI, A., MEHRABANI-ZEINABAD, A., BOOZARJOMEHRY, R. B., 2008, “Data reconciliation: Development of an object-oriented software tool”, *Korean Journal of Chemical Engineering*, v. 25, n. 5, pp. 955–965.
- FARZI, A., MEHRABANI-ZEINABAD, A., BOZORGMEHRY BOOZARJOMEHRY, R., 2009, “On-Line Nonlinear Dynamic Data Reconciliation Using Extended Kalman Filtering: Application to a Distillation Column and a CSTR”, *Iranian Journal of Chemistry and Chemical Engineering (IJCCE)*, v. 28, n. 3, pp. 1–14.
- FEITAL, T., PINTO, J. C., 2015, “Use of variance spectra for in-line validation of process measurements in continuous processes”, *The Canadian Journal of Chemical Engineering*, v. 93, n. 8, pp. 1426–1437.
- FEITAL, T., PRATA, D. M., PINTO, J. C., 2014, “Comparison of methods for estimation of the covariance matrix of measurement errors”, *The Canadian Journal of Chemical Engineering*, v. 92, n. 12, pp. 2228–2245.
- FERNANDES, F. A. N., PIZZO, S. M., MORAES, J. D., 2006, “Apostila de Termodinâmica Química”. 1ª Edição, Engenharia Química – Universidade Federal do Ceará, Brasil.
- FINLAYSON, B., SCRIVEN, L., 1966, “The method of weighted residuals—a review”, *Appl. Mech. Rev.*, v. 19, n. 9, pp. 735–748.
- FISCHER, H., 2010, *A history of the central limit theorem: From classical to modern probability theory*. Springer Science & Business Media.

- FRANCIS, J. B., 1909, *Lowell Hydraulic Experiments: Being a Selection from Experiments on Hydraulic Motors, on the Flow of Water Over Weirs, in Open Canals of Uniform Rectangular Section, and Through Submerged Orifices and Diverging Tubes, Made at Lowell, Massachusetts*. D. Van Nostrand.
- FRIDAY, J. R., SMITH, B. D., 1964, “An analysis of the equilibrium stage separations problem—formulation and convergence”, *AIChE journal*, v. 10, n. 5, pp. 698–707.
- GARY, F., 2016. “Reduce column energy consumption: Focus on five costly distillation errors and how to prevent them”. Disponível em: <<https://www.chemicalprocessing.com/articles/2008/123/>>.
- GERTLER, J., ALMASY, G., 1973, “Balance calculations through dynamic system modelling”, *Automatica*, v. 9, n. 1, pp. 79–85.
- GRACIANO, J. E. A., 2016, *Real time optimization in chemical process: evaluation of strategies, improvements and industrial application*. Tese de Doutorado, Universidade de São Paulo, São Paulo, SP.
- GUIMARÃES, D. A., 2007, “Probabilidade, Estatística e Processos Estocásticos”. Notas de Aulas, Instituto Nacional de Telecomunicações-Inatel.
- GUPTA, G., NARASIMHAN, S., 1993, “Application of neural networks for gross error detection”, *Industrial & engineering chemistry research*, v. 32, n. 8, pp. 1651–1657.
- HALVORSEN, I. J., SKOGESTAD, S., 2011, “Energy efficient distillation”, *Journal of Natural Gas Science and Engineering*, v. 3, n. 4, pp. 571–580.
- HAMPEL, F. R., 1968, *Contribution to the theory of robust estimation*. Tese de Doutorado, University of California, Berkeley.
- HAMPEL, F. R., 1974, “The influence curve and its role in robust estimation”, *Journal of the american statistical association*, v. 69, n. 346, pp. 383–393.
- HAMPEL, F. R., RONCHETTI, E. M., ROUSSEUW, P. J., et al., 1986, *Robust statistics: the approach based on influence functions*. John Wiley & Sons.
- HENIS, J. M., TRIPODI, M. K., 1981, “Composite hollow fiber membranes for gas separation: the resistance model approach”, *Journal of Membrane Science*, v. 8, n. 3, pp. 233–246.

- HIMMELBLAU, D. M., 1978, *Fault detection and diagnosis in chemical and petrochemical processes*, v. 8. Elsevier Science Ltd.
- HLAVÁČEK, V., 1977, “Analysis of a complex plant-steady state and transient behavior”, *Computers & Chemical Engineering*, v. 1, n. 1, pp. 75–100.
- HOAGLIN, D. C., MOSTELLER, F., TUKEY, J. W., 1983, *Understanding robust and exploratory data analysis*, v. 3. Wiley New York.
- HODOUIN, D., MAKNI, S., 1996, “Real-time reconciliation of mineral processing plant data using bilinear material balance equations coupled to empirical dynamic models”, *International Journal of Mineral Processing*, v. 48, n. 3-4, pp. 245–264.
- HOLLAND, C. D., 1981, *Fundamentals of multicomponent distillation*. McGraw-Hill.
- HOLLAND, C. D., LIAPIS, A. I., 1983, *Computer methods for solving dynamic separation problems*. McGraw-Hill.
- HONEYWELL, 2018. “Honeywell’s UniSim Design Suite”. Disponível em: <https://www.honeywellprocess.com/en-US/explore/products/advanced-applications/unisim/Pages/unisim-design-suite.aspx>.
- HONEYWELL, 2021. “Production Balance”. Honeywell. Disponível em: <https://www.honeywellprocess.com/library/marketing/notes/Production-Balance-pin.pdf>.
- HOSEINNEZHAD, R., BAB-HADIASHAR, A., 2011, “An M-estimator for high breakdown robust estimation in computer vision”, *Computer Vision and Image Understanding*, v. 115, n. 8, pp. 1145–1156.
- HUBER, P. J., 1964, “Robust estimation of a location parameter”, *The Annals of Mathematical Statistics*, v. 35, n. 1, pp. 73–101.
- HUBER, P. J., 1981, *Robust Statistics*. John Wiley & Sons.
- HUBER, P. J., 1984, “Finite sample breakdown of M-and P-estimators”, *The Annals of Statistics*, pp. 119–126.
- HUBER, P. J., RONCHETTI, E. M., 2009, *Robust Statistics*, v. 10. A John Wiley & Sons, Inc., Publication.

- HUCKABA, C., FRANKE, F., MAY, F., et al., 1965, “Experimental Confirmation of a Predictive Model for Dynamic Distillation”. In: *Chem. Eng. Prog. Symp. Ser.*, v. 61, p. 126.
- HUMPHREY, J. L., SIEBERT, A. F., 1992, “Separation technologies; An opportunity for energy savings”, *Chemical Engineering Progress*, v. 88, n. 3.
- JEFFREYS, H., 1932, “An alternative to the rejection of observations”, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, v. 137, n. 831, pp. 78–87.
- JIN, S., LI, X., HUANG, Z., et al., 2012, “A New Target function for robust data reconciliation”, *Industrial & Engineering Chemistry Research*, v. 51, n. 30, pp. 10220–10224.
- JOHNSTON, L. P., KRAMER, M. A., 1995, “Maximum likelihood data rectification: Steady-state systems”, *AIChE Journal*, v. 41, n. 11, pp. 2415–2426.
- JUREČKOVA, J., 1984, “Rates of Consistency of Classical One-Side Tests”. In: *Robustness of Statistical Methods and Nonparametric Statistics*, Springer, pp. 60–62.
- KALMAN, R. E., 1960, “A new approach to linear filtering and prediction problems”, *Journal of basic Engineering*, v. 82, n. 1, pp. 35–45.
- KIENLE, A., 2000, “Low-order dynamic models for ideal multicomponent distillation processes using nonlinear wave propagation theory”, *Chemical Engineering Science*, v. 55, n. 10, pp. 1817–1828.
- KIM, I.-W., LIEBMAN, M. J., EDGAR, T. F., 1991, “A sequential error-in-variables method for nonlinear dynamic systems”, *Computers & chemical engineering*, v. 15, n. 9, pp. 663–670.
- KISAKÜREK, B., 1983, “A Predictive Model for Dynamic Distillation”, *Chemical Engineering Communications*, v. 20, n. 1-2, pp. 63–79.
- KISTER, H., 1992, *Distillation design*. McGraw-Hill Professional.
- KNEPPER, J. C., GORMAN, J. W., 1980, “Statistical analysis of constrained data sets”, *AIChE Journal*, v. 26, n. 2, pp. 260–264.

- KOOIJMAN, H. A., TAYLOR, R., 1995, "A nonequilibrium model for dynamic simulation of tray distillation columns", *AIChE Journal*, v. 41, n. 8, pp. 1852–1863.
- KORPELA, T., SUOMINEN, O., MAJANNE, Y., et al., 2016, "Robust data reconciliation of combustion variables in multi-fuel fired industrial boilers", *Control Engineering Practice*, v. 55, pp. 101–115.
- KOVVALI, A. S., VEMURY, S., ADMASSU, W., 1994, "Modeling of multicomponent countercurrent gas permeators", *Industrial & engineering chemistry research*, v. 33, n. 4, pp. 896–903.
- KRISHNAMURTHY, R., TAYLOR, R., 1985, "Simulation of packed distillation and absorption columns", *Industrial and engineering chemistry. Process design and development*, v. 24, n. 3, pp. 513–524.
- KUEHN, D. R., DAVIDSON, H., 1961, "Computer control II. Mathematics of control", *Chemical Engineering Progress*, v. 57, n. 6, pp. 44–47.
- KUMAR, A., DAOUTIDIS, P., 2003, "Nonlinear model reduction and control for high-purity distillation columns", *Industrial & engineering chemistry research*, v. 42, n. 20, pp. 4495–4505.
- LASHKARI, S., KRUCZEK, B., 2012, "Reconciliation of membrane properties from the data influenced by resistance to accumulation of gasses in constant volume systems", *Desalination*, v. 287, pp. 178–189.
- LEI, Z., CHEN, B., DING, Z., 2005, *Special distillation processes*. Elsevier.
- LEWIS, W., MATHESON, G., 1932, "Studies in distillation", *Industrial & Engineering Chemistry*, v. 24, n. 5, pp. 494–498.
- LID, T., SKOGESTAD, S., 2008a, "Data reconciliation and optimal operation of a catalytic naphtha reformer", *Journal of Process Control*, v. 18, n. 3-4, pp. 320–331.
- LID, T., SKOGESTAD, S., 2008b, "Scaled steady state models for effective on-line applications", *Computers & Chemical Engineering*, v. 32, n. 4-5, pp. 990–999.
- LIEBMAN, M. J., EDGAR, T. F., LASDON, L. S., 1992, "Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques", *Computers & chemical engineering*, v. 16, n. 10-11, pp. 963–986.

- LIEBMAN, M. J., 1991, *Reconciliation of process measurements using statistical and nonlinear programming techniques*. Tese de Doutorado, University of Texas at Austin.
- LIEBMAN, M., EDGAR, T., 1988, “Data reconciliation for nonlinear processes”. In: *AIChE Annual Meeting, Washington, DC*.
- LINGKE, Z., HONGYE, S., JIAN, C., 2006, “A New Method to Solve Robust Data Reconciliation in Nonlinear Process”, *Chinese Journal of Chemical Engineering*, v. 14, n. 3, pp. 357–363.
- LIU, J., SU, Q., MORENO, M., et al., 2018, “Robust state estimation of feeding–blending systems in continuous pharmaceutical manufacturing”, *Chemical Engineering Research and Design*, v. 134, pp. 140–153.
- LIU, W., POKHAREL, P. P., PRINCIPE, J. C., 2006, “Error entropy, correntropy and m-estimation”. In: *Machine Learning for Signal Processing, 2006. Proceedings of the 2006 16th IEEE Signal Processing Society Workshop on*, pp. 179–184. IEEE.
- LIU, W., POKHAREL, P. P., PRÍNCIPE, J. C., 2007, “Correntropy: Properties and applications in non-Gaussian signal processing”, *IEEE Transactions on Signal Processing*, v. 55, n. 11, pp. 5286–5298.
- LLANOS, C. E., SANCHÉZ, M. C., MARONNA, R. A., 2015, “Robust estimators for data reconciliation”, *Industrial & Engineering Chemistry Research*, v. 54, n. 18, pp. 5096–5105.
- LLANOS, C. E., SÁNCHEZ, M. C., MARONNA, R. A., 2017, “Classification of systematic measurement errors within the framework of robust data reconciliation”, *Industrial & Engineering Chemistry Research*, v. 56, n. 34, pp. 9617–9628.
- LLNL, 2020. “Energy Flow Charts: Charting the Complex Relationships among Energy, Water, and Carbon”. Lawrence Livermore National Laboratory & U.S. Department of Energy. Disponível em: <<https://flowcharts.llnl.gov/>>.
- LSU, 2021. “Interactive On-Line Optimization”. Louisiana State University. Disponível em: <<https://www.lsu.edu/mpri/computerprograms/on-lineopt.php>>.

- LUYBEN, W. L., VERNEUIL, V. S., GERSTER, J., 1964, “Experimental transient response of a pilot-plant distillation column: Part IV. Response of a ten-tray column”, *AIChE Journal*, v. 10, n. 3, pp. 357–364.
- MAH, R. S., STANLEY, G. M., DOWNING, D. M., 1976, “Reconciliation and rectification of process flow and inventory data”, *Industrial & Engineering Chemistry Process Design and Development*, v. 15, n. 1, pp. 175–183.
- MAH, R. S. H., 1990, *Chemical process structures and information flows*. Elsevier.
- MAH, R., TAMHANE, A., 1982, “Detection of gross errors in process data”, *AIChE Journal*, v. 28, n. 5, pp. 828–830.
- MAKNI, S., HODOUIN, D., BAZIN, C., 1995, “A recursive node imbalance method incorporating a model of flowrate dynamics for on-line material balance of complex flowsheets”, *Minerals Engineering*, v. 8, n. 7, pp. 753–766.
- MANUAL, H. U., 1999, “Hyprotech Ltd”, *Calgary, Canada*.
- MARSHALL, W. R., PIGFORD, R. L., 1947. “The application of differential equations to chemical engineering problems”. University of Delaware. Disponível em: <<https://catalog.hathitrust.org/Record/000616805>>.
- MCBRAYER, K. F., EDGAR, T. F., 1995, “Bias detection and estimation in dynamic data reconciliation”, *Journal of Process Control*, v. 5, n. 4, pp. 285–289.
- MIGON, H., 2007. “Análise de séries temporais”. Departamento de Métodos Estatísticos, Instituto de Matemática, UFRJ. Disponível em: <<http://www.dme.ufrj.br/dani/pdf/slidespartefrequentista.pdf>>.
- MINGFANG, K., BINGZHEN, C., BO, L., 2000, “An Integral approach to dynamic data rectification”, *Computers & Chemical Engineering*, v. 24, n. 2-7, pp. 749–753.
- MODLA, G., LÁNG, P., 2017, “Decrease of the energy demand of distillation with vapour recompression”, *International scientific conference on advances in mechanical engineering*, pp. 339–344.
- MOKHATAB, S., POE, W. A., MAK, J. Y., 2018, *Handbook of natural gas transmission and processing: principles and practices*. Gulf professional publishing.

- MORAES, M. B. D. C., 2017, “Métodos Quantitativos Avançados”, Disponível em: <https://slideplayer.com.br/slide/12756512/>.
- MORETTIN, P. A., TOLOI, C. M., 2018, *Análise de séries temporais: modelos lineares univariados*. Editora Blucher.
- MUNOZ, J. C., CHEN, J., 2012, “Removal of the effects of outliers in batch process data through maximum correntropy estimator”, *Chemometrics and Intelligent Laboratory Systems*, v. 111, n. 1, pp. 53–58.
- NAPHTALI, L. M., SANDHOLM, D. P., 1971, “Multicomponent separation calculations by linearization”, *AIChE Journal*, v. 17, n. 1, pp. 148–153.
- NARASIMHAN, S., MAH, R., 1987, “Generalized likelihood ratio method for gross error identification”, *AIChE Journal*, v. 33, n. 9, pp. 1514–1521.
- NARASIMHAN, S., JORDACHE, C., 1999, *Data reconciliation and gross error detection: An intelligent use of process data*. Gulf Professional Publishing.
- NASCIMENTO, A. J. V., 2013a, *Análise de modelos reduzidos de colunas de destilação para aplicações em tempo real*. Tese de Mestrado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, a.
- NASCIMENTO, W. S., 2013b, *Sobre algumas características da entropia de Shannon para sistemas atômicos confinados*. Tese de Mestrado, Universidade Federal da Bahia, Instituto de Física, b.
- NASRI, Z., BINOUS, H., 2009, “Rigorous distillation dynamics simulations using a computer algebra”, *Computer Applications in Engineering Education*, v. 20, n. 2, pp. 193–202.
- NEVES, J. F. M., 2011, *A Programação Linear no Ensino Secundário*. Tese de Mestrado, Departamento de Matemática, Universidade de Aveiro, Aveiro, Portugal.
- NICHOLSON, B., LÓPEZ-NEGRETE, R., BIEGLER, L. T., 2014, “On-line state estimation of nonlinear dynamic systems with gross errors”, *Computers & Chemical Engineering*, v. 70, pp. 149–159.
- NICULESCU, A., CONSTANTIN, T., ANA, G., et al., 2017, “Dynamic simulation of a multicomponent distillation column for DT separation”, *Fusion Engineering and Design*.

- NOORAI, A., ROMAGNOLI, J. A., 1998, "Implementation of advanced operational and control techniques for a pilot distillation column within a DCS environment", *Computers & chemical engineering*, v. 22, n. 4-5, pp. 695-708.
- NORONHA, F. B., PINTO, J. C., MONTEIRO, J. L., et al., 1993, "ESTIMA: Um Pacote Computacional para Estimaco de Parmetros e Projeto de Experimentos, Relatório Técnico PEQ/COPPE", *Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil*.
- NYQUIST, H., 1928, "Certain topics in telegraph transmission theory", *Transactions of the American Institute of Electrical Engineers*, v. 47, n. 2, pp. 617-644.
- OLIVEIRA JR., A. M. D., 2006, *Estimaco de parmetros em modelos de processo usando dados industriais e tcnica de reconciliao de dados*. Tese de Doutorado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- PAI, C. C., FISHER, G. D., 1988, "Application of Broyden's method to reconciliation of nonlinearly constrained data", *AIChE Journal*, v. 34, n. 5, pp. 873-876.
- PARZEN, E., 1962, "Extraction and detection problems and reproducing kernel Hilbert spaces", *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, v. 1, n. 1, pp. 35-62.
- PEISER, A., GROVER, S., 1962, "Dynamic simulation of a distillation tower", *Chemical Engineering Progress*, v. 58, n. 9, pp. 65.
- PETLYUK, F. B., 2004, *Distillation theory and its application to optimal design of separation units*. Cambridge University Press.
- PINEDA, J. O. D. C., 2006, *A entropia segundo Claude Shannon: o desenvolvimento do conceito fundamental da teoria da informao*. Tese de Mestrado, Pontifícia Universidade Católica de São Paulo, São Paulo, SP.
- PINTO, J., BISCAIA, E., 1988, "Order reduction strategies for models of staged separation systems", *Computers & chemical engineering*, v. 12, n. 8, pp. 821-831.
- PLUS, A., 2003, "Aspen Plus user guide", *Aspen Technology Limited, Cambridge, Massachusetts, United States*.

- POLING, B. E., PRAUSNITZ, J. M., O'CONNELL, J. P., et al., 2001, *The properties of gases and liquids*, v. 5. Mcgraw-hill New York.
- POLO, J. C. F., 2017, *Estudo das características estocásticas do roll paramétrico*. Tese de Doutorado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- PRATA, D. M., 2005, *Reconciliação de Dados em um Reator de Polimerização*. Tese de Mestrado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- PRATA, D. M., 2009, *Reconciliação robusta de dados para monitoramento em tempo real*. Tese de Doutorado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- PRATA, D. M., LIMA, E. L., PINTO, J. C., 2008a, "In-Line Monitoring of Bulk Polypropylene Reactors Based on Data Reconciliation Procedures", *Macromolecular symposia*, v. 271, n. 1, pp. 26–37.
- PRATA, D. M., PINTO, J. C., LIMA, E. L., 2008b, "Comparative analysis of robust estimators on nonlinear dynamic data reconciliation", *Computer Aided Chemical Engineering*, v. 25, pp. 501–506.
- PRATA, D. M., LIMA, E. L., PINTO, J. C., 2009, "Nonlinear dynamic data reconciliation in real time in actual processes", *Computer Aided Chemical Engineering*, v. 27, pp. 47–54.
- PRATA, D. M., SCHWAAB, M., LIMA, E. L., et al., 2009, "Nonlinear dynamic data reconciliation and parameter estimation through particle swarm optimization: Application for an industrial polypropylene reactor", *Chemical Engineering Science*, v. 64, n. 18, pp. 3953–3967.
- PRATA, D. M., SCHWAAB, M., LIMA, E. L., et al., 2010, "Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor", *Chemical Engineering Science*, v. 65, n. 17, pp. 4943–4954.
- PRAUSNITZ, J. M., LICHTENTHALER, R. N., DE AZEVEDO, E. G., 1998, *Molecular thermodynamics of fluid-phase equilibria*. Pearson Education.
- QUELHAS, A. D., DE JESUS, N. J. C., PINTO, J. C., 2013, "Common vulnerabilities of RTO implementations in real chemical processes", *The Canadian Journal of Chemical Engineering*, v. 91, n. 4, pp. 652–668.

- RACKLEY, S. A., 2017, *Carbon capture and storage*. Butterworth-Heinemann.
- RAGOT, J., CHADLI, M., MAQUIN, D., 2005, “Mass balance equilibration: a robust approach using contaminated distribution”, *AIChE Journal*, v. 51, n. 5, pp. 1569–1575.
- RAMAMURTHI, Y., BEQUETTE, B. W., 1990, “Data reconciliation of systems with unmeasured variables using nonlinear programming techniques”. In: *American Institute of Chemical Engineers Spring National Meeting*.
- RAO, R. R., NARASIMHAN, S., 1996, “Comparison of techniques for data reconciliation of multicomponent processes”, *Industrial & engineering chemistry research*, v. 35, n. 4, pp. 1362–1368.
- REILLY, P., CARPANI, R., 1963, “Application of statistical theory of adjustment to material balances”. In: *Proceedings of the 13th Canadian chemical engineering conference*, pp. 21–23.
- REY, W., 1983, *Introduction to robust and quasi-robust statistical methods*. Springer Science & Business Media.
- RIBEIRO, L., SECHI, A., BISCAIA, E., 2010, “A New Technique of Model Order Reduction Based on Weighted Residuals in Discrete Domain”. In: *Proceedings of the 20th European Symposium on Computer Aided Process Engineering, Naples, Italy*, pp. 6–9.
- RIBEIRO, L. D., SECCHI, A. R., BISCAIA JR, E. C., 2015, “Highly-accurate model order reduction technique on a discrete domain”, *Brazilian Journal of Chemical Engineering*, v. 32, n. 3, pp. 767–779.
- RIBEIRO, L. D., 2011, *Nova técnica de redução de ordem de modelo baseada em resíduos ponderados no domínio discreto*. Tese de Mestrado, COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- ROLLINS, D., DAVIS, J., 1992, “Unbiased estimation of gross errors in process measurements”, *AIChE Journal*, v. 38, n. 4, pp. 563–572.
- ROMAGNOLI, J. A., SANCHEZ, M. C., 1999, *Data processing and reconciliation for chemical process operations*, v. 2. Academic Press.
- ROUSSEEUW, P., YOHAI, V., 1984, “Robust regression by means of S-estimators”. In: *Robust and nonlinear time series analysis*, Springer, pp. 256–272.

- ROUSSEEUW, P. J., CROUX, C., 1992, “Explicit scale estimators with high breakdown point”, *L1-Statistical analysis and related methods*, v. 1, pp. 77–92.
- ROUSSEEUW, P. J., CROUX, C., 1993, “Alternatives to the median absolute deviation”, *Journal of the American Statistical association*, v. 88, n. 424, pp. 1273–1283.
- ROUSSEEUW, P., LEROY, A., 1987, *Robust Regression and Outlier Detection*.
- SANTAMARÍA, I., POKHAREL, P. P., PRINCIPE, J. C., 2006, “Generalized correlation function: definition, properties, and application to blind equalization”, *IEEE Transactions on Signal Processing*, v. 54, n. 6, pp. 2187–2197.
- SARRUF, I. C. P., 2014, *Reconciliação robusta de dados utilizando o simulador EMSO*. Tese de Mestrado, Universidade Federal Fluminense, Niterói, RJ.
- SCHLADT, M., HU, B., 2007, “Soft sensors based on nonlinear steady-state data reconciliation in the process industry”, *Chemical Engineering and Processing: Process Intensification*, v. 46, n. 11, pp. 1107–1115.
- SEADER, J. D., HENLEY, E. J., ROPER, D. K., 2010, *Separation process principles*. Wiley New York, 3rd Edition.
- SEBORG, D. E., MELLICHAMP, D. A., EDGAR, T. F., et al., 2010, *Process dynamics and control*. John Wiley & Sons.
- SECCHI, A., 1988, *Controle Adaptativo de Colunas de Destilação Utilizando Modelos Reduzidos*. Tese de Mestrado, Dissertação de Mestrado, COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- SEFERLIS, P., HRYMAK, A., 1994, “Adaptive collocation on finite elements models for the optimization of multistage distillation units”, *Chemical engineering science*, v. 49, n. 9, pp. 1369–1382.
- SERTH, R., HEENAN, W., 1986, “Gross error detection and data reconciliation in steam-metering systems”, *AIChE Journal*, v. 32, n. 5, pp. 733–742.
- SHANNON, C. E., 1948, “A mathematical theory of communication, Part I, Part II”, *Bell Syst. Tech. J.*, v. 27, pp. 623–656.
- SHANNON, C. E., WEAVER, W., 1964, *The mathematical theory of communication*. University of Illinois press.

- SHANNON, C. E., 1949, “Communication in the presence of noise”, *Proceedings of the IRE*, v. 37, n. 1, pp. 10–21.
- SHEVLYAKOV, G. L., VILCHEVSKI, N. O., 2001, *Robustness in Data Analysis: criteria and methods*. Walter de Gruyter.
- SIGMAFINE, 2021. “Sigmafine”. Pimsoft Inc. Disponível em: <<https://sigmafine.pimsoftinc.com/wp-content/uploads/2017/05/Sigmafine-Overview-SF.pdf>>.
- SIMÕES, A. R., 2000, *Modelos de ordem reduzida e simulação em colunas de destilação*. Tese de Mestrado, Universidade Estadual de Campinas, Campinas, SP.
- SINGH, V., JAIN, P., DAS, C., 2013, “Performance of spiral wound ultrafiltration membrane module for with and without permeate recycle: experimental and theoretical consideration”, *Desalination*, v. 322, pp. 94–103.
- SMITH, J., VAN NESS, H., ABBOTT, M., 2005, *Introdução à Termodinâmica da Engenharia Química*. 7^a edição. Rio de Janeiro: LTC editora.
- SOARES, R. D. P., SECCHI, A. R., 2003, “EMSO: A new environment for modelling, simulation and optimisation”, *Computer Aided Chemical Engineering*, v. 14, pp. 947–952.
- SOARES, R. D. P., 2003, *Desenvolvimento de um simulador genérico de processos dinâmicos*. Tese de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS.
- SODERSTROM, T. A., EDGAR, T. F., RUSSO, L. P., et al., 2000, “Industrial application of a large-scale dynamic data reconciliation strategy”, *Industrial & Engineering Chemistry Research*, v. 39, n. 6, pp. 1683–1693.
- SPEIGHT, J. G., 2018, *Natural gas: a basic handbook*. Gulf Professional Publishing.
- STANLEY, G., MAH, R., 1977, “Estimation of flows and temperatures in process networks”, *AIChE Journal*, v. 23, n. 5, pp. 642–650.
- STANLEY, G., MAH, R., 1981a, “Observability and redundancy classification in process networks: Theorems and algorithms”, *Chemical Engineering Science*, v. 36, n. 12, pp. 1941–1954.
- STANLEY, G., MAH, R., 1981b, “Observability and redundancy in process data estimation”, *Chemical Engineering Science*, v. 36, n. 2, pp. 259–272.

- STATHAKI, A., MELLICHAMP, D. A., SEBORG, D. E., 1985, “Dynamic simulation of a multicomponent distillation column with asymmetric dynamics”, *The Canadian Journal of Chemical Engineering*, v. 63, n. 3, pp. 510–518.
- STEWART, W. E., LEVIEN, K. L., MORARI, M., 1985, “Simulation of fractionation by orthogonal collocation”, *Chemical Engineering Science*, v. 40, n. 3, pp. 409–421.
- STROMBERG, A., OTHERS, 2004, “Why write statistical software? The case of robust statistical methods”, *Journal of Statistical Software*, v. 10, n. 5, pp. 1–8.
- SU, Q., BOMMIREDDY, Y., SHAH, Y., et al., 2019, “Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing”, *International Journal of Pharmaceutics*, v. 563, pp. 259–272.
- SWARTZ, C., 1989, “Data reconciliation for generalized flowsheet applications”. In: *ABSTRACTS OF PAPERS OF THE AMERICAN CHEMICAL SOCIETY*, v. 197, pp. 58–IAEC. AMER CHEMICAL SOC 1155 16TH ST, NW, WASHINGTON, DC 20036.
- SZKLO, A., ULLER, V. C., 2008, *Fundamentos do refino de petróleo: tecnologia e economia*. Interciência, 2^a edição.
- TERRY, P. A., HIMMELBLAU, D. M., 1993, “Data rectification and gross error detection in a steady state process via artificial neural networks”, *Industrial & engineering chemistry research*, v. 32, n. 12, pp. 3020–3028.
- THIELE, E., GEDDES, R., 1933, “Computation of distillation apparatus for hydrocarbon mixtures”, *Industrial & Engineering Chemistry*, v. 25, n. 3, pp. 289–295.
- TJOA, I.-B., BIEGLER, L. T., 1991, “Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems”, *Computers & chemical engineering*, v. 15, n. 10, pp. 679–690.
- TONG, H., CROWE, C. M., 1995, “Detection of gross errors in data reconciliation by principal component analysis”, *AIChE Journal*, v. 41, n. 7, pp. 1712–1722.
- TUKEY, J. W., 1960, “A survey of sampling from contaminated distributions”, *Contributions to probability and statistics*, v. 2, pp. 448–485.

- UEHARA, M., 2011, *Evolução dos Microprocessadores Utilizados nos Computadores Pessoais*. Programa de Graduação em Processamento de Dados, FATEC, São Paulo. Monograph.
- VACHHANI, P., RENGASWAMY, R., VENKATASUBRAMANIAN, V., 2001, “A framework for integrating diagnostic knowledge with nonlinear optimization for data reconciliation and parameter estimation in dynamic systems”, *Chemical Engineering Science*, v. 56, n. 6, pp. 2133–2148.
- VALDETARO, E. D., SCHIRRU, R., 2011, “Simultaneous model selection, robust data reconciliation and outlier detection with swarm intelligence in a thermal reactor power calculation”, *Annals of Nuclear Energy*, v. 38, n. 9, pp. 1820–1832.
- VALLURU, J., PUROHIT, J. L., PATWARDHAN, S. C., et al., 2016, “Development of a Moving Window Maximum Likelihood Parameter Estimator and its Application on Ideal Reactive Distillation System”, *IFAC-PapersOnLine*, v. 49, n. 7, pp. 484–489.
- VALLURU, J., PATWARDHAN, S. C., BIEGLER, L. T., 2018, “Development of robust extended Kalman filter and moving window estimator for simultaneous state and parameter/disturbance estimation”, *Journal of Process Control*, v. 69, pp. 158–178.
- VAN DE VUSSE, J., 1964, “Plug-flow type reactor versus tank reactor”, *Chemical Engineering Science*, v. 19, n. 12, pp. 994–996.
- VEVERKA, V. V., MADRON, F., 1997, *Material and energy balancing in the process industries: From microscopic balances to large plants*, v. 7. Elsevier.
- VICHARE, N. S., 1993, *Robust Mahalanobis distances in power system state estimation*. Tese de Doutorado, Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA.
- VRTECH, 2018. “iiSE”. Disponível em: <<http://vrtech.com.br/pt/simulador-iise/>>.
- VRTECH, 2017. “iiSE - Industrial Integrated Simulation Environment”. vrtech.com.br/pt. Disponível em: <<http://vrtech.com.br/pt/simulador-iise/>>.
- WANG, J., HENKE, G., 1966, “Tridiagonal matrix for distillation”, *Hydrocarbon Processing*, v. 45, n. 8, pp. 155.

- WANG, X., MAHALEC, V., QIAN, F., 2017, “Globally optimal dynamic real time optimization without model mismatch between optimization and control layer”, *Computers & Chemical Engineering*, v. 104, pp. 64–75.
- WANG, X., 2000, *Research on data reconciliation technique—theory and its applications*. Tese de Doutorado, Ph. D. Thesis, Zhejiang University, Hangzhou.
- WIJN, E., 1999, “Weir flow and liquid height on sieve and valve trays”, *Chemical Engineering Journal*, v. 73, n. 3, pp. 191–204.
- WILLIAMS, T. J., OTTO, R. E., 1960, “A generalized chemical processing model for the investigation of computer control”, *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics*, v. 79, n. 5, pp. 458–473.
- WONG, K. T., LUUS, R., 1980, “Model reduction of high-order multistage systems by the method of orthogonal collocation”, *The Canadian Journal of Chemical Engineering*, v. 58, n. 3, pp. 382–388.
- WONGRAT, W., SRINOPHAKUN, T., SRINOPHAKUN, P., 2005, “Modified genetic algorithm for nonlinear data reconciliation”, *Computers & chemical engineering*, v. 29, n. 5, pp. 1059–1067.
- WU, S., XU, J., LIU, W., et al., 2017, “Data Reconciliation Based on an Improved Robust Estimator and NT-MT for Gross Error Detection”. In: *Advanced Computational Methods in Life System Modeling and Simulation*, Springer, pp. 400–409.
- XIE, S., YANG, C., YUAN, X., et al., 2019, “A novel robust data reconciliation method for industrial processes”, *Control Engineering Practice*, v. 83, pp. 203–212.
- YANG, Y., TEN, R., JAO, L., 1995, “A study of gross error detection and data reconciliation in process industries”, *Computers & chemical engineering*, v. 19, pp. 217–222.
- YAO, M., WANG, H., 2015, “On-line monitoring of batch processes using generalized additive kernel principal component analysis”, *Journal of Process Control*, v. 28, pp. 56–72.
- YNOGUTI, C. A., 2011. “Probabilidade, Estatística e Processos Estocásticos”. Disponível em: <<https://www.linux.ime.usp.br/~daniLOSS/antes-2012/Apostila%20TP501%20-%20Ynoguti,%202011.pdf>>.

- YOUNG, S., 1903, *Fractional distillation*. Macmillan and Company, limited.
- YU, W.-C., JOHNSON, A. E. J., PLANCHARD, J. A., 1975. “Dynamic simulation of a multicomponent distillation column”. Louisiana State University. Disponível em: <http://www.informs-sim.org/wsc75papers/1975_0092.pdf>.
- ZHANG, J., LI, G., OTHERS, 1998, “Breakdown properties of location M -estimators”, *The Annals of Statistics*, v. 26, n. 3, pp. 1170–1189.
- ZHANG, Z., PIKE, R. W., HERTWIG, T. A., 1995, “Source reduction from chemical plants using on-line optimization”, *Waste Management*, v. 15, n. 3, pp. 183–191.
- ZHANG, Z., CHEN, J., 2014, “Simultaneous data reconciliation and gross error detection for dynamic systems using particle filter and measurement test”, *Computers & Chemical Engineering*, v. 69, pp. 66–74.
- ZHANG, Z., CHEN, J., 2015, “Correntropy based data reconciliation and gross error detection and identification for nonlinear dynamic processes”, *Computers & Chemical Engineering*, v. 75, pp. 120–134.
- ZHANG, Z., SHAO, Z., CHEN, X., et al., 2010, “Quasi-weighted least squares estimator for data reconciliation”, *Computers & chemical engineering*, v. 34, n. 2, pp. 154–162.
- ZHANG, Z., CHUANG, Y.-Y., CHEN, J., 2014, “Methodology of data reconciliation and parameter estimation for process systems with multi-operating conditions”, *Chemometrics and Intelligent Laboratory Systems*, v. 137, pp. 110–119.
- ZOUBIR, A. M., KOIVUNEN, V., OLLILA, E., et al., 2018, *Robust statistics for signal processing*. Cambridge University Press.
- ÖZYURT, D. B., PIKE, R. W., 2004, “Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes”, *Computers & chemical engineering*, v. 28, n. 3, pp. 381–402.

Apêndice A

Artigo Publicado - Modeling of Spiral Wound Membranes for Gas Separations—Part II: Data Reconciliation for Online Monitoring

Resumo: O presente trabalho apresenta uma metodologia baseada na reconciliação de dados para monitorar os processos de separação por membrana de forma confiável, online e em tempo real pela primeira vez. A metodologia proposta foi implementada de acordo com as seguintes etapas: aquisição de dados; pré-tratamento de dados; caracterização de dados; reconciliação de dados; detecção de erros grosseiros; e avaliação crítica dos dados medidos com um sensor virtual. A aquisição de dados constituiu a etapa mais lenta do processo de monitoramento, conforme esperado em aplicações de tempo real. A etapa de pré-tratamento foi fundamental para garantir a robustez do código e a caracterização inicial dos dados coletados foi realizada offline. A caracterização dos dados mostrou que a modelagem em regime permanente do processo seria adequada, permitindo também a implementação de procedimentos numéricos mais rápidos para a etapa de reconciliação dos dados. A etapa de reconciliação de dados teve um bom desempenho, de forma rápida e consistente. Assim, a reconciliação de dados permitiu estimar variáveis não medidas, desempenhando o papel de sensor virtual e permitindo a futura instalação de um gêmeo digital. Além disso, o monitoramento dos *bias* de medição constituiu uma ferramenta para o diagnóstico de medição. Conforme mostrado no manuscrito, a metodologia proposta pode ser implementada com sucesso online e em tempo real para monitoramento dos processos de separação por membrana, conforme mostrado por um painel desenvolvido em aplicativo-web.

Article

Modeling of Spiral Wound Membranes for Gas Separations—Part II: Data Reconciliation for Online Monitoring

Diego Queiroz Faria de Menezes ^{1,*}, Marília Caroline Cavalcante de Sá ¹,
Tahyná Barbalho Fontoura ¹, Thiago Koichi Anzai ², Fabio Cesar Diehl ²,
Pedro Henrique Thompson ² and Jose Carlos Pinto ¹

¹ Programa de Engenharia Química/COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro CEP 21941-972, RJ, Brazil; marilia@peq.coppe.ufrj.br (M.C.C.d.S.); tahyna@peq.coppe.ufrj.br (T.B.F.); pinto@peq.coppe.ufrj.br (J.C.P.)

² Centro de Pesquisas Leopoldo Américo Miguez de Mello—CENPES, Petrobras—Petróleo Brasileiro SA, Rio de Janeiro CEP 21941-915, RJ, Brazil; tanzai@petrobras.com.br (T.K.A.); fabio.diehl@petrobras.com.br (F.C.D.); pedrothompson@petrobras.com.br (P.H.T.)

* Correspondence: dmenezes@coppe.ufrj.br; Tel.: +55-21-98807-7489

Received: 24 July 2020; Accepted: 19 August 2020; Published: 25 August 2020



Abstract: The present work presents a methodology based on data reconciliation to monitor membrane separation processes reliably, online and in real time for the first time. The proposed methodology was implemented in accordance with the following steps: data acquisition; data pre-treatment; data characterization; data reconciliation; gross error detection; and critical evaluation of measured data with a soft sensor. The acquisition of data constituted the slowest stage of the monitoring process, as expected in real-time applications. The pre-treatment stage was fundamental to assure the robustness of the code and the initial characterization of collected data was carried out offline. The characterization of the data showed that steady-state modeling of the process would be appropriate, also allowing the implementation of faster numerical procedures for the data reconciliation step. The data reconciliation step performed well, quickly and consistently. Thus, data reconciliation allowed the estimation of unmeasured variables, playing the role of a soft sensor and allowing the future installation of a digital twin. Additionally, monitoring of measurement bias constituted a tool for measurement diagnosis. As shown in the manuscript, the proposed methodology can be successfully implemented online and in real time for monitoring of membrane separation processes, as shown through a real dashboard web application developed for monitoring of an actual industrial site.

Keywords: membrane; data reconciliation; real-time; online; monitoring

1. Introduction

A common problem in oil production is the excess of CO₂ gas present in natural gas streams. The first and most notorious issue is related to the emission of this gas into the environment. However, in addition to the possible environmental problems, the excess of CO₂ in oil streams can cause problems in the process plant, such as freezing due to pressure drop in compression and cooling sections of the plant and corrosion of metal pipelines [1]. According to an ANP (Brazilian National Agency of Petroleum, Natural Gas and Biofuels) resolution, commercial natural gas must contain a maximum of 3% (mol) of CO₂ [2]. Therefore, a possible solution to deal with the produced CO₂ is the reinjection of CO₂ into the oil well, which may also allow the increase of the productivity of the well. This can certainly minimize environmental impacts and problems in natural gas process plants.

Therefore, the CO₂ separation constitutes a fundamental step during the treatment of natural gas in oil production fields.

Different physical/chemical processes can be used to separate CO₂ from natural gas, such as cryogenic distillation, absorption, or membrane reverse osmosis processes [1,3]. Particularly, the removal of CO₂ from natural gas with help of membrane separation processes has been used since 1981 [4]. However, applications were initially limited because of intrinsic economic risks associated with the oil production activity and operation constraints related to membrane separations. Nevertheless, the scenario has been changing due to advantages related to the lower energy consumption, low capital investment, low operating costs, and more compact nature of these pieces of equipment [5–7].

Given the increase of the industrial importance of membrane separation processes, demands for development of mathematical modeling, simulation, optimization, control, statistical data treatment, and online monitoring procedures have also increased, as these techniques are fundamental for design and monitoring of chemical processes. As a consequence, the performance of the analyzed process can be evaluated more precisely and monitored, allowing the detection of failures in line and in real time. Based on these technologies, risks and time required for decision-making can be minimized [8].

Based on the previous paragraphs, the main objective of the present work is to develop and implement a web application that makes possible the online and real-time monitoring of membrane CO₂ separation processes on an industrial scale for the first time, based on rigorous numerical and statistical procedures. The application can also be used to provide information about unmeasured variables (soft sensor) and to diagnose the occurrence of gross error measurements and instrument malfunctioning. The proposed methodology comprises the following stages: (i) pre-treatment and characterization of process data; (ii) data reconciliation of process data to minimize measurement uncertainties, with the aid of mass balance equations; (iii) detection of systematic deviations for identification of process malfunctions; and (iv) observation of unmeasured variables (soft sensor or digital twin). Finally, the proposed data acquisition and visualization system is implemented online for successful monitoring of an actual industrial membrane separation site in real time for the first time.

1.1. Data Rectification

Technological and computer advancement have allowed the wide, easy, and fast access to process data of industrial plants. As a matter of fact, access to actual data are extremely important for real-time monitoring and optimization of production units [9]. The dynamic monitoring of a plant, unit or industrial equipment is increasingly necessary to improve product quality, enhance process safety, and reduce energy costs and waste; however, the acquired information must be reliable and validated with physical process constraints, as the reliability of the data are of paramount importance for any decision-making related to the analyzed process [10]. Nevertheless, process measurements are subject to errors and fluctuations due to intrinsic imprecision, degradation, malfunction, improper installation, poor calibration, and failure of measurement instruments. Additionally, human errors associated with operation and calibration, or gross errors related to the operation of the process, can result in data that do not represent the process reliably. Consequently, measured data are not expected to satisfy physical constraints precisely and are not expected to comply with conservation laws (mass, momentum, and energy balances) [11]. For these reasons, process controllers and data acquisition systems, if not treated properly, can cause the plant to operate at sub-optimal or unsafe operation conditions. In addition, decision-making based on unreliable data can lead to the occurrence of industrial accidents, reduction of product quality, and financial losses [12]. Therefore, the use of data rectification procedures can be essential to improve the quality of the information contained in the data, and consequently provide a margin of reliability for the control and optimization of the process in real time.

Data rectification procedures usually comprise three steps: variable classification; gross error detection (GED); and data reconciliation (DR). Among these three steps, DR and GED are the ones studied most often and applied more frequently in data rectification schemes [13].

1.1.1. Data Reconciliation and Gross Error Detection

The variable classification step determines whether the available information is sufficient to solve the DR problem and identify the sets of observable variables (measured and unmeasured variables, which can be estimated using the other measured variables and the process constraints) and unobservable variables (unmeasured variables that cannot be estimated). This way, the variable classification step makes possible data set size reduction in order to include only the relevant variables that can be observed and used to build the mathematical model of the process, reducing the size of the process database [14].

The GED step is performed to identify and/or eliminate/compensate for the occurrence of deviations that do not follow the admitted statistical distribution of errors. Gross errors can be caused by poor calibration of the measuring instruments, deterioration of the sensors, power surges, among other causes described previously. However, in order to obtain accurate estimates of parameters and variables, the negative influence of gross errors must be minimized or eliminated. The use of robust estimators has been frequently suggested to eliminate the negative effect of gross errors, often implemented simultaneously with DR to avoid the use of iterative and computationally intensive numerical procedures [11].

In the DR stage, measured data are adjusted in a statistically coherent manner by an estimator, which frequently is based on the maximum likelihood principle, with the support of a statistical distribution admitted a priori for measurement fluctuations. According to the DR technique, adjusted data must satisfy the conservation laws and other constraints imposed on the system, maximizing the probability of occurrence of that particular measurement and, simultaneously, respecting the mathematical model of the process. Thus, more reliable estimates can be obtained for the variables and parameters of the process [13,15]. Traditionally, the normal distribution is assumed to be valid, which results in the Weighted Least Squares (WLS) estimator [16]. For illustrative purposes, data rectification applications in the industry can be implemented as shown in Figure 1.

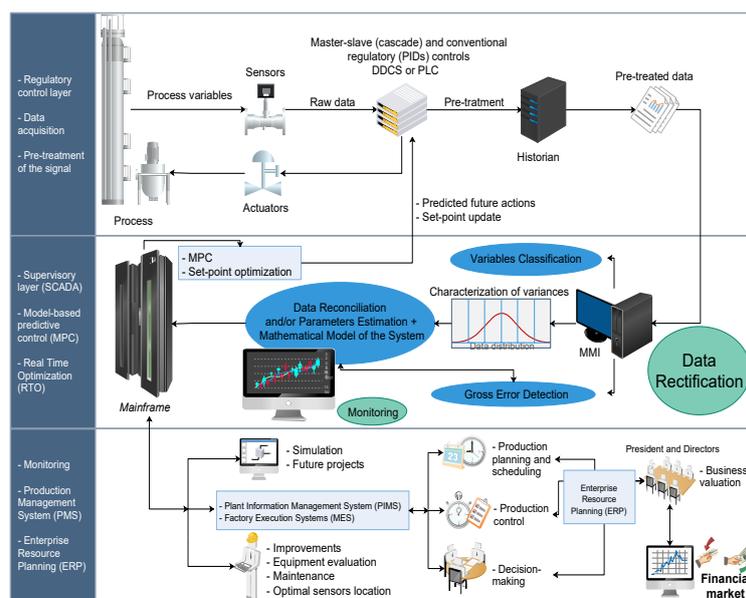


Figure 1. Illustrative representation of industrial data rectification applications [12,17].

Kuehn and Davidson [18] were the pioneers in using DR procedures in chemical engineering processes. Since then, many works have proposed the use of DR procedures for monitoring of industrial processes, although the vast majority of the published material investigates simulated processes that operate at steady-state conditions. Investigations of actual dynamic processes in real time and using actual data are scarce and have never been performed for industrial membrane separations [14,19]. Therefore, the present work contributes to the development of DR procedures through the successful implementation of an original application in an actual industrial environment and using real data in real time to perform the proposed analyses. In addition, the present work shows that similar DR applications can be implemented in many industrial membrane separation environments using simple computational resources in real time.

1.2. Membrane Separation Process

Membranes constitute excellent alternatives for gas separations due to their low installation and maintenance costs. In the industrial environment, membranes are usually organized in modules with spiral-wound or hollow fiber geometries. Hollow fiber separation units are normally applied to relatively smaller fluxes when compared to spiral-wound modules. On the other hand, spiral-wound modules are cheaper, capable of handling higher operating pressures, and are more resistant to scaling, as particles present in the feed gas stream can block the fine membrane fibers [20,21]. Because of that, the spiral-wound units are largely used in industrial gas separation processes. In a previous work of our group, a mathematical model based on a phenomenological approach for a leaf of a spiral-wound membrane was developed. The model was validated in four case studies of common gas separations, with very good performance and robustness. Furthermore, it allowed the prediction of flow rates and concentrations along the membrane leaf, which are important features for the understanding of membrane operation processes. In addition, a discretization method was proposed to solve the model, which proved to be faster and more efficient than the shooting method [22]. It is also important to emphasize that industrial spiral-wound membrane separation units for CO₂ applications are formed by several leaves, which are wound onto a central perforated collecting tube, forming one modular separation element. These elements can be arranged in series to build a membrane separation tube. Then, tubes can be organized in parallel to form a bank. Finally, the banks can be aligned in parallel to compose a membrane separation train, while the trains can be arranged in parallel to form a stage [21]. Figure 2 shows an example of this kind of unit.

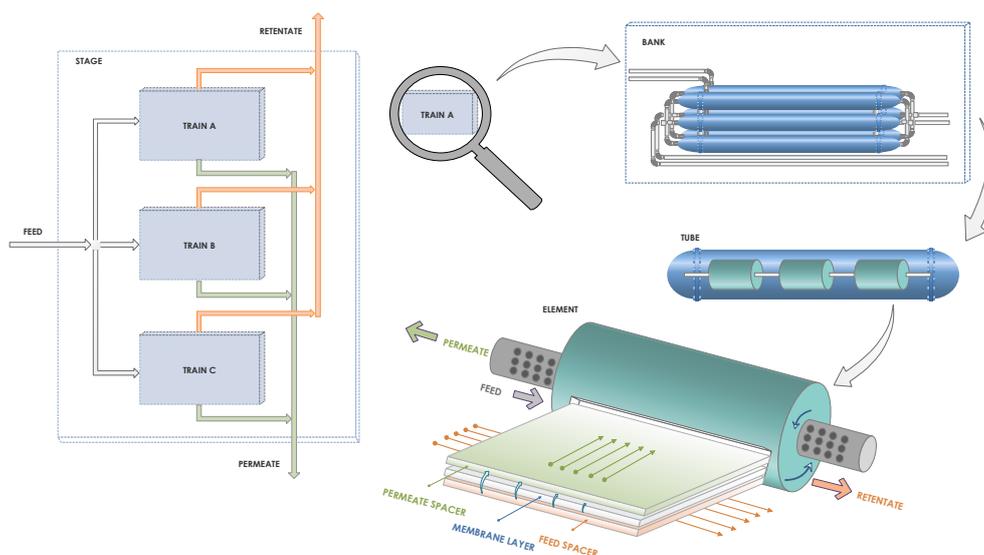


Figure 2. Schematic representation of a spiral-wound membrane unit.

1.2.1. Data Reconciliation in the Membrane Separation Process

Currently, very few papers are somehow involved with Data Reconciliation in Membrane Separation Processes. However, the work developed by Lashkari and Kruczek [23] showed that it is possible to reconcile the properties of the membrane separation process, using data related to the process dead time, which is affected by the resistance that the membrane offers to the gas flow. In fact, this work developed procedures to estimate effective permeabilities and diffusion coefficients, allowing the modeling of position-dependent resistance effects, using a lab-scale unit as an example.

Crivellari [19] proposed a model to simulate the separation of CO₂ from natural gas through a counter-current spiral wound polymer membrane. The model was used to analyze the influence of distinct variables on process operation conditions and was based on phenomenological balance equations. The model was validated with data collected from the literature and some industrial sites. Moreover, a DR procedure was used to treat the available data and estimate the model parameters; however, the study was implemented offline and did not allow any sort of real-time monitoring of the process operation.

Based on the previous discussions, the present work pioneers the use of DR procedures to monitor membrane separation processes reliably, online and in real time. For this purpose, a membrane plant located in one of the Petrobras Offshore Units was used during this paper. Finally, it is worth noting that, for reasons of industrial confidentiality, numerical results are presented in normalized form.

2. Methods

The methodology implemented in the present work comprises six stages: data acquisition; data pre-treatment; data characterization; data reconciliation; gross error detection, and process monitoring (soft sensor or digital twin). The first stages of the implemented procedure involved pre-treatment and characterization of the data. As a matter of fact, proper understanding of some characteristics of the data are fundamental for adequate implementation of the data reconciliation stage [24]. The initial characterization of the data was performed offline and using historical data available in the data acquisition system of the industrial site. The available data were used to determine appropriate sampling periods (based on process response times) and calculate measurement variances (used to formulate the estimation problem) and variable correlations (to characterize independence of measuring devices). Variable classification was also performed to determine the sets of observable and unobservable variables (with the help of the proposed model, as described below) [25].

Using the available data and the model equations, the DR procedure (as described in the following paragraphs) was solved offline to validate the proposed procedure and determine some performance indexes. Particularly, a statistical metrics was used to describe the magnitudes of the deviations between measured and reconciled variables. Then, the model was used offline for calculation of unmeasured variables, providing the soft sensor (or digital twin) response. Finally, the proposed procedures were implemented online and in real time.

The numerical procedures and codes were developed and implemented in Python 3.7.6 (Python Software Foundation, Beaverton, OR, USA) and the details of the proposed methodology are explained in the following sections.

2.1. Data Acquisition

Data acquisition was performed through direct access to an industrial database, using standard Plant Information (PI) resources. After performing the numerical operations, a file was saved with the measured, reconciled, estimated, and calculated variables. Storage was performed during monitoring, to avoid accumulation of data in the computer memory and save the relevant information in real time.

The “pandaspi” library was utilized to provide communication between Python and PI, transferring the information directly to a data frame [26,27]. By using these resources, the data acquisition process became very simple and practical, as access to the data depended only on the login,

password, tags of the desired variables, the size of the sample window, and the sampling frequency. The time interval selected for offline analyses was equivalent to two weeks with a sampling frequency of 5 min, which provided a sufficiently high number of points for the execution of the pre-treatment step. An additional number of data points did not provide any significant improvement of the preliminary analyses in the considered case so that this should not be regarded as a drawback of the proposed analysis.

2.2. Data Pre-Treatment

Data pre-treatment comprised the following steps: reading the Excel Workbookfiles in UTF-8 encoding; pre-treatment of raw data spreadsheets; and data storage in Hierarchy Data Format version 5 (HDF5).

During real-time monitoring, the data reading step was performed as described in the previous section. Data storage in HDF5 format was also performed as described previously. When the file is saved in HDF5 format, reading presents better performance, since reading data directly from the Excel spreadsheet can be too slow for real-time applications [28,29].

The pre-treatment stage organized the raw data from the PI into data frames (Pandas library) [26,27]. The treatment followed the following steps: standardize the indices (day/hour/minute of the samples); variables that contain some string must be replaced (such as on/off by 0/1 and error notices by NaN, not a number); chronologically sort the data and replace missing data (NaN) with neighbors (back and forward fill). This last step is crucial to assure that the acquired data window does not contain missing data, which can make the calculation of variances difficult in the acquisition window. Unit conversions, normalization of concentrations and calculation of standard deviations and variances completed this step.

Data storage must be carried out after data reconciliation and energy balance calculations. The file was saved with the measured, reconciled, estimated, and calculated variables. As already explained, storage was performed during monitoring, to avoid accumulation of data in the computer memory and save the relevant information in real time.

2.3. Data Characterization

During this step, the procedures that must precede the data reconciliation task, such as the visualization and statistical characterization of the data, must be carried out. The characterization of the data was performed in accordance with the following steps:

- Visualization of variables;
- Selection of variables of interest;
- Quantification and visualization of missing data (NaN);
- Construction of the boxplots of the variables of interest [30];
- Analysis of the variance spectra [31];
- Calculation of variances and correlations.

The main pursued objectives during this stage were the proper characterization of the data quality, the analysis of the operation dynamics, and the characterization of the stationarity of the phenomenological process model. The analyses of missing data and boxplot properties can indicate the quality of collected data during the selected time period and the number of gross errors in the data base.

The boxplot is a graph used to assess the empirical distribution of data. The boxplot is a non-parametric analytical technique, which shows the measurement variations within a statistical population without making any assumption about the underlying statistical distribution. The box is usually built with the first (Q_1) and third (Q_3) quartiles (50% of the data) and the median (Q_2). The lower and upper lines extend, respectively, from the lower quartile to the lowest value not lower

than the lower limit (LL), and from the upper quartile to the highest value not higher than the upper limit (UL). The limits can be calculated according to Equations (1) and (2):

$$LL = \max [\min(data), \quad Q_1 - 1.5(Q_3 - Q_1)] \quad (1)$$

$$UL = \min [\max(data), \quad Q_1 + 1.5(Q_3 - Q_1)] \quad (2)$$

The value 1.5 is tuned to capture 99.7% of the data between the lower and upper limits, assuming the normal distribution [30]. In summary, the boxplot identifies the regions in the variable domain where 50% and 99.7% of the data are located. The points outside these limits are tagged as outliers. Boxplots were plotted using the Seaborn library available for Python [32].

The analysis of the variance spectra shows how the process variance depends on the size of the sampling window. This type of spectrum provides information about the various sources that contribute to the signal of a variable, including noise/measurement errors (short window sizes) and intrinsic process variations (large window sizes). The variance spectrum can be defined as a set of variances calculated while some variable related to them evolves [31]. The spectrum of variances for short sampling windows are controlled by the variances of the measuring instrument. This way, the best estimate for the variance of the measuring device can be calculated using the variance spectrum with short sampling windows. However, the use of very short sampling windows may not reveal the actual variability of the data, due to poor measurement quality. The spectrum for sufficiently large sampling windows captures the variability of the entire process, including operational changes. More details about the usefulness of this technique for characterization of process data are provided by Feital and Pinto [31].

Analyzing the correlations, autocorrelations, and cross-correlations between pairs of variables of interest allows the characterization of stationarity, seasonality, and regions of operation of the process [33,34]. Observing correlations between input and output flows can indicate process stationarity. The absence of correlation can indicate the occurrence of significant nonlinearity or dynamics in the process response [35]. For dynamic responses, cross-correlation can capture the lags between the actions on the input variables and the steady-state of the output variables. Obviously, the identification of dynamic correlations among the many variables of the system may indicate the necessity to build and implement dynamic models for more accurate representation of the available data. For this reason, proper characterization of stationarity can be important for more successful implementations of monitoring procedures [36].

2.4. Data Reconciliation

The DR procedure consists of solving an optimization problem characterized by an Objective Function (OF) that must be minimized while respecting certain restrictions (model). The OF of the DR is often proposed as the maximum likelihood estimator resulting from a statistical distribution of measurement errors, which is commonly adopted as the normal distribution. After application of the principle of maximum likelihood, the normal distribution results in the WLS estimator. This way, the problem originally formulated by Kuehn and Davidson [18] can be written in accordance with Equations (3)–(7) [17]:

$$\hat{z} = \min_{\hat{z}} \frac{1}{2} [z - \hat{z}]^T \underline{\underline{V}}^{-1} [z - \hat{z}] \quad (3)$$

subject to:

$$\underline{h}(\underline{z}, \underline{u}) = \underline{0} \quad (4)$$

$$\underline{g}(\underline{z}, \underline{u}) \geq \underline{0} \quad (5)$$

$$\underline{z}^L \leq \underline{z} \leq \underline{z}^U \quad (6)$$

$$\underline{u}^L \leq \underline{u} \leq \underline{u}^U \quad (7)$$

where \underline{z} is the vector of the reconciled variables; \underline{z} is the vector of the measured variables; \underline{V} is the matrix of variances for measurement errors; \underline{u} is the vector of the unmeasured variables (observable); $\underline{h}()$ is the vector of linear or nonlinear algebraic constraint equations; $\underline{g}()$ is the vector of the inequalities of linear or nonlinear algebraic restrictions; \underline{z}^L and \underline{z}^U are the upper and lower parameter vectors of the \underline{z} vector and \underline{u}^L and \underline{u}^U are the upper and lower parameter vectors of the \underline{u} vector.

For DR problems where the model constraints are linear and all variables are measured, the analytical resolution of the problem can be obtained with help of Lagrange Multipliers [37]. However, the set of individual mass balance equations generate a nonlinear system of equations that involve unmeasured variables, as described below. For this reason, a successive linearization procedure was used to solve this problem [38].

The observability analysis of the system can also be performed during the successive linearization procedure, using QR factorization. The procedure consists of describing the nonlinear model in terms of two matrices: a matrix related to the measured variables and a second matrix related to the unmeasured variables. Therefore, the system is observable if the rank of the matrix of unmeasured variables is equal to the number of unmeasured variables [39].

The following points were considered during the implementation of the proposed model in the industrial site, illustrated in Figure 3:

- Period for preliminary characterization of database: two weeks with sample frequency of 5 min;
- Measured variables (z):
 - Flowrates: Total Feed (F), Total Retentate (R), Train Retentate A , B and C (R_A , R_B and R_C) [kNm^3/h];
 - Components: C_1 (methane), C_2 (ethane), C_3 (propane), C_6 (hexane), C_7 (heptane), C_8 (octane), CO_2 (carbon dioxide), iC_4 (i -butane), iC_5 (i -pentane), N_2 (nitrogen), nC_4 (n -butane) and nC_5 (n -pentane) in the 3 streams.
- Unmeasured variables (u):
 - Flowrate: Total permeate (P) [kNm^3/h].

Model (Component Mass Balance—Steady-State):

$$\underline{h}(\underline{z}, \underline{u}) = \begin{cases} 0 = F\underline{y}_F - R\underline{y}_R - P\underline{y}_P \\ 0 = R - R_A - R_B - R_C \\ 0 = 1 - \sum_{i=1}^{nc} \underline{y}_{F,i} \\ 0 = 1 - \sum_{i=1}^{nc} \underline{y}_{R,i} \\ 0 = 1 - \sum_{i=1}^{nc} \underline{y}_{P,i} \end{cases} \quad (8)$$

- Number of points in the study phase (nt) = 3457
- Number of components (nc) = 12
- Number of measured variables at each sampling point (Nm) = $3nc + 5 = 41$
- Number of unmeasured variables at each sampling point (Nu) = 1
- Number of total equations at each sampling point (Nv): $Nm + Nu = 42$

- Number of constraint equations (N_{ce}) = $nc + 4 = 16$
- Number of optimization variables at each sampling point (N_{opt}): $Nv - N_{ce} = 26$
- Degrees of freedom at each sampling point (DF): $Nm - N_{opt} = Nm - (Nm + Nu - N_{ce}) = 15$

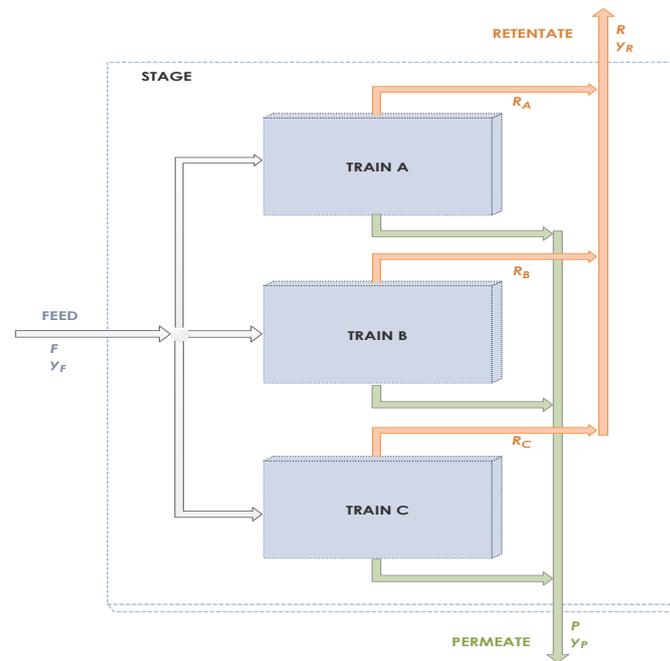


Figure 3. Individual mass balances in the envelope around the process.

2.5. Gross Error Detection

Gross errors are those originating from non-random events, having little or no connection with the measured value. They may be related to measurements (such as malfunctioning of instruments) or process (such as leaks). Consequently, gross errors invalidate the classic statistical basis of traditional DR methods and undermine the systematic analysis of the data, demanding the implementation of Gross Error Detection procedures for compensation or elimination of gross errors, which may precede the DR step. Gross errors can be further classified as **bias**, **outliers**, and **drifts** [40].

The metrics applied in this step to identify the occurrence of gross errors is related to the statistical test of standard hypotheses for GED [12]. This test is widely used and simple, and also forms the basis for all other classic GED procedures. The test is applied to the data set to determine whether the measurements follow a symmetrical distribution of the measurement errors.

Reilly and Carpani [41] were the first to study the detection of gross errors in process engineering. The authors proposed the use of a statistical test of the type χ^2 (chi-square) based on the residues of the process model, which was called the Global Test (GT).

In the present work, GED was based on the GT, using the sampling window to observe samples and identify the occurrence of gross measurement errors. According to this procedure, the data in the moving window are used to compare the errors of the Weighted Least Squares estimator with the deviations from median values within the window and using the χ^2 function to characterize significant deviations between these computed variances [42].

The outlier effect during DR is removed through compensation. After identification of a possible outlier, the variable value was manipulated to adjust the expected variance. Thus, a moving variance window was implemented to monitor changes in operation, measurement errors, and failures. The statistical test was performed in sampling windows containing at least 20 samples. With this, the median and the standard deviation were calculated and compared with the value of the variable at the current point. Therefore, if the value was more than seven standard deviations apart from

the median, the measurement would be regarded as an outlier and the variance adjustment would be performed.

To monitor the possible occurrence of bias, a dynamic bar graph illustrating the magnitude of errors for each variable was implemented, using the following metrics:

$$Bias_{i,t} = \frac{med(|z_{i,t} - \hat{z}_{i,t}|)}{dp_{i,t}} \quad (9)$$

$$dp_{i,t} = NMAD(|z_{i,t} - \hat{z}_{i,t}|) \quad (10)$$

where *NMAD* is Normalized Median Absolute Deviation.

2.6. Monitoring

Instrumenting the whole process can be very costly and may lead to acquisition of unnecessary and obsolete information. In addition, in some cases, it can be impossible to measure the desired variable. For this reason, some information that is essential for process monitoring can be assessed through models. Therefore, a soft sensor can estimate variables with a mathematical model, in real time, using available plant data, as measured by existing instrumentation. Particularly, the use of plant variables can provide opportunities to improve the performance of a plant [43].

In the analyzed process, the permeate flow was not measured. It must be emphasized that this is not unusual at real industrial sites. However, through DR, this information can be obtained with the help of the model, after the application of the DR procedure. The complete set of temperature and pressure was not available either, which is not unusual at the plant site. In the analyzed case, the pressure of the feed stream and the temperature of the permeate stream were not measured. Therefore, the full implementation of the energy balance in the proposed DR procedure was not viable due to the lack of observability. For this reason, the pressure of the feed stream was evaluated through calculation of the pressure loss in the separation stage, based on design data and the pressure data of the retained stream. In this case, after characterization of the pressure loss, the energy balance equation can be used to estimate the permeate stream temperature, with the aid of Equation (11), calculating the enthalpies of the process streams with the Peng–Robinson equation of state [44]:

$$0 = F.H(T_F, P_F, \underline{y}_F) - R.H(T_R, P_R, \underline{y}_R) - P.H(T_P, P_P, \underline{y}_P) \quad (11)$$

It must be noted that the execution of the numerical procedure in this analyzed case was extremely fast (order of milliseconds), so that the computer hardware exerted little influence on the application. The slowest step of the numerical procedure was data acquisition (and the bottleneck was data transfer and connection speed), so that instrumentation hardware and data handling software constituted the most sensitive parts for this particular real-time application. Despite that, it took only few seconds for data downloading to be complete; consequently, the online and real-time implementation could be performed in a standard notebook equipped with an Intel Core i7 8th gen processor (Intel Corporation, Santa Clara, CA, USA).

3. Results and Discussion

3.1. Data Characterization

The Missingno library was used to analyze the missing data [45]. In Figure 4, one can observe the data density, visualizing completely the pattern of missing data in the whole set, with the columns representing the variables and lines representing data points of the time series. In addition, a frequency bar that indicates the number of variables measured at each particular timeline can be seen on the right side of the graph.

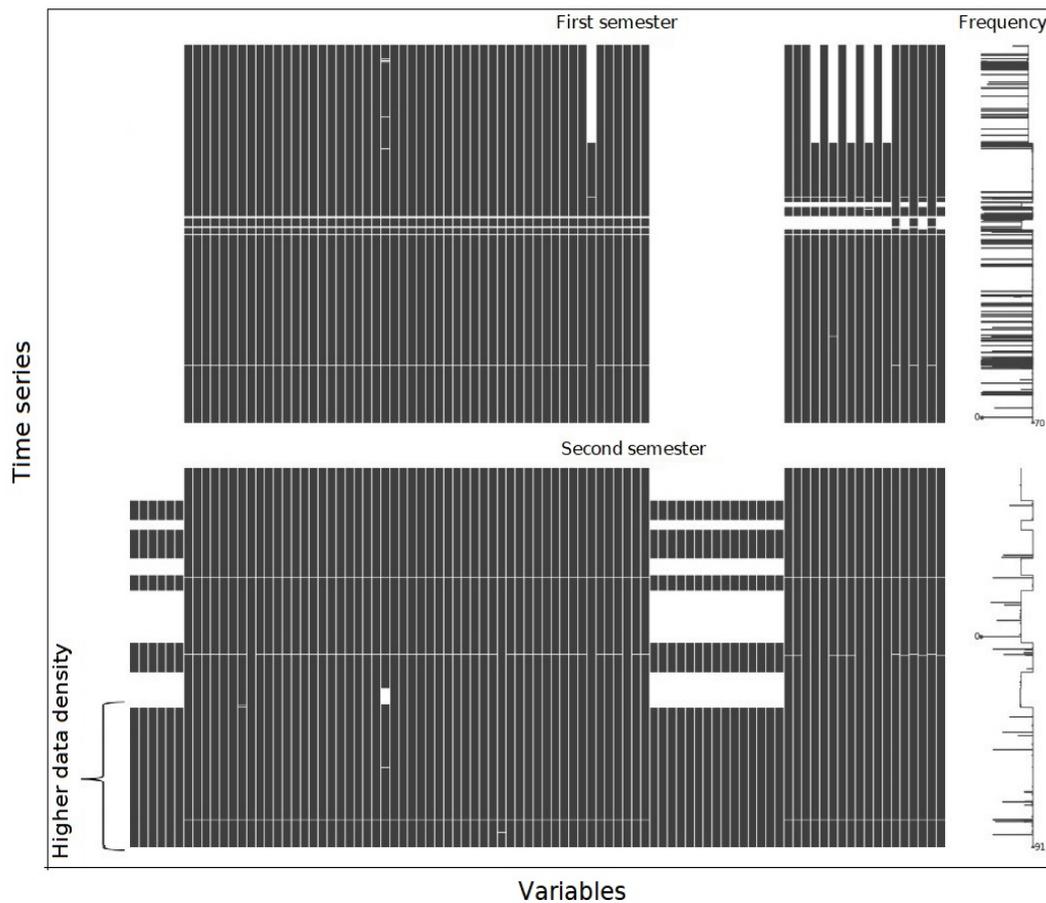


Figure 4. Missing data analysis.

The analysis of Figure 4 illustrates how one can select the best data period for the execution of the initial data processing stage. The period with the highest data density allows the analysis of the plant operation with better reliability, avoiding blind periods of instrumentation or shutdown of operation. Therefore, the selection of the period is fundamental for the data characterization stage.

Figure 5 illustrates the boxplot of chromatographic analyses associated with the four main components of the feed stream. This analysis allows the preliminary evaluation of the precision of the instrumentation and/or the variability of the operation. A process with distinct operating points generates multimodal distributions, which requires more involving analysis to qualify the precision of the measurement, such as violin plot (a combination of boxplot and kernel density estimate) [46]. However, during stationary operation periods, the boxplot analysis showed that chromatographic measurements presented good precision, which is also illustrated by the small number of outliers with respect to the total number of 3673 samples.

Figure 6 shows that the “gross errors” followed a downward trend, with highest concentration below the modal value. This is because a drop in the process flowrate was observed during this period. Therefore, in this case, the analysis interprets that the operating changes are “gross error”, when they are not. On the other hand, the good behavior of the data indicates that the flowrate was measured with good precision. Instrumentation accuracy was also analyzed after DR, with help of the bias analysis.

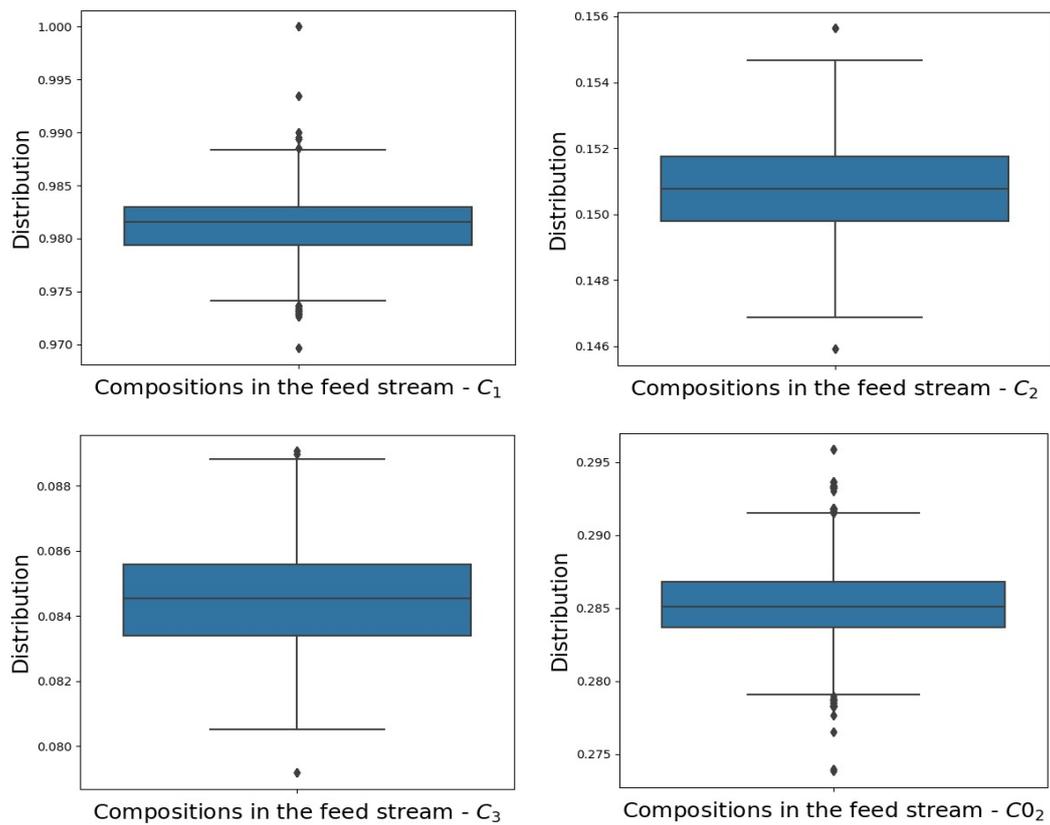


Figure 5. Analysis of boxplots for compositions C₁, C₂, C₃, and CO₂ in the feed stream.

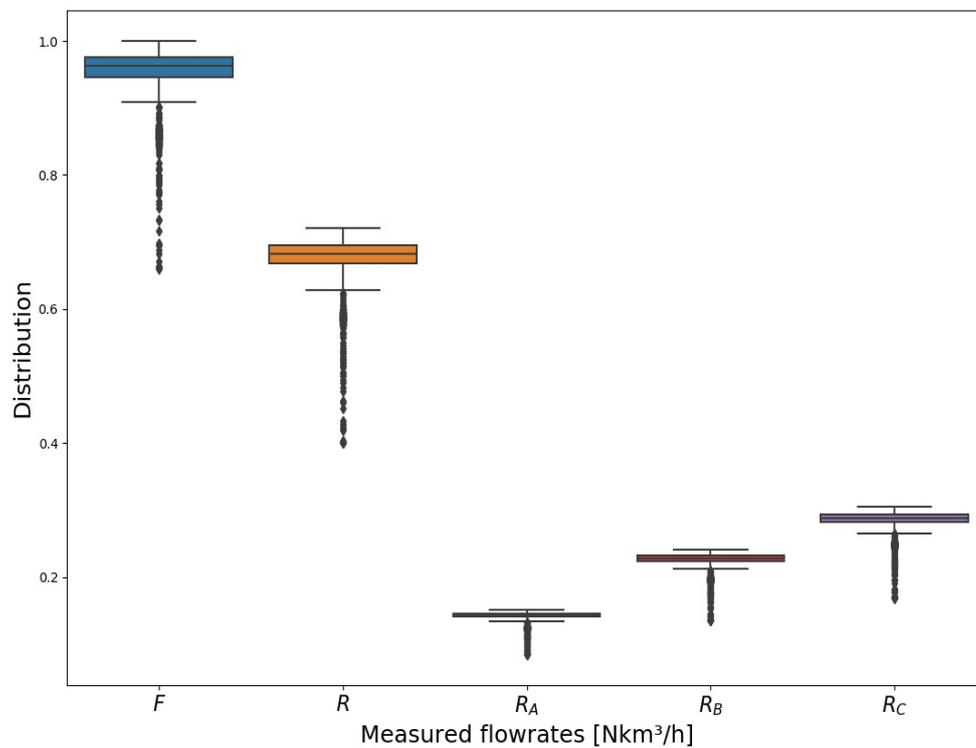


Figure 6. Boxplot analysis of flowrates.

Figure 7 illustrates the time series for the compositions of C₁, C₂, C₃, and CO₂ in the feed stream, while Figure 8 shows the feed and retentate flowrates, for the same time ranges analyzed in the boxplot.

Therefore, it becomes evident in the case of flowrates that the supposed gross errors actually indicated a change in operation and not a failure of the sensor. In the case of composition, outliers can possibly be assigned to gross errors, although only the DR can allow the proposition of reliable statements about the alleged gross errors.

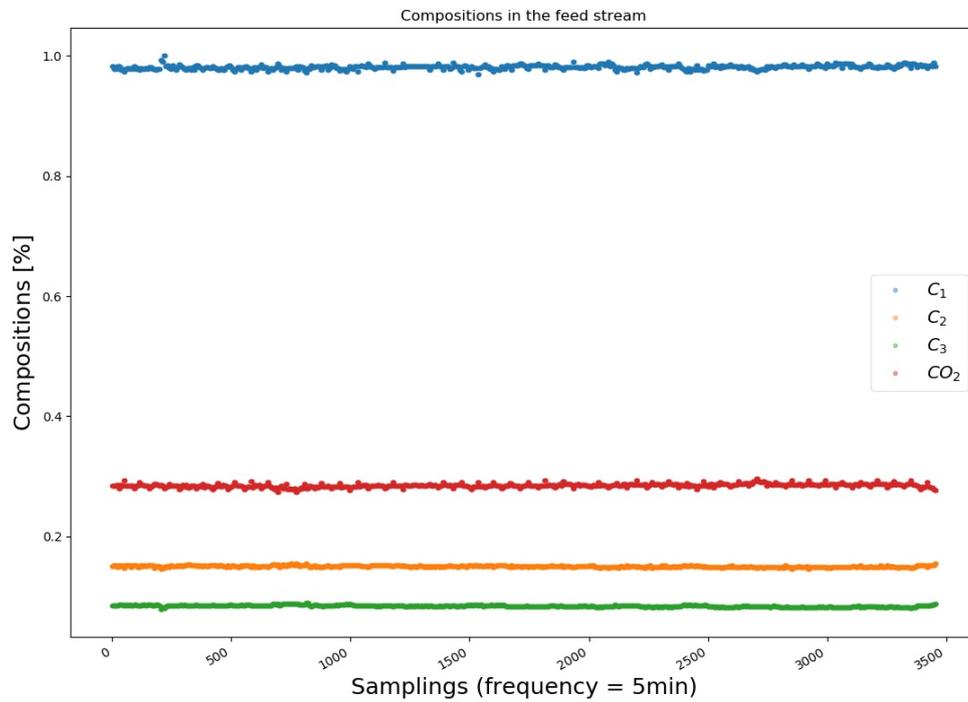


Figure 7. C_1 , C_2 , C_3 , and CO_2 compositions in the feed stream.

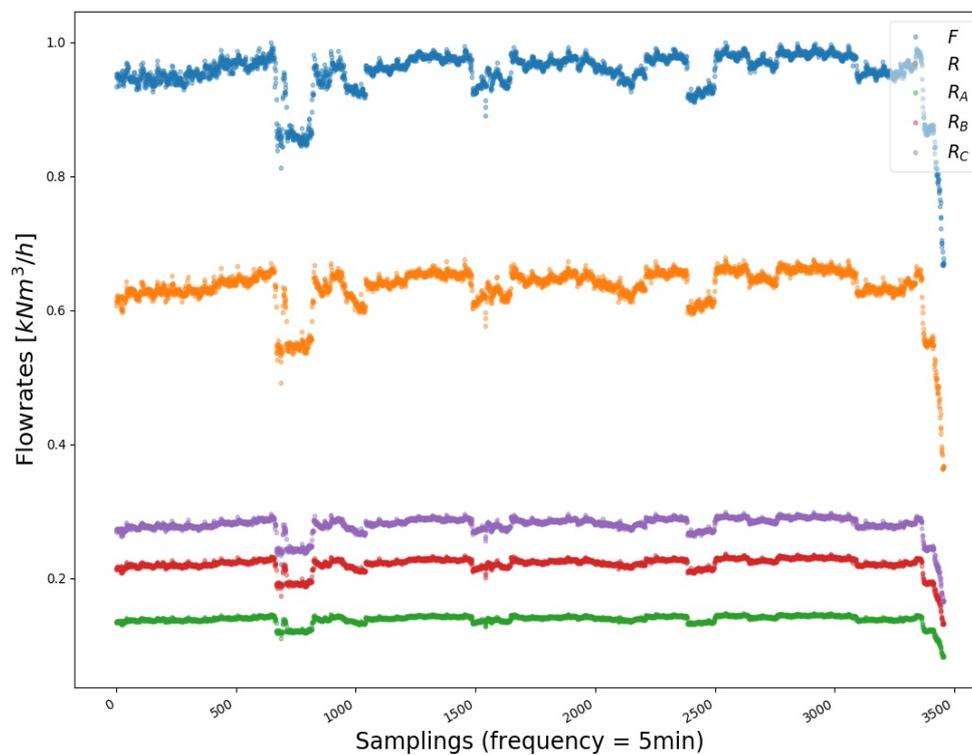


Figure 8. Measured flowrates.

An important analysis is related to the observation of the correlations between pairs of variables. Correlations can indicate absence or presence of process stationarity. Figure 9 shows a strong linear correlation between feed, retentate, and permeate flowrates. Strong linear correlation between inlet and outlet flowrates can be an indication of process stationarity [47].

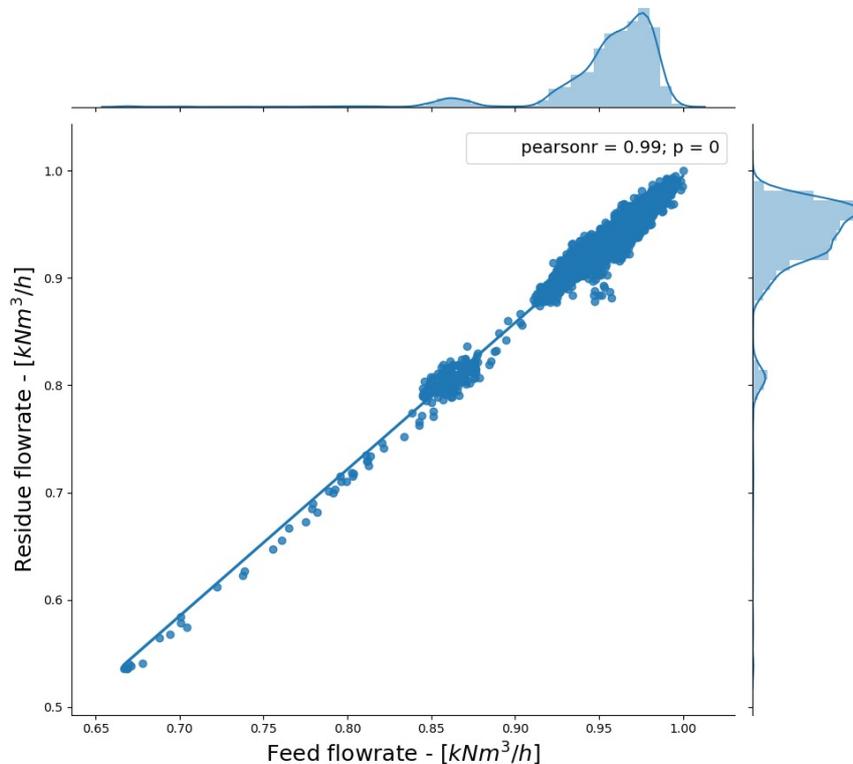


Figure 9. Correlation analysis for feed and retentate flowrates.

Figure 10 illustrates the Autocorrelation Function (ACF) and Partial Autocorrelation Functions (PACF) for feed and residue streams. This analysis provides the diagnosis of temporal dependence between the lags of individual variables, which in this case were evaluated for lags ranging from 0 to 50 lags. As shown in Figure 10, the ACF decayed continuously and just one lag caused the appearance of strong correlation (close to 1) in the PACF. Therefore, the process presents very short dynamic memory, indicating the quasi steady-state behavior and constituting an auto-regressive process of order 1 [33].

Figure 11 illustrates the Cross-Correlations Function (CCF) between feed and retentate flowrates up to 50 lag. Cross-correlations decayed slowly for different pairs of variables, indicating that the process operated at quasi steady-state conditions and that responses were much faster than the characteristic sampling times. Therefore, the analyzed membrane separation process could be considered to operate at steady-state. This validated the use of the steady-state mass and energy balance equations in the DR problem. Given the small volumes of most membrane separation modules and the large flowrates of typical industrial plants, this conclusion can probably be extended to other industrial sites, allowing the more general use of the analysed procedures in other industrial facilities.

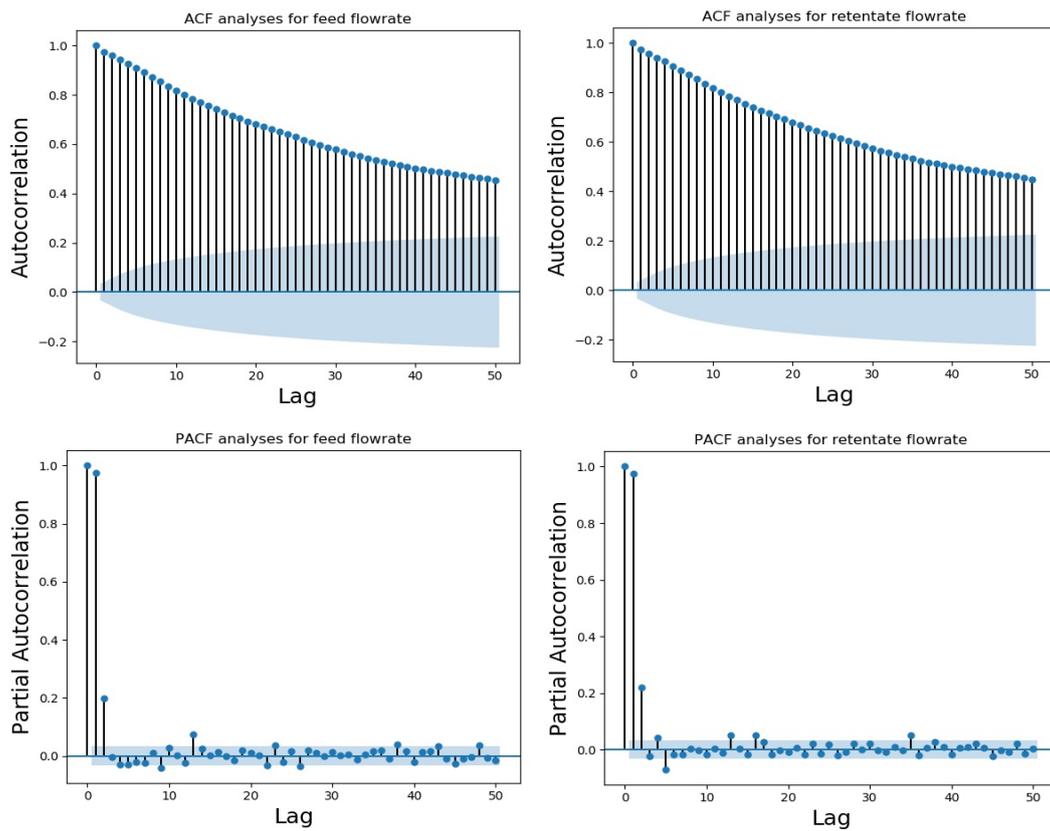


Figure 10. ACF (Autocorrelation Function) and PACF analyses (Partial Autocorrelation Functions) for feed and retentate flowrates.

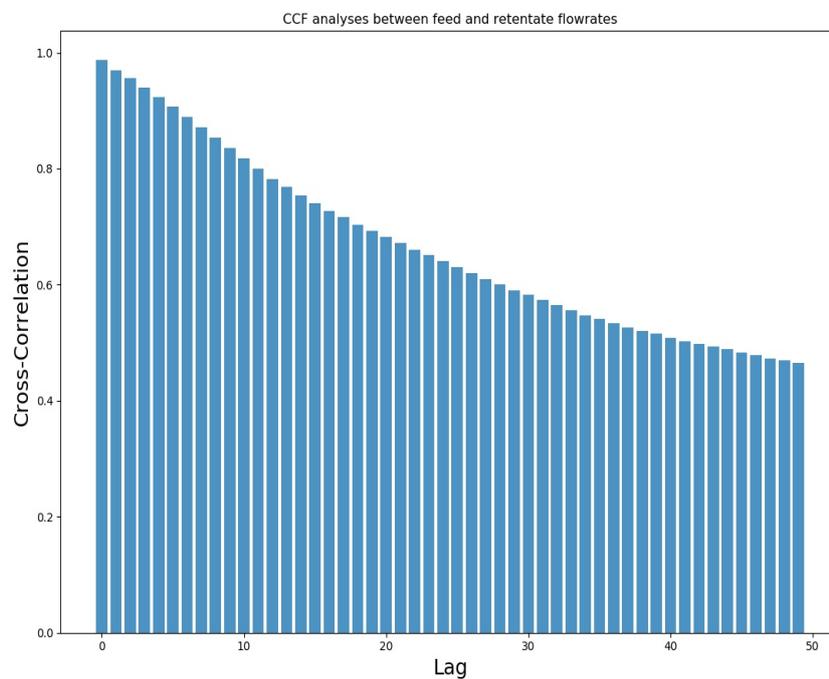


Figure 11. CCF analyses between feed and retentate flowrates.

3.2. Data Reconciliation

Before starting the DR procedure, the system was analyzed with the help of Variable Classification techniques and the system was classified as observable, indicating that the measured variables can be reconciled and that unmeasured variables (permeate flowrate) can be estimated.

The first DR results were obtained through offline simulations, using a sampling period of two weeks with a sampling interval of 5 min. The data reconciliation performed very well and the problem was solved at average computational speed of 1.7 ms/sample. This result clearly showed that the application could be implemented online and in real time due to the sampling interval of 5 min.

Figure 12 illustrates, as an example, the measured and reconciled data for chromatographic measurements of the four main components in the feed stream. One of the advantages of DR is to restore the resolution of the amplitude signal of the measured value, which can be seen in Figure 12, especially for samplings of the C_2 component.

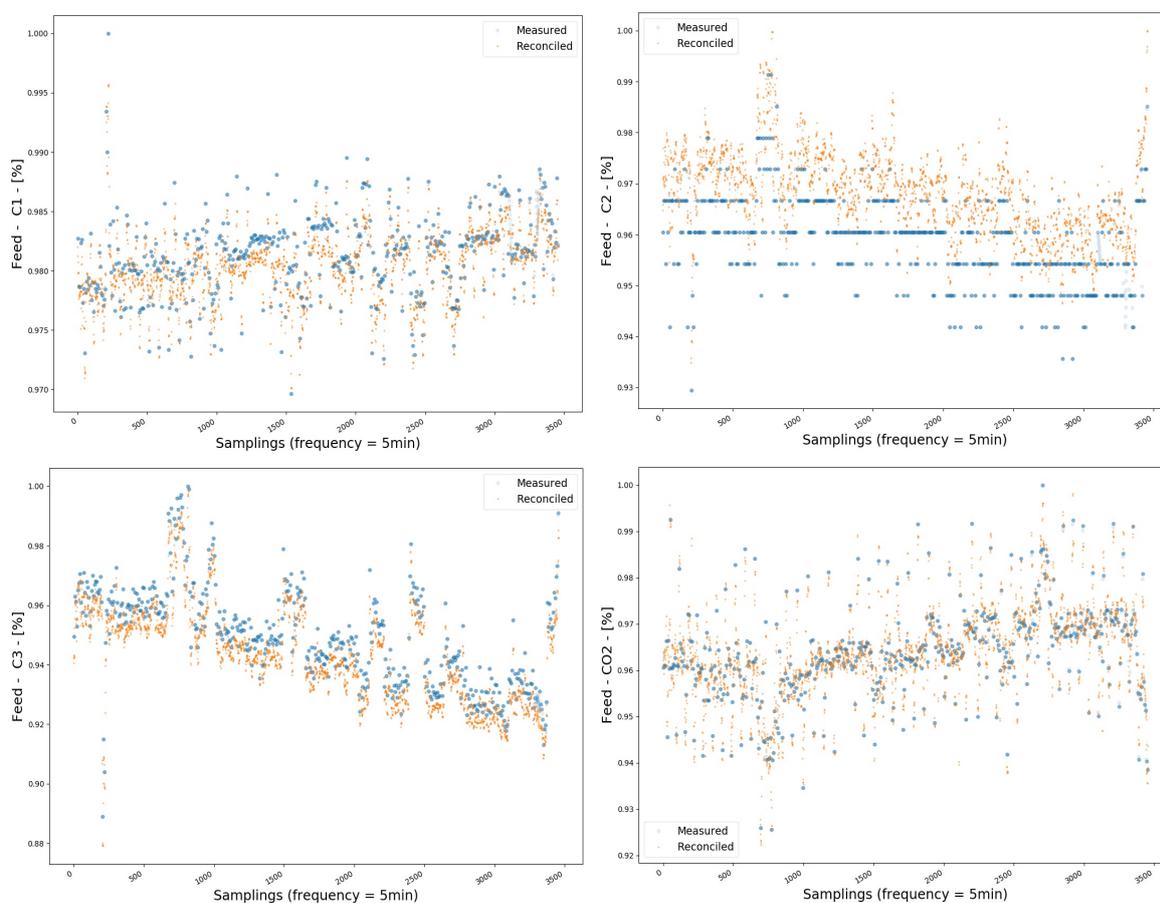


Figure 12. Offline data reconciliation for compositions C_1 , C_2 , C_3 , and CO_2 in the feed stream.

Figure 13 illustrates the measured, reconciled, and estimated data for the flowrates. Another major advantage of DR is to identify the occurrence of systematic deviations in measurements, often caused by miscalibrated instruments. Figure 13 shows the occurrence of bias for the feed and residue flowrates.

Figure 14 shows the sum of residuals with respect to the mathematical model. The total squared residual is a measure of corrections that were needed in order to reconcile variables to satisfy the mass and energy balance equations.

Based on the previous results, it could be concluded that the DR procedure presented good performance and advantageous aspects for monitoring of the process. Monitoring processes with statistically treated information, detection of measurement bias, and identification of poor

instrumentation performances constitute good tools for diagnosing the states of the analyzed process and respective instrumentation.

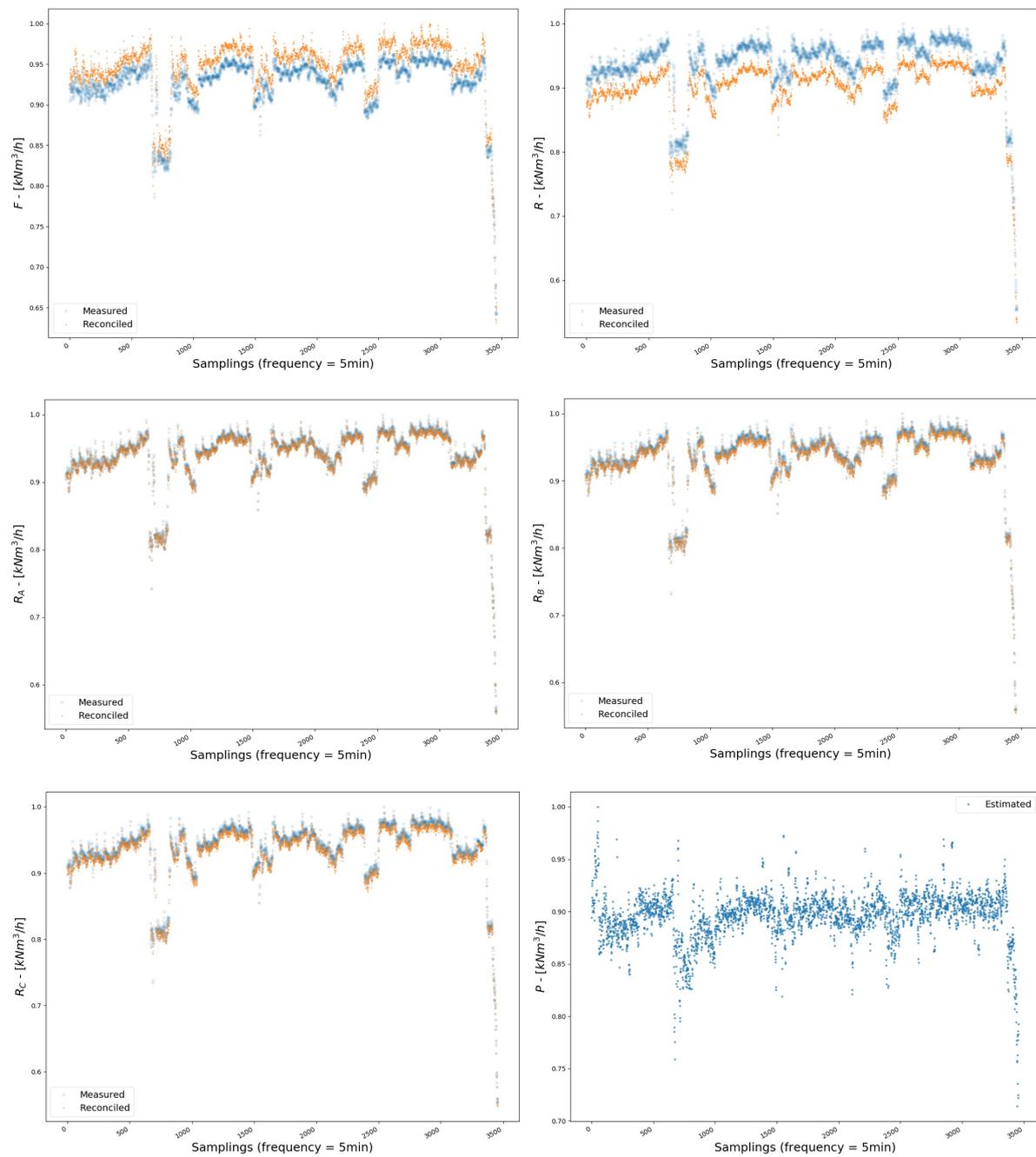


Figure 13. Offline data reconciliation of flowrates.

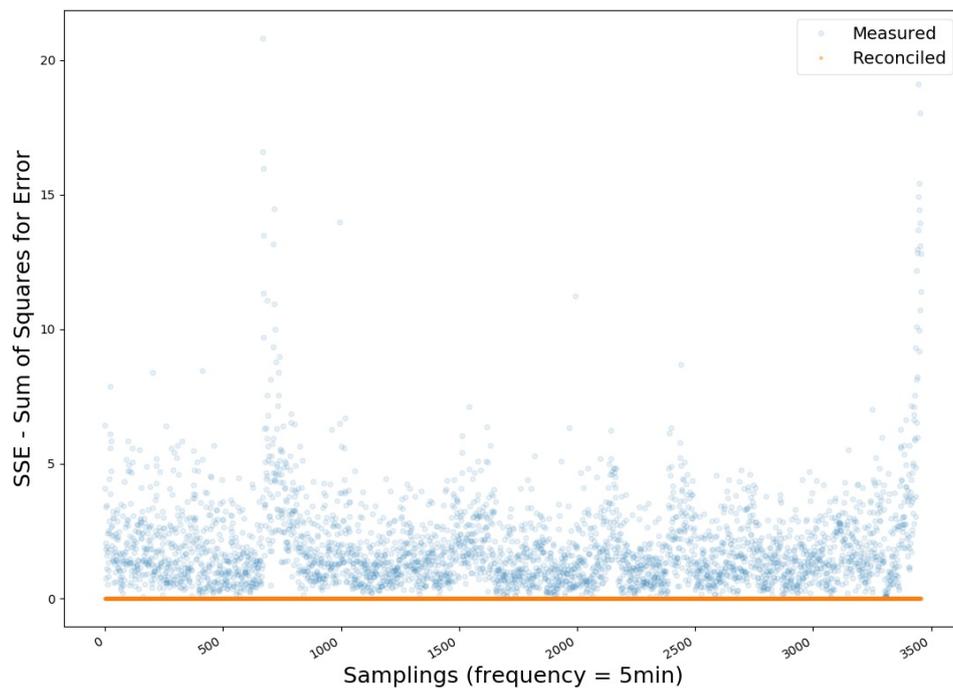


Figure 14. Sum of squares of model deviations (residuals) during offline data reconciliation.

3.3. Gross Error Detection

Procedures for removal of gross errors were implemented for online and real-time DR. These procedures were based on statistical tests using a moving variance window. Figures 15–20 illustrate a case of gross errors in which the procedure proved to be robust. However, it is possible to observe that the gross errors that affected the compositions of the retentate flowrate influenced the reconciliation of the feed flowrate. This occurred because of the well-known “smearing effect” when the DR procedure was performed with the WLS estimator (non-robust), even when variance adjustment was performed [48]. As gross error measurements were observed for a short period of time, it was not possible to detect the main source of the problem in the analyzed data set. Nevertheless, the occurrence of the problem was reported for maintenance teams for evaluation of measurement consistency.

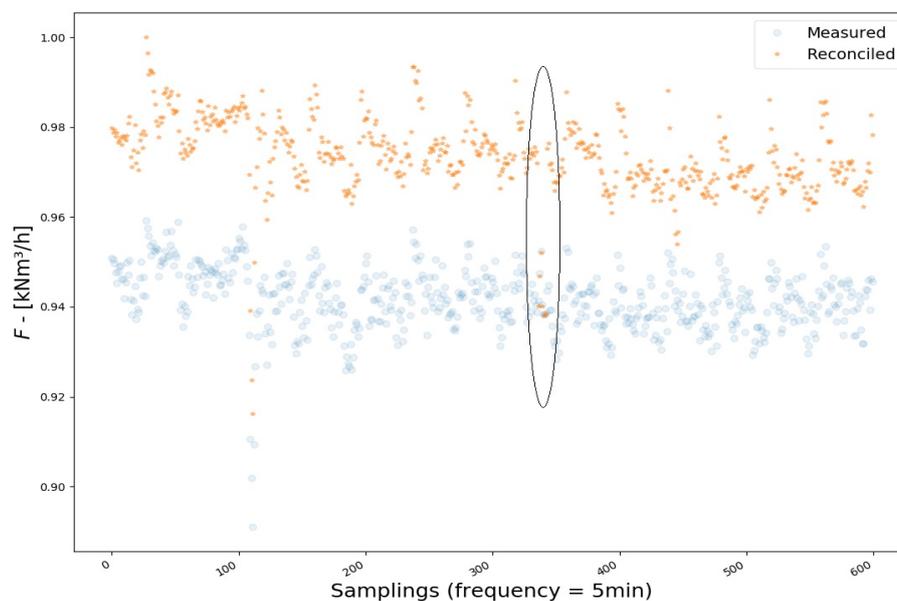


Figure 15. “Smearing” effect during the DR of feed flowrates.

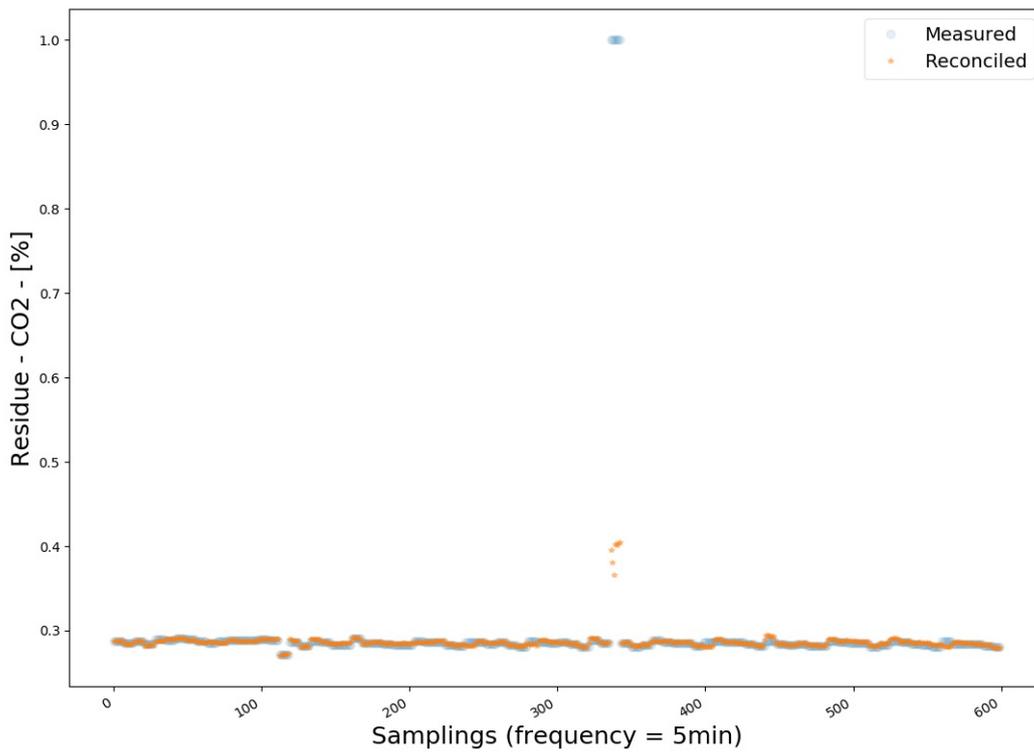


Figure 16. Gross Error Detection—CO₂ in the feed stream.

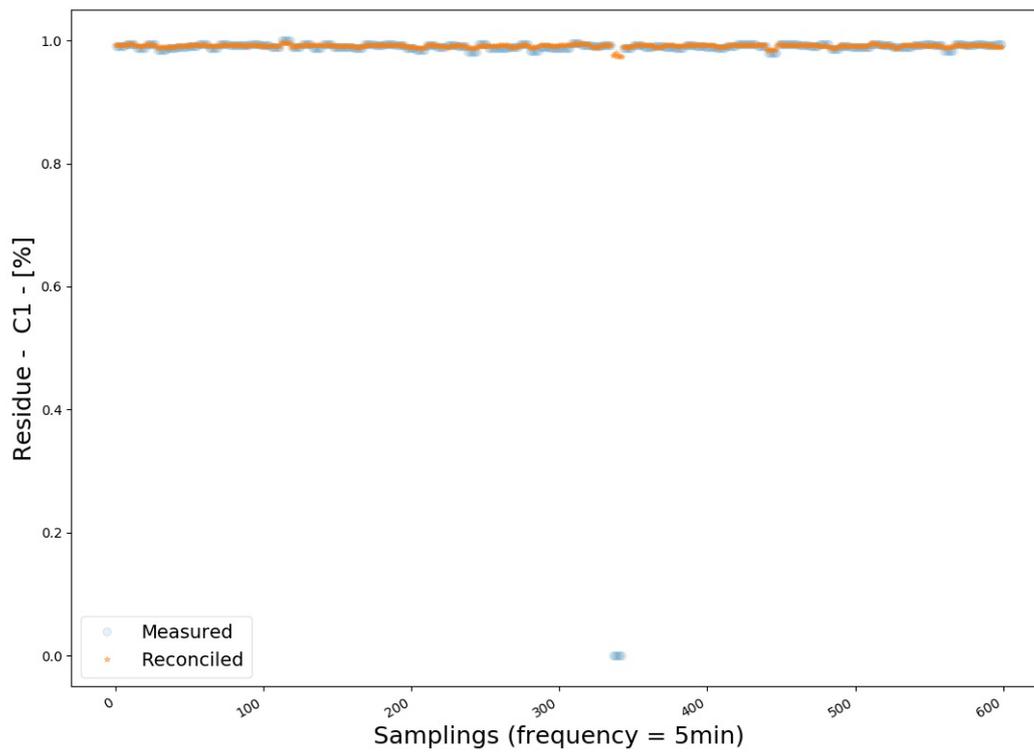


Figure 17. Gross Error Detection—C₁ in the feed stream.

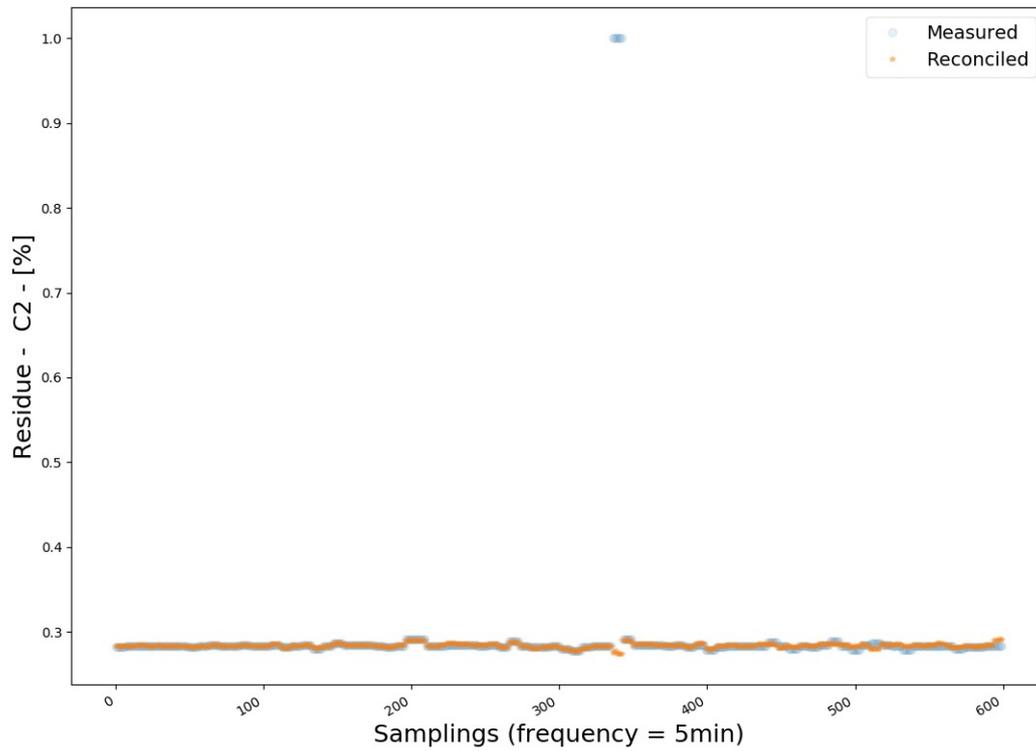


Figure 18. Gross Error Detection—C₂ in the feed stream.

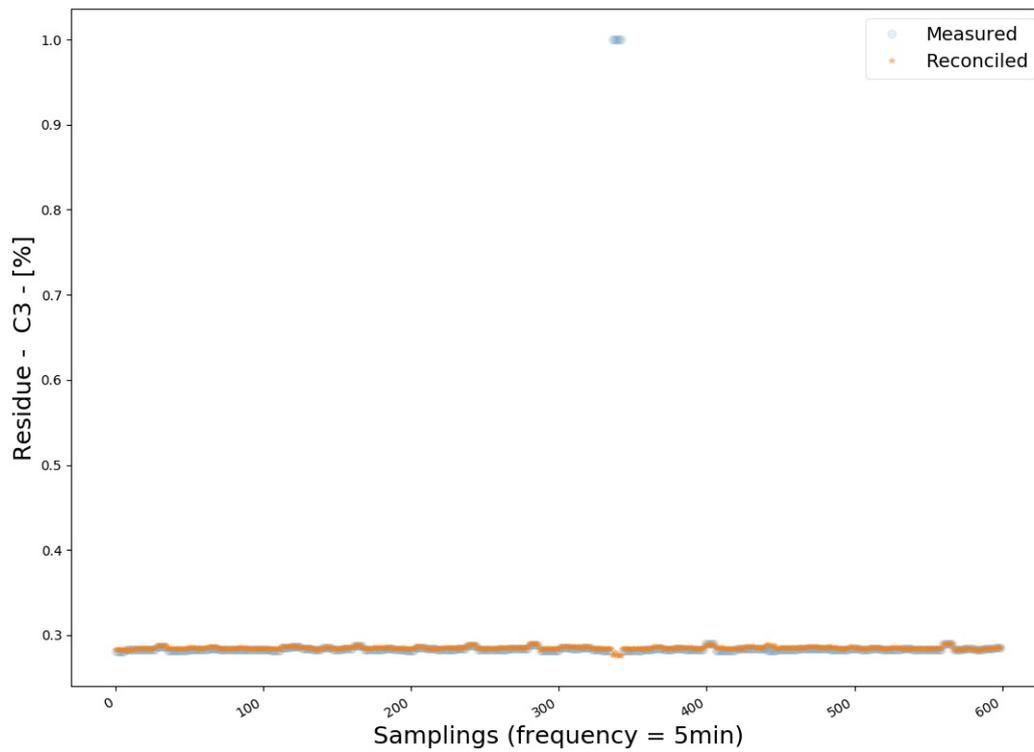


Figure 19. Gross Error Detection—C₃ in the feed stream.

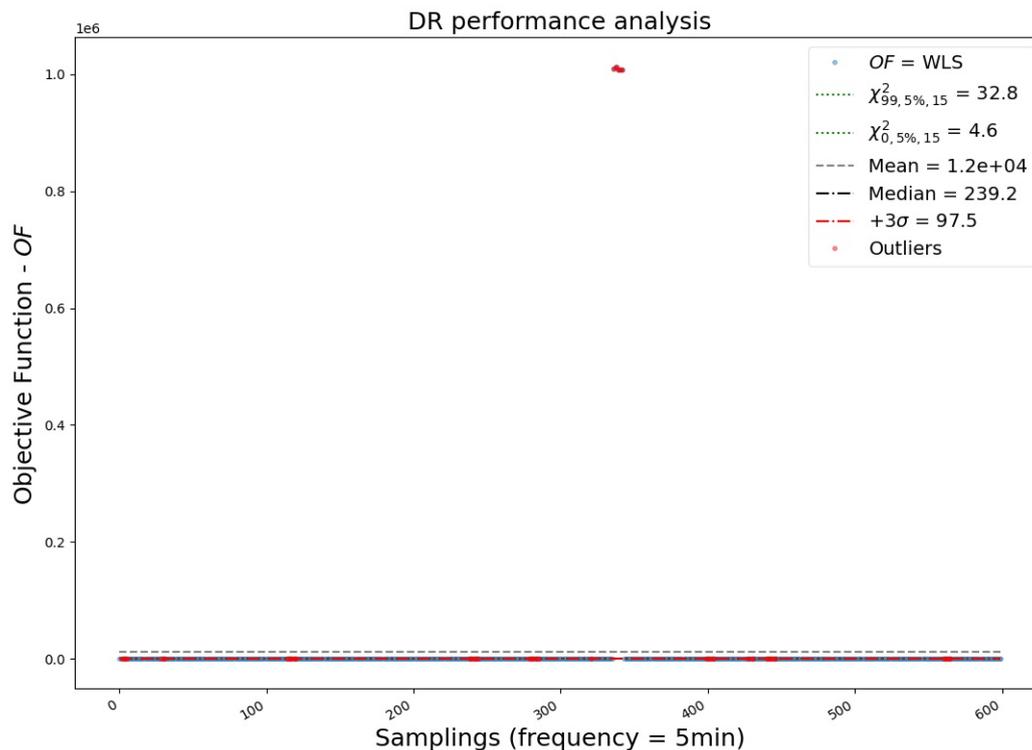


Figure 20. Gross Error Detection—DR performance analysis.

It is important to note that the statistical tests were implemented only for the compositions. The fact is that operational changes hindered the test because in many cases the test interpreted operational changes as outliers. Figure 21 illustrates that data reconciliation was effective and performed well after several operational changes.

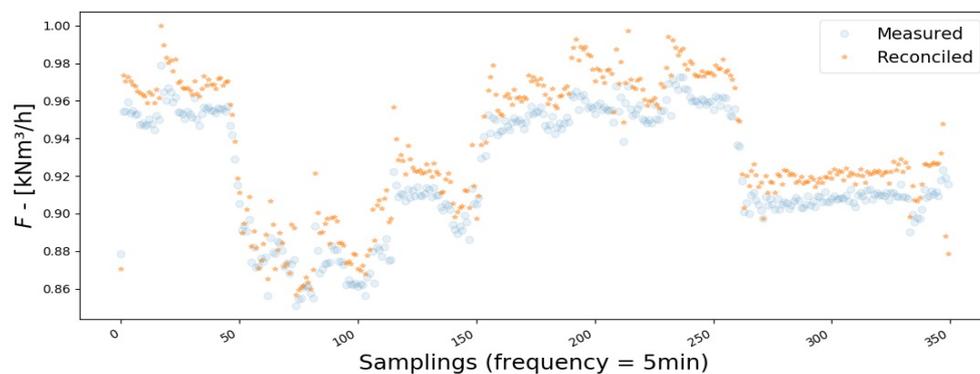


Figure 21. Monitoring through data reconciliation.

An important advantage of the moving variance window was to avoid the interruption of the online DR procedure due to measurement problems. These failures occur more frequently with compositions measured online through gas chromatography. These measurement failures cause missing data and, consequently, occurrence of series of constant values. As a result, the signal loses variability, preventing the realization of DR. Figures 22 and 23 illustrate cases of variable freezing caused by missing data.

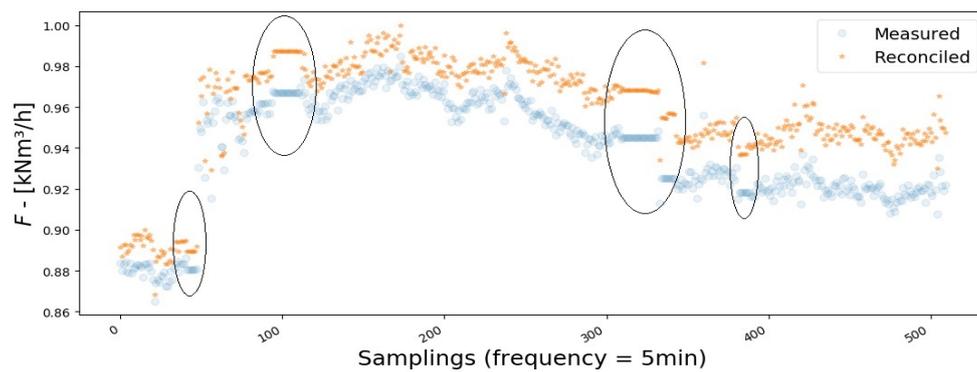


Figure 22. Measurement failures: missing data and frozen values of feed flowrate.

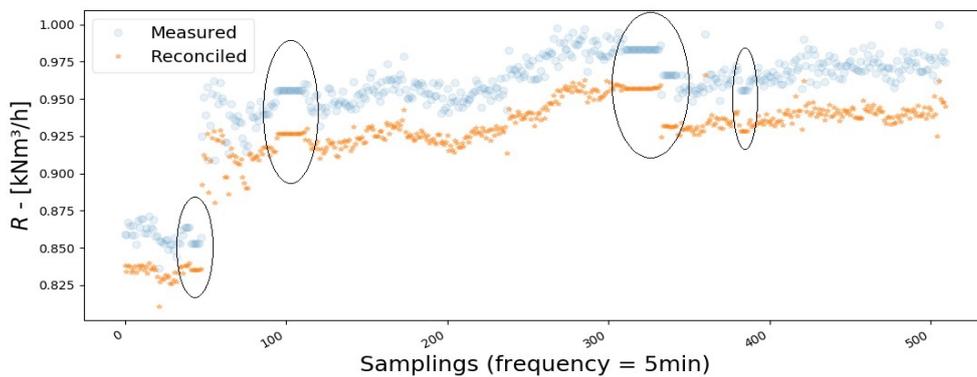


Figure 23. Measurement failures: missing data and frozen values of residue flowrate.

The analysis of bias can be performed through dynamic bar graph monitoring, illustrating the magnitude of the errors of each variable. Figure 24 informs the magnitudes of the systematic deviations from the median, that is, how many times the reconciled variable deviated from the measured median value. Systematic deviations that are larger than three times the value of the standard deviation can be regarded as a bias. Therefore, analyzing Figure 24, five variables with measurement biases could be observed: N2 (feed); N2 (residue); C8 (permeate); feed flowrate; and residue flowrate. Generally, biases can indicate the occurrence of unbalanced measurements, calibration problems, and instrument malfunctioning. For this reason, the obtained results were relayed to maintenance teams for evaluation of the instrumentation performances.

Figure 25 analyzes the performance of the DR, presenting the value of the OF, which represents the degree of correction of the reconciliation. The two green dotted lines represent the region where a normal distribution is expected for the errors of all measured variables. The region above the red dotted line indicates the samples that were subject to large corrections during the reconciliation step. Therefore, it is reasonable to consider the possible occurrence of outliers when the obtained value of OF deviated more than three standard deviations from the median value.

A test to observe the influences of biases on the analysis and performance of DR was performed. Figures 26 and 27 illustrate the same analyses performed for the same time window, as presented in Figures 24 and 25, but after identification and compensation of outliers. It can be observed that outliers significantly affected the average OF value. Based on Figures 24–27, it can be said that the analysis of bias and outliers performed very well. Figure 27 also illustrates the benefits of bias and outlier adjustments, as objective function values were reduced significantly and shed light on the existence of persistent outlier measurements. This reinforces the importance of bias and gross error diagnosing and the necessity to involve maintenance teams for evaluation of the instrumentation performances.

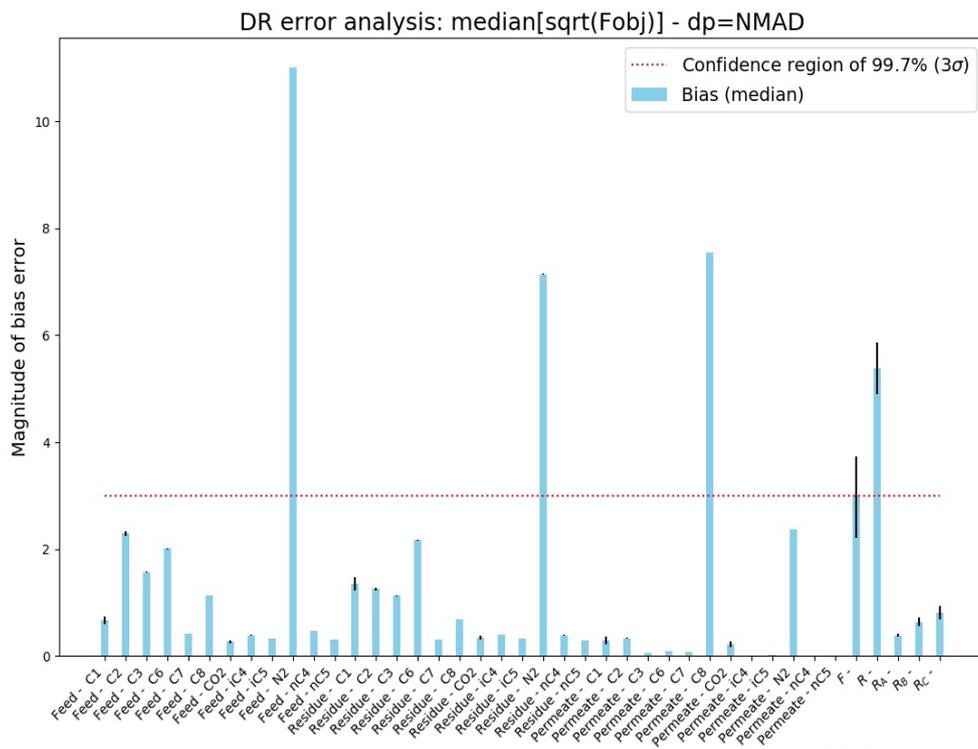


Figure 24. DR analysis without bias compensation.

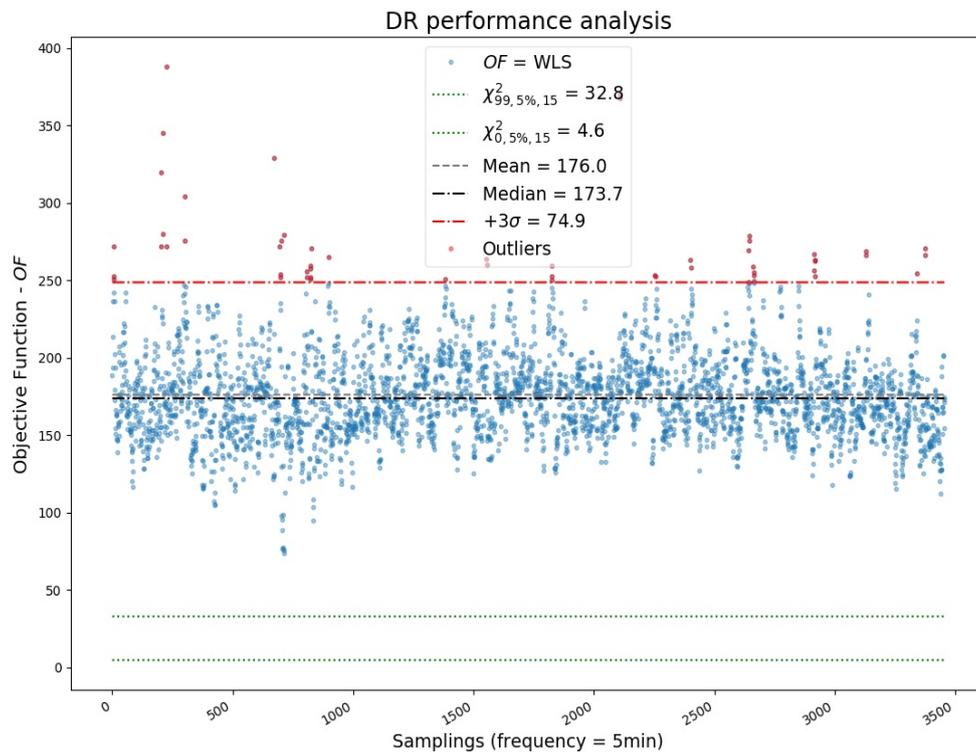


Figure 25. DR performance analysis without bias compensation.

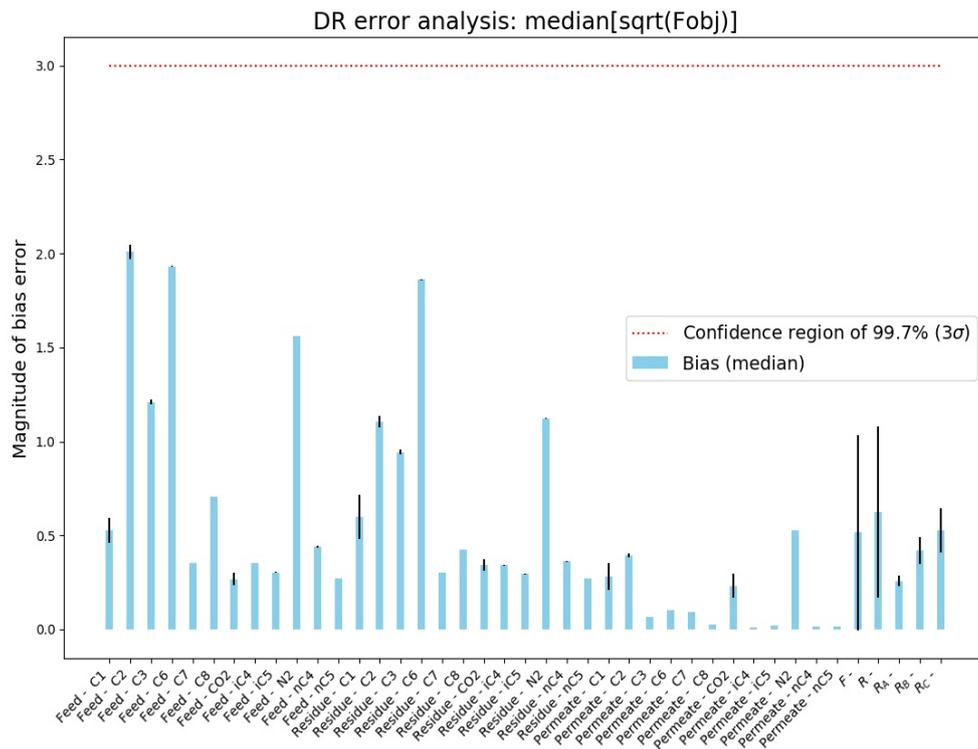


Figure 26. DR analysis with bias compensation.

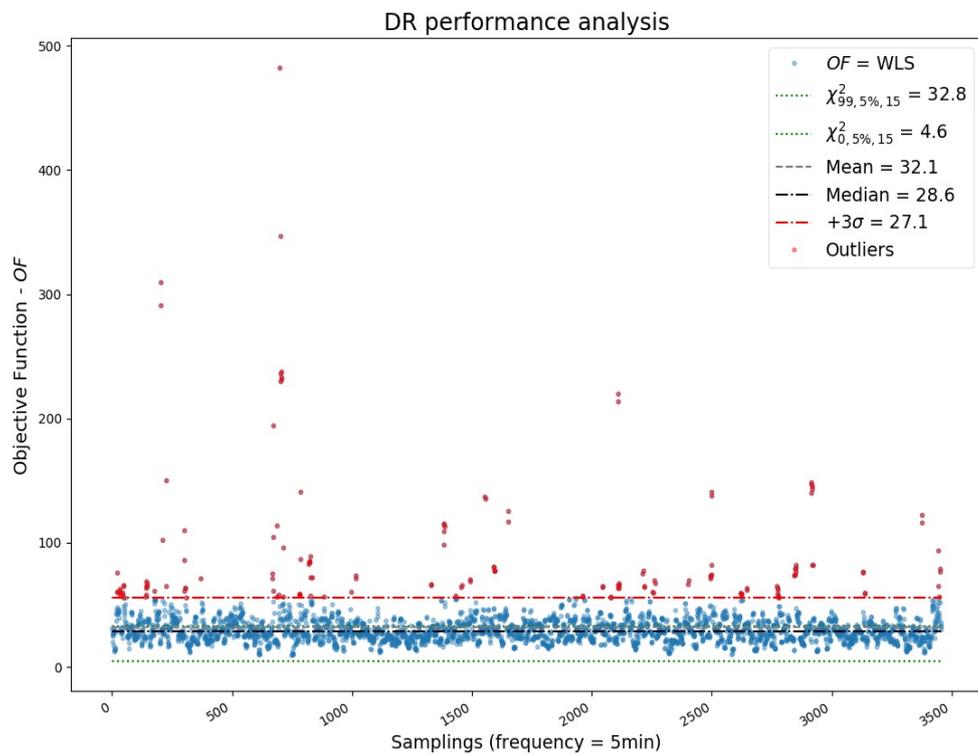


Figure 27. DR performance analysis with bias compensation.

3.4. Monitoring

In each new cycle of data acquisition, the code runs in sequence the pre-treatment, statistical tests and outlier compensation, data reconciliation with the permeate flowrate estimation, and finally the energy balance to calculate the temperature of the permeate flowrate. The first inferred variable was the permeate flowrate, estimated within the data reconciliation procedure. Figure 28 illustrates the real-time monitoring of the inferred variable.

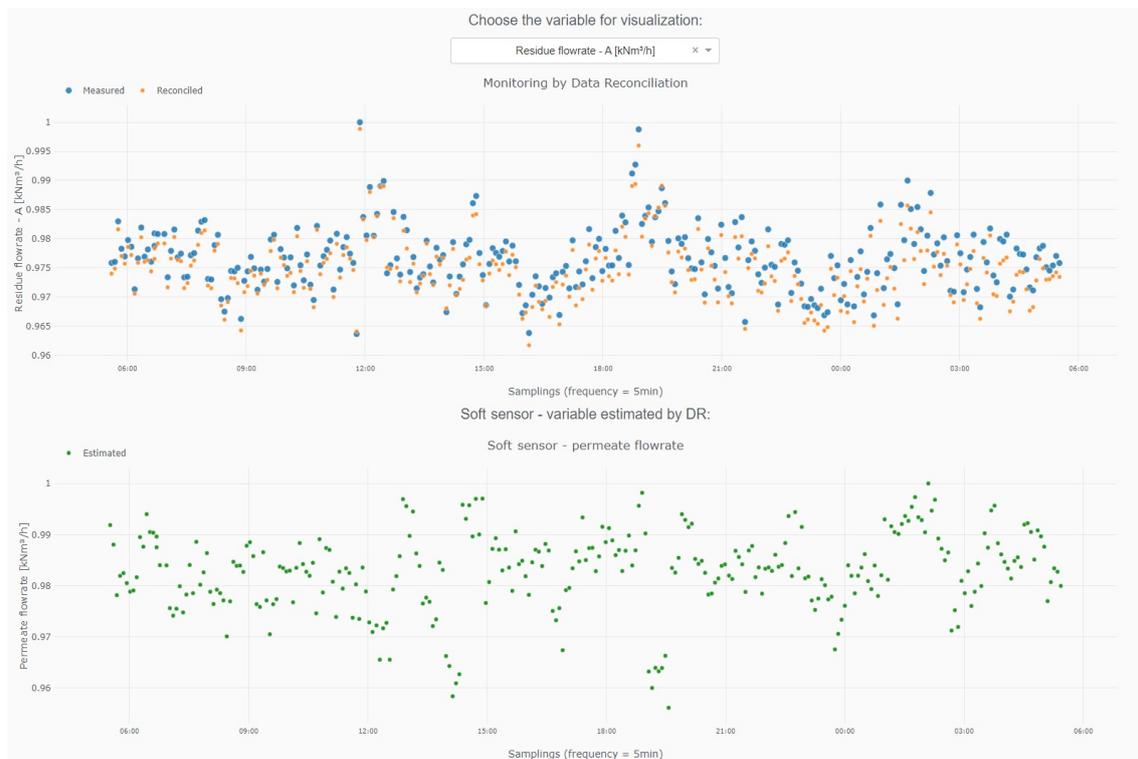


Figure 28. Real-time soft sensor—permeate flowrate.

The second inferred variable was the temperature of the permeate. At this stage, it was not possible to reconcile data due to the lack of redundancy of measured variables. Thus, this variable was inferred without the proper statistical treatment by the DR stage. Figure 29 illustrates part of the web application (web-app) where the user interacts with the interface. The variables can be selected through a dropdown menu. In addition, the application provides graphs of gross error analysis (Figure 30), visualization of inferred variables, and a button to start and stop monitoring. Figure 31 illustrates the three temperatures of each stream, during the testing period of the web-app. The permeate temperature was calculated with help of the energy balance and the regions without data are the days when the web-app was paused. All monitoring data are saved and can be read and analyzed offline, as in the case of Figure 31.

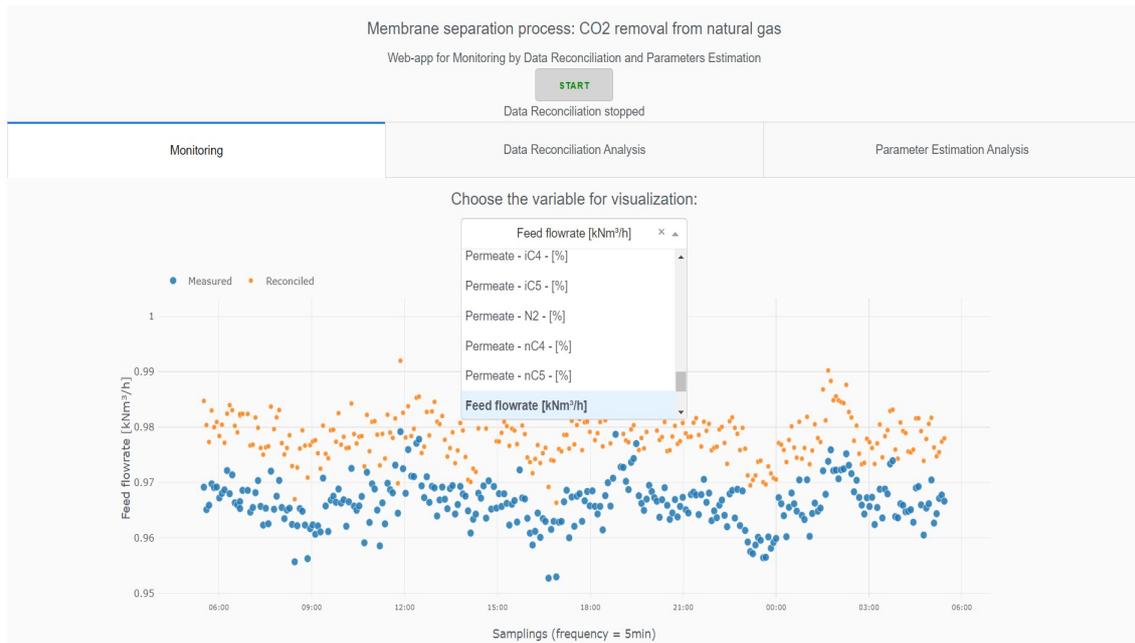


Figure 29. Part of the web-app: variables measured and reconciled.

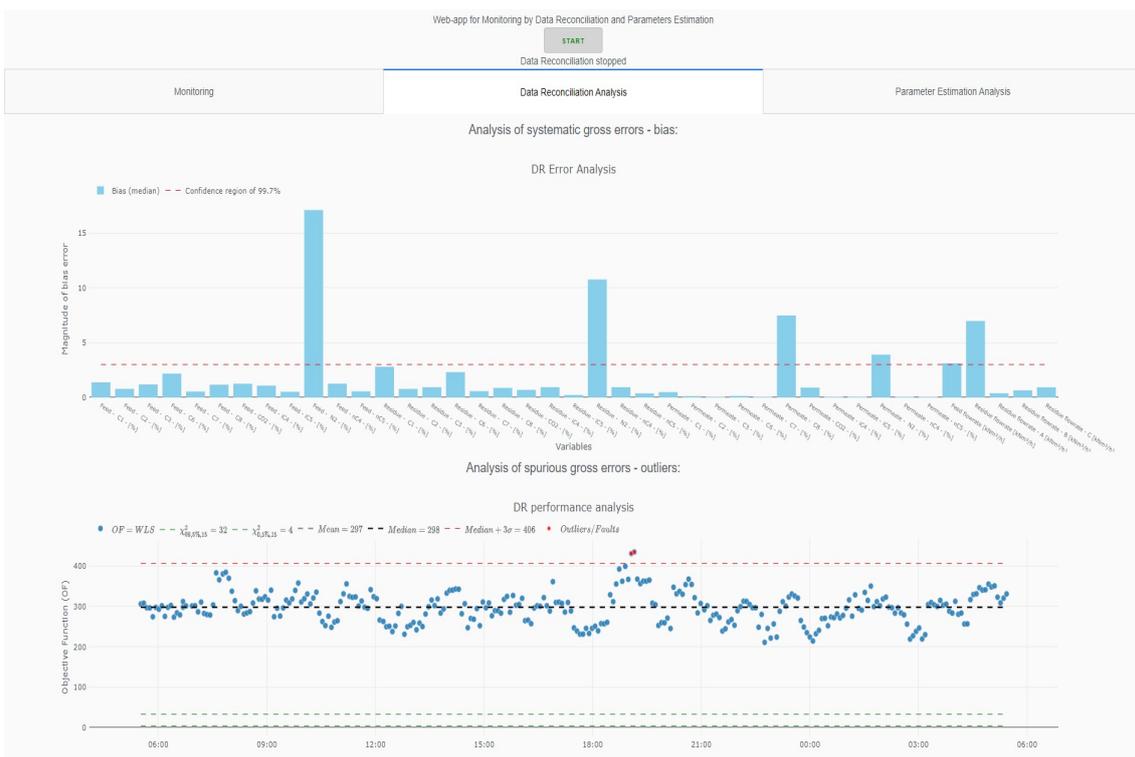


Figure 30. Part of the web-app: data reconciliation analysis.

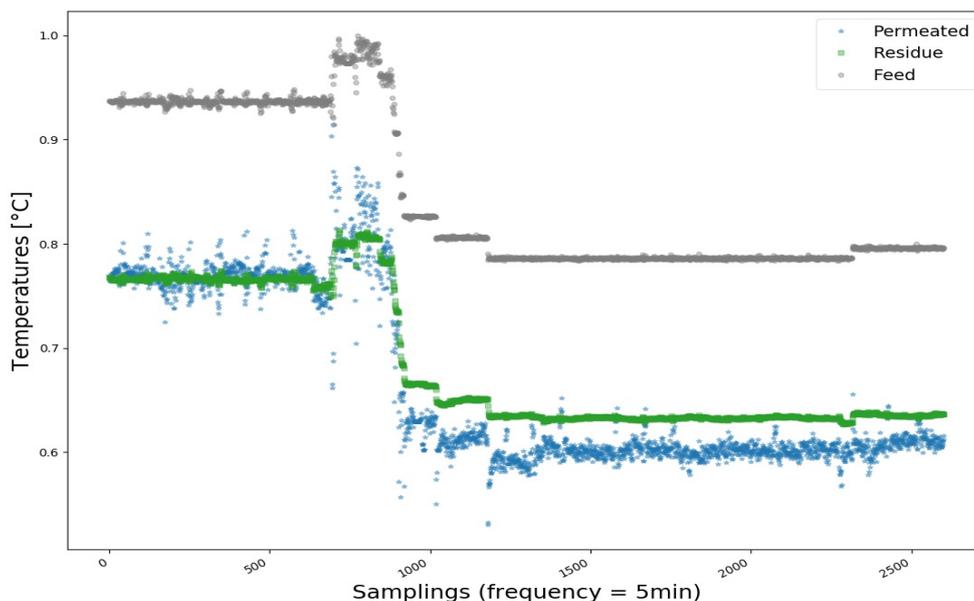


Figure 31. Offline data analysis: Temperatures in the testing period.

Therefore, the temperature inferred by the calculation of the energy balance showed good accuracy in relation to reported offline measurements, which demonstrates the importance of DR for the treatment of the variables used in the soft sensor.

4. Conclusions

A methodology was developed and implemented for the first time in the form of a web application to allow the monitoring of membrane separation processes online and in real time, making use of statistical techniques for treatment of process data. The proposed methodology comprises the following stages: (i) pre-treatment and characterization of process data; (ii) data reconciliation of process data to minimize measurement uncertainties, with the aid of mass balance equations; (iii) detection of systematic deviations for identification of process malfunctions; and (iv) observation of unmeasured variables (working as a soft sensor or digital twin). The pre-treatment and data characterization steps were fundamental for the understanding and correct formulation of the problem. The characterization step can find wide application, as this procedure can be applied in any chemical process. This step is essential for the appropriate selection of data reconciliation techniques and gross error detection procedures. After that, the proposed data reconciliation and gross error detection steps showed robustness, good performance, and speed. The proposed scheme was based on detailed steady-state balance equations, validated after proper characterization of actual operation data. The numerical procedures were validated offline and then implemented online and in real time for the first time, allowing the successful identification of measurement biases and outliers and providing estimates for unmeasured data. The developed procedures can be used for online and real-time detection of process faults and process diagnosing. In addition, the procedure provides reliable data for future stages of simulations and parameter estimation, allowing the implementation of digital twins, as the model proposed in part I of this research project. Production Management System and Enterprise Resource Planning steps can also benefit from availability of more reliable data, and variables inferred by a soft sensor. Therefore, the main advantages of the procedure are reliable data handling, diagnosis of gross errors/failures, and real-time monitoring of the process.

Author Contributions: D.Q.F.d.M.: Conceptualization, Methodology, Software, Validation, Data curation, Writing—original draft, Writing—review and editing; M.C.C.d.S.: Software, Validation, Formal analysis; T.B.F.: Software, Validation, Formal analysis; T.K.A.: Resources and Funding acquisition; F.C.D.: Resources and Funding acquisition; P.H.T.: Resources and Funding acquisition; J.C.P.: Supervision, Project administration, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Petróleo Brasileiro SA (Petrobras); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Acknowledgments: The authors thank Petrobras, CNPq, and CAPES for the financial support to this work, as well as for covering the costs to publish in open access.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations

The following abbreviations are used in this manuscript:

GED	Gross error detection
DR	Data reconciliation
WLS	Weighted least squares
PI	Plant information
HDF5	Hierarchy Data Format version 5
NaN	Not a number
LL	Lower limit
UL	Upper limit
OF	Objective function
GT	Global test
NMAD	Normalized median absolute deviation
ACF	Autocorrelation functions
PACF	Partial autocorrelation functions
CCF	Cross-Correlations Function

References

1. Mokhatab, S.; Poe, W.A.; Mak, J.Y. *Handbook of Natural Gas Transmission and Processing: Principles and Practices*; Gulf Professional Publishing: Houston, TX, USA, 2018.
2. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. *ANP, ANP RESOLUTION N° 16, D*; Diário Oficial da União: Brasília, Brazil, 2008. Available online: <http://www.anp.gov.br/> (accessed on 11 July 2020).
3. Speight, J.G. *Natural Gas: A Basic Handbook*; Gulf Professional Publishing: Houston, TX, USA, 2018.
4. Henis, J.M.; Tripodi, M.K. Composite hollow fiber membranes for gas separation: The resistance model approach. *J. Membr. Sci.* **1981**, *8*, 233–246. [[CrossRef](#)]
5. Al-Obaidi, M.A.; Kara-Zaïtri, C.; Mujtaba, I.M. Simulation and sensitivity analysis of spiral wound reverse osmosis process for the removal of dimethylphenol from wastewater using 2D dynamic model. *J. Clean. Prod.* **2018**, *193*, 140–157. [[CrossRef](#)]
6. Singh, V.; Jain, P.; Das, C. Performance of spiral wound ultrafiltration membrane module for with and without permeate recycle: Experimental and theoretical consideration. *Desalination* **2013**, *322*, 94–103. [[CrossRef](#)]
7. Kovvali, A.S.; Vemury, S.; Admassu, W. Modeling of multicomponent countercurrent gas permeators. *Ind. Eng. Chem. Res.* **1994**, *33*, 896–903. [[CrossRef](#)]
8. Soares, R.d.P. Desenvolvimento de um Simulador Genérico de Processos Dinâmicos. Master's Thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil, 2003.
9. Nicholson, B.; López-Negrete, R.; Biegler, L.T. On-line state estimation of nonlinear dynamic systems with gross errors. *Comput. Chem. Eng.* **2014**, *70*, 149–159. [[CrossRef](#)]
10. Câmara, M.M.; Soares, R.M.; Feital, T.; Anzai, T.K.; Diehl, F.C.; Thompson, P.H.; Pinto, J.C. Numerical Aspects of Data Reconciliation in Industrial Applications. *Processes* **2017**, *5*, 56. [[CrossRef](#)]
11. Prata, D.M.; Schwaab, M.; Lima, E.L.; Pinto, J.C. Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor. *Chem. Eng. Sci.* **2010**, *65*, 4943–4954. [[CrossRef](#)]
12. Farias, A.C. Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros. Master's Thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil, 2009.

13. de Menezes, D.Q.F. Reconciliação Robusta de Dados em Colunas de Destilação. Master's Thesis, Universidade Federal Fluminense, Niterói, RJ, Brazil, 2015.
14. Prata, D.M. Reconciliação Robusta de Dados para Monitoramento em Tempo Real. Ph.D. Thesis, COPPE—Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil, 2009.
15. Prata, D.M.; Schwaab, M.; Lima, E.L.; Pinto, J.C. Nonlinear dynamic data reconciliation and parameter estimation through particle swarm optimization: Application for an industrial polypropylene reactor. *Chem. Eng. Sci.* **2009**, *64*, 3953–3967. [[CrossRef](#)]
16. Benqlilou, C. Data Reconciliation as a Framework for Chemical Processes Optimization and Control. Ph.D. Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2004.
17. Narasimhan, S.; Jordache, C. *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*; Gulf Professional Publishing: Houston, TX, USA, 1999.
18. Kuehn, D.R.; Davidson, H. Computer control II. Mathematics of control. *Chem. Eng. Prog.* **1961**, *57*, 44–47.
19. Crivellari, G.P. Modelagem Matemática e Simulação de um Permeador de Gases para Separação de CO₂ de Gás Natural. Master's Thesis, Universidade de São Paulo, São Paulo, SP, Brazil, 2016.
20. Dortmund, D.; Doshi, K. *Recent Developments in CO₂ Removal Membrane Technology*; UOP LLC: Des Plaines, IL, USA, 1999; pp. 1–30.
21. Rackley, S.A. *Carbon Capture and Storage*; Butterworth-Heinemann: Oxford, UK, 2017.
22. Dias, A.C.S.; De Sá, M.C.C.; Fontoura, T.B.; Menezes, D.Q.; Anzai, T.K.; Diehl, F.C.; Thompson, P.H.; Pinto, J.C. Modeling of spiral wound membranes for gas separations. Part I: An iterative 2D permeation model. *J. Membr. Sci.* **2020**, *612*, 118278. [[CrossRef](#)]
23. Lashkari, S.; Kruczek, B. Reconciliation of membrane properties from the data influenced by resistance to accumulation of gasses in constant volume systems. *Desalination* **2012**, *287*, 178–189. [[CrossRef](#)]
24. Feital, T.; Prata, D.M.; Pinto, J.C. Comparison of methods for estimation of the covariance matrix of measurement errors. *Can. J. Chem. Eng.* **2014**, *92*, 2228–2245. [[CrossRef](#)]
25. Stanley, G.; Mah, R. Observability and redundancy classification in process networks: Theorems and algorithms. *Chem. Eng. Sci.* **1981**, *36*, 1941–1954. [[CrossRef](#)]
26. McKinney, W. Pandas, Python Data Analysis Library. 2020. Available online: <http://pandas.pydata.org> (accessed on 11 July 2020).
27. McKinney, W. Pandas: A foundational Python library for data analysis and statistics. *Python High Perform. Sci. Comput.* **2011**, *14*. Available online: https://www.dlr.de/sc/en/desktopdefault.aspx/tabid-7649/13008_read-32724/ (accessed on 11 July 2020).
28. Kuriakose, J. Using HDF5 with Python. 2017. Available online: <https://medium.com/> (accessed on 11 July 2020).
29. Zaitsev, I. The Best Format to Save Pandas Data. 2019. Available online: <https://towardsdatascience.com/> (accessed on 11 July 2020).
30. Galarnyk, M. Understanding Boxplots. 2018. Available online: <https://towardsdatascience.com/> (accessed on 11 July 2020).
31. Feital, T.; Pinto, J.C. Use of variance spectra for in-line validation of process measurements in continuous processes. *Can. J. Chem. Eng.* **2015**, *93*, 1426–1437. [[CrossRef](#)]
32. Waskom, M. Seaborn: Statistical Data Visualization. Python 2.7 and 3.5. 2020. Available online: <https://seaborn.pydata.org> (accessed on 11 July 2020).
33. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
34. Peixeiro, M. The Complete Guide to Time Series Analysis and Forecasting. 2019. Available online: <https://towardsdatascience.com> (accessed on 11 July 2020).
35. Evsukoff, A.G. *Inteligência Computacional—Fundamentos e Aplicações*; E-papers: Rio de Janeiro, RJ, Brazil, 2020.
36. Vanhatalo, E.; Kulahci, M.; Bergquist, B. On the structure of dynamic principal component analysis used in statistical process monitoring. *Chemom. Intell. Lab. Syst.* **2017**, *167*, 1–11. [[CrossRef](#)]
37. Himmelblau, D.M. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*; Elsevier Science Ltd.: Amsterdam, The Netherlands, 1978; Volume 8.
38. Romagnoli, J.A.; Sanchez, M.C. *Data Processing and Reconciliation for Chemical Process Operations*; Academic Press: Cambridge, MA, USA, 1999; Volume 2.

39. Veverka, V.V.; Madron, F. *Material and Energy Balancing in the Process Industries: From Microscopic Balances to Large Plants*; Elsevier: Amsterdam, The Netherlands, 1997; Volume 7.
40. Chen, J.; Romagnoli, J.A. A strategy for simultaneous dynamic data reconciliation and outlier detection. *Comput. Chem. Eng.* **1998**, *22*, 559–562. [CrossRef]
41. Reilly, P.; Carpani, R. Application of statistical theory of adjustment to material balances. In Proceedings of the 13th Canadian Chemical Engineering Congress, Montreal, QC, Canada, 1963; pp. 21–23.
42. Schwaab, M. *Análise de Dados Experimentais: I. Fundamentos de Estatística e Estimação de Parâmetros*; E-papers: Rio de Janeiro, RJ, Brazil, 2007.
43. Lotufo, F.A.; Garcia, C. Sensores Virtuais ou Soft Sensors: Uma Introdução. In Proceedings of the 7th Brazilian Conference on Dynamics, Control and Applications, Presidente Prudente, Brazil, 5–9 May 2008; pp. 1–9.
44. Seader, J.D.; Henley, E.J.; Roper, D.K. *Separation Process Principles*, 3rd ed.; Wiley: New York, NY, USA, 2010.
45. Bilogur, A. Missingno: A missing data visualization suite. *J. Open Source Softw.* **2018**, *3*, 547. [CrossRef]
46. Lewinson, E. Violin Plots Explained—Learn How to Use Violin Plots and What Are Their Advantages over Box Plots! 2019. Available online: <https://towardsdatascience.com/> (accessed on 11 July 2020).
47. Câmara, M.M.; Quelhas, A.D.; Pinto, J.C. Performance evaluation of real industrial RTO systems. *Processes* **2016**, *4*, 44. [CrossRef]
48. Özyurt, D.B.; Pike, R.W. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. *Comput. Chem. Eng.* **2004**, *28*, 381–402. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Apêndice B

Artigo Aceito - A Review on Robust M-Estimators for Regression Analysis



Computers & Chemical Engineering

Available online 5 February 2021, 107254

In Press, Journal Pre-proof 



Review

A Review on Robust M-Estimators for Regression Analysis

D.Q.F. de Menezes ^a, D.M. Prata ^b, A.R. Secchi ^a, J.C. Pinto ^a

^a Programa de Engenharia Química, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, CEP 21941-972, Brasil

^b Departamento de Engenharia Química e de Petróleo, Universidade Federal Fluminense, Niterói, RJ, CEP 24210-240, Brasil

Received 11 August 2020, Revised 12 December 2020, Accepted 2 February 2021, Available online 5 February 2021.

Show less 

+ Add to Mendeley  Share  Cite

<https://doi.org/10.1016/j.compchemeng.2021.107254>

[Get rights and content](#)

Resumo: A análise de regressão constitui uma ferramenta importante para investigar o efeito das variáveis explicativas nas variáveis de resposta. Quando erros do tipo *outliers* e tipo *bias* estão presentes, o estimador de mínimos quadrados ponderados pode ter um desempenho insatisfatório. Por esse motivo, técnicas alternativas robustas têm sido estudadas em diversas áreas da ciência. No entanto, muitas vezes essas diferentes comunidades científicas estão desconectadas umas das outras, culminando na escassez de troca de conhecimento entre essas áreas. Assim, este artigo apresenta uma revisão sobre estimadores-M robustos em várias áreas do conhecimento. 50 (48 robustos) estimadores-M são ilustrados, incluindo o estimador de mínimos quadrados ponderados (não robusto), o estimador normal contaminado (quasi-robusto), o estimador de Huber (monótono), o estimador de correntropia (*soft-redescending*), o Estimador de Smith (*hard-redescending*) e os estimadores generalizados e adaptativos de Barron e Distribuição-T Generalizada. São apresentadas as funções matemáticas que descrevem os estimadores e suas respectivas formas gráficas. As constantes de ajuste de todos esses estimadores, para os níveis de eficiência relativa de 90%, 95%, 98% e 99% em relação à distribuição Normal, também são apresentadas. Além disso, o estudo apresenta uma tabela guia para ajudar o leitor a selecionar o estimador-M robusto mais adequado para seu problema de análise de regressão.

A Review on Robust M-Estimators for Regression Analysis

de Menezes, D. Q. F.^a, Prata, D. M.^b, Secchi, A. R.^a, Pinto, J. C.^a

^a*Programa de Engenharia Química, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, CEP 21941-972, Brasil*

^b*Departamento de Engenharia Química e de Petróleo, Universidade Federal Fluminense, Niterói, RJ, CEP 24210-240, Brasil*

Abstract

Regression analysis constitutes an important tool for investigating the effect of explanatory variables on response variables. When outliers and bias errors are present, the weighted least squares estimator can perform poorly. For this reason, alternative robust techniques have been studied in several areas of science. However, often these different scientific communities are disconnected from each other, culminating in the scarcity of knowledge exchange among these areas. Thus, this paper presents a review on robust M-estimators in various knowledge areas. 50 (48 robust) M-estimators are illustrated, including the Weighted Least Squares estimator (non-robust), the Contaminated Normal estimator (quasi-robust), the Huber estimator (monotone), the Correntropy estimator (soft-re-descending), the Smith estimator (hard-re-descending), and the adaptive Barron and Generalized T-distribution. The mathematical functions that describe the estimators and their respective graphical forms are presented. The tuning constants of all these estimators, for 90%, 95%, 98%, and 99% relative efficiency levels in respect to the Normal distribution are also presented.

Keywords: Regression analysis; data reconciliation; parameter estimation;

Email address: dmenezes@coppe.ufrj.br (de Menezes, D. Q. F.)

robust statistic; M-estimator.

1. Introduction

Regression analysis is a statistical tool widely used in almost all areas of knowledge that seeks to fit a mathematical model to an experimental data set. This is particularly true for engineering, especially chemical process engineering. In these fields, the term “regression analysis” usually refers to problems of data reconciliation, parameter estimation, or both. Although there are several methods for obtaining the statistically coherent value of variables and estimating the parameters of the mathematical model based on available data, the weighted least squares (WLS) method is the one used most frequently and assumes that measurement errors follow the Normal distribution model [1, 2, 3]. The Normal distribution model, developed in 1733 by Abraham de Moivre, and later attributed to C. F. Gauss, is used most often because of its general acceptance, elegant statistical properties, and ease of calculation [4, 5].

Bard [6], in his classic book on parameter estimation, points out the main reasons for using the Normal distribution model:

- Generally describes well the fluctuations of large number of experimental measurements;
- As the number of experimental disturbances increases, many distributions approach the Normal distribution (Central Limit Theorem);
- Having defined the mean and variance of a set of measurements and using the concepts of variational calculus, it becomes possible to show that the Normal distribution is the one that demands the input of minimum amount of additional information about a particular problem;
- Easy mathematical treatment, allowing the theoretical derivation of well-known sample statistics, including the t-Student, Chi-square (χ^2) and F-Fisher distributions.

56 According to Rey [7] and Stigler [13], the term “robustness” for use in the
57 field of statistical mathematics was introduced by G.E.P. Box in 1953. Qualita-
58 tively, it can be understood as “insensitivity to small deviations from idealized
59 hypotheses” for which the estimator is optimized [14]. The word “small” can have
60 two different interpretations: the first refers to a small number of considerable
61 divergences (spurious values) and the second to a considerable number of small
62 divergences (non-normal distribution). Stigler [9] commented that the term
63 “robustness” refers to the sensitivity of procedures to deviations from idealized
64 hypotheses, particularly the hypothesis of normality - provided it is employing
65 well-defined procedures.

66 Robust statistics aims to provide methods that emulate conventional sta-
67 tistical methods but are not unduly affected by spurious values or other small
68 deviations from the reference statistical distribution model. It is particularly
69 important to note that, according to Box [15]:

70 *“All models are wrong, but some are useful”.*

71 The three main classes of robust estimators are:

- 72 • L-estimators (linear combinations of order statistics of the observations);
- 73 • R-estimators (estimator based on waste ranking);
- 74 • M-estimators (generalizations of a Maximum Likelihood estimator).

75 Although there are other classes of robust estimators (D, S, τ , CM, GM,
76 GS, MM, RM, LMS, and TS), these classes have been rarely applied to regression
77 analyses in chemical engineering problems [16, 17].

78 L-estimators are usually extremely simple and have robust statistics, such
79 as the sample median and trimmed mean (α -trimmed mean). Astronomers
80 from ancient Egypt and Rome usually discarded extreme data to average ob-
81 servations, indicating that they were the first to use the L-estimator α -trimmed

82 mean [18]. R-estimators involve the classification of residues. The classification
83 of a sample is a mapping of n real numbers, assuming the ordering of the data.
84 Thus, the calculation of weights and interpretations are facilitated and gener-
85 ally resistant to outliers. The L- and R-estimators are therefore useful in robust
86 statistics, such as descriptive statistics, statistical education, and when calculat-
87 ing the full statistical distribution of a particular problem becomes very costly.
88 For these reasons, the L- and R-estimators play key roles in many nonpara-
89 metric statistical approaches. However, currently, M-estimators are preferred,
90 although they are computationally more expensive. M-estimators are simpler
91 to handle because the shape of the estimator is fixed by a function. Indeed,
92 definition of the robust properties of the L- and R-estimators are difficult to
93 obtain *a priori*. However, this problem does not exist with M-estimators. If
94 one decides to calculate a robust estimate with good model efficiency over a
95 known distribution, then one must select a limited robust M-estimator with
96 previously determined good efficiency over the known distribution. If one sus-
97 pects that spurious values can be present with large deviations, then one must
98 select a strongly robust M-estimator for correction or removal of these spurious
99 measurements [14, 19].

100 According to Albuquerque and Biegler [20] and Prata et al. [3], an advan-
101 tage of M-estimators is that outlier detection and regression problems can be
102 solved simultaneously, unlike sequential approaches based on regression residues.
103 M-estimators are robust because of their intrinsic mathematical structure, which
104 renders the estimator less sensitive to spurious deviations [7, 14]. These esti-
105 mators tend to value the majority of the data located around the mean and
106 ignore the influence of spurious values (usually located far from the mean) si-
107 multaneously. Thus, an accurate regression can be performed even if nothing
108 is known *a priori* about outliers or the structure of data errors. In addition,
109 nothing prevents exploratory methods from being used to detect gross errors
110 and obtain more information about the statistical properties of the data once
111 residues are estimated. Although the main concepts of robust statistics have

112 been formally developed recently, the first robust M-estimators were proposed
113 during the early development of statistics as presented in the studies of Stephen
114 Mack Stigler, who analyzed the history of statistics before the 1900's:

- 115 • Roger Joseph Boscovich (1757) used the absolute minimum value esti-
116 mator, resulting from the hypothesis that measurement errors follow a
117 Laplacian distribution [21] for linear regression of observed data (apud
118 [22]);
- 119 • Daniel Bernoulli (1785), based on the unpublished paper of 1769, pro-
120 posed the iterative mean reweighting algorithm, using a semi-circle func-
121 tion (apud [13]);
- 122 • Simon Newcomb (1886) used combinations of Normal distributions to solve
123 regression problems (apud [9]);
- 124 • Smith (1888) developed the first robust M-estimator, the Smith estimator,
125 that fits completely into the modern notation (apud [23]).

126 However, it was G. Galileo in 1632 who first reported that measurement
127 errors deserve systematic and scientific treatment [5]. This shows that the dis-
128 cussion regarding the statistical treatment of real data has been in progress
129 since the beginning of statistic. The literature shows that robust M-estimators
130 constitute powerful tools for development of sampling strategies and solution of
131 laboratory or industrial data regression problems (parameter estimation, data
132 reconciliation, or both simultaneously), when idealized hypotheses are inaccu-
133 rate, as they are capable of ignoring atypical values (spurious values), due to
134 their mathematical structure. For these reasons, robust M-estimators have been
135 applied in several fields, including:

- 136 • Electrical, electronic and telecommunications engineering (Merrill and
137 Schweppe [24]);
- 138 • Image processing (Charbonnier et al. [25], Stewart [26], Arya et al. [27]);

- 139 • Econometrics and finance (Ronchetti and Trojani [28], Martin and Simin
140 [29]);
- 141 • Civil engineering, meteorology and geodesy (Krarup [30], Berberan [31],
142 Wieser and Brunner [32]);
- 143 • Chemical engineering (Özyurt and Pike [33], Prata et al. [1]);
- 144 • Astronomy (Wu and Wu [34]);
- 145 • Mechanical engineering (Pennacchi [35]);
- 146 • Petrochemical engineering (Prata et al. [36]);
- 147 • Nuclear engineering (Valdetaro and Schirru [37]);
- 148 • Industrial pharmaceuticals (Liu et al. [38], Su et al. [39]);
- 149 • Medical, biomedical and biotechnological applications (Tabatabai et al.
150 [40]).

151 Based on the previous paragraphs, main purpose of the present work is
152 surveying the applications of robust M-estimators for data regression problems
153 (data reconciliation and parameter estimation) that are subject to gross errors.
154 A collection of 48 robust models from numerous areas of science and covering
155 a long period of time (from 1888 to 2019) are disclosed, including presenta-
156 tion of their main characteristics and performances when applied to different
157 applications, mainly in the chemical engineering field. Besides, the underlying
158 theory is presented, with emphasis on concepts associated with robust statistics,
159 including discussions about the influence function (M-estimator derivative), dis-
160 continuous estimators, relationship between robustness and relative efficiency in
161 respect to a reference distribution (generally Normal) and tuning of estimator
162 efficiency. In particular, parameter tuning is discussed in depth (including the
163 numerical stability), so that parameter values are presented for levels of 90%,
164 95%, 98% and 99% efficiency in respect to the Normal distribution for all ana-
165 lyzed M-estimators for the first time, including the computational codes used for

166 tuning computations. This article is therefore expected to provide the scientific
167 community with opportune directions when it comes to selecting appropriate
168 M-estimators for a particular application.

169 **2. Foundations of robust M-estimators**

170 There are many classes of robust estimators. The most important class
171 for the field of data rectification, as described in Albuquerque and Biegler [20],
172 Özyurt and Pike [33] and Prata et al. [36], is the class of M-estimators, which
173 corresponds to a generalization of the Maximum Likelihood Estimator. As
174 described by [16], no other class of robust estimators has been used so far in the
175 field of data rectification.

176 An estimator can be developed as the result of the Maximum Likelihood
177 formulation, after assuming the validity of a certain statistical distribution for
178 measurement errors, as described below.

179 *2.1. Maximum Likelihood Formulation*

180 The Maximum Likelihood formulation was introduced by Fisher in a se-
181 ries of articles between 1912 and 1922 [41], although some of the basic ideas had
182 been previously discussed by Daniel Bernoulli [42] and Gauss [43], and is the
183 procedure used most often to obtain estimators for data regression problems in
184 different environments (experimental, laboratory or industrial).

185 Initially, it is necessary to assume that a population of measured data
186 with random errors, ε_i (white Gaussian noise) related to the population of sam-
187 pled measurements, follow the probability function f_i (where $z_i = \hat{z}_i + \varepsilon_i$, z_i are
188 the measured variables, and \hat{z}_i are the estimated variables or reconciled vari-
189 ables). If the random errors ε_i are uncorrelated, measurements are independent
190 and the respective covariance matrix, which contains the variance values of the
191 measured data, becomes diagonal. In this case, the joint probability functions

192 f of observed fluctuations can be calculated as the product of the individual
 193 probability functions f_i . Thus, the maximum likelihood estimate can be de-
 194 fined as the one that maximizes the probability of observing experimental data
 195 f , according to Equation (1).

$$\max_{\hat{z}} f = \max_{\hat{z}} \prod_{i=1}^n f_i \quad (1)$$

196 However, in practice, it is more convenient to work with the minimization
 197 of the likelihood function [44], according to Equation (2).

$$\max_{\hat{z}} \prod_{i=1}^n f_i = - \min_{\hat{z}} \prod_{i=1}^n f_i \quad (2)$$

198 Equation (2) can be rewritten with help of the logarithm operator as
 199 presented in Equation (3).

$$\min_{\hat{z}} - \ln \left[\prod_{i=1}^n f_i \right] = \min_{\hat{z}} - \sum_{i=1}^n \ln [f_i] \quad (3)$$

200 Equation (3) can be used to build different data regression estimators,
 201 including the most popular ones, as shown below.

202 **Normal** \rightarrow **WLS**:

$$\max_{\hat{z}} \prod_{i=1}^n \left\{ \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] \right\} \quad (4)$$

$$\min_{\hat{z}} - \sum_{i=1}^n \ln \left\{ \exp \left[-\frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \right] \right\} \quad (5)$$

$$\min_{\hat{z}} \sum_{i=1}^n \frac{1}{2} \frac{(z_i - \hat{z}_i)^2}{\sigma_i^2} \quad (6)$$

$$\min_{\hat{z}} \sum_{i=1}^n \frac{1}{2} \xi_i^2 \quad (7)$$

203 The standardized residue ξ_i corresponds to the difference between the
 204 measured value z_i and the estimated (reconciled) value \hat{z}_i weighted by the stan-
 205 dard deviation σ_i . Optimal \hat{z}_i values are those that maximize the Likelihood
 206 function.

207 The Normal distribution results in the WLS estimator, presented in Equa-
 208 tion (7), also known as an L_2 estimator [7]. The WLS estimator can usually be
 209 regarded as a suitable estimator when measurements are free of spurious values.

210 **Contaminated Normal:**

$$\min_{\hat{z}} \sum_{i=1}^n -\ln \left[\frac{(1-p) \exp\left(-\frac{\xi_i^2}{2}\right)}{\sigma_i \sqrt{2\pi}} + \frac{p \exp\left(-\frac{\xi_i^2}{2}\right)}{b\sigma_i \sqrt{2\pi}} \right] \quad (8)$$

211 Equation (8) presents the estimator based on the Contaminated Normal
 212 distribution, also known as the bivariate distribution [45], where p is the prob-
 213 ability of finding spurious values (with $0 < p \leq 0.5$), and $b^2\sigma_i^2$ is the variance of
 214 contamination by spurious values (with $b > 1$).

215 **Laplacian \rightarrow LAV:**

$$\min_{\hat{z}} \sum_{i=1}^n |\xi_i| \quad (9)$$

216 The Laplacian distribution, also known as the double exponential dis-
 217 tribution, results in the Least Absolute Value (LAV) estimator, presented in
 218 Equation (9), also known as the L_1 estimator [7].

219 **Cauchy:**

$$\min_{\hat{z}} \sum_{i=1}^n \ln(1 + \xi_i^2) \quad (10)$$

220 Equation (10) presents the estimator based on the Cauchy distribution.

221 Figure 1 illustrates the probability distribution functions - $f(\xi)$ - for the
 222 functions presented in this section: Normal, Contaminated Normal, Laplacian,
 223 and Cauchy, respectively.

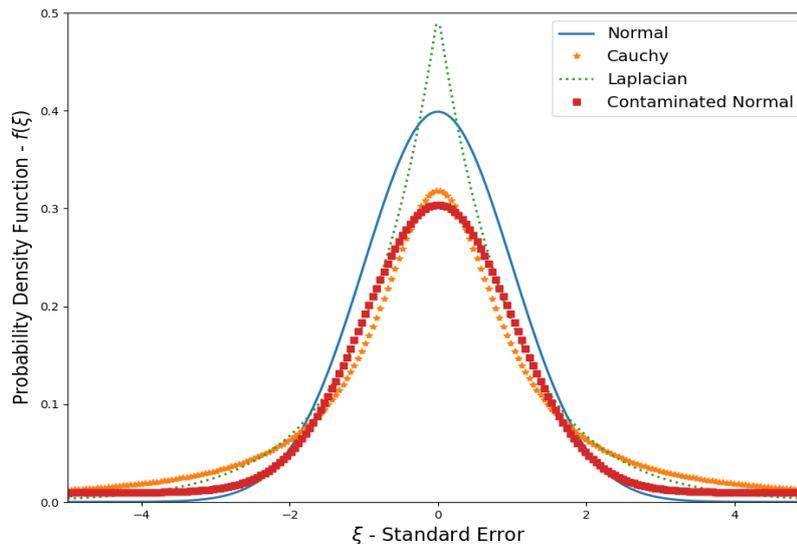


Figure 1: Probability Distribution Functions $f(\xi)$ of some common estimators.

2.2. Generalization of the Maximum Likelihood Objective Function

225 Huber [11] proposed a generalization of the maximum likelihood objec-
 226 tive function. In 1973 [46], Huber extended the idea of using M-estimators for
 227 solution of regression problems through minimization of a smooth, symmetrical
 228 and reasonably monotonic function of residuals - ρ - over optimal estimates, as
 229 presented in Equation (11).

$$\min_{\hat{\underline{z}}} \sum_{i=1}^n \rho \left(\frac{z_i - \mu_i}{\sigma_i} \right) = \min_{\hat{\underline{z}}} \sum_{i=1}^n \rho(\xi_i) \quad (11)$$

230 It is important to emphasize that M-estimators should not be independent
 231 of the dispersion measure, σ_i (scale parameter) while they attempt to estimate

232 the “mean” of the sample set of the variable i , μ_i (location parameter), with help
233 of some sort of iterative procedure [5]. In order to estimate the means simply
234 and robustly, the median ($\mathbf{med}(\mathbf{z})$) is frequently used to represent the average.
235 In Equation (11) it is assumed that the measurement errors are independent
236 and not correlated, with diagonal covariance-variance matrix.

237 The evaluation of the scale can be performed with traditional statistics
238 [47] or robustly, using the M-estimators iteratively or the robust L-estimators
239 MADn (Normalized Median Absolute Deviation) [48, 5], Tn [49] or Sn and Qn
240 [50], with help of direct or indirect methods. Indirect methods make use of
241 the process model during the estimation of standard deviations, based on the
242 available data set and assuming that deviations from the reference model pro-
243 vide appropriate estimates for measurement errors. However, finding a global
244 optimum for the Location and Scale of non-Gaussian Maximum Likelihood func-
245 tions still remains an open problem [5]. Figure 2 illustrates the M-Estimators
246 - $\rho(\xi)$ - for the previously analyzed functions: Normal, Contaminated Normal,
247 Laplacian, and Cauchy, respectively.

248 It must be pointed out that the development of most M-estimators is not
249 based on a well-defined and previously known probability distribution. Accord-
250 ing to Rey [7], the development of most estimators is based on a convenient
251 mathematical structure, such as the Fair estimator [51].

252 Some common properties of ρ are:

- 253 • ρ is continuous, to render the numerical manipulation easier, although
254 this is not needed for successful use and numerical implementation of the
255 estimator;
- 256 • ρ is symmetrical ($\rho(\xi) = \rho(-\xi)$), although this does not necessarily rep-
257 resent the real behavior of measurement errors, which can be distributed
258 asymmetrically around the mean;
- 259 • ρ must be strictly positive ($\rho(\xi) \geq 0$) and integrable in order to allow the

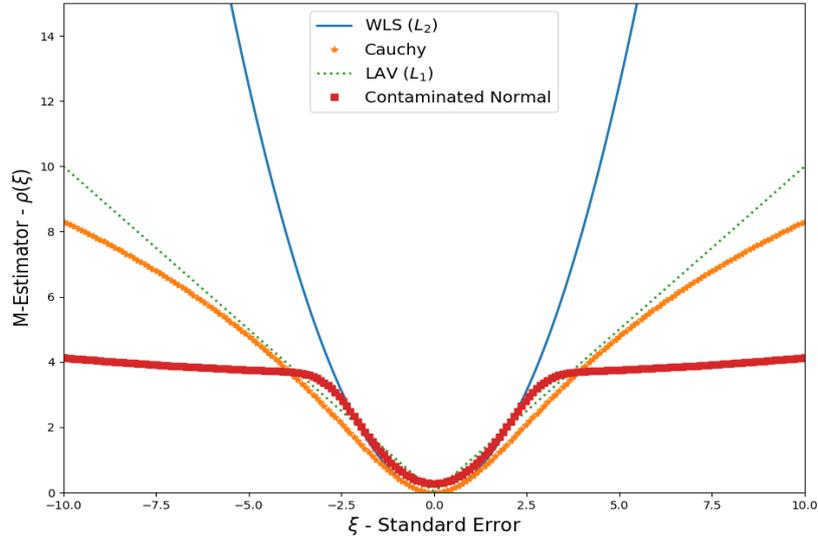


Figure 2: Graphical representation of some estimators $\rho(\xi)$.

260 statistical interpretation of the problem;

- 261 • ρ should be an increasing monotonous function of ξ_i ($\rho(\xi_i) \geq \rho(\xi_j)$, for
- 262 $|\xi_i| > |\xi_j|$), indicating that the probability of measuring large errors de-
- 263 crease with the increase of the magnitude of the error, although this is not
- 264 needed for successful use and numerical implementation of the estimator;
- 265 • $\rho(0) = 0$; (It is recommended that the tuning constants satisfy this con-
- 266 straint, Hoaglin et al. [52, p. 366])
- 267 • As a consequence of the previous properties, ρ should be preferentially
- 268 convex, although this is not needed for successful use and implementation
- 269 of the estimator.

270 The convexity of ρ guarantee uniqueness of the obtained solution (global

271 optimal) for problems described by linear models [53, 7], or nonlinear problems

272 that can be appropriately described by linear constraints after linearization.

273 When ρ is pseudo-convex or quasi-convex (generally called "non-convex"),
 274 Huber [53] suggests the implementation of an iterative procedure that initiates
 275 the optimization procedure with another convex ρ function and then switches
 276 conveniently to the desired ρ function. This is the procedure used most often by
 277 statisticians, although the increase of the computational power and advances in
 278 mathematics, including the development of deterministic and non-deterministic
 279 global search methods, are exerting enormous influence on the evolution of nu-
 280 merical techniques proposed for solution of data regression problems. Particu-
 281 larly, in most cases data regression performed with robust estimators can only
 282 be regarded as robust when global optimization methods are used. This point
 283 has been formally discussed by Baselga [54] using non-deterministic simulated
 284 annealing and genetic algorithm procedures as numerical tools for minimization
 285 of robust objective functions. For these reasons, as already said, the convexity
 286 of the M-estimator and the continuity of the derivatives are not mandatory, as
 287 pointed out and illustrated by Pennacchi [35].

288 *2.3. Influence Function*

289 The concept of influence function (or curve) was introduced by Hampel
 290 [12] and constitutes an important measure of robustness, being generally clas-
 291 sified as a measure of qualitative robustness [55, 48, 53]. For example, for an
 292 estimator to be considered robust, its influence function must be limited. In
 293 simple words, the influence function ψ corresponds to the weight (influence/im-
 294 pact) given to the effect of the magnitude of a spurious value (almost always
 295 measured in terms of multiples of the standardized residue ξ) on the obtained
 296 estimates. This function can be formally defined by Equation (12), and is char-
 297 acterized by the first derivative of the objective function, ρ , in respect to the
 298 standardized residue, ξ .

$$\psi(\xi) = \frac{\partial \rho(\xi)}{\partial \xi} \quad (12)$$

299 The estimator WLS $[\frac{\xi^2}{2}]$ is not robust because its influence function is
300 ξ , meaning that the influence of spurious values on the estimates is unlimited
301 (lack of robustness) and increases proportionally with the increasing magnitude
302 of the spurious value. On the other hand, the Welsch [56, 7] and Smith [57, 23]
303 estimators present decreasing influence functions that can remove the negative
304 effects of spurious values on estimates, even when the magnitudes of spurious
305 values increase, explaining why these estimators are classified as **redescending**.

306 According to Holland and Welsch [58], redescending estimators can be
307 classified as **soft-redescending** (pseudo-convex) and **hard-redescending** (quasi-
308 convex), depending whether the respective influence functions are approximately
309 null and exactly null for spurious values of high magnitude. In the case of hard-
310 redescending estimators, ψ is usually built with discontinuous segments and
311 makes use of “if” clauses, having a finite point of rejection, such as the three-
312 part Hampel estimator [12].

313 The Fair estimator, on the other hand, can be influenced by spurious val-
314 ues, even if only to a limited extent, and is classified as a **monotone** estimator.
315 However, the influence of spurious values in this case is much smaller than in the
316 case of the WLS estimator. One must note that the ρ function of a monotonous
317 estimator is convex. Rey [7, p. 101] proposed the term “quasi-robust” to refer
318 to estimators that are more robust than the traditional WLS estimator without
319 being strictly robust.

320 Thus, based on the previous discussion, M-estimators can be classified in
321 terms of the mathematical properties of the respective influence functions as:

- 322 • Non-robust (as the WLS estimator);
- 323 • quasi-Robust (as the Contaminated Normal estimator);
- 324 • Robust-Monotonous (as the Fair estimator);
- 325 • Robust-Soft-Redescending (as the Welsh estimator);

326 • Robust-Hard-Redescending (as the Smith estimator).

327 An illustrative graphical comparison of the mentioned influence functions
328 is presented in Figure 3.

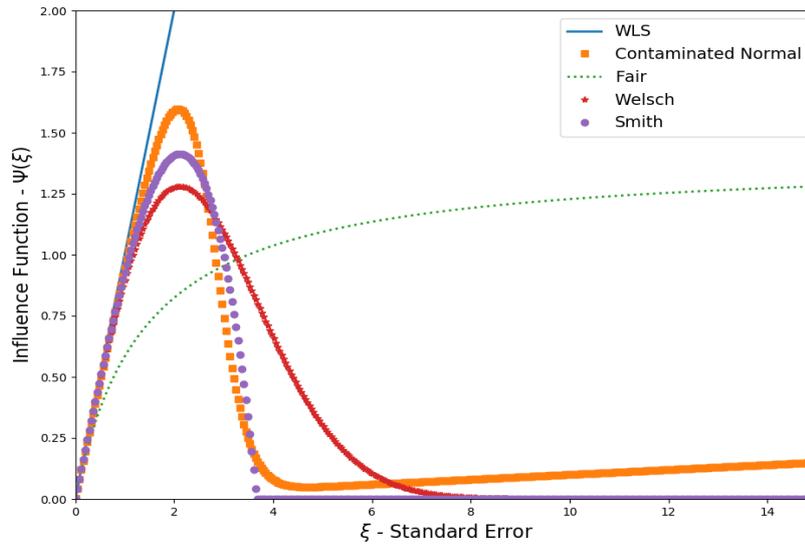


Figure 3: Graphical representation of the influence functions of some analyzed estimators.

329 It can be observed in Figure 3 that the Contaminated Normal (quasi-
330 robust) estimator has its influence function limited to approximately $\xi = 4$
331 (actually, $\xi = 4.742$); after that, the estimator becomes unlimited and behaves
332 as a pseudo-robust estimator.

333 Based on the structure of the influence function, three other properties
334 can be defined and discussed [59]:

335 • Gross-error sensitivity - γ - defined as the supreme (maximum) of the ab-
336 solute value of the influence function. γ measures the maximum (negative)
337 effect that a possible contamination on the measured value can exert on
338 the estimator [52, p. 358].

- 339 • Local-shift sensitivity - λ - defined as the derivative of the influence func-
 340 tion at the analyzed point. λ measures the limit of influence on the esti-
 341 mator if a measured value deviates from its original value (leverage point)
 342 and evaluates the possible occurrence of numerical sensitivity and discon-
 343 tinuities in ψ .

- 344 • Rejection point - r_p - defined as the point where the influence function
 345 becomes exactly null. This is related to the complete nullification of the
 346 (negative) effect of the high magnitude spurious value on parameter and/or
 347 variable estimates. As explained previously, this can only be defined for
 348 hard-re-descending estimators. In soft-re-descending estimators, this value
 349 can be calculated as an approximation, if some sort of threshold value is
 350 defined for the influence function [33].

351 Figure 4 presents an illustrative graphical explanation of these properties
 352 based on a hypothetical influence function.

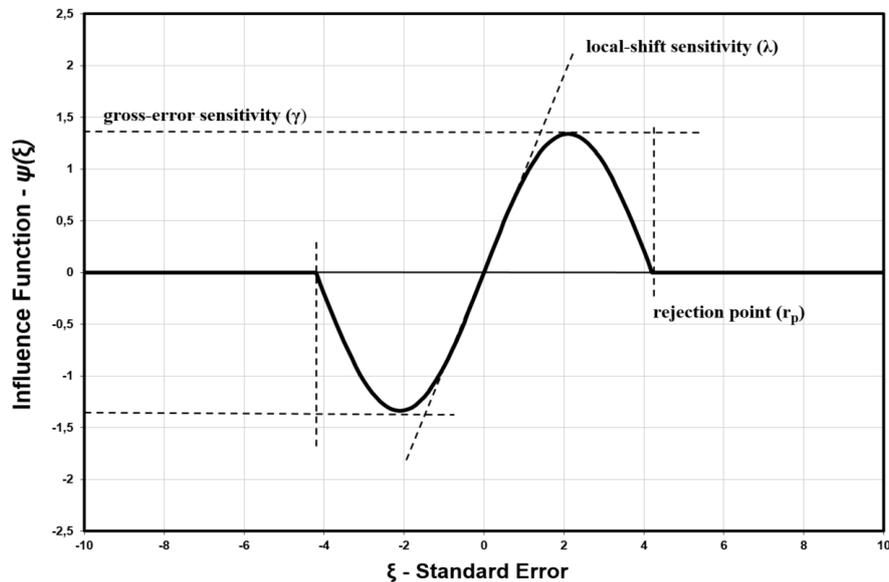


Figure 4: Properties of an Influence Function [59].

353 Some common properties of ψ are [52, p. 365]:

- 354 • ψ is limited, which characterizes robustness (finite gross error sensitivity
355 γ);
- 356 • ψ is continuous or piecewise continuous (finite local-shift sensitivity λ),
357 to facilitate the implementation of numerical procedures, although this is
358 not required for successful use and numerical implementation of robust
359 estimators;
- 360 • ψ is an odd function ($\psi(-\xi) = -\psi(\xi)$), which is related to the sym-
361 metrical nature of ρ , although measurement errors are not necessarily
362 symmetrical;
- 363 • ψ is approximately linear in the vicinities of the origin ($\psi(\xi) \approx k.\xi$, $k \neq 0$,
364 for small ξ), which is related to the quadratic nature of $\rho(\xi)$ in the vicinities
365 of the origin, although this property is not necessary for successful use and
366 numerical implementation of robust estimators;
- 367 • the rejection point of ψ is finite (if resistance to large spurious values is
368 desired).

369 2.4. First Derivative of the Influence Function

370 Another very important function can be defined as the derivative of the
371 influence function ψ in respect to the standardized residue ξ . This function
372 corresponds to the second derivative of ρ in respect to the standardized residue
373 ξ , as presented in Equation (13).

$$\psi'(\xi) = \frac{\partial\psi(\xi)}{\partial\xi} = \frac{\partial^2\rho(\xi)}{\partial\xi^2} \quad (13)$$

374 Thus, ψ' corresponds to the Hessian matrix of $\rho(\xi)$. According to Hoaglin
375 et al. [52], ψ' has been rarely reported in the literature in the field of robust

376 statistics, since the vast majority of analyzed regression problems have been
377 solved with deterministic methods and making use of numerical approximations
378 of the Hessian.

379 When ψ' is discontinuous, there is a lack of robustness concerning classical
380 (deterministic) optimization procedures. To deal with possible discontinuities
381 of ψ' , it is possible to:

- 382 • Modify the discontinuous form of ψ' with smoothing functions using de-
383 terministic optimization procedures, as performed by Arora and Biegler
384 [60] for the Hampel's hard redescending estimator.
- 385 • Use non-deterministic optimization procedures.

386 2.5. Robustness and Relative Efficiency

387 The use of M-estimators in data regression problems usually involves the
388 competitive relationship between efficiency and robustness [20], namely:

- 389 • (Relative) Efficiency: refers to the quality of the fit performed with an es-
390 timator when errors actually follow another distribution (reference), often
391 assuming the Normal distribution as the appropriate reference.
- 392 • Robustness: refers to estimator performance over a variety of (nonnormal)
393 error distributions.

394 This relationship can usually be manipulated through specification of
395 tuning constants. Attention should be paid to the important statement by Al-
396 buquerque and Biegler [20]: “*The more robust an estimator is, the less efficient*
397 *it is*”.

398 It is particularly relevant to report the tunable parameter values when one
399 intends to compare the performances of distinct robust estimators. As a matter
400 of fact, comparisons must be performed for similar levels of relative efficiency

401 and reference functions. If calculation of tunable parameter values results in
402 lower relative efficiencies (that is, larger robustness), the analysis of removal of
403 spurious values may become biased. This occurred recently with Xie et al. [61]
404 during tuning of the Xie estimator with relative efficiency of 95% in respect
405 to the normal distribution. The value attributed to the tunable parameter was
406 equal to 1.7134, resulting in relative efficiency of 91% instead of 95% (the correct
407 value corresponds to 1.9597, as shown in Table 3). As a consequence, the Xie
408 estimator presented the best overall performance, because it was more robust
409 than the other analyzed estimators (tuned for relative efficiencies of 95%).

410 According to Prata [16] and Prata et al. [36], robust estimators are de-
411 signed to perform very well when the distribution is contaminated and corrupted
412 by spurious values, with a small loss of efficiency when the distribution is Normal
413 and free of spurious values, which makes the WLS estimator the appropriate
414 optimal estimator.

415 Recently, Valdetaro and Schirru [37] proposed a method based on the
416 Akaike criterion for tuning robust estimators based on the quality of the data
417 set. This approach can be particularly useful in real-time applications of dy-
418 namic systems. As already said, the performance comparison of robust estima-
419 tors should be based on the same relative efficiency, which sheds light on the
420 importance of tuning methods [7, 33].

421 Equation (14) presents the mathematical definition of relative efficiency.

$$E_{ff}[\psi(\xi), f(\xi)] = \frac{V_f[\psi_f(\xi), f(\xi)]}{V[\psi(\xi), f(\xi)]} \quad (14)$$

422 where V_f is the asymptotic variance of the reference estimator and V is the
423 asymptotic variance attributed to a particular M-estimator [20]. On the other
424 hand, M-estimators are consistent and asymptotically normally distributed,
425 with asymptotic variance V defined in Equation (15) [53, p. 103], [59, p. 103],
426 [62]:

$$V[\psi(\xi), f(\xi)] = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[\int_{-\infty}^{+\infty} \psi'(\xi) f(\xi) d\xi \right]^2} \quad (15)$$

427 where f is the true error distribution, ψ is the influence function and ψ' is the
428 first derivative of the influence function.

429 Some desirable properties of ψ , f , ψ' and f' for calculation of relative
430 efficiencies are [62], [59, p. 126]:

- 431 • f should be symmetrical and unimodal;
- 432 • f should be continuously differentiable and present a finite support $\{\xi :$
433 $f(\xi) > 0\} = (-l, l)$ with $f(l) = f'(l) = 0$;
- 434 • For relative efficiency in respect to a reference distribution with density
435 f , ψ should be proportional to $-[\ln(f)]' = -\frac{f'}{f}$, i.e., $\psi \propto \psi_f$ [52, p. 365];
- 436 • $\int_{-\infty}^{+\infty} \psi(\xi) f(\xi) d\xi = 0$ (i.e., ψ should be an odd function);
- 437 • $\int_{-\infty}^{+\infty} \psi^2(\xi) d\xi < \infty$ (i.e., ψ should be asymptotically or identically null
438 when it tends to $\pm\infty$);
- 439 • $0 < \int_{-\infty}^{+\infty} \psi'(\xi) f(\xi) d\xi = - \int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi < \infty$ (i.e., ψ' and f' should
440 be limited).

441 As ψ' may be discontinuous, it may be convenient to rewrite Equation
442 (15) in the form of Equation (16), making use of the previously defined desirable
443 properties:

$$V[\psi(\xi), f(\xi)] = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[- \int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi \right]^2} = \frac{\int_{-\infty}^{+\infty} \psi^2(\xi) f(\xi) d\xi}{\left[\int_{-\infty}^{+\infty} \psi(\xi) f'(\xi) d\xi \right]^2} \quad (16)$$

444 *2.6. Breakdown Point*

445 The Breakdown Point (BP) concept was introduced by Hampel [12] and
446 is another important measure of robustness, being used as a measure of quan-
447 titative robustness [53]. The BP can be defined as the proportion of incorrect
448 (measured) observations (spurious values) in a dataset that the robust regression
449 technique can tolerate and still be used successfully. The higher the BP of an
450 estimator, the more robust it is. Theoretically, the highest possible BP value is
451 0.5 (50%) [63], because if more than half of the measurements are corrupted, it
452 is not possible to distinguish between the true distribution and the distribution
453 contaminated by spurious values. The BP of an M-estimator does not depend
454 on its probability distribution [63, p. 10], [52, p. 370]. Importantly, the WLS
455 estimator has BP close to zero; therefore, a single spurious value causes sig-
456 nificant deviations on the obtained estimates. This again reflects the extreme
457 sensitivity of the WLS estimator to spurious values [63]. The BP of a very
458 robust M-estimator is expected to be 0.5 [64], as these estimators can handle
459 approximately 50% of spurious values in the data set. This has been asymp-
460 totically illustrated through simulation for the Biweight, Hampel, Andrews and
461 Hyperbolic Tangent M-estimators [65].

462 **3. Application of Robust M-estimators in Data Reconciliation**

463 It is important to emphasize that the literature on Data Reconciliation
464 (DR) and Multiple Gross Errors Detection (MGED) has usually recommended
465 redescending M-estimators as the best options to reject the negative effects
466 of spurious values (large deviations) on estimates of variables and parameters
467 [60, 33], as these M-estimators showed better results in comparative studies.
468 However, the use of robust M-estimators for DR and MGED is still being con-
469 solidated, as these estimators have been increasingly employed by researchers.
470 Applications include analyses of very distinct problems, ranging from evalua-
471 tion of linear models used to fit simulated data sets at steady state conditions

472 to evaluation of highly nonlinear model constraints used to fit real dynamic
473 industrial data in real time, for purposes of process monitoring and real time
474 optimization. Table 1 summarizes some works that used robust M-estimators
475 to solve DR and MGED problems.

Table 1: Examples of robust estimators applied in the DR literature - part 1/3

Reference	System*	M-Estimators	Process(es)
[45]	S-SS	Contaminated Normal	Heat exchanger
[66]	S-SS	Contaminated Normal and Lorentzian	Heat exchanger and flow network
[67]	R-SS	Contaminated Normal	Sulfuric acid plant
[20]	S-D	Contaminated Normal and Fair	Heat exchanger, tanks and hydrolysis
[68]	R-SS	Fair and Lorentzian	Chemical reactor
[69]	R-SS	Contaminated Normal	Multistage flash desalination plants
[70]	S-SS	Kong	Adiabatic CSTR
[60]	S-SS/D	Fair and Hampel Contaminated	Steam metering and tanks Chemical reactor, steam
[33]	S/R-SS	Normal, Cauchy, Fair, Hampel, Logistic e Lorentzian	metering, metallurgical grinding, heat-exchanger, sulfuric acid and alkylation
[71]	S-SS	Contaminated Normal	Component material balances
[72]	S-SS	Hampel	Pai and Fisher [73] problem
[74]	S-SS	Huber and Kong	Pai and Fisher [73] problem
[75, 76]	R-SS	Kong	Coke-oven-gas purification
[77]	R-SS	Contaminated Normal	Chemical reaction and distillation column
[78]	S-SS	Contaminated Normal	Mineral processing plant

***S** - simulated data; **R** - real data; **SS** - steady state e **D** - dynamic.

Table 1: Examples of robust estimators applied in the DR literature - part 2/3

Reference	System*	M-Estimators	Process(es)
[79, 80]	S/R-SS	Contaminated Normal	Catalytic reformer, pipe model and flash
		Contaminated	
[81]	S-D	Normal, Cauchy, Fair, Hampel, Logistic, Lorentzian e Welsch	Liebman et al. [82] reactor (CSTR)
[36]	R-D	Welsch	Industrial polypropylene reactor
[83]	S-SS	Zhang	Atmospheric tower, ethylene separation, air separation and steam metering
[37]	S/R-SS	Hampel	Thermal reactor power
		Jin, Cauchy and Huber	Measurement network and Pai and Fisher [73] problem
[84]	S-SS		
[85]	S-SS	Zhang, Fair, Hampel and Correntropy	Atmospheric distillation tower and steam metering
[86]	R-SS	Correntropy	Air separation
[87]	S-D	Correntropy	Polymerization of styrene
[88]	S-SS	Zhang, Biweight, Welsch and Correntropy	Steam metering and Pai and Fisher [73] problem
[89]	S/R-SS	Welsch	Multi-fuel fired industrial boilers
[90]	R-D	Welsch	Polymerizations of methyl methacrylate

***S** - simulated data; **R** - real data; **SS** - steady state e **D** - dynamic.

Table 1: Examples of robust estimators applied in the DR literature - part 3/3

Reference	System*	M-Estimators	Process(es)
[91]	S-SS	Contaminated Normal, Bell, Huber, Lorentzian, Hampel, Welsch, Andrews, Smith, Jin, Biweight and Correntropy	Van de Vusse [92] reactor
[93]	S-SS	Wu	Crude oil distillation set
[94]	S-SS	Biweight, Huber, Hampel, Welsch and Correntropy	Steam metering and heat exchanger network
[95]	S/R-D	Fair and Hampel	Williams and Otto [96] reactor and reactive distillation
[97]	S-SS/D	Contaminated Normal, Cauchy, Fair, Hampel, Logistic, Lorentzian and Zhang (quasi-Weighted)	Collection of benchmark test problems
[38]	R-D	Fair, Logistic, Lorentzian and Welsch	Pharmaceutical manufacturing
[39]	R-D	Welsch	Pharmaceutical manufacturing
[61]	S/R-SS	Fair, Welsch, Cauchy e Xie	Measurement network and industrial evaporation
[98]	S-SS	16 robust M-estimators	Chemical reactors problems

***S** - simulated data; **R** - real data; **SS** - steady state e **D** - dynamic.

476 4. The 48 Robust M-Estimators Used in Data Regression Problems

477 Many M-estimators have been proposed in the robust statistics literature
478 and some of these have been used in data regression problems to minimize the
479 negative influence of less frequent spurious values on estimates of variables,
480 parameters or both together. Table 2 shows the analyzed robust estimators
481 divided into 3 categories, with the respective reference numbers used in the
482 present work.

483 It is important to emphasize that the present work displays the most
484 complete set of robust estimators described in the literature and that no arti-
485 cles or books already published in the literature review more than 10 examples
486 of robust M-estimators. Moreover, in the vast majority of the previously pub-
487 lished articles, M-estimators are presented only in terms of the function ρ , or
488 the function ψ , although frequently in terms of the **weight function** w , where
489 $w(\xi) = \psi(\xi)/\xi$. Additionally, presentation of tunable parameters for different
490 levels of relative efficiency is rare, as parameters are usually tuned to 95% effi-
491 ciency in respect to the Normal distribution. Moreover, some estimators have
492 seemingly never been tuned at all and available tuned parameter values are scat-
493 tered in the literature. The presentation of illustrative plots for the ρ , ψ , and
494 ψ' functions of the M-estimators has also been rare in the literature, making
495 more difficult the numerical validation of M-estimators and respective tunable
496 parameters.

497 Based on the previous remarks, the present review aims to facilitate the
498 access to the M-estimator families of functions (ρ , ψ and ψ'), respective graphs
499 and associated tuning parameters. Another objective of the present work is to
500 motivate the use of some M-estimators that have been reportedly and success-
501 fully used in few areas of knowledge to solve data regression problems. Thus, we
502 believe that the present review can facilitate the exchange of knowledge between
503 the areas of statistics, geodesy, meteorology, engineering, economics, image pro-
504 cessing, astronomy, among others. Additionally, this article aims to fill this gap

505 by presenting and discussing the properties and use of 50 different M-Estimators
506 (WLS, Contaminated Normal and 48 robust estimators).

Table 2: Analyzed Robust M-Estimators

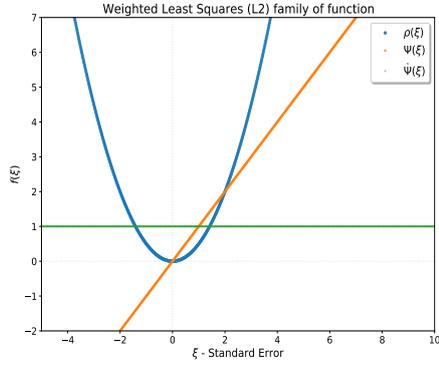
#- Monotone	#- Soft-re-descending	#- Hard-re-descending
1- LAV L_1	11- Alamgir	31- Asad
2- $L_1 - L_2$ (Charbonnier)	12- Bab-Hadiashar	32- Andrews
3- Fair	13- Bell	33- Biweight (Tukey-biweight)
4- Huber	14- Blake-Zisserman	34- Collins
5- Modified Huber	15- Cauchy (Lorentz)	35- Hampel
6- Kong	16- Correntropy	36- Hiperbolic Tangent
7- LnCosh	17- Danish	37- LQQ (Linear Quadratic Quadratic)
8- Logistic	18- Geman-McClure	38- Modified Asad-Qadir I
9- Müller	19- GGW (Generalized Gauss Weight)	39- Modified Asad-Qadir II
10- Zhang	20- Insha	40- Optimal
	21- Jin	41- Qadir
	22- Kumar	42- Smith
	23- Lorentzian	43- Talwar (Huber type-skipped mean)
	24- Merril-Schweppe (BDS)	44- Uk
	25- Ramsay	45- Yang I
	26- Sech	46- Yang II
	27- Welsch	
	28- Wu	
	29- Xie	
	30- Youssef	

507 Besides the estimators shown in Table 2, the 47-Barron and 48-GT (Gen-
508 eralized t-distribution) estimators, which belong to the category of **General**
509 **and Adaptive Robust Estimators**, are also considered [99, 100]. The prop-
510 erties of these two adaptive estimators depend on certain tunable parameters
511 so that the estimator obtains characteristics of monotonous, soft-re-descending
512 or hard-re-descending estimators. These estimators are not subject to formal ef-
513 ficiency tuning, because the variation of the parameters can completely modify
514 the characteristics of the estimator, instead of the simpler asymptotic adjust-
515 ment to the variance of the reference distribution.

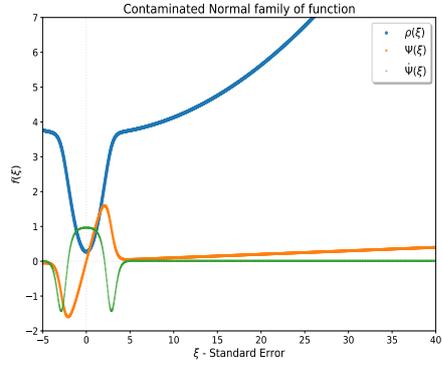
516 Tables C.5 and C.6 in Appendix C present the ρ functions (estimator)
517 and ψ functions (Influence Function) as functions of the standardized residue ξ
518 and their tuning constants, for all estimators mentioned previously. For better
519 visualization of these functions, the following figures were elaborated:

- 520 • Figures 5 to 12 illustrate graphically the families of ρ , ψ and ψ' functions
521 of all analyzed M-estimators;
- 522 • Figure 13 illustrates the ρ and ψ functions of the Barron (General and
523 Adaptive Loss Function) and GT (Generalized t-distribution) estimators.

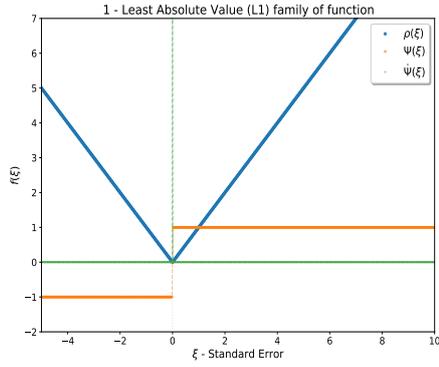
524 Properties of the analyzed monotonous and re-descending robust M-estimators
525 discussed previously and based on Figures 3 and 4 remain valid for all estimators
526 represented in the present section. A detailed summary of the main character-
527 istics of each analyzed M-estimator is provided in Table 4. Except for the WLS,
528 Charbonnier, LAV, Barron and GT M-estimators, all figures were built for Rela-
529 tive Efficiency of 95% in respect to the Normal distribution, while the respective
530 tuning constants are presented in Section 5.



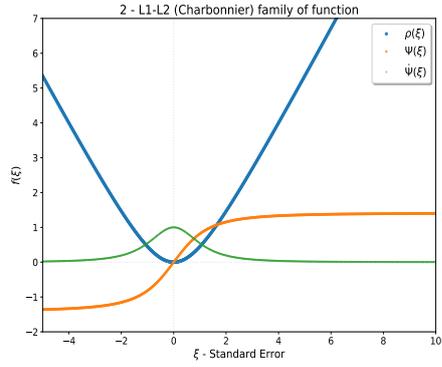
(a) WLS (L_2).



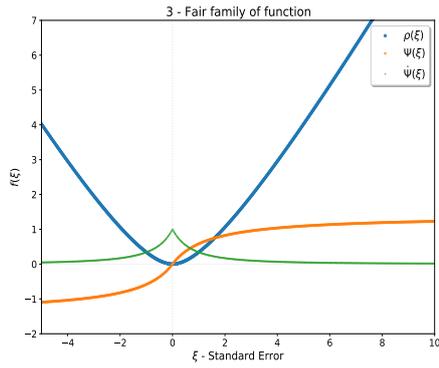
(b) 0- Contaminated Normal.



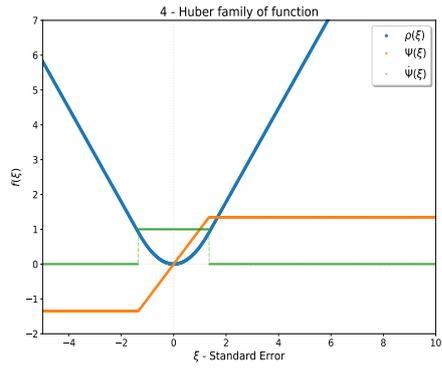
(c) 1- LAV (L_1).



(d) 2- Charbonnier.

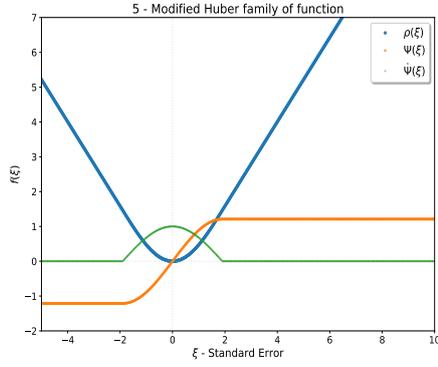


(e) 3- Fair.

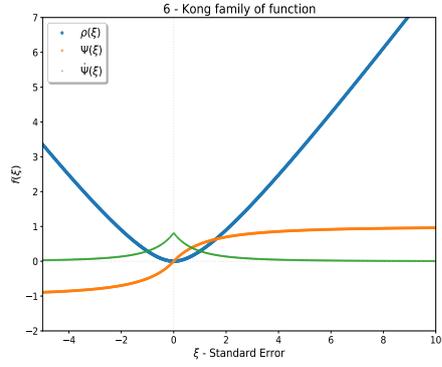


(f) 4- Huber.

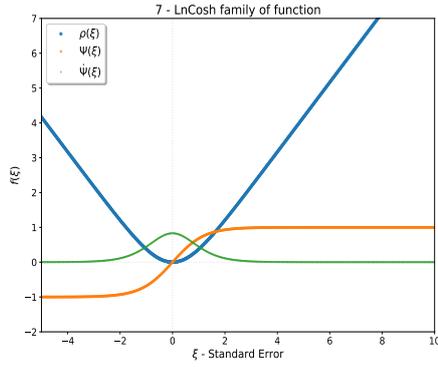
Figure 5: Families of analyzed functions for distinct estimators (ρ , ψ , ψ') – part 1/8.



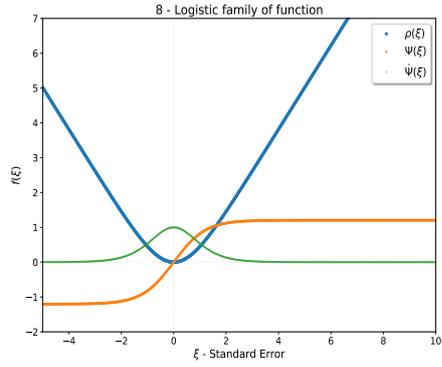
(a) 5- Modified Huber.



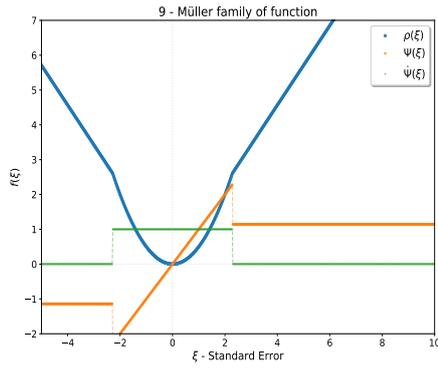
(b) 6- Kong.



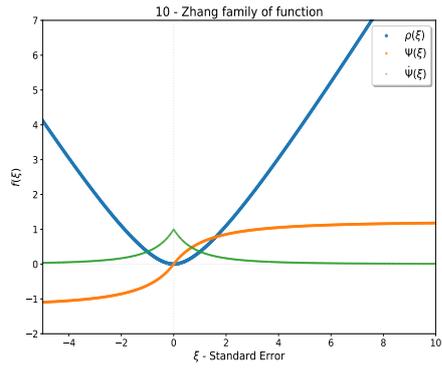
(c) 7- LnCosh.



(d) 8- Logistic.

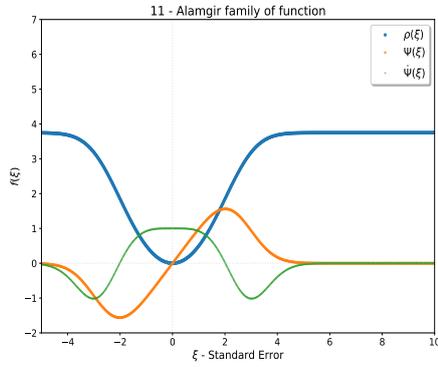


(e) 9- Müller.

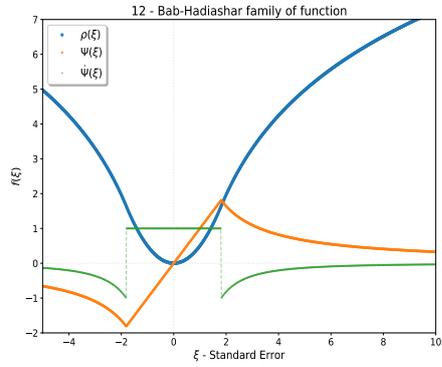


(f) 10- Zhang.

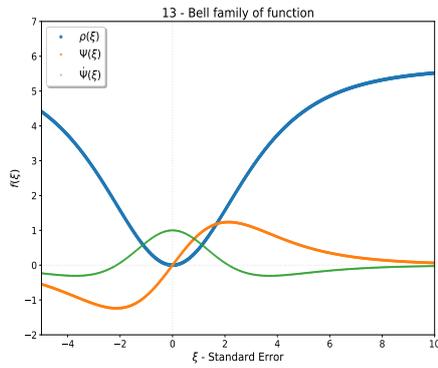
Figure 6: Family of function (ρ, ψ, ψ') – part 2/8.



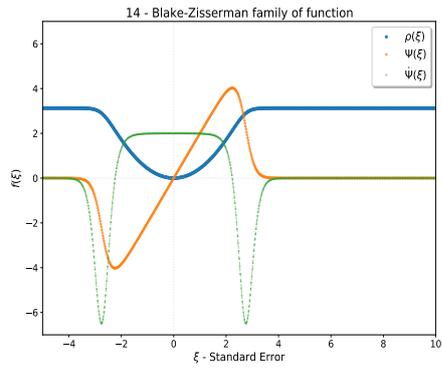
(a) 11- Alamgir.



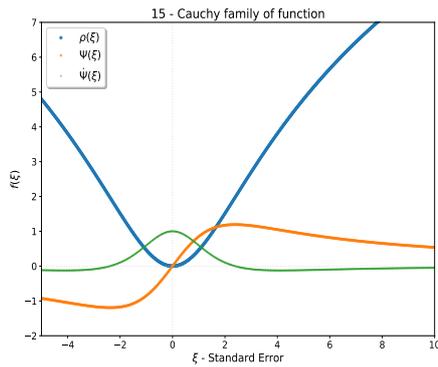
(b) 12- Bab-Hadiashar.



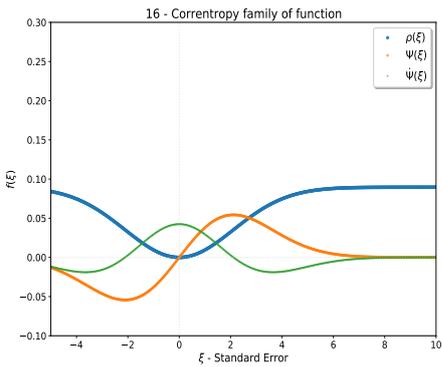
(c) 13- Bell.



(d) 14- Blake-Zirsserman.

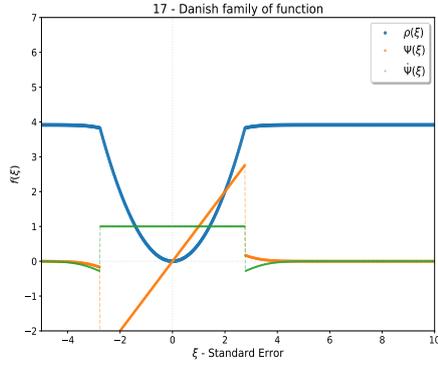


(e) 15- Cauchy.

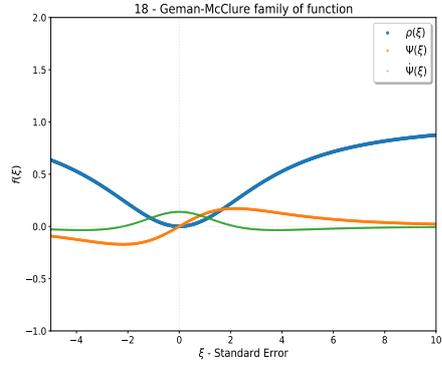


(f) 16- Correntropy.

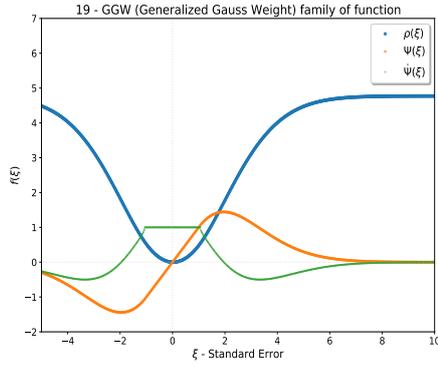
Figure 7: Families of analyzed functions for distinct estimators (ρ, ψ, ψ') – part 3/8.



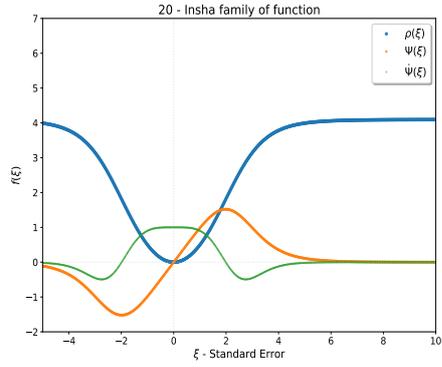
(a) 17- Danish.



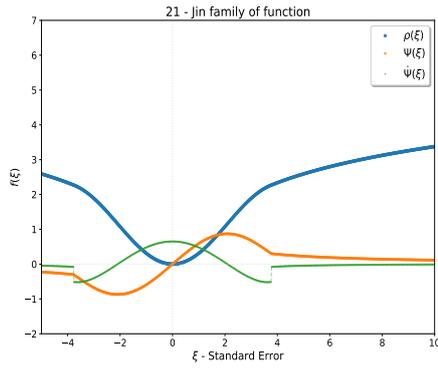
(b) 18- Geman-McClure.



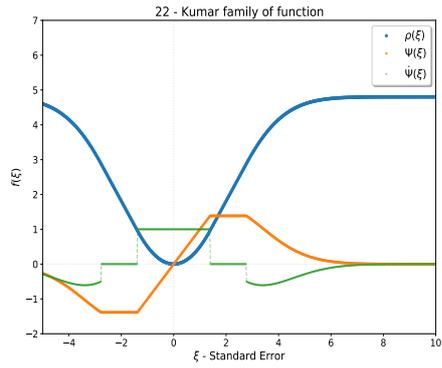
(c) 19- GGW.



(d) 20- Insha.

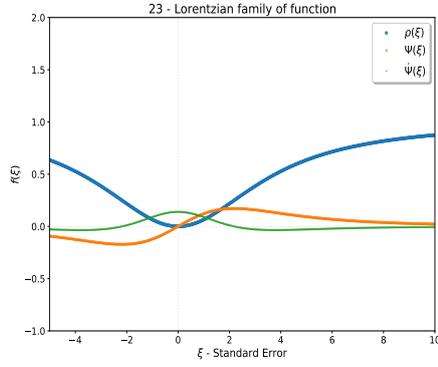


(e) 21- Jin.

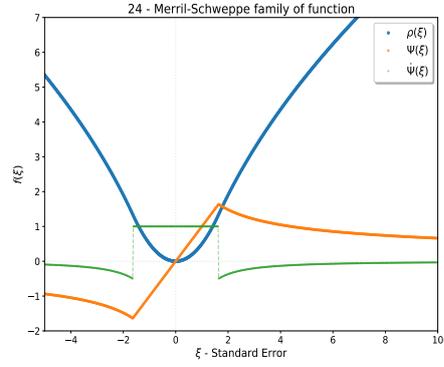


(f) 22- Kumar.

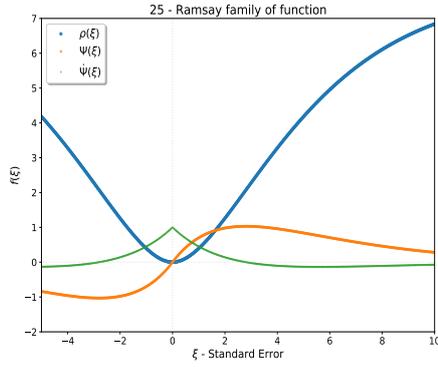
Figure 8: Families of analyzed functions for distinct estimators (ρ , ψ , ψ') – part 4/8.



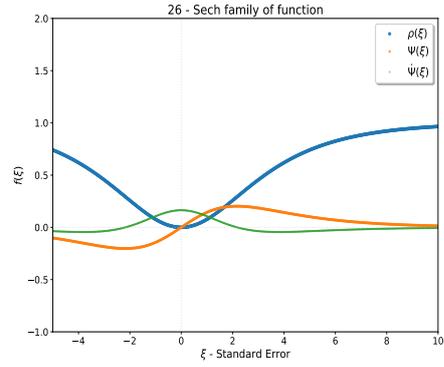
(a) 23- Lorentzian.



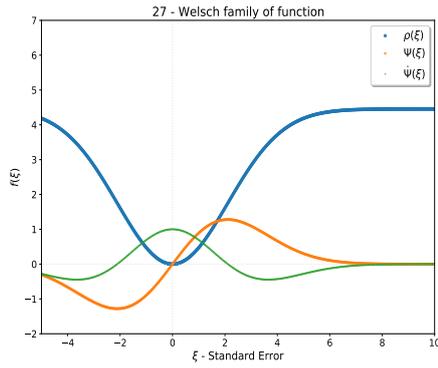
(b) 24- Merril-Scheppe.



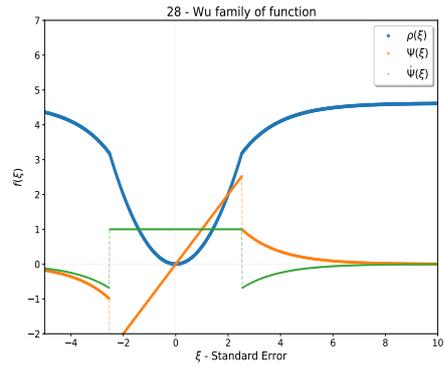
(c) 25- Ramsay.



(d) 26- Sech.

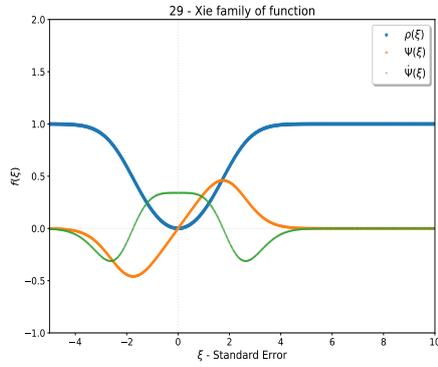


(e) 27- Welsh.

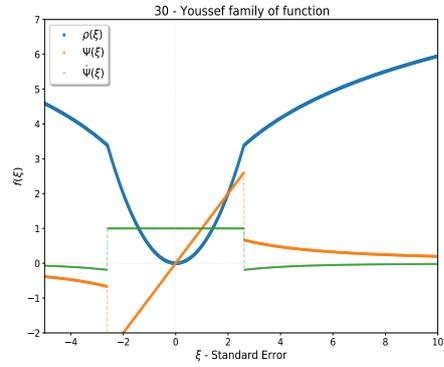


(f) 28- Wu.

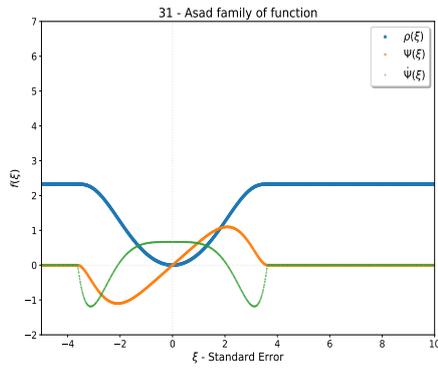
Figure 9: Families of analyzed functions for distinct estimators (ρ , ψ , ψ') – part 5/8.



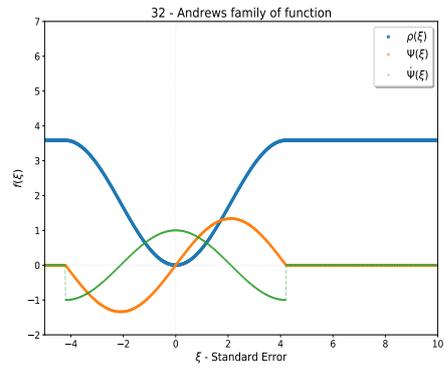
(a) 29- Xie.



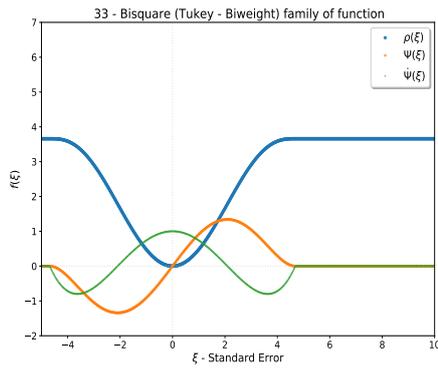
(b) 30- Youssef.



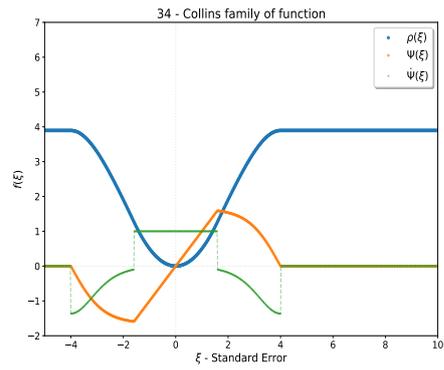
(c) 31- Asad.



(d) 32- Andrews.

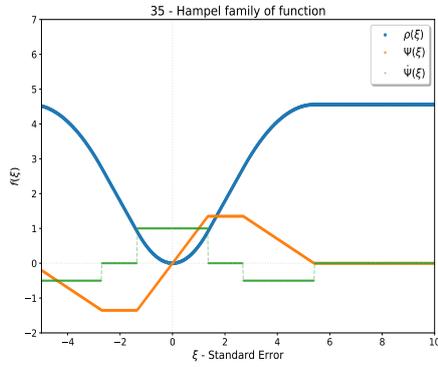


(e) 33- Biweight.

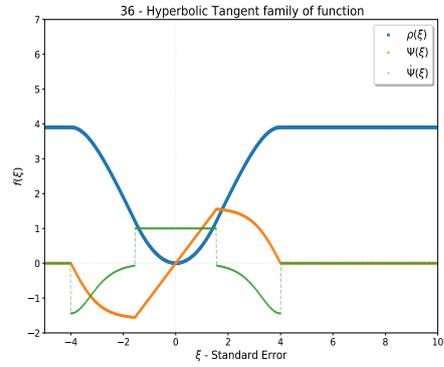


(f) 34- Collins.

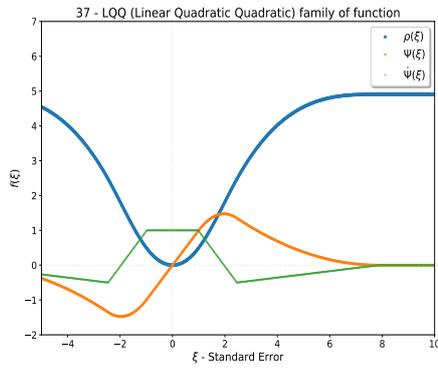
Figure 10: Families of analyzed functions for distinct estimators (ρ , ψ , ψ') – part 6/8.



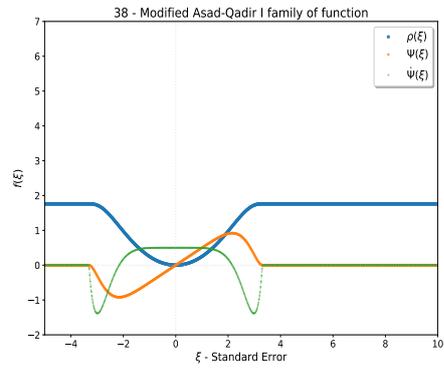
(a) 35- Hampel.



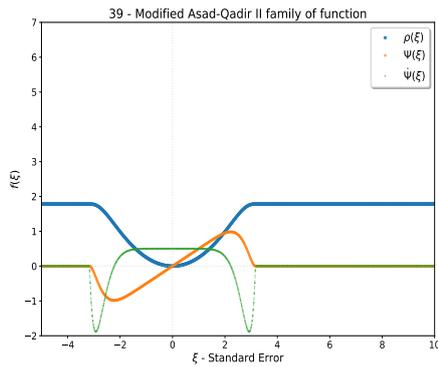
(b) 36- Hyperbolic Tangent.



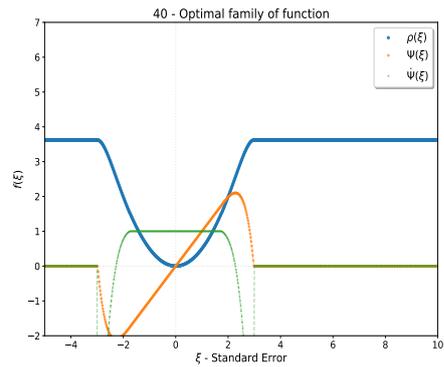
(c) 37- LQQ.



(d) 38- Modified Asad-Qadir I.

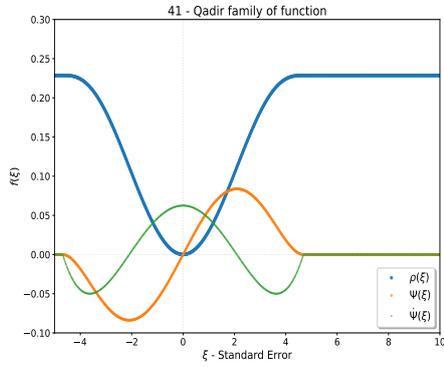


(e) 39- Modified Asad-Qadir II.

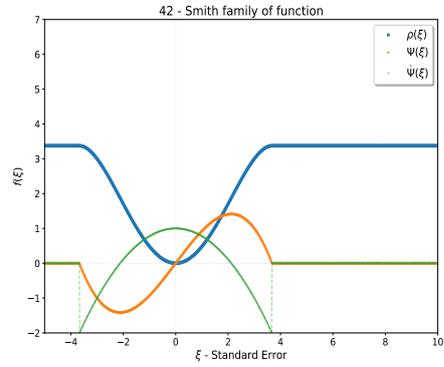


(f) 40- Optimal.

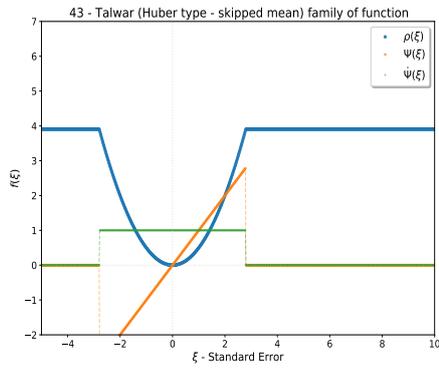
Figure 11: Families of analyzed functions for distinct estimators (ρ , ψ , ψ') – part 7/8.



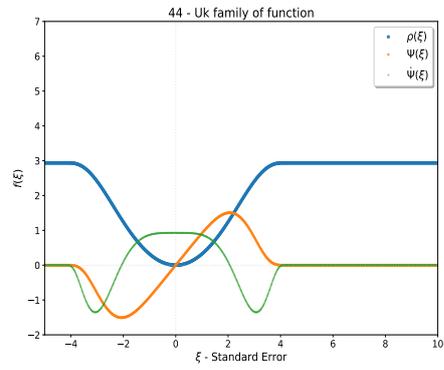
(a) 41- Qadir.



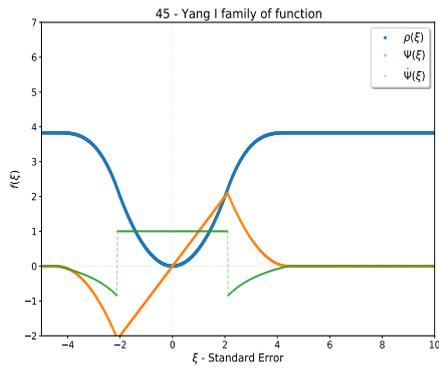
(b) 42- Smith.



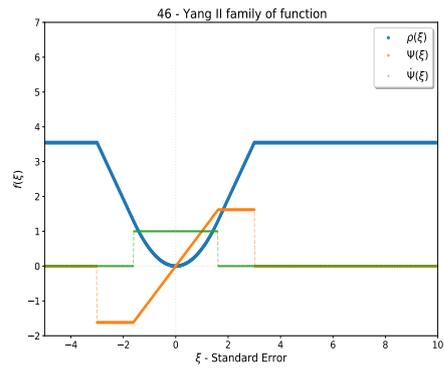
(c) 43- Talwar.



(d) 44- Uk.

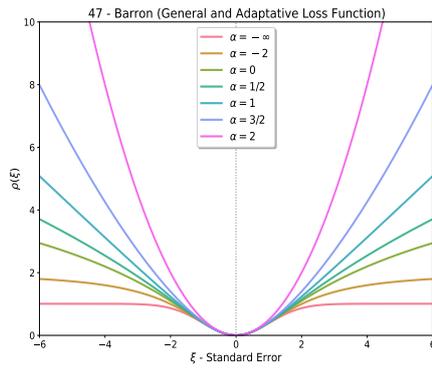


(e) 45- Yang I.

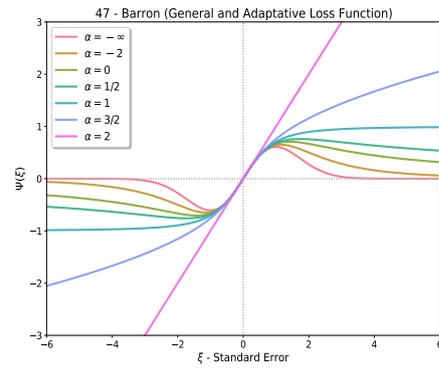


(f) 46- Yang II.

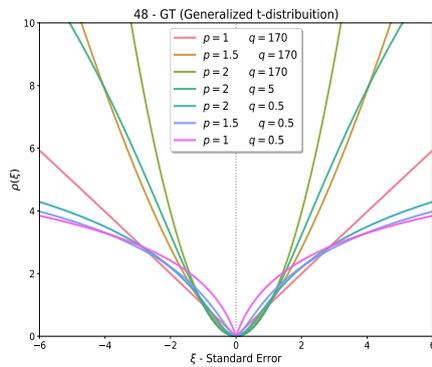
Figure 12: Families of analyzed functions for distinct estimators (ρ , ψ , ψ') – part 8/8.



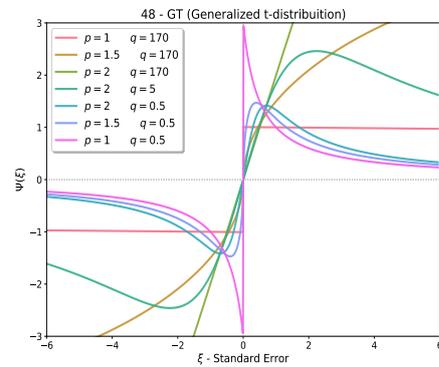
(a) 47- Barron - ρ functions.



(b) 47- Barron - ψ functions.



(c) 48- GT - ρ functions.



(d) 48- GT - ψ functions.

Figure 13: General and Adaptive Robust Estimators.

531 **5. The Tuning Parameters of the M-estimators**

532 Table 3 presents the values of the tuning parameters for the analyzed M-
 533 estimators for a given relative efficiency in respect to the Normal distribution.
 534 The references for these parameters and the introductory reference for each
 535 analyzed M-estimator are also presented.

Table 3: Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 1/5

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
0	Contaminated Normal	$\begin{cases} b_{CN} = 10 \\ p_{CN} = 0.235^{[1]} \\ p_{CN} = 0.26449^{[2]} \end{cases}$	[33] ^[1] ; This work ^[2]	[101]
1	LAV (L_1)	No tuning parameters. $E_{ff} = \pi/2 \approx 64\%$		[21]
2	$L_1 - L_2$	No tuning parameters. $E_{ff} = 95.75\%$		[25]
3	Fair	$c_F = 1.3998$	[7, 58, 33]	[51]
4	Huber	$c_{Hu} = 1.345$	[7, 58]	[11]
5	Modified Huber	$c_{MH} = 1.2107$	[7, 35]	[7]
6	Kong	$\begin{cases} c_k = 0.2^{[3]} \\ c_k = 2.4541^{[4]} \end{cases}$	[102] ^[3] ; This work ^[4] .	[102]
7	LnCosh	$c_{LC} = 0.83$	This work	[103]
8	Logistic	$c_L = 1.205$	[58]	[104]
9	Müller	$c_{Mu} = 2.28302$	This work	[18]
10	Zhang	$c_Z = 0.814$	This work	[83]

E_{ff} : [1]=95.53%; [2]=95.00%; [3]=74.03%; [4]=95.00%.

Table 3: Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 2/5

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
11	Alamgir	$c_{AI} = 2.37111$	This work	[105]
12	Bab-Hadiashar	$c_{BH} = 1.81206$	[106]	[106]
13	Bell	$c_{Bl} = 2.1522$	This work	[107]
14	Blake-Zisserman	$c_{BZ} = 0.00075$	This work	[108]
				Poisson
15	Cauchy	$c_C = 2.3849$	[7, 33]	(1824) (Apud [109])
16	Correntropy	$c_{Co} = 2.1105$	This work	[110]
17	Danish	$\begin{cases} c_{Da} = 2.76705 \\ D_{Da} = 3.91523 \end{cases}$	This work	[31]
18	Geman-McClure	$c_{GM} = 3.787376$	This work	[111]
19	GGW	$\begin{cases} a_{GGW} = 1.38636 \\ b_{GGW} = 1.50 \\ c_{GGW} = 1.06282 \end{cases}$	[112]	[112]
20	Insha	$c_I = 3.2296$	This work	[113]
21	Jin	$\begin{cases} a_J = 0.65 \\ c_J = 3.77312 \end{cases}$	This work	[84]

Table 3: Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 3/5

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
22	Kumar	$\begin{cases} a_{Ku} = 1.3860 \\ b_{Ku} = 2.7720 \\ c_{Ku} = 4.80284 \end{cases}$	This work	[114]
23	Lorentzian	$\begin{cases} c_{Lz} = 2.6^{[5]} \\ c_{Lz} = 2.678^{[6]} \end{cases}$	[33] ^[5] ; This work ^[6]	[66]
24	Merril- Schweppe	$c_{MS} = 1.637$	This work	[24]
25	Ramsay	$c_R = 0.3569$	This work	[115]
26	Sech	$c_{Sh} = 0.40497$	This work	[40]
27	Welsch	$c_W = 2.9846$	[7, 58]	[56]
28	Wu	$\begin{cases} a_{Wu} = 0.5 \\ c_{Wu} = 2.521 \end{cases}$	This work	[93]
29	Xie	$\begin{cases} c_X = 1.7134^{[7]} \\ c_X = 1.9598^{[8]} \end{cases}$	[61] ^[7] ; This work ^[8]	[61]
30	Youssef	$c_Y = 2.6006$	This work	[116]
31	Asad	$c_{As} = 3.6175$	This work	[117]
32	Andrews	$c_{An} = 1.338$	[7, 58]	[118]
33	Biweight	$c_{Bi} = 4.6851$	[7, 35]	[119]

E_{ff} : [5]=94.58%; [6]=95.00%; [7]=90.97%; [8]=95.00%.

Table 3: Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 4/5

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
34	Collins	$\left\{ \begin{array}{l} a_{Cl} = 1.6515 \\ r_{Cl} = 4.0 \\ q_{Cl} = 1.5908 \\ D_{Cl} = 3.8948 \end{array} \right.$	This work	[120]
35	Hampel	$\left\{ \begin{array}{l} a_H = 1.35 \\ b_H = 2.70 \\ c_H = 5.40 \end{array} \right.$	[33]	[118]
36	Hyperbolic Tangent	$\left\{ \begin{array}{l} A = 0.7261 \\ B = 0.8234 \\ k = 4.5 \\ r = 4.0 \\ q = 1.5562 \\ D = 3.9053 \end{array} \right.$	This work	[121]
37	LQQ	$\left\{ \begin{array}{l} b = 1.4734 \\ c = 0.9823 \\ s = 1.5 \\ a = 5.4025 \\ D_1 = -2.5802 \\ D_2 = 4.9047 \end{array} \right.$	[112]	[112]
38	Modified Asad-Qadir I	$a_{AQI} = 3.3094$	This work	[117]

Table 3: Tuning parameters for $E_{ff} = 95\%$ in respect to the Normal distribution - part 5/5

#	M-Estimators	Tuning Parameters	Reference (parameters)	Reference (estimator)
39	Modified Asad-Qadir II	$a_{AQII} = 3.1665$	This work	[117]
40	Optimal*	$\left\{ \begin{array}{l} a_{Op} = 0.3225 \\ b_{Op} = 3.0263 \\ c_{Op} = 8.2006e^{-4} \\ d_{Op} = 5.0349e^{-3} \\ e_{Op} = 1.8418e^{-3} \end{array} \right.$	This work	[122]
41	Qadir	$a_Q = 4.6851$	This work	[123]
42	Smith	$a_S = 3.6732$	This work	[57]
43	Talwar	$a_T = 2.7955$	[58]	[124]
44	Uk	$a_{Uk} = 4.1323$	This work	[125]
45	Yang I	$\left\{ \begin{array}{l} c_{YaI} = 2.3708 \\ a_{YaI} = 3.5 \\ D_{YaI} = 3.7026 \end{array} \right.$	This work	[126]
46	Yang II	$\left\{ \begin{array}{l} c_{YaII} = 1.3884 \\ a_{YaII} = 3.5 \end{array} \right.$	This work	[127]
47	Barron	Tuning not applied in this case. Shape parameters in Figure 14		[99]
48	GT	Tuning not applied in this case. Shape parameters in Figure 14		[100]

* polynomial approximation.

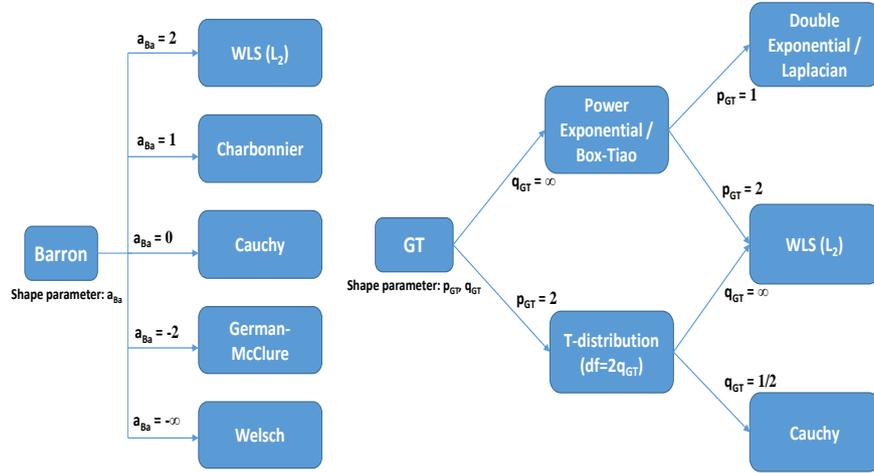


Figure 14: Shape parameter of Barron and GT M-estimators.

536 Looking at Figures 13 and 14, we can understand why these estimators
 537 are classified as General and Adaptive. The Barron estimator can take the
 538 shape of 5 different estimators, ranging from the non-robust shape of the WLS
 539 to the soft-re-descending Welsch, through the monotone Charbonnier ($L_1 - L_2$).
 540 The GT estimator can be adapted into 3 different estimators: LAV, WLS, and
 541 Cauchy.

542 As noted in Table 3, many of the M-estimators have two or more tuning
 543 parameters, so it is necessary to know the relationship among them for appro-
 544 priate tuning. It is important to remember that the desired properties presented
 545 by the ρ , ψ and ψ' functions are fundamental for correct tuning. An estimator
 546 that is not well aligned with the desired properties may be tuned incorrectly, so
 547 that the desired relationship among some of the parameters must be known and
 548 kept constant for correct tuning. The desired properties of the functions that
 549 characterize the analyzed M-estimators, such as the Gross-Error Sensitivity (γ)
 550 and the Rejection Point (r_p) (Figure 4), are fundamental for definition of the
 551 geometric properties of the considered function and calculation of its Relative
 552 Efficiency. Therefore, these properties may depend on one or more parameters,

553 particularly in the case of hard-re-descending M-estimators. In the present work,
 554 the Maple software version 15 was used for tuning the estimators. An illustrative
 555 example of the tuning procedure is detailed in Appendix A (analytic procedure
 556 for LAV estimator) and Appendix B (symbolic algebraic procedure), illustrat-
 557 ing the tuning of the Fair, Welsch, and Hampel estimators. For M-estimators
 558 with two or more parameters, the tuning was performed according to:

- 559 • Normal Contaminated: The constant b_{CN} represents the contamination
 560 of the variance by spurious values. The p_{CN} constant represents the prob-
 561 ability of contaminated distribution. They have no fixed relationship with
 562 each other. Therefore, the estimator can be tuned by fixing one of the
 563 constants, in the case of this work the parameter $b_{CN} = 10$ was fixed.

- 564 • Danish: The D_{Da} constant is an auxiliary constant, and depends on the
 565 tuning constant c_{Da} . The estimators that comprise two or more functions
 566 must be adjusted in order to remove any discontinuity, according to the
 567 desired properties. For that, it is necessary to add auxiliary constants to
 568 correct these discontinuities. For this estimator, the auxiliary constant
 569 follows the following relationship: $D_{Da} = \rho_1(c_{Da}) - \rho_2(c_{Da})$.

- 570 • GGW: This estimator has 3 tuning parameters, but no definition has been
 571 found for the relationships among these parameters. Thus, 2 parameters
 572 were selected for tuning. The c_{GGW} parameter is the parameter that limits
 573 two parts of the characteristic estimator functions and was manipulated
 574 to keep the shape of the function approximately the same. The parameter
 575 $b_{GGW} = 1.5$ is fixed and the parameter a_{GGW} is used for tuning. It is
 576 important to remember that this estimator does not present an analytical
 577 integral for the second part of the Influence Function (ψ_2). The graph of
 578 the ρ function of this estimator was built numerically.

- 579 • Jin: The parameter a_J was fixed at $a_J = 0.65$ as recommended by the
 580 authors of the original work. Therefore, the parameter c_J was used to
 581 tune the estimator.

- 582 • Kumar: This estimator presents 3 parameters and 3 parts of functions.
583 One of them is the auxiliary parameter to adjust the desired properties, the
584 parameter $D_{Ku} = \rho_2(b_{Ku}) - \rho_3(b_{Ku})$, which is a function of the parameters
585 a_{Ku} and b_{Ku} . The a_{Ku} parameter is related to the b_{Ku} parameter in order
586 to keep the geometrical properties of the function, defined in the present
587 work in the form $a_{Ku} = b_{Ku}/2$.

- 588 • Wu: This estimator presents 2 tuning parameters, the parameters a_{Wu}
589 and c_{Wu} . $0 < a_{Wu} < 1$ according to the reference work. Therefore,
590 $a_{Wu} = 0.5$ was set and the parameter c_{Wu} was used for tuning.

- 591 • Collins: This estimator presents 4 parameters and 3 parts of functions.
592 The D_{Cl} parameter is the auxiliary parameter, with $D_{Cl} = \rho_1(q_{Cl}) -$
593 $\rho_2(q_{Cl})$. The q_{Cl} parameter represents the gross-error sensitivity, in addi-
594 tion to adjusting the continuity of the estimator. The calculation of this
595 parameter constitutes an intermediate step within the tuning procedure,
596 where q_{Cl} is the positive root of the equation $(\psi_2 - \psi_1 = 0)$, ψ_i being the
597 i^{th} part of influence function. The r_{Cl} parameter (which represents the
598 Rejection Point (r_p) of the estimator) must follow the geometric propor-
599 tionality, varying according to the Relative Efficiency and fixed for tuning.
600 Thus, for this work the tuning parameter was a_{Cl} .

- 601 • Hampel: This estimator presents 3 parameters and 4 parts of functions.
602 The 3 parameters (a_H, b_H, c_H) have a geometric relationship widely used
603 in the literature, namely: $c_H = 4a_H$ and $b_H = 2a_H$. The parameters must
604 follow the relation $a_H < b_H < c_H$ to satisfy the geometric proportional-
605 ity, therefore the relation most used in the literature and in this work is
606 adequate. Thus, for this work the tuning parameter used was a_H .

- 607 • Hyperbolic Tangent: This estimator presents 6 parameters and 3 parts
608 of functions. The D parameter is the auxiliary parameter, with $D =$
609 $\rho_1(q) - \rho_2(q)$. The q parameter represents the gross-error sensitivity, in
610 addition to adjusting the continuity of the estimator. The calculation of

611 this parameter constitutes an intermediate step within the tuning proce-
612 dure, where q is the positive root of the equation $(\psi_2 - \psi_1 = 0)$, ψ_i being
613 the i^{th} part of influence function. The parameter r (which represents the
614 point of rejection of the estimator) must follow the geometric proportional-
615 ity, varying according to the Relative Efficiency and fixed for a particular
616 tuning. However, parameter r is related to parameter k , where $r \leq k$.
617 Parameters A and B must obey the inequalities $0 < A < B < 1$. Thus,
618 fixing k and B , and respecting the proposed relationships, only parameter
619 A is left for tuning.

- LQQ: This estimator has 6 parameters and 4 parts of functions. The 3 auxiliary parameters are a , D_1 and D_2 , being:

$$a = \frac{bs - 2b - 2c}{1 - s} \quad (17)$$

$$D_1 = \frac{(-b^3 + (-a - 3c)b^2 - 3c(a + c)b - 3ac^2 - c^3)s + (c + b)^3}{6a} \quad (18)$$

$$D_2 = -\frac{1}{3}a^2s - \frac{1}{2}abs + ac + bc - \frac{1}{6}b^2s + \frac{1}{2}c^2 + \frac{1}{3}a^2 + ab + \frac{1}{2}b^2 \quad (19)$$

620 The parameter s is fixed at $s = 1.5$, where the geometric relationship
621 of the parameter is $s = 1 - \inf(\psi'(\xi))$. Parameter c follows geometric
622 proportionality, varying according to Relative Efficiency, as it is associated
623 with sensitivity to gross error. Therefore, the tuning parameter is b .

- Optimal: This estimator has 5 parameters and 3 parts of function. The
624 5 parameters are the parameters of the polynomial interpolation of the
625 original estimator, because the analytical function of the estimator was
626 not found. The 5 parameters must follow the relationships that suit the
627 desired properties of the characteristic functions of the M-estimator, such
628 as continuity. This estimator can be tuned iteratively. Initially, c_{Op} and
629 d_{Op} are estimated at $c_{Op} = 2e^{-4}$ and $d_{Op} = 2e^{-3}$. The a_{Op} and b_{Op}
630 parameters depend on the other parameters through the following rela-
631 tionships:
632

1. a_{Op} is the smallest positive root of the equation $(\psi_2 - \psi_1 = 0)$, ψ_i

634 being the i^{th} part of influence function;

635 2. b_{Op} is the positive root of the equation ($\psi_2 = 0$), ψ_i being the i^{th}
636 part of influence function.

637 Thus, the parameters c_{Op} and d_{Op} must be readjusted, because they de-
638 pend on a_{Op} and b_{Op} . The adjustments follow the relation:

639 1. $c_{Op} = \frac{7a_{Op}^6 e_{Op} - 5a_{Op}^4 d_{Op}}{-3a_{Op}^2}$;

640 2. $d_{Op} = \frac{b_{Op}^7 e_{Op} + b_{Op}^3 c_{Op} - b_{Op}}{b_{Op}^5}$.

641 Thus, the only possible parameter for tuning is the e_{Op} parameter.

642 • Yang I: This estimator presents 3 parameters and 3 parts of functions.
643 The D_{YaI} parameter is the auxiliary parameter, with $D_{YaI} = \rho_1(c_{YaI}) -$
644 $\rho_2(c_{YaI})$. The parameter a_{YaI} (which represents the point of rejection of
645 the estimator) must follow the geometric proportionality, varying accord-
646 ing to the Relative Efficiency and fixed for a tuning. Thus, for this work
647 the tuning parameter used was c_{YaI} .

648 • Yang II: This estimator presents 2 parameters and 3 parts of functions.
649 The parameter a_{YaII} (which represents the point of rejection of the es-
650 timator) must follow the geometric proportionality, varying according to
651 the Relative Efficiency and fixed for a tuning. Thus, for this work the
652 tuning parameter used was c_{YaII} .

653 Additional information regarding the tuning constants for each estimator
654 can be found in the respective original works, as shown in Table 3. Table D.7
655 in Appendix D shows the tunings for the Relative Efficiencies of 90, 98 and
656 99%, in respect to the Normal distribution. Most of these tunings are proposed
657 here for the first time, especially for the 98 and 99% efficiencies. All parameters
658 presented in Table D.7 were tuned in the present work, even when they were
659 available in the literature, to assure the consistency of the reported data.

660 It can be seen in Figure 15 that the ρ functions (objective function) tend
661 to become more convex when the relative efficiency increases for the estimators:
662 (i) Contaminated Normal (Figure 15a) - quasi-robust; (ii) Fair (Figure 15b) -
663 monotonous; (iii) Welsch (Figure 15c) - soft-re-descending; and (iv) Smith (Fig-
664 ure 15d) - hard-re-descending. This is because these estimators were tuned in
665 respect to the Normal distribution, which is strictly convex. Notably, the Con-
666 taminated Normal (quasi-robust) estimator is more convex at the 99% efficiency
667 level than at the 90% efficiency level (with $\xi = 3$), as its structure is based on
668 the Normal distribution. It can also be observed that the robustness decreases
669 when the relative efficiency increases, as indicated by the ψ functions (influence
670 function). Again, this is because these estimators were tuned in respect to the
671 Normal distribution (which is not robust). This is particularly evident in the
672 case of the Smith estimator (hard-re-descending and non-convex), as the rejection
673 point (r_p) increases at high efficiency levels; that is, it takes longer (in terms
674 of the magnitude of the standard deviation) to remove the effect of a spurious
675 value (less robust) in this case. Otherwise, the rejection point decreases at low
676 efficiency levels; that is, the estimator descends steeply to nullify the effect of a
677 spurious value, so that the estimator becomes more robust and less similar to
678 the Normal distribution. For example, selecting a spurious value of magnitude
679 $\xi = 4$, this estimator nullifies the negative effect on the estimates for the levels
680 of 90-95% of relative efficiency, which does not occur for the 98-99% efficiency
681 levels.

682 Therefore, as previously shown in Table 1 (which presents 37 M-estimator
683 applications in the chemical engineering field) and based on the results reported
684 in these works, it becomes possible to say that soft-re-descending M-estimators
685 described by exponential terms (without “if” clause) have received more at-
686 tention from researchers so far. Besides, as reported by those authors, these
687 estimators have presented the best overall performances in comparative studies
688 (Welsch - Prata et al. [81]; Correntropy - Chen et al. [85]; Correntropy/Bi-
689 weight - Llanos et al. [88, 94]; Welsch/Correntropy/Bell - da Cunha et al. [91];

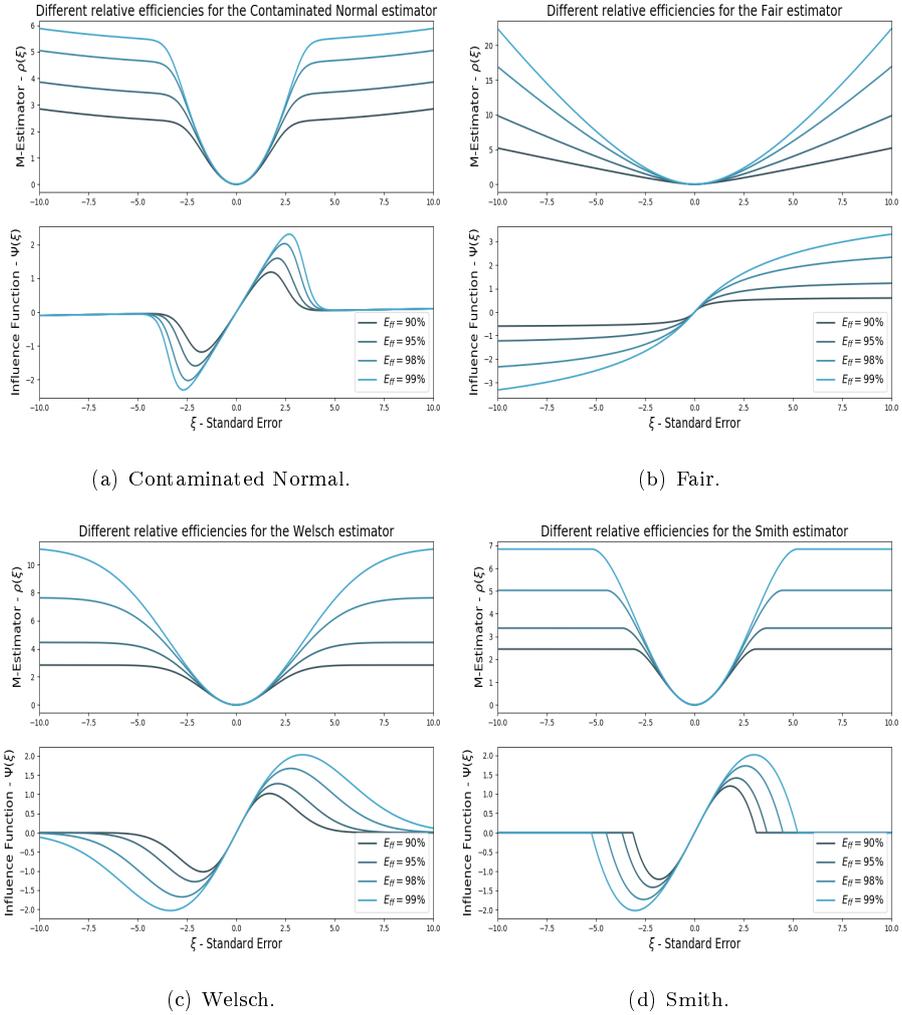


Figure 15: Characteristic functions of different estimators for distinct relative efficiency levels.

690 Xie/Welsch - Xie et al. [61]). Additionally, these estimators also performed well
 691 when the data sets were contaminated with random errors (at 95% level of ef-
 692 ficiency in respect to the Normal distribution). As shown in Table 4 (which
 693 presents a summary of the properties and characteristics of the M-estimators
 694 presented in this work), the use of these estimators is also attractive because
 695 the respective IF's (ψ) are continuous and redescending, nullifying the negative

696 effects of spurious values over the estimates. Besides, the derivatives of the IF's
697 (ψ') are also continuous in these cases and do not require the use of smoothing
698 functions when deterministic optimizations are carried out [60, 72]. In addition,
699 these estimators depend on a single tuning parameter, which can be useful for
700 actual real time applications, when it is desired to adjust the tuning constant
701 appropriately to data sets that may contain gross error frequencies that vary
702 with the operation conditions [36, 37]. However, although many redescending
703 robust M-estimators are available and have been used so far, characterization of
704 relative advantages through detailed comparative studies involving benchmark-
705 ing problems (particularly when nonlinear constraints and dynamic problems
706 are considered) have not been developed yet. This indicates that special atten-
707 tion must be given to the redescending M-estimators presented in the present
708 work and to the main features that may encourage the development of new
709 ones in the near future (based on the desired properties that are revised and
710 discussed here).

Table 4: Summary of properties and characteristics of analyzed robust M-estimators

	Monotonous		Soft-redescending		Hard-redescending
	Without "if" clause	With "if" clause	Without "if" clause	With "if" clause	Always with "if" clause
Estimator (#)	#1, 2, 3, 6, 7, 8, 10	#4, 5, 9	#11, 13, 14, 15, 16, 18, 20, 23, 25, 26, 27, 29	#12, 17, 19, 21, 22, 24, 28, 30	#31 - 46
Function (ρ)	Convex		pseudo-convex (except #19 does not exist ρ form; it was built from ψ function)		quasi-convex
IF (ψ) - gross errors influence	moderately increases and asymptote to a constant value (except #1)		asymptote to zero		is exactly zero
Rejection Point (r_P) in ψ	No		Yes (can be approximated)		Yes (exactly defined by the last parameter)
ψ discontinuous (Figures 5 to 12)	#1	#9	Continuous	#17, 28, 30	#43, 46
ψ' discontinuous (Figures 5 to 12)	#1	#4, 9	Continuous	#12, 17, 21, 22, 24, 28, 30	#32, 34, 35, 36, 42, 43, 45, 46
Tuning parameters	Just one, Except #1, 2	Just one	Just one	One, two (#17, 21, 28), three (#22)	One, two (#46), three (#35), four (#34), five(#40), six(#36, 37)
Difficulty of tuning	Easy	Easy, Moderate (#9)	Easy/Difficult (#19)	Easy, Moderate (#17, 21, 28, 30), Difficult (#22)	Easy, Moderate (#43, 35, 46), Difficult (#34), Hard (#40, 36, 37)

711 6. Conclusions

712 Although robust M-estimators have been applied in many fields of knowl-
713 edge, including the fields of pure and applied statistics, and outperform the
714 classical WLS estimator, robust regression methods have not been widely used
715 yet. Several reasons may possibly help to understand this point:

- 716 • Robust estimation is computationally more intense than classical method
717 [59]; however, due to the increasing computation capacity this possible
718 reason is becoming less pertinent;
- 719 • For many robust estimators, especially those classified as redescending,
720 there is a need for use of global optimization methods, as some of these
721 estimators are pseudo-convex or quasi-convex (both generally called “non-
722 convex” in literature) [54, 3, 36];
- 723 • The vast majority of popular statistical computational packages have not
724 yet introduced robust methods [128];
- 725 • The opinion of many statisticians and researchers that classical statistical
726 methods are robust enough (although they are not).

727 Most of the present review was devoted to presentation and use of M-
728 estimators, because even though some applications of these estimators are fo-
729 cused specifically on robust statistics, electrical/electronic engineering, and sig-
730 nal processing, other areas of knowledge have also developed, studied and found
731 use for M-estimators, particularly for solution of regression problems in the
732 field of chemical engineering. For this reason, the present review presents for
733 the first time a compilation of 50 (48 robust) estimators, with the respective
734 characteristic functions (the ρ function, the ψ function and the ψ' function) and
735 tunable parameters for efficiency levels of 90%, 95%, 98% and 99% in respect
736 to the normal distribution. Importantly, many of these estimators had never
737 been presented previously in graphical form, making more difficult the correct

738 implementation and validation of the estimators. Besides, many of the analyzed
739 estimators had been used few times (sometimes a single time), showing the im-
740 portance of tuning and introducing them to the scientific communities. As a
741 whole, this paper presents a collection of robust M-estimators that are spread
742 across several areas of science that study data regression problems. Thus, the
743 present work fills a gap and allows the knowledge exchange among these di-
744 verse scientific communities, providing a a large and well-documented number
745 of case studies, results, and conclusions about the use of robust M-estimators
746 are presented. Finally, it should be noted through Table 1, that the Contam-
747 inated Normal (quasi-robust), Welsch, Hampel, Fair, Lorentzian, Correntropy,
748 and Cauchy M-estimators were the most used for regression analysis in chem-
749 ical engineering problems, showing that the preference in this area represents
750 a reduced set in the universe of M-estimators, such as the 49 M-estimators
751 (Contaminated Normal + 48 robust estimators) presented in this work.

752 **Appendix A. Relative Efficiency of LAV Estimator**

753 To illustrate the calculation of Relative Efficiency (Eff), the M-LAV esti-
 754 mator is selected because it does not contain any tuning constant. This estima-
 755 tor has an efficiency equal to $2/\pi$, approximately equal to 64% (Huber, 1981)
 756 in respect to the Normal distribution. Based on the following settings:

$$f_{Normal} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) \quad (\text{A.1})$$

$$f'_{Normal} = -\xi \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) \quad (\text{A.2})$$

$$\psi_{Normal}(\xi) = \frac{d\rho(\xi)}{d\xi} = \frac{d}{d\xi} \left(\frac{1}{2}\xi^2\right) = \xi \quad (\text{A.3})$$

$$\psi_{LAV}(\xi) = \text{sgn}(\xi) \quad (\text{A.4})$$

757 and using Equations (14) and (16), it is possible to calculate the Efficiency of
 758 the LAV M-estimator in respect to to the Normal distribution, as presented in
 759 Equation (A.5) through (A.8).

$$E_{ff}[\psi, f]_{LAV} = \frac{V_f[\psi_{Normal}, f_{Normal}]}{V[\psi_{LAV}, f_{Normal}]} \quad (\text{A.5})$$

$$E_{ff}[\psi, f]_{LAV} = \frac{\int_{-\infty}^{+\infty} \psi_{Normal}^2(\xi) f_{Normal}(\xi) d\xi}{\left[- \int_{-\infty}^{+\infty} \psi_{Normal}(\xi) f'_{Normal}(\xi) d\xi \right]^2} \quad (A.6)$$

$$\frac{\int_{-\infty}^{+\infty} \psi_{LAV}^2(\xi) f_{Normal}(\xi) d\xi}{\left[- \int_{-\infty}^{+\infty} \psi_{LAV}(\xi) f'_{Normal}(\xi) d\xi \right]^2}$$

$$E_{ff}[\psi, f]_{LAV} = \frac{\frac{1}{\sigma\sqrt{2\pi}} \int_0^{+\infty} \xi^2 \exp(-\frac{1}{2}\xi^2) d\xi}{\left(\frac{1}{\sigma\sqrt{2\pi}} \right)^2 \left[\int_0^{+\infty} \xi^2 \exp(-\frac{1}{2}\xi^2) d\xi \right]^2} \quad (A.7)$$

$$\frac{\frac{1}{\sigma\sqrt{2\pi}} sgn^2(\xi) \int_0^{+\infty} \exp(-\frac{1}{2}\xi^2) d\xi}{\left(\frac{1}{\sigma\sqrt{2\pi}} \right)^2 sgn^2(\xi) \left[\int_0^{+\infty} \xi \exp(-\frac{1}{2}\xi^2) d\xi \right]^2}$$

$$E_{ff}[\psi, f]_{LAV} = \frac{\frac{\sqrt{2\pi}}{2}}{\left[\frac{\sqrt{2\pi}}{2} \right]^2} = \frac{1}{\frac{\sqrt{2\pi}}{2}} = \frac{1}{\sqrt{2\pi}} \frac{1}{\frac{\sqrt{2\pi}}{2}} = \frac{4}{2\pi} = \frac{2}{\pi} \quad (A.8)$$

$$\frac{\frac{2}{[1]^2}}{\frac{2}{2}}$$

760 It is important to observe that this is the basis for calculation of tuning
761 constants of all estimators. The relative efficiency (90, 95, 98 or 99%) and
762 the reference distribution (usually Normal distribution) are fixed. Numerical
763 methods can be used to solve the resulting equation.

Fair

Tuning

```

> restart
> Tuning := proc( cF )
  local  $\psi^N, f, f2, Vf, \psi^F, V, Eff$ ;


$$\psi^N := \xi: f := \frac{e^{-\frac{1}{2}\xi^2}}{\sqrt{2\cdot\pi}}: f2 := \frac{d}{d\xi} f: Vf := \frac{2\cdot\int_0^{+\infty} \psi^N{}^2 \cdot f d\xi}{\left(2\cdot\int_0^{+\infty} \psi^N \cdot \xi \cdot f d\xi\right)^2}:$$



$$\psi^F := \frac{cF \cdot \xi}{cF + |\xi|}:$$



$$V := \text{evalf}\left(\frac{2\cdot\left(\int_0^{+\infty} \psi^F{}^2 \cdot f d\xi\right)}{\left(-2\cdot\left(\int_0^{+\infty} \psi^F \cdot f2 d\xi\right)\right)^2}\right):$$



$$Eff := \frac{Vf}{V}:$$


  return Eff; end proc;
> #Digits:=30:
> f := simplify(evalf(ℜ(Tuning(cF))), 'assume = positive'):
> cF := fsolve(f - 0.95, cF, 0..4)
cF := 1.399776831
> Tuning(cF)
0.9500000000

```

Figure B.16: Tuning procedure for the Fair estimator.

Welsch

Tuning

```

> restart
> Tuning := proc(cW)
  local psiN, f, f2, Vf, psi, V, Eff;

  psiN := xi: f :=  $\frac{e^{-\frac{1}{2}\xi^2}}{\sqrt{2\pi}}$ : f2 :=  $\frac{d}{d\xi} f$ : Vf :=  $\frac{2 \cdot \int_0^{+\infty} \psi N^2 \cdot f d\xi}{\left(2 \cdot \int_0^{+\infty} \psi N \cdot \xi \cdot f d\xi\right)^2}$ :

  psi :=  $\xi e^{-\left(\frac{\xi}{cW}\right)^2}$ :
  V := evalf  $\left( \frac{2 \cdot \left(\int_0^{+\infty} \psi^2 \cdot f d\xi\right)}{\left(-2 \cdot \left(\int_0^{+\infty} \psi \cdot f2 d\xi\right)\right)^2} \right)$ :

  Eff :=  $\frac{Vf}{V}$ :
  return Eff; end proc:
> #Digits := 30:
> f := simplify(evalf(Tuning(cW)), 'assume = real'):
> cW := fsolve(f - 0.95, cW, 0..5)
cW := 2.984637176
> Tuning(cW)
0.9500000000

```

Figure B.17: Tuning procedure for the Welsch estimator.

Hampel

Tuning

```

> restart
> Tuning := proc(aH, bH, cH)
  local psiN, f, f2, Vf, psi, V, Eff;

  psiN := xi: f :=  $\frac{e^{-\frac{1}{2}\xi^2}}{\sqrt{2\cdot\pi}}$ : f2 :=  $\frac{d}{d\xi} f$ : Vf :=  $\frac{2\cdot\int_0^{+\infty} \psi N^2 \cdot f d\xi}{\left(2\cdot\int_0^{+\infty} \psi N \cdot \xi \cdot f d\xi\right)^2}$ :

  psi := piecewise(|xi| < aH, xi, aH < |xi| and |xi| < bH, aH·signum(xi), bH < |xi| and |xi| < cH,
    aH  $\left(\frac{cH - |xi|}{cH - bH}\right)$ ·signum(xi), cH < |xi|, 0):

  V := evalf  $\left(\frac{2\cdot\left(\int_0^{+\infty} \psi^2 \cdot f d\xi\right)}{\left(-2\cdot\left(\int_0^{+\infty} \psi \cdot f2 d\xi\right)\right)^2}\right)$ :

  Eff :=  $\frac{Vf}{V}$ :

  return Eff: end proc:
> #Digits:=30:
> f := simplify(evalf(ℜ(Tuning(aH, 2·aH, 4·aH))), 'assume = real'):
> aH := fsolve(f - 0.90, aH, 1..3)
aH := 1.105233041
> Tuning(aH, 2·aH, 4·aH)
0.9000000009

```

Figure B.18: Tuning procedure for the Hampel estimator.

Table C.5: M-Estimators: $\rho(\xi)$ - part 1/5

#	$\rho(\xi)$	Range
1	$ \xi $	$ \xi \leq \infty$
2	$2 \left(\sqrt{1 + \frac{\xi^2}{2}} - 1 \right)$	$ \xi \leq \infty$
3	$c_F^2 \left[\frac{ \xi }{c_F} - \ln \left(1 + \frac{ \xi }{c_F} \right) \right]$	$ \xi \leq \infty$
4	$\begin{cases} \frac{\xi^2}{2} \\ c_{Hu} \xi - \frac{c_{Hu}^2}{2} \end{cases}$	$ \xi \leq c_{Hu}$ $ \xi > c_{Hu}$
5	$\begin{cases} c_{MH}^2 \left[1 - \cos \left(\frac{\xi}{c_{MH}} \right) \right] \\ c_{MH} \xi + c_{MH}^2 \left(1 - \frac{\pi}{2} \right) \end{cases}$	$ \xi \leq \frac{\pi}{2} c_{MH}$ $ \xi > \frac{\pi}{2} c_{MH}$
6	$\frac{\xi^2}{c_K + \xi }$	$ \xi \leq \infty$
7	$\frac{\ln(\cosh(c_{LC}\xi))}{c_{LC}}$	$ \xi \leq \infty$
8	$c_L^2 \ln \left[\cosh \left(\frac{\xi}{c_L} \right) \right]$	$ \xi \leq \infty$
9	$\begin{cases} \frac{\xi^2}{2} \\ \frac{c_M \xi }{2} \end{cases}$	$ \xi \leq c_M$ $ \xi > c_M$
10	$\frac{\xi^2}{2 + c_Z \xi }$	$ \xi \leq \infty$
11	$\frac{4c_{Al}^2 \left[1 + 3 \exp \left(-\frac{\xi^2}{c_{Al}^2} \right) \right]}{3 \left[1 + \exp \left(-\frac{\xi^2}{c_{Al}^2} \right) \right]^3} - \frac{2c_{Al}^2}{3}$	$ \xi \leq \infty$
12	$\begin{cases} \frac{\xi^2}{2} \\ c_{BH}^2 \left[\ln \left(\frac{ \xi }{c_{BH}} \right) + \frac{1}{2} \right] \end{cases}$	$ \xi \leq c_{BH}$ $ \xi > c_{BH}$
13	$\frac{5c_{Be}^2}{4} \left[1 - \left(1 + \frac{\xi^2}{5c_{Be}^2} \right)^{-2} \right]$	$ \xi \leq \infty$
14	$\log \left[\frac{1 + c_{BZ}}{\exp(-\xi^2) + c_{BZ}} \right]$	$ \xi \leq \infty$
15	$\frac{c_C^2}{2} \ln \left(1 + \frac{\xi^2}{c_C^2} \right)$	$ \xi \leq \infty$

Table C.5: M-Estimators: $\rho(\xi)$ - part 2/5

#	$\rho(\xi)$	Range
16	$\frac{1}{c_{Co}\sqrt{2\pi}} \left[1 - \exp\left(-\frac{\xi^2}{2c_{Co}^2}\right) \right]$	$ \xi \leq \infty$
17	$\begin{cases} \frac{\xi^2}{2} \\ -\frac{c_{Da}}{2} \exp\left(\frac{-\xi^2}{c_{Da}}\right) + D_{Da} \end{cases}$	$ \xi \leq c_{Da}$ $ \xi > c_{Da}$
18	$\frac{\xi^2}{c_{GM}^2 + \xi^2}$	$ \xi \leq \infty$
19	*There is no analytical integral of the influence function	
20	$\frac{c_I^2}{4} \left[\arctan\left(\frac{\xi^2}{c_I^2}\right) + \frac{c_I^2 \xi^2}{c_I^4 + \xi^4} \right]$	$ \xi \leq \infty$
21	$\begin{cases} \frac{c_J^2}{6} \left[1 - \left(1 - a_J \frac{\xi^2}{c_J^2}\right)^3 \right] \\ \frac{c_J^2 a_J}{6} \left[3(1 - a_J)^2 \ln\left(\frac{\xi^2}{c_J^2}\right) + a_J^2 - 3a_J + 3 \right] \end{cases}$	$ \xi \leq c_J$ $ \xi > c_J$
22	$\begin{cases} \frac{\xi^2}{2} \\ a_{Ku} \xi - \frac{a_{Ku}^2}{2} \\ -\frac{a_{Ku} b_{Ku}}{2} \exp\left(1 - \frac{\xi^2}{b_{Ku}^2}\right) + D_{Ku} \end{cases}$	$ \xi \leq a_{Ku}$ $a_{Ku} < \xi \leq b_{Ku}$ $ \xi > b_{Ku}$
23	$1 - \frac{1}{1 + \left(\frac{\xi^2}{2c_{Lz}^2}\right)}$	$ \xi \leq \infty$
24	$\begin{cases} \frac{\xi^2}{2} \\ 2c_{MS}^{3/2} \sqrt{ \xi } - 1.5c_{MS}^2 \end{cases}$	$ \xi \leq c_{MS}$ $ \xi > c_{MS}$
25	$\frac{1}{c_R^2} [1 - (1 + c_R \xi) \exp(-c_R \xi)]$	$ \xi \leq \infty$
26	$1 - \operatorname{sech}(c_{Sh}\xi)$	$ \xi \leq \infty$
27	$\frac{c_W^2}{2} \left[1 - \exp\left(-\frac{\xi^2}{c_W^2}\right) \right]$	$ \xi \leq \infty$
28	$\begin{cases} \frac{\xi^2}{2} \\ \frac{a_{Wu}^{ \xi - c_{Wu}}}{\ln(a_{Wu})} - \frac{1}{\ln(a_{Wu})} + \frac{c_{Wu}^2}{2} \end{cases}$	$ \xi \leq c_{Wu}$ $ \xi > c_{Wu}$

Table C.5: M-Estimators: $\rho(\xi)$ - part 3/5

#	$\rho(\xi)$	Range
29	$\frac{1 - \exp\left(-\frac{\xi^2}{c_X^2}\right)}{1 + \exp\left(-\frac{\xi^2}{c_X^2}\right)}$	$ \xi \leq \infty$
30	$\begin{cases} \frac{\xi^2}{2} \\ \ln(\xi^2 + 1) + \left[\frac{c_Y^2}{2} - \ln(c_Y^2 + 1)\right] \end{cases}$	$ \xi \leq c_Y$ $ \xi > c_Y$
31	$\begin{cases} \frac{\xi^2}{45c_{As}^8} (3\xi^8 - 10c_{As}^4\xi^4 + 15c_{As}^8) \\ \frac{8c_{As}^2}{45} \end{cases}$	$ \xi \leq c_{As}$ $ \xi > c_{As}$
32	$\begin{cases} c_{An}^2 \left[1 - \cos\left(\frac{\xi}{c_{An}}\right)\right] \\ 2c_{An}^2 \end{cases}$	$ \xi \leq \pi c_{An}$ $ \xi > \pi c_{An}$
33	$\begin{cases} \frac{c_{Bi}^2}{6} \left[1 - \left(1 - \frac{\xi^2}{c_{Bi}^2}\right)^3\right] \\ \frac{c_{Bi}^2}{6} \end{cases}$	$ \xi \leq c_{An}$ $ \xi > c_{An}$
34	$\begin{cases} \frac{\xi^2}{2} \\ -2 \left\{ \log \left[\cosh \left(\frac{a_{Cl}}{2} (\xi - r_{Cl}) \right) \right] \right\} + d_{Cl} \\ d_{Cl} \end{cases}$	$ \xi \leq q_{Cl}$ $q_{Cl} < \xi \leq r_{Cl}$ $ \xi > r_{Cl}$
35	$\begin{cases} \frac{\xi^2}{2} \\ a_H \xi - \frac{a_H^2}{2} \\ a_H b_H - \frac{a_H^2}{2} + \frac{a_H(c_H - b_H)}{2} \left[1 - \left(\frac{c_H - \xi }{c_H - b_H} \right)^2 \right] \\ a_H b_H - \frac{a_H^2}{2} + \frac{a_H(c_H - b_H)}{2} \end{cases}$	$ \xi \leq a_H$ $a_H < \xi \leq b_H$ $b_H < \xi \leq c_H$ $ \xi > c_H$
36	$\begin{cases} \frac{\xi^2}{2} \\ -\frac{2A}{B} \left\{ \log \left[\cosh \left(\frac{(k-1)^{1/2} B}{2A^{1/2}} (\xi - r) \right) \right] \right\} + D \\ D \end{cases}$	$ \xi \leq q$ $q < \xi \leq r$ $ \xi > r$

Table C.5: M-Estimators: $\rho(\xi)$ - part 4/5

#	$\rho(\xi)$	Range
37	$\frac{\xi^2}{2}$	$ \xi \leq c$
	$\frac{1}{6b} [3b\xi^2 + s(c^3 - 3c^2 \xi + 3c\xi^2 - \xi ^3)]$	$c < \xi \leq c + b$
	$\frac{1}{6a} \{ (s-1) \xi ^3 - 3(s-1)(s_2)\xi^2 +$ $+ [(3b^2 + (3a+6c)b + 6ac + 3c^2)s +$ $-3(c+b)^2] \xi \} + D_1$	$c + b < \xi \leq s_2$
	D_2	$ \xi > s_2$
where: $s_2 = a + b + c$		
38	$\frac{\xi^2}{56a_{AQI}^{12}} (14a_{AQI}^{12} - 7a_{AQI}^6\xi^6 + 2\xi^{12})$	$ \xi \leq a_{AQI}$
	$\frac{9a_{AQI}^2}{56}$	$ \xi > a_{AQI}$
39	$\frac{\xi^2}{180a_{AQII}^{16}} (45a_{AQII}^{16} - 18a_{AQII}^8\xi^8 + 5\xi^{16})$	$ \xi \leq a_{AQII}$
	$\frac{8a_{AQII}^2}{45}$	$ \xi > a_{AQII}$
40	$\frac{\xi^2}{2}$	$ \xi \leq a_{Op}$
	$\frac{\xi^2}{2} - \frac{c_{Op}\xi^4}{4} + \frac{d_{Op}\xi^6}{6} - \frac{e_{Op}\xi^8}{8}$	$a_{Op} < \xi \leq b_{Op}$
	$\frac{b_{Op}^2}{2} - \frac{c_{Op}b_{Op}^4}{4} + \frac{d_{Op}b_{Op}^6}{6} - \frac{e_{Op}b_{Op}^8}{8}$	$ \xi > b_{Op}$
41	$\frac{\xi^2}{96c_Q^4} (\xi^4 - 3c_Q^2\xi^2 + 3c_Q^4)$	$ \xi \leq c_Q$
	$\frac{c_Q^2}{96}$	$ \xi > c_Q$
42	$\frac{c_S^2}{4} \left[1 - \left(1 - \frac{\xi^2}{c_S^2} \right)^2 \right]$	$ \xi \leq c_S$
	$\frac{c_S^2}{4}$	$ \xi > c_S$
43	$\frac{\xi^2}{2}$	$ \xi \leq c_T$
	$\frac{c_T^2}{2}$	$ \xi > c_T$

Table C.5: M-Estimators: $\rho(\xi)$ - part 5/5

#	$\rho(\xi)$	Range
44	$\begin{cases} \frac{3}{2} \sin\left(\frac{4}{9}\right) \left[\frac{\xi^{10}}{10c_{Uk}^8} - \frac{\xi^6}{3c_{Uk}^4} + \frac{\xi^2}{2} \right] \\ \frac{3}{2} \sin\left(\frac{4}{9}\right) c_{Uk}^2 \end{cases}$	$ \xi \leq c_{Uk}$ $ \xi > c_{Uk}$
45	$\begin{cases} \frac{\xi^2}{2} \\ -\frac{c_{YaI}(a_{YaI}- \xi)^3}{3(c_{YaI}-a_{YaI})^2} + D_{YaI} \\ D_{YaI} \end{cases}$	$ \xi \leq c_{YaI}$ $c_{YaI} < \xi \leq a_{YaI}$ $ \xi > a_{YaI}$
46	$\begin{cases} \frac{\xi^2}{2} \\ c_{YaII} \xi + \frac{c_{YaII}^2}{2} - c_{YaII} c_{YaII} \\ c_{YaII} a_{YaII} + \frac{c_{YaII}^2}{2} - c_{YaII} c_{YaII} \end{cases}$	$ \xi \leq c_{YaII}$ $c_{YaII} < \xi \leq a_{YaII}$ $ \xi > a_{YaII}$
47	$\frac{ 2-a_{Ba} }{a_{Ba}} \left[\left(\left(\frac{c_{Ba}}{ \xi } \right)^2 + 1 \right)^{\frac{a_{Ba}}{2}} - 1 \right]$	$ \xi \leq \infty$
48	$\ln \left(1 + \frac{ \xi ^{p_{GT}}}{q_{GT}} \right) \left(q_{GT} + \frac{1}{p_{GT}} \right)$	$ \xi \leq \infty$

Table C.6: Influence Function: $\psi(\xi)$ - part 1/5

#	$\psi(\xi)$	Range
1	$\text{sign}(\xi)$	$ \xi \leq \infty$
2	$\frac{\xi}{\sqrt{1+\frac{\xi^2}{2}}}$	$ \xi \leq \infty$
3	$\frac{c_F \xi}{c_F + \xi }$	$ \xi \leq \infty$
4	$\begin{cases} \xi & \xi \leq c_{Hu} \\ c_{Hu} \text{sign}(\xi) & \xi > c_{Hu} \end{cases}$	$ \xi \leq c_{Hu}$ $ \xi > c_{Hu}$
5	$\begin{cases} c_{MH} \sin\left(\frac{\xi}{c_{MH}}\right) & \xi \leq \frac{\pi}{2} c_{MH} \\ c_{MH} \text{sign}(\xi) & \xi > \frac{\pi}{2} c_{MH} \end{cases}$	$ \xi \leq \frac{\pi}{2} c_{MH}$ $ \xi > \frac{\pi}{2} c_{MH}$
6	$\frac{\xi(\xi +2c_K)}{(c_K+ \xi)^2}$	$ \xi \leq \infty$
7	$\tanh(c_L c \xi)$	$ \xi \leq \infty$
8	$c_L \tanh\left(\frac{\xi}{c_L}\right)$	$ \xi \leq \infty$
9	$\begin{cases} \xi & \xi \leq c_M \\ \frac{c_M \text{sign}(\xi)}{2} & \xi > c_M \end{cases}$	$ \xi \leq c_M$ $ \xi > c_M$
10	$\frac{\xi(4+c_Z \xi)}{(2+c_Z \xi)^2}$	$ \xi \leq \infty$
11	$\frac{16\xi \left[\exp\left(-\frac{\xi^2}{c_{At}^2}\right) \right]^2}{\left[1 + \exp\left(-\frac{\xi^2}{c_{At}^2}\right) \right]^4}$	$ \xi \leq \infty$
12	$\begin{cases} \xi & \xi \leq c_{BH} \\ \frac{c_{BH}^2}{\xi} & \xi > c_{BH} \end{cases}$	$ \xi \leq c_{BH}$ $ \xi > c_{BH}$
13	$\xi \left(1 + \frac{\xi^2}{5c_{Be}^2} \right)^{-3}$	$ \xi \leq \infty$
14	$\frac{2\xi \exp(-\xi^2)}{[\exp(-\xi^2) + c_{BZ}] \ln(10)}$	$ \xi \leq \infty$
15	$\frac{\xi}{1+\frac{\xi^2}{c^2}}$	$ \xi \leq \infty$

Table C.6: Influence Function: $\psi(\xi)$ - part 2/5

#	$\psi(\xi)$	Range
16	$\frac{\xi}{c_{Co}^3 \sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2c_{Co}^2}\right)$	$ \xi \leq \infty$
17	$\begin{cases} \xi \\ \xi \exp\left(\frac{-\xi^2}{c_{Da}}\right) \end{cases}$	$ \xi \leq c_{Da}$ $ \xi > c_{Da}$
18	$\frac{2\xi c_{GM}^2}{(c_{GM}^2 + \xi^2)^2}$	$ \xi \leq \infty$
19	$\begin{cases} \xi \\ \xi \exp\left[-\frac{(\xi - c_{GGW})^{b_{GGW}}}{2a_{GGW}}\right] \end{cases}$	$ \xi \leq c_{GGW}$ $ \xi > c_{GGW}$
20	$\xi \left(1 + \frac{\xi^4}{c_I^4}\right)^{-2}$	$ \xi \leq \infty$
21	$\begin{cases} a_J \xi \left(1 - a_J \frac{\xi^2}{c_J^2}\right)^2 \\ \frac{c_J^2 \cdot a_J (1 - a_J)^2}{\xi} \end{cases}$	$ \xi \leq c_J$ $ \xi > c_J$
22	$\begin{cases} \xi \\ a_{Ku} \text{sign}(\xi) \\ \frac{a_{Ku} \xi}{b_{Ku}} \exp\left(1 - \frac{\xi^2}{b_{Ku}^2}\right) \end{cases}$	$ \xi \leq a_{Ku}$ $a_{Ku} < \xi \leq b_{Ku}$ $ \xi > b_{Ku}$
23	$\frac{\xi}{c_{Lz}^2 \left(1 + \frac{\xi^2}{2c_{Lz}^2}\right)^2}$	$ \xi \leq \infty$
24	$\begin{cases} \xi \\ \frac{c_{MS}^{3/2} \text{sign}(\xi)}{\sqrt{ \xi }} \end{cases}$	$ \xi \leq c_{MS}$ $ \xi > c_{MS}$
25	$\xi \exp(-c_R \xi)$	$ \xi \leq \infty$
26	$\frac{c_{Sh} \sinh(c_{Sh} \xi)}{\cosh(c_{Sh} \xi)^2}$	$ \xi \leq \infty$
27	$\xi \exp\left(-\frac{\xi^2}{c_W^2}\right)$	$ \xi \leq \infty$
28	$\begin{cases} \xi \\ \text{sign}(\xi) a_{Wu} (\xi - c_{Wu}) \end{cases}$	$ \xi \leq c_{Wu}$ $ \xi > c_{Wu}$

Table C.6: Influence Function: $\psi(\xi)$ - part 3/5

#	$\psi(\xi)$	Range
29	$\frac{4\xi \exp\left(-\frac{\xi^2}{c_X^2}\right)}{c_X^2 \left[1 + \exp\left(-\frac{\xi^2}{c_X^2}\right)\right]^2}$	$ \xi \leq \infty$
30	$\begin{cases} \xi \\ \frac{2\xi}{\xi^2+1} \end{cases}$	$ \xi \leq c_Y$ $ \xi > c_Y$
31	$\begin{cases} \frac{2}{3}\xi \left(1 - \frac{\xi^4}{c_{As}^4}\right)^2 \\ 0 \end{cases}$	$ \xi \leq c_{As}$ $ \xi > c_{As}$
32	$\begin{cases} c_{An} \sin\left(\frac{\xi}{c_{An}}\right) \\ 0 \end{cases}$	$ \xi \leq \pi c_{An}$ $ \xi > \pi c_{An}$
33	$\begin{cases} \xi \left(1 - \frac{\xi^2}{c_{Bi}^2}\right)^2 \\ 0 \end{cases}$	$ \xi \leq c_{Bi}$ $ \xi > c_{Bi}$
34	$\begin{cases} \xi \\ a_{Cl} \tanh\left[\frac{a_{Cl}}{2}(rc_l - \xi)\right] \text{sign}(\xi) \\ 0 \end{cases}$	$ \xi \leq qc_l$ $qc_l < \xi \leq rc_l$ $ \xi > rc_l$
35	$\begin{cases} \xi \\ a_H \text{sign}(\xi) \\ \frac{a_H(c_H - \xi)}{c_H - b_H} \text{sign}(\xi) \\ 0 \end{cases}$	$ \xi \leq a_H$ $a_H < \xi \leq b_H$ $b_H < \xi \leq c_H$ $ \xi > c_H$
36	$\begin{cases} \xi \\ [A(k-1)]^{1/2} \tanh\left[\frac{(k-1)^{1/2}B}{2A^{1/2}}(r - \xi)\right] \text{sign}(\xi) \\ 0 \end{cases}$	$ \xi \leq q$ $q < \xi \leq r$ $ \xi > r$

Table C.6: Influence Function: $\psi(\xi)$ - part 4/5

#	$\psi(\xi)$	Range
37	ξ	$ \xi \leq c$
	$\text{sign}(\xi) \left[\xi - \frac{s}{2b} (\xi - c)^2 \right]$	$c < \xi \leq c + b$
	$\text{sign}(\xi) \left[c + b - \frac{bs}{2} + \right.$ $\left. + \frac{(s-1)}{a} \left(\frac{1}{2} (\xi - b - c)^2 - a(\xi - b - c) \right) \right]$	$c + b < \xi \leq s_2$
	0 where: $s_2 = a + b + c$	$ \xi > s_2$
38	$\left\{ \begin{array}{l} \frac{\xi}{2} \left(1 - \frac{\xi^6}{a_{AQI}^6} \right)^2 \\ 0 \end{array} \right.$	$ \xi \leq a_{AQI}$ $ \xi > a_{AQI}$
	39	$\left\{ \begin{array}{l} \frac{\xi}{2} \left(1 - \frac{\xi^8}{a_{AQII}^8} \right)^2 \\ 0 \end{array} \right.$
40		$\left\{ \begin{array}{l} \xi \\ \xi - c_{Op}\xi^3 + d_{Op}\xi^5 - e_{Op}\xi^7 \\ 0 \end{array} \right.$
	41	$\left\{ \begin{array}{l} \frac{\xi}{16} \left(1 - \frac{\xi^2}{c_Q^2} \right)^2 \\ 0 \end{array} \right.$
42		$\left\{ \begin{array}{l} \xi \left(1 - \frac{\xi^2}{c_S^2} \right) \\ 0 \end{array} \right.$
	43	$\left\{ \begin{array}{l} \xi \\ 0 \end{array} \right.$

Table C.6: Influence Function: $\psi(\xi)$ - part 5/5

#	$\psi(\xi)$	Range
44	$\begin{cases} \frac{3}{2}\xi \left(1 - \frac{\xi^4}{c_{Uk}^4}\right)^2 \sin \left[\frac{2}{3} \left(1 - \frac{\xi^4}{c_{Uk}^4}\right)^2\right] \\ 0 \end{cases}$	$ \xi \leq c_{Uk}$ $ \xi > c_{Uk}$
45	$\begin{cases} \frac{\xi^2}{2} \\ \frac{\xi \cdot c_{YaI} (a_{YaI} - \xi)^2}{ \xi (a_{YaI} - c_{YaI})^2} \\ 0 \end{cases}$	$ \xi \leq c_{YaI}$ $c_{YaI} < \xi \leq a_{YaI}$ $ \xi > a_{YaI}$
46	$\begin{cases} \xi \\ \frac{c_{YaII} \cdot \xi}{ \xi } \\ 0 \end{cases}$	$ \xi \leq c_{YaII}$ $c_{YaII} < \xi \leq a_{YaII}$ $ \xi > a_{YaII}$
47	$\frac{\xi}{c_{Ba}^2} \left[\left(\frac{\xi}{c_{Ba}} \right)^2 \frac{1}{ 2 - a_{Ba} } + 1 \right]^{\frac{a_{Ba}}{2} - 1}$	$ \xi \leq \infty$
48	$\frac{(p_{GT} \cdot q_{GT} + 1) \xi ^{p_{GT} - 1} \text{sign}(\xi)}{q_{GT} + \xi ^{p_{GT}}}$	$ \xi \leq \infty$

766 **Appendix D. Tuning parameters for $E_{ff} = 90, 98, \text{ and } 99\%$**

Table D.7: Tuning parameters for $E_{ff} = 90, 98 \text{ and } 99\%$ - part 1/4

#	M-Estimators	Tuning Parameters	90%	98%	99%
0	Contaminated	$\begin{cases} b_{CN} \\ p_{CN} \end{cases}$	10	10	10
	Normal		0.514	0.09647	0.0441
3	Fair	c_F	0.6351	3.0421	4.958
4	Huber	c_{Hu}	0.9818	1.7459	2.0102
5	Modified Huber	c_{MH}	0.8590	1.6543	2.0062
6	Kong	c_k	1.221	4.998	7.915
7	LnCosh	c_{LC}	1.2778	0.5475	0.423
8	Logistic	c_L	0.7826	1.8265	2.364
9	Müller	c_{Mu}	1.9127	2.6861	2.951
10	Zhang	c_Z	1.638	0.4002	0.252
11	Alamgir	c_{Al}	2.031	2.819	3.165
12	Bab-Hadiashar	c_{BH}	1.5139	2.1586	2.3948
13	Bell	c_{Bl}	1.6703	2.8928	3.5533
14	Blake-Zisserman	c_{BZ}	$3.9e^{-3}$	$9.3e^{-5}$	$2.0e^{-5}$
15	Cauchy	c_C	1.7249	3.3962	4.2904
16	Correntropy	c_{Co}	1.6851	2.7651	3.35219
17	Danish	c_{Da}	2.4583	3.1180	3.3544
		D_{Da}	3.1268	4.93	5.6845

Table D.7: Tuning parameters for $E_{ff} = 90, 98$ and 99% - part 2/4

#	M-Estimators	Tuning Parameters	90%	98%	99%
18	Geman-McClure	c_{GM}	2.8937	5.1597	6.3809
19	GGW	$\begin{cases} a_{GGW} \\ b_{GGW} \\ c_{GGW} \end{cases}$	$\begin{matrix} 1.0283 \\ 1.5 \\ 0.8709 \end{matrix}$	$\begin{matrix} 1.9523 \\ 1.5 \\ 1.3 \end{matrix}$	$\begin{matrix} 2.3777 \\ 1.5 \\ 1.5 \end{matrix}$
20	Insha	c_I	2.7331	3.8834	4.3882
21	Jin	$\begin{cases} a_J \\ c_J \end{cases}$	$\begin{matrix} 0.65 \\ 3.11 \end{matrix}$	$\begin{matrix} 0.65 \\ 4.7735 \end{matrix}$	$\begin{matrix} 0.65 \\ 5.677 \end{matrix}$
22	Kumar	$\begin{cases} a_{Ku} \\ b_{Ku} \\ D_{Ku} \end{cases}$	$\begin{matrix} 1.1156 \\ 2.2311 \\ 3.1111 \end{matrix}$	$\begin{matrix} 1.7521 \\ 3.5042 \\ 7.6745 \end{matrix}$	$\begin{matrix} 2.0115 \\ 4.0230 \\ 10.1154 \end{matrix}$
23	Lorentzian	c_{Lz}	2.0462	3.6485	4.5120
24	Merril-Schweppe	c_{MS}	1.3209	1.9999	2.2452
25	Ramsay	c_R	0.5266	0.2189	0.1527
26	Sech	c_{Sh}	0.5271	0.2985	0.2418
27	Welsch	c_W	2.3831	3.9104	4.7407
28	Wu	$\begin{cases} a_{Wu} \\ c_{Wu} \end{cases}$	$\begin{matrix} 0.5 \\ 2.1320 \end{matrix}$	$\begin{matrix} 0.5 \\ 2.9261 \end{matrix}$	$\begin{matrix} 0.5 \\ 3.1885 \end{matrix}$
29	Xie	c_X	1.6706	2.3409	2.6359
30	Youssef	c_Y	2.1966	3.0050	3.2637

Table D.7: Tuning parameters for $E_{ff} = 90, 98$ and 99% - part 3/4

#	M-Estimators	Tuning Parameters	90%	98%	99%
31	Asad	c_{As}	3.1576	4.2103	4.6664
32	Andrews	c_{An}	1.1117	1.6930	2.0170
33	Biweight	c_{Bi}	3.8827	5.9207	7.0414
34	Collins	a_{Cl}	1.5112	1.9016	2.1827
		r_{Cl}	3.5	4.5	5.0
		q_{Cl}	1.3913	1.8759	2.1736
		D_{Cl}	2.8493	5.3769	7.1493
35	Hampel	a_H	1.1052	1.7515	2.0114
		b_H	2.2105	3.5030	4.0228
		c_H	4.4209	7.00609	8.0456
36	Hyperbolic Tangent	A	0.4174	0.8038	0.9173
		B	0.70	0.90	0.95
		k	4.0	5.0	5.5
		r	4.0	5.0	5.5
		q	1.1091	1.7874	2.0290
		D	3.0285	6.12261	7.7731
37	LQQ	b	1.2314	1.8262	1.9945
		c	0.8	1.2	1.4
		s	1.5	1.5	1.5
		a	4.4314	6.6262	7.5945
		D_1	-1.7555	-3.9058	-4.9878
		D_2	3.3206	7.4040	9.5731

Table D.7: Tuning parameters for $E_{ff} = 90, 98$ and 99% - part 4/4

#	M-Estimators	Tuning Parameters	90%	98%	99%
38	Modified Asad-Qadir I	a_{AQI}	2.9260	3.7751	4.1102
39	Modified Asad-Qadir II	a_{AQII}	2.8134	3.5855	3.8791
40	Optimal	a_{Op}	0.3279	0.3211	0.3206
		b_{Op}	2.7024	3.4216	3.7025
		c_{Op}	$8.1e^{-4}$	$4.9e^{-4}$	$3.3e^{-4}$
		d_{Op}	$5.0e^{-3}$	$3.0e^{-3}$	$2.0e^{-3}$
		e_{Op}	$3.2e^{-3}$	$8.8e^{-3}$	$5.3e^{-3}$
41	Qadir	a_Q	3.8827	5.9207	7.0414
42	Smith	a_S	3.1316	4.4864	5.2317
43	Talwar	a_T	2.5003	3.1365	3.3682
44	Uk	a_{Uk}	3.5890	4.8374	5.3812
45	Yang I	c_{YaI}	2.1622	2.6646	2.8260
		a_{YaI}	3.0	4.0	4.5
		D_{YaI}	2.9414	4.7362	5.5699
46	Yang II	c_{YaII}	1.1018	1.7619	2.0140
		a_{YaII}	3.0	4.0	4.5

767 **References**

- 768 [1] D. M. Prata, E. L. Lima, J. C. Pinto, In-Line Monitoring of Bulk
769 Polypropylene Reactors Based on Data Reconciliation Procedures, Macro-
770 molecular symposia 271 (1) (2008) 26–37.
- 771 [2] D. M. Prata, M. Schwaab, E. L. Lima, J. C. Pinto, Nonlinear dynamic
772 data reconciliation and parameter estimation through particle swarm op-
773 timization: Application for an industrial polypropylene reactor, Chemical
774 Engineering Science 64 (18) (2009) 3953–3967.
- 775 [3] D. M. Prata, E. L. Lima, J. C. Pinto, Nonlinear dynamic data reconcil-
776 iation in real time in actual processes, Computer Aided Chemical Engi-
777 neering 27 (2009) 47–54.
- 778 [4] D. R. Bellhouse, C. Genest, et al., Maty’s biography of Abraham de
779 Moivre, translated, annotated and augmented, Statistical science 22 (1)
780 (2007) 109–136.
- 781 [5] A. M. Zoubir, V. Koivunen, E. Ollila, M. Muma, Robust statistics for
782 signal processing, Cambridge University Press, 2018.
- 783 [6] Y. Bard, Nonlinear parameter estimation, Academic press, 1974.
- 784 [7] W. Rey, Introduction to robust and quasi-robust statistical methods,
785 Springer Science & Business Media, 1983.
- 786 [8] J. W. Tukey, A survey of sampling from contaminated distributions, Con-
787 tributions to probability and statistics 2 (1960) 448–485.
- 788 [9] S. M. Stigler, Simon Newcomb, Percy Daniell, and the history of robust
789 estimation 1885–1920, Journal of the American Statistical Association
790 68 (344) (1973) 872–879.
- 791 [10] J. W. Tukey, The future of data analysis, The annals of mathematical
792 statistics 33 (1) (1962) 1–67.

- 793 [11] P. J. Huber, Robust estimation of a location parameter, *The Annals of*
794 *Mathematical Statistics* 35 (1) (1964) 73–101.
- 795 [12] F. R. Hampel, Contribution to the theory of robust estimation, Ph.D.
796 thesis, University of California, Berkeley, 1968.
- 797 [13] S. M. Stigler, The changing history of robustness, *The American Statisti-*
798 *cian* 64 (4) (2010) 277–281.
- 799 [14] P. J. Huber, E. M. Ronchetti, *Robust Statistics*, vol. 10, A John Wiley &
800 Sons, Inc., Publication, 2009.
- 801 [15] G. E. Box, Non-normality and tests on variances, *Biometrika* 40 (3/4)
802 (1953) 318–335.
- 803 [16] D. M. Prata, Reconciliação robusta de dados para monitoramento em
804 tempo real, Ph.D. thesis, COPPE - Universidade Federal do Rio de
805 Janeiro, Rio de Janeiro, RJ, 2009.
- 806 [17] H. Kodamana, B. Huang, R. Ranjan, Y. Zhao, R. Tan, N. Sammakne-
807 jad, Approaches to robust process identification: A review and tutorial of
808 probabilistic methods, *Journal of Process Control* 66 (2018) 68–83.
- 809 [18] N. S. Vichare, Robust Mahalanobis distances in power system state es-
810 timation, Ph.D. thesis, Faculty of the Virginia Polytechnic Institute and
811 State University, Blacksburg, Virginia, USA, 1993.
- 812 [19] J. Jurečkova, Rates of Consistency of Classical One-Side Tests, in: *Ro-*
813 *bustness of Statistical Methods and Nonparametric Statistics*, Springer,
814 60–62, 1984.
- 815 [20] J. S. Albuquerque, L. T. Biegler, Data reconciliation and gross-error de-
816 tection for dynamic systems, *AIChE journal* 42 (10) (1996) 2841–2856.
- 817 [21] P. de Laplace, Mémoire sur les suites récurro-récurrentes et sur leurs us-
818 ages dans la théorie des hasards, *Mém. Acad. Roy. Sci. Paris* 6 (1774)
819 353–371.

- 820 [22] S. M. Stigler, Studies in the history of probability and statistics xl
821 boscovich, simpson and a 1760 manuscript note on fitting a linear relation,
822 *Biometrika* 71 (3) (1984) 615–620.
- 823 [23] S. M. Stigler, Studies in the history of probability and statistics XXXVIII:
824 RH Smith, a Victorian interested in robustness, *Biometrika* 67 (1) (1980)
825 217–221.
- 826 [24] H. M. Merrill, F. C. Schweppe, Bad data suppression in power system
827 static state estimation, *IEEE Transactions on Power Apparatus and Sys-*
828 *tems* (6) (1971) 2718–2725.
- 829 [25] P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud, Deterministic
830 edge-preserving regularization in computed imaging, *IEEE Transactions*
831 *on image processing* 6 (2) (1997) 298–311.
- 832 [26] C. V. Stewart, Robust parameter estimation in computer vision, *SIAM*
833 *review* 41 (3) (1999) 513–537.
- 834 [27] K. Arya, P. Gupta, P. K. Kalra, P. Mitra, Image registration using robust
835 M-estimators, *Pattern Recognition Letters* 28 (15) (2007) 1957–1968.
- 836 [28] E. Ronchetti, F. Trojani, Robust inference with GMM estimators, *Journal*
837 *of econometrics* 101 (1) (2001) 37–69.
- 838 [29] R. D. Martin, T. T. Simin, Outlier-resistant estimates of beta, *Financial*
839 *Analysts Journal* 59 (5) (2003) 56–69.
- 840 [30] T. Krarup, Gotterdammerung over least squares adjustment, in: *Proc.*
841 *14th Congress of the International Society of Photogrammetry*, vol. 3,
842 369–378, 1980.
- 843 [31] A. Berberan, Multiple outlier detection. A real case study, *Survey review*
844 33 (255) (1995) 41–49.

- 845 [32] A. Wieser, F. K. Brunner, Robust estimation applied to correlated GPS
846 phase observations, in: First International Symposium on Robust Statis-
847 tics and Fuzzy Techniques in Geodesy and GIS, Zurich, Switzerland, 2001.
- 848 [33] D. B. Özyurt, R. W. Pike, Theory and practice of simultaneous data
849 reconciliation and gross error detection for chemical processes, *Computers
850 & chemical engineering* 28 (3) (2004) 381–402.
- 851 [34] H.-y. Wu, L.-d. Wu, A robust estimation method in orbit improvement,
852 *Chinese Astronomy and Astrophysics* 29 (4) (2005) 430–437.
- 853 [35] P. Pennacchi, Robust estimate of excitations in mechanical systems using
854 M-estimators—theoretical background and numerical applications, *Journal
855 of Sound and Vibration* 310 (4-5) (2008) 923–946.
- 856 [36] D. M. Prata, M. Schwaab, E. L. Lima, J. C. Pinto, Simultaneous robust
857 data reconciliation and gross error detection through particle swarm op-
858 timization for an industrial polypropylene reactor, *Chemical Engineering
859 Science* 65 (17) (2010) 4943–4954.
- 860 [37] E. D. Valdetaro, R. Schirru, Simultaneous model selection, robust data
861 reconciliation and outlier detection with swarm intelligence in a thermal
862 reactor power calculation, *Annals of Nuclear Energy* 38 (9) (2011) 1820–
863 1832.
- 864 [38] J. Liu, Q. Su, M. Moreno, C. Laird, Z. Nagy, G. Reklaitis, Robust state es-
865 timation of feeding–blending systems in continuous pharmaceutical manu-
866 facturing, *Chemical Engineering Research and Design* 134 (2018) 140–153.
- 867 [39] Q. Su, Y. Bommireddy, Y. Shah, S. Ganesh, M. Moreno, J. Liu, M. Gonz-
868 lez, N. Yazdanpanah, T. O’Connor, G. V. Reklaitis, et al., Data reconcil-
869 iation in the Quality-by-Design (QbD) implementation of pharmaceutical
870 continuous tablet manufacturing, *International Journal of Pharmaceutics*
871 563 (2019) 259–272.

- 872 [40] M. Tabatabai, J. Kengwoung-Keumo, W. Eby, S. Bae, U. Manne,
873 M. Fouad, K. Singh, A new robust method for nonlinear regression, Jour-
874 nal of biometrics & biostatistics 5 (5) (2014) 211.
- 875 [41] J. Aldrich, et al., RA Fisher and the making of maximum likelihood 1912-
876 1922, Statistical Science 12 (3) (1997) 162–176.
- 877 [42] M. G. Kendall, Daniel Bernoulli on maximum likelihood, Biometrika
878 48 (1) (1961) 1–2.
- 879 [43] A. Hald, et al., On the history of maximum likelihood in relation to inverse
880 probability and least squares, Statistical Science 14 (2) (1999) 214–222.
- 881 [44] P. Reilly, R. Carpani, Application of statistical theory of adjustment to
882 material balances, in: Proceedings of the 13th Canadian chemical engi-
883 neering conference, 21–23, 1963.
- 884 [45] I.-B. Tjoa, L. T. Biegler, Simultaneous strategies for data reconciliation
885 and gross error detection of nonlinear systems, Computers & chemical
886 engineering 15 (10) (1991) 679–690.
- 887 [46] P. J. Huber, et al., Robust regression: asymptotics, conjectures and Monte
888 Carlo, The Annals of Statistics 1 (5) (1973) 799–821.
- 889 [47] T. Feital, D. M. Prata, J. C. Pinto, Comparison of methods for estimation
890 of the covariance matrix of measurement errors, The Canadian Journal of
891 Chemical Engineering 92 (12) (2014) 2228–2245.
- 892 [48] F. R. Hampel, The influence curve and its role in robust estimation, Jour-
893 nal of the american statistical association 69 (346) (1974) 383–393.
- 894 [49] P. J. Rousseeuw, C. Croux, Explicit scale estimators with high breakdown
895 point, L1-Statistical analysis and related methods 1 (1992) 77–92.
- 896 [50] P. J. Rousseeuw, C. Croux, Alternatives to the median absolute deviation,
897 Journal of the American Statistical association 88 (424) (1993) 1273–1283.

- 898 [51] R. C. Fair, On the robust estimation of econometric models, in: Annals of
899 Economic and Social Measurement, Volume 3, number 4, NBER, 667–677,
900 1974.
- 901 [52] D. C. Hoaglin, F. Mosteller, J. W. Tukey, Understanding robust and ex-
902 ploratory data analysis, vol. 3, Wiley New York, 1983.
- 903 [53] P. J. Huber, Robust Statistics, John Wiley & Sons, 1981.
- 904 [54] S. Baselga, Global optimization solution of robust estimation, Journal of
905 surveying Engineering 133 (3) (2007) 123–128.
- 906 [55] F. R. Hampel, A general qualitative definition of robustness, The Annals
907 of Mathematical Statistics (1971) 1887–1896.
- 908 [56] J. E. Dennis Jr, R. E. Welsch, Techniques for nonlinear least squares and
909 robust regression, Communications in Statistics-Simulation and Compu-
910 tation 7 (4) (1978) 345–359.
- 911 [57] R. H. Smith, True average of observations?, Nature 37 (959) (1888) 464–
912 464.
- 913 [58] P. W. Holland, R. E. Welsch, Robust regression using iteratively
914 reweighted least-squares, Communications in Statistics-theory and Meth-
915 ods 6 (9) (1977) 813–827.
- 916 [59] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel, Robust
917 statistics: the approach based on influence functions, John Wiley & Sons,
918 1986.
- 919 [60] N. Arora, L. T. Biegler, Redescending estimators for data reconciliation
920 and parameter estimation, Computers & Chemical Engineering 25 (11)
921 (2001) 1585–1599.
- 922 [61] S. Xie, C. Yang, X. Yuan, X. Wang, Y. Xie, A novel robust data recon-
923 ciliation method for industrial processes, Control Engineering Practice 83
924 (2019) 203–212.

- 925 [62] G. Shevlyakov, S. Morgenthaler, A. Shurygin, Redescending M -estimators,
926 Journal of Statistical Planning and Inference 138 (10) (2008) 2906–2917.
- 927 [63] P. Rousseeuw, A. Leroy, Robust Regression and Outlier Detection, 1987.
- 928 [64] P. J. Huber, Finite sample breakdown of M -and P -estimators, The Annals
929 of Statistics (1984) 119–126.
- 930 [65] J. Zhang, G. Li, et al., Breakdown properties of location M -estimators,
931 The Annals of Statistics 26 (3) (1998) 1170–1189.
- 932 [66] L. P. Johnston, M. A. Kramer, Maximum likelihood data rectification:
933 Steady-state systems, AIChE Journal 41 (11) (1995) 2415–2426.
- 934 [67] Z. Zhang, R. W. Pike, T. A. Hertwig, Source reduction from chemical
935 plants using on-line optimization, Waste Management 15 (3) (1995) 183–
936 191.
- 937 [68] J. Chen, A. Bandoni, J. A. Romagnoli, Outlier detection in process plant
938 data, Computers & chemical engineering 22 (4-5) (1998) 641–646.
- 939 [69] M. Bourouis, L. Pibouleau, P. Floquet, S. Domenech, D. M. K. Al-Gobaisi,
940 Simulation and data validation in multistage flash desalination plants,
941 Desalination 115 (1) (1998) 1–14.
- 942 [70] K. Mingfang, C. Bingzhen, L. Bo, An Integral approach to dynamic data
943 rectification, Computers & Chemical Engineering 24 (2-7) (2000) 749–753.
- 944 [71] J. Ragot, M. Chadli, D. Maquin, Mass balance equilibration: a robust
945 approach using contaminated distribution, AIChE Journal 51 (5) (2005)
946 1569–1575.
- 947 [72] W. Wongrat, T. Srinophakun, P. Srinophakun, Modified genetic algorithm
948 for nonlinear data reconciliation, Computers & chemical engineering 29 (5)
949 (2005) 1059–1067.

- 950 [73] C. C. Pai, G. D. Fisher, Application of Broyden's method to reconciliation
951 of nonlinearly constrained data, *AIChE Journal* 34 (5) (1988) 873–876.
- 952 [74] Z. Lingke, S. Hongye, C. Jian, A New Method to Solve Robust Data Rec-
953 onciation in Nonlinear Process, *Chinese Journal of Chemical Engineering*
954 14 (3) (2006) 357–363.
- 955 [75] R. Faber, B. Li, P. Li, G. Wozny, Data reconciliation for real-time op-
956 timization of an industrial coke-oven-gas purification process, *Simulation*
957 *Modelling Practice and Theory* 14 (8) (2006) 1121–1134.
- 958 [76] R. Faber, H. Arellano-Garcia, P. Li, G. Wozny, An optimization frame-
959 work for parameter estimation of large-scale systems, *Chemical Engineer-
960 ing and Processing: Process Intensification* 46 (11) (2007) 1085–1095.
- 961 [77] M. Schladt, B. Hu, Soft sensors based on nonlinear steady-state data rec-
962 onciation in the process industry, *Chemical Engineering and Processing:
963 Process Intensification* 46 (11) (2007) 1107–1115.
- 964 [78] M. Alhaj-Dibo, D. Maquin, J. Ragot, Data reconciliation: A robust ap-
965 proach using a contaminated distribution, *Control Engineering Practice*
966 16 (2) (2008) 159–170.
- 967 [79] T. Lid, S. Skogestad, Data reconciliation and optimal operation of a cat-
968 alytic naphtha reformer, *Journal of Process Control* 18 (3-4) (2008) 320–
969 331.
- 970 [80] T. Lid, S. Skogestad, Scaled steady state models for effective on-line ap-
971 plications, *Computers & Chemical Engineering* 32 (4-5) (2008) 990–999.
- 972 [81] D. M. Prata, J. C. Pinto, E. L. Lima, Comparative analysis of robust esti-
973 mators on nonlinear dynamic data reconciliation, *Computer Aided Chem-
974 ical Engineering* 25 (2008) 501–506.
- 975 [82] M. J. Liebman, T. F. Edgar, L. S. Lasdon, Efficient data reconciliation
976 and estimation for dynamic processes using nonlinear programming tech-
977 niques, *Computers & chemical engineering* 16 (10-11) (1992) 963–986.

- 978 [83] Z. Zhang, Z. Shao, X. Chen, K. Wang, J. Qian, Quasi-weighted least
979 squares estimator for data reconciliation, *Computers & chemical engi-*
980 *neering* 34 (2) (2010) 154–162.
- 981 [84] S. Jin, X. Li, Z. Huang, M. Liu, A New Target function for robust data rec-
982 onciation, *Industrial & Engineering Chemistry Research* 51 (30) (2012)
983 10220–10224.
- 984 [85] J. Chen, Y. Peng, J. C. Munoz, Correntropy estimator for data reconcili-
985 ation, *Chemical Engineering Science* 104 (2013) 1019–1027.
- 986 [86] Z. Zhang, Y.-Y. Chuang, J. Chen, Methodology of data reconciliation and
987 parameter estimation for process systems with multi-operating conditions,
988 *Chemometrics and Intelligent Laboratory Systems* 137 (2014) 110–119.
- 989 [87] Z. Zhang, J. Chen, Correntropy based data reconciliation and gross error
990 detection and identification for nonlinear dynamic processes, *Computers*
991 *& Chemical Engineering* 75 (2015) 120–134.
- 992 [88] C. E. Llanos, M. C. Sánchez, R. A. Maronna, Robust estimators for
993 data reconciliation, *Industrial & Engineering Chemistry Research* 54 (18)
994 (2015) 5096–5105.
- 995 [89] T. Korpela, O. Suominen, Y. Majanne, V. Laukkanen, P. Lautala, Robust
996 data reconciliation of combustion variables in multi-fuel fired industrial
997 boilers, *Control Engineering Practice* 55 (2016) 101–115.
- 998 [90] J. C. Coimbra, P. A. Melo, D. M. Prata, J. C. Pinto, On-Line Dy-
999 namic Data Reconciliation in Batch Suspension Polymerizations of Methyl
1000 Methacrylate, *Processes* 5 (3) (2017) 51.
- 1001 [91] A. S. da Cunha, L. S. Santos, F. C. Peixoto, D. M. Prata, Robust data
1002 reconciliation in a chemical reactor through simulated annealing optimiza-
1003 tion, *Latin American Applied Research* 47 (2017) 131–136.

- 1004 [92] J. Van de Vusse, Plug-flow type reactor versus tank reactor, *Chemical*
1005 *Engineering Science* 19 (12) (1964) 994–996.
- 1006 [93] S. Wu, J. Xu, W. Liu, X. Wu, X. Gu, Data Reconciliation Based on an
1007 Improved Robust Estimator and NT-MT for Gross Error Detection, in:
1008 *Advanced Computational Methods in Life System Modeling and Simula-*
1009 *tion*, Springer, 400–409, 2017.
- 1010 [94] C. E. Llanos, M. C. Sánchez, R. A. Maronna, Classification of systematic
1011 measurement errors within the framework of robust data reconciliation,
1012 *Industrial & Engineering Chemistry Research* 56 (34) (2017) 9617–9628.
- 1013 [95] J. Valluru, S. C. Patwardhan, L. T. Biegler, Development of robust ex-
1014 tended Kalman filter and moving window estimator for simultaneous state
1015 and parameter/disturbance estimation, *Journal of Process Control* 69
1016 (2018) 158–178.
- 1017 [96] T. J. Williams, R. E. Otto, A generalized chemical processing model for
1018 the investigation of computer control, *Transactions of the American Insti-*
1019 *tute of Electrical Engineers, Part I: Communication and Electronics* 79 (5)
1020 (1960) 458–473.
- 1021 [97] E. C. do Valle, R. de Araújo Kalid, A. R. Secchi, A. Kiperstok, *Collection*
1022 *of Benchmark Test Problems for Data Reconciliation and Gross Error*
1023 *Detection and Identification*, *Computers & Chemical Engineering* .
- 1024 [98] A. S. da Cunha, F. C. Peixoto, D. M. Prata, Robust Data Reconciliation
1025 in Chemical Reactors, *Computers & Chemical Engineering* (2020) 107170.
- 1026 [99] J. T. Barron, A general and adaptive robust loss function, in: *Proceedings*
1027 *of the IEEE Conference on Computer Vision and Pattern Recognition*,
1028 4331–4339, 2019.
- 1029 [100] J. B. McDonald, W. K. Newey, Partially adaptive estimation of regression
1030 models via the generalized t distribution, *Econometric theory* 4 (3) (1988)
1031 428–457.

- 1032 [101] H. Jeffreys, An alternative to the rejection of observations, Proceedings
1033 of the Royal Society of London. Series A, Containing Papers of a Mathe-
1034 matical and Physical Character 137 (831) (1932) 78–87.
- 1035 [102] M. Kong, B. Chen, B. Li, Simultaneous gross error detection and data rec-
1036 onciliation based on the robust estimation principle, Journal of Tsinghua
1037 University 02.
- 1038 [103] O. Karal, Maximum likelihood optimal and robust Support Vector Re-
1039 gression with Incosh loss function, Neural networks 94 (2017) 1–12.
- 1040 [104] P. F. Verhulst, Notice sur la loi que la population poursuit dans son ac-
1041 croissement, Correspondance mathématique et physique 10 (1838) 113–
1042 121.
- 1043 [105] A. A. Alamgir, S. A. Khan, D. M. Khan, U. Khalil, A new efficient re-
1044 descending M-estimator: Alamgir Redescending M-estimator, Research
1045 Journal of Recent Sciences 2 (8) (2013) 79–91.
- 1046 [106] A. Bab-Hadiashar, D. Suter, R. Hesami, Robust model fitting in pattern
1047 recognition, in: International Conference on Digital Image Computing
1048 Techniques and Applications 2002, Australian Pattern Recognition Soci-
1049 ety (APRS), 358–363, 2002.
- 1050 [107] R. M. Bell, An adaptive choice of the scale parameter for M-estimators .
- 1051 [108] R. Hartley, A. Zisserman, Multiple view geometry in computer vision,
1052 Cambridge university press, 2003.
- 1053 [109] S. M. Stigler, Studies in the History of Probability and Statistics. XXXIII
1054 Cauchy and the witch of Agnesi: An historical note on the Cauchy distri-
1055 bution, Biometrika (1974) 375–380.
- 1056 [110] W. Liu, P. P. Pokharel, J. C. Principe, Error entropy, correntropy and m-
1057 estimation, in: Machine Learning for Signal Processing, 2006. Proceedings

- 1058 of the 2006 16th IEEE Signal Processing Society Workshop on, IEEE,
1059 179–184, 2006.
- 1060 [111] S. Geman, D. E. McClure, Statistical methods for tomographic image
1061 reconstruction, *Bulletin of the International Statistical Institute* 52 (4)
1062 (1987) 5–21.
- 1063 [112] M. Koller, W. A. Stahel, Sharpening wald-type inference in robust regres-
1064 sion for small samples, *Computational Statistics & Data Analysis* 55 (8)
1065 (2011) 2504–2515.
- 1066 [113] I. Ullah, M. F. Qadir, A. Ali, Insha’s redescending M-estimator for ro-
1067 bust regression: A comparative study, *Pakistan journal of statistics and*
1068 *operation research* 2 (2) (2006) 135–144.
- 1069 [114] T. A. Kumar, K. D. Rao, A new M-estimator based robust multiuser de-
1070 tection in flat-fading non-Gaussian channels, *IEEE Transactions on com-*
1071 *munications* 57 (7) (2009) 1908–1913.
- 1072 [115] J. O. Ramsay, A comparative study of several robust estimates of slope,
1073 intercept, and scale in linear regression, *Journal of the American Statisti-*
1074 *cal Association* 72 (359) (1977) 608–615.
- 1075 [116] A. H. Youssef, A. A. El-Sheikh, E. T. H. Mohammed, New M-Estimator
1076 Objective Function in Simultaneous Equations Model (A Comparative
1077 Study), in: *International Mathematical Forum*, vol. 8, Citeseer, 1007–
1078 1022, 2013.
- 1079 [117] A. Ali, M. F. Qadir, A Modified M-estimator for the Detection of Outliers,
1080 *Pakistan journal of statistics and operation research* 1 (1).
- 1081 [118] D. F. Andrews, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers,
1082 J. W. Tukey, *Robust estimates of location: Survey and advances*, Prince-
1083 ton University Press, 1972.

- 1084 [119] A. E. Beaton, J. W. Tukey, The fitting of power series, meaning polynomi-
1085 als, illustrated on band-spectroscopic data, *Technometrics* 16 (2) (1974)
1086 147–185.
- 1087 [120] J. R. Collins, Robust estimation of a location parameter in the presence
1088 of asymmetry, *The Annals of Statistics* (1976) 68–85.
- 1089 [121] F. R. Hampel, P. J. Rousseeuw, E. Ronchetti, The change-of-variance
1090 curve and optimal re-descending M-estimators, *Journal of the American*
1091 *Statistical Association* 76 (375) (1981) 643–648.
- 1092 [122] V. J. Yohai, R. H. Zamar, Optimal locally robust M-estimates of regres-
1093 sion, *Journal of Statistical Planning and Inference* 64 (2) (1997) 309–323.
- 1094 [123] A. Ali, M. F. Qadir, A. Q. Salahuddin, Regression Outliers: New M-
1095 Class Ψ Functions based on Winsor’s Principle with Improved Asymptotic
1096 Efficiency, in: *Proceedings of the 8 th Islamic countries conference on*
1097 *statistical sciences in*, Citeseer, 2005.
- 1098 [124] M. J. Hinich, P. P. Talwar, A simple method for robust regression, *Journal*
1099 *of the American Statistical Association* 70 (349) (1975) 113–119.
- 1100 [125] U. Khalil, A new robust m-estimator with an application to non-stationary
1101 time series forecasting, Ph.D. thesis, University of Peshawar, Pakistan,
1102 2012.
- 1103 [126] Y. Yang, Robust estimation of geodetic datum transformation, *Journal of*
1104 *geodesy* 73 (5) (1999) 268–274.
- 1105 [127] Y. Yang, Robust estimation for dependent observations, *Manuscripta geo-*
1106 *daetica* 19 (1) (1994) 10–17.
- 1107 [128] A. Stromberg, et al., Why write statistical software? The case of robust
1108 statistical methods, *Journal of Statistical Software* 10 (5) (2004) 1–8.