

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA (PEQ-UFS)

REGINA LUANA SANTOS DE FRANÇA

ANÁLISE DE FUNÇÕES ROBUSTAS NA RECONCILIAÇÃO DE DADOS NÃO LINEAR
EM PROCESSOS QUÍMICOS

São Cristóvão (SE)

2015

REGINA LUANA SANTOS DE FRANÇA

ANÁLISE DE FUNÇÕES ROBUSTAS NA RECONCILIAÇÃO DE DADOS NÃO
LINEAR EM PROCESSOS QUÍMICOS

Dissertação de Mestrado apresentado ao
programa de Pós-Graduação em Engenharia
Química, como requisito parcial à obtenção do
título de Mestre em Engenharia Química.

Orientador: Prof. Dr. Antônio Martins de Oliveira Júnior

Coorientador: Prof. Dr. Domingos Fabiano Santana de Souza

São Cristóvão (SE)

2015

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE

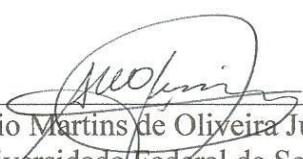
F814a	França, Regina Luana Santos de Análise de funções robustas na reconciliação de dados não linear em processos químicos / Regina Luana Santos de França ; orientador Antônio Martins de Oliveira Júnior. – São Cristóvão, 2015. 120 f. : il. Dissertação (Mestrado em Engenharia Química) – Universidade Federal de Sergipe, 2015. 1. Análise Funcional. 2. Processos químicos. 3. Sistemas lineares. 4. Programação não-linear. 5. Programação quadrática. I. Oliveira Júnior, Antônio Martins de. II. Título. CDU 517.98:543.062
-------	--

REGINA LUANA SANTOS DE FRANÇA

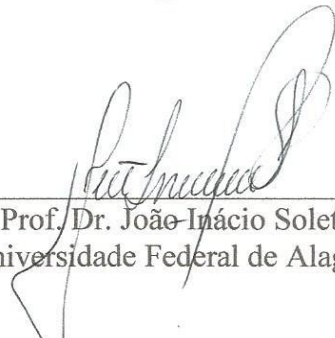
ANÁLISE DE FUNÇÕES ROBUSTAS NA RECONCILIAÇÃO DE DADOS NÃO
LINEAR EM PROCESSOS QUÍMICOS

Dissertação de Mestrado aprovada no Programa de Pós-Graduação em Engenharia Química
da Universidade Federal de Sergipe em 26 de Fevereiro de 2015.

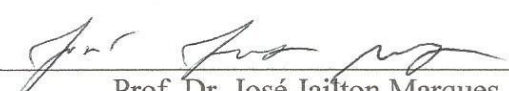
BANCA EXAMINADORA



Prof. Dr. Antônio Martins de Oliveira Júnior - Orientador
Universidade Federal de Sergipe



Prof. Dr. João Inácio Soletti
Universidade Federal de Alagoas



Prof. Dr. José Jailton Marques
Universidade Federal de Sergipe

Quando o SENHOR trouxe do cativeiro os que voltaram a Sião, estávamos como os que sonham.

Então a nossa boca se encheu de riso e a nossa língua de cântico; então se dizia entre os gentios: Grandes coisas fez o Senhor a estes.

Grandes coisas fez o Senhor por nós, pelas quais estamos alegres.

Traze-nos outra vez, ó Senhor, do cativeiro, como as correntes das águas no sul.

Os que semeiam em lágrimas segarão com alegria.

Aquele que leva a preciosa semente, andando e chorando, voltará, sem dúvida, com alegria, trazendo consigo os seus molhos.

Salmos 126:1-6

AGRADECIMENTOS

Passados 2 anos de árduos momentos de estudo, a meu entender não há nada mais sublime que agradecer. Então lá vai!!!

A MEU DEUS E SENHOR (em maiúsculo mesmo!!!) por mais essa promessa cumprida, pois, mesmo sem capacidade alguma, ELE me cercou de anjos valentes que me ajudaram a chegar até ao fim.

Minha belíssima família por tanto me chamar de louca por continuar a estudar, mas sempre me apoiar. O guerreiro Sr. Romeu Cardoso de França e a sua “Julieta” Sr^a Luiza Santos de França. Minha reverência!!!!

Ao meu amigo-orientador Professor Antônio Martins de Oliveira Júnior pelas palavras sinceras, pelo apoio e esforço contínuo em me ajudar a navegar na praia dele de Otimização e Controle de Processos, me apresentando a uma beleza chamada Reconciliação de Dados.

Ao Professor Domingos Fabiano Souza que chegou para ser Coorientador e pegou na minha mão e me ensinou a ser pesquisadora mestre em programação. Sou a melhor!!!(kkkk)

Ao corpo docente do PEQ, em especial aos Professores Luanda Gimeno, Roberto Rodrigues, Pedro Leite, José da Paixão pelos grandes ensinamentos. E mais ainda ao Professor Manoel Prado por ter me mostrado uma capacidade que eu não imaginava que possuía, e mesmo com tanta deficiência, ele acreditou e apostou suas fichas em mim. Obrigada mesmo!

Aos Professores que passaram pelas bancas Jailton Marques, Gabriel Souza, Oscar Sotomayor e João Solletti, que agregaram conhecimento e contribuíram substancialmente para essa dissertação.

Aos amigos conquistados ao longo do mestrado, Polena (Doido!) pelos biscoitos bono e companhia na fila enooooorme do resun. Raimundo (Prof. Dr. Dep.) digníssimo amigo que nunca se deve perguntar se as coisas estão bem, porque recebe um: Tudo péssimo, uma m*\$#D@!!! E aos demais, Kryslaine, Patrícia Rabelo, Patrícia Maria, Gleidinei, Livia, Tiago e Greicy.

E ao melhor encontro que o mestrado em Engenharia Química me proporcionou, meu companheiro, amigo e amor Clécio Martinho Rosário dos Santos que começou me ensinando Fenômenos de Transporte e além de me apaixonar por FT, me encantei por ele por tabela. E se o stress do mestrado não nos separou nada mais o fará!!!!

Um simples conselho aos que querem o título de Mestre: Não deixem o mar vos engolir. Não somos perfeitos, mas podemos ser tão “Mestres” quanto qualquer outro. Quem finge ser o que não é sempre acaba em descrédito. #FicaaDica

RESUMO

As funções robustas pertencem a famílias de estimadores que possuem a capacidade de atenuar os erros quando estes estão presentes nas medidas. A literatura expõe inúmeros trabalhos teóricos relacionados à utilização de funções robustas para reconciliação, quase todos direcionados a processos estacionários representados por sistemas lineares, no entanto, poucos estudos aprofundados são direcionados a processos estacionários não lineares, bem como a dados reais de plantas industriais. Além disso, questões referentes a: (i) definição da função; (ii) técnica de resolução do problema de reconciliação e, (iii) capacidade de predição de funções robustas na presença de erros grosseiros, ainda representam um desafio a ser explorado e que motivou o delineamento deste trabalho, o qual teve como objetivo avaliar a aptidão de algumas funções robustas na resolução de problemas de reconciliação de dados em processos químicos estacionários representados por sistemas lineares e não lineares. Inicialmente, as tradicionais funções robustas *Cauchy*, *Fair*, *Normal Contaminada* e *Logística* foram utilizadas no problema de reconciliação, tendo suas estimativas sido comparadas com as obtidas com o uso de funções mais recentes, como a *New Target* e *Alarm*. Para tal propósito, foi utilizado o método numérico de programação não linear, em particular, o “*Sequential Quadratic Programming*” (SQP), que se encontra implementado em ambiente computacional. Como critérios de desempenho, aplicaram-se o erro médio relativo, o número de iterações da função objetivo e o ajuste dos dados reais contaminados com erros. Observou-se que a função *Mínimos Quadrados Ponderados* apresentou um número reduzido de iterações em quase todos os estudos de casos realizados. As funções *Cauchy* e *Normal Contaminada* apresentaram bons resultados quanto ao número de iterações, inclusive no estudo de caso usando dados reais. Entretanto, em um dos casos testados, a função *Normal Contaminada* apresentou um problema de convergência. Já a função *Alarm* apresentou erro de convergência em uma das variáveis estimadas do estudo de caso com dados reais. Em sistemas não lineares contendo um único erro grosseiro, as funções *New Target*, *Alarm* e *Cauchy* apresentaram os bons índices de desempenho, principalmente em termos do erro médio relativo sendo que a última apresentou menor número de iterações. Na reconciliação dos dados industriais aplicando desvios sistemáticos em uma variável para comparar as eficiências das funções, a *Alarm* e a *Normal Contaminada* exibiram melhores ajustes.

Palavras chaves: Funções robustas; reconciliação de dados; erros grosseiros; programação não linear.

ABSTRACT

The functions robust estimators belong to families that have the ability to mitigate errors when they are present in the measurements. Literature expose numerous theoretical works related to the use of robust functions for reconciliation, almost all directed at stationary processes represented by linear systems, however, few in-depth studies are directed to nonlinear stationary processes, as well as the actual data of industrial plants. In addition, issues related to: (i) defining the function; (ii) resolution of the technical issue of reconciliation and, (iii) prediction capacity of robust functions in the presence of gross errors, still represent a challenge to be explored and that motivated the design of this study, which aimed to assess the suitability some function in resolving robust data reconciliation problems steady state chemical processes represented by linear and nonlinear systems. Initially, the traditional robust functions Cauchy, Fair, Contaminated Normal and Logistics were used in the issue of reconciliation, and their estimates have been compared with those obtained with the use of the latest features, such as New Target and Alarm. For this purpose, the numerical method used was nonlinear programming, in particular the "Sequential Quadratic Programming" (SQP), which is implemented in the computing environment. As performance criteria, applied the average relative error, the number of iterations of the objective function and the adjustment of the actual data contaminated with errors. With the results, it was observed that the Weighted Least Squares function showed a reduced number of iterations in almost all cases studies performed. The Cauchy and Normal Contaminated functions showed good results as the number of iterations, including the case study using real data. However, in one of the cases tested, the Contaminated Normal function presented a problem of convergence. Already the Alarm function showed convergence error in one of the variables estimated the case study with real data. In nonlinear systems containing a single gross error, the functions New Target, Alarm and Cauchy had good levels of performance, especially in terms of the average relative error and the last presented fewer iterations. Reconciliation of industrial data by applying systematic deviations in a variable to compare the efficiencies of the functions, the Alarm and Normal Contaminated exhibited optimal settings.

Key words: robust functions; data reconciliation; gross errors; nonlinear programming.

LISTA DE TABELAS

Tabela 2.1 Dados medidos da unidade	27
Tabela 3.1 Dados de fluxo molar medido e desvio padrão	36
Tabela 3.2 Dados medidos e desvio padrão das medidas.....	38
Tabela 3.3 Valor das constantes para diferentes funções com valores de eficiência de 95% .	41
Tabela 3.4 Dados do trocador de calor com bypass	44
Tabela 3.5 Dados do trocador de calor com válvula bypass	45
Tabela 3.6 Variâncias para as variáveis medidas do processo de polimerização.....	48
Tabela 3.7 Valores dos parâmetros fixos do modelo.....	49
Tabela 4.1 Análise do método de Gauss-Newton para o estudo de caso 1	52
Tabela 4.2 Análise do método SQP para o estudo de caso 1.....	53
Tabela 4.3 Análise do método de Gauss Newton para o estudo de caso 2.....	54
Tabela 4.4 Análise do método SQP para o estudo de caso 2.....	54
Tabela 4.5 Análise do método de Gauss Newton para o estudo de caso 3.....	56
Tabela 4.6 Análise do método SQP para o estudo de caso 3.....	56
Tabela 4.7 Dados reconciliados utilizando funções robustas: DEG linear.....	58
Tabela 4.8 Resultados das iterações	58
Tabela 4.9 Dados reconciliados utilizando funções robustas: DEG não linear.....	62
Tabela 4.10 Resultado das iterações: Não linear	62

LISTA DE FIGURAS

Figura 2.1 Condicionamento de dados e suas aplicações	5
Figura 2.2 Características de conjuntos de dados.....	6
Figura 2.3 Fluxograma representativo.....	20
Figura 2.4 Matriz representativa.....	21
Figura 2.5 Fluxograma operacional de três unidades	27
Figura 3.1 Fluxograma do método SQP	33
Figura 3.2 Espécies observadas no reator de cloração	34
Figura 3.3 Reação de cloração do metano	35
Figura 3.4 Reação de substituição para formação de diclorometano	35
Figura 3.5 Fluxograma do trocador de calor	37
Figura 3.6 Comportamento das funções robustas em relação ao erro	41
Figura 3.7 Comportamento das funções de influência em relação ao erro.....	42
Figura 3.8 Trocador de calor com válvula de desvio (bypass).....	43
Figura 3.9 Trocador de calor com válvula bypass e balanço de energia	44
Figura 3.10 Processo de polimerização simplificado	46
Figura 4.1 Análise do desempenho das funções robustas em função da redução do erro relativo para o Caso 4.	59
Figura 4.2 Avaliação do desempenho das funções robustas em relação a redução do erro relativo para o Caso 5	63
Figura 4.3 Variação da vazão medida de propeno e reconciliada através da função WLS.....	65
Figura 4.4 Variação da vazão medida de propeno e reconciliada através da função CAUCHY	65
Figura 4.5 Variação da vazão medida de propeno e reconciliada através da função FAIR	65
Figura 4.6 Variação da vazão medida de propeno e reconciliada através da função LOGÍSTICA	65
Figura 4.7 Variação da vazão medida de propeno e reconciliada através da função ALARM	66
Figura 4.8 Variação da vazão medida de propeno e reconciliada através da função NEW TARGET	66
Figura 4.9 Variação da vazão medida de propeno e reconciliada através da função NORMAL CONTAMINADA	66
Figura 4.10 Variação da vazão medida de polímero e reconciliada através da função WLS	67
Figura 4.11 Variação da vazão medida de polímero e reconciliada através da função ALARM	67
Figura 4.12 Variação da vazão medida de polímero e reconciliada através da função CAUCHY	67
Figura 4.13 Variação da vazão medida de polímero e reconciliada através da função FAIR.....	67
Figura 4.14 Variação da vazão medida de polímero e reconciliada através da função LOGÍSTICA	68
Figura 4.15 Variação da vazão medida de polímero e reconciliada através da função NEW TARGET	68

Figura 4.16 Variação da vazão medida de polímero e reconciliada através da função NORMAL CONTAMINADA.....	68
Figura 4.17 Variação do nível medido do reator e reconciliado através da função WLS	69
Figura 4.18 Variação do nível medido do reator e reconciliado através da função CAUCHY	69
Figura 4.19 Variação do nível medido do reator e reconciliado através da função FAIR	69
Figura 4.20 Variação do nível medido do reator e reconciliado através da função LOGÍSTICA	69
Figura 4.21 Variação do nível medido do reator e reconciliado através da função ALARM ..	70
Figura 4.22 Variação do nível medido do reator e reconciliado através da função NEW TARGET	70
Figura 4.23 Variação do nível medido do reator e reconciliado através da função NORMAL CONTAMINADA	70
Figura 4.24 Variação da vazão medida contaminada com erro e reconciliada com WLS	73
Figura 4.25 Variação da vazão medida contaminada com erro e reconciliada com ALARM .	73
Figura 4.26 Variação da vazão medida contaminada com erro e reconciliada com CAUCHY	73
Figura 4.27 Variação da vazão medida contaminada com erro e reconciliada com FAIR	73
Figura 4.28 Variação da vazão medida contaminada com erro e reconciliada com LOGÍSTICA	74
Figura 4.29 Variação da vazão medida contaminada com erro e reconciliada com NEW TARGET	74
Figura 4.30 Variação da vazão medida contaminada com erro e reconciliada com NORMAL CONTAMINADA	74
Figura 4.31 Variação da vazão medida contaminada com erro e reconciliada com ALARM .	75
Figura 4.32 Variação da vazão medida contaminada com erro e reconciliada com CAUCHY	75
Figura 4.33 Variação da vazão medida contaminada com erro e reconciliada com FAIR	75
Figura 4.34 Variação da vazão medida contaminada com erro e reconciliada com LOGÍSTICA	75
Figura 4.35 Variação da vazão medida contaminada com erro e reconciliada com NEW TARGET	76
Figura 4.36 Variação da vazão medida contaminada com erro e reconciliada com NORMAL CONTAMINADA	76
Figura 4.37 Variação da vazão medida contaminada com erro e reconciliada com WLS	76
Figura 4.38 Análise do desempenho das funções robustas em relação a redução do erro relativo.	77
Figura 4.39 Análise do desempenho das funções robustas em relação a redução do erro relativo	78
Figura 4.40 Número de iterações em relação a reconciliação de dados com as funções robustas	79

SUMÁRIO

SUMÁRIO.....	XII
1 INTRODUÇÃO	1
1.1 OBJETIVO GERAL	3
1.1.1 OBJETIVOS ESPECÍFICOS.....	3
2 REVISÃO DA LITERATURA.....	4
2.1 RECONCiliaÇÃO DE DADOS.....	4
2.1.1 HISTÓRICO DA RECONCiliaÇÃO DE DADOS.....	7
2.2 MÉTODOS PARA A RECONCiliaÇÃO DE DADOS.....	14
2.2.1 RECONCiliaÇÃO DE DADOS EM ESTADO ESTACIONÁRIO.....	14
2.3 CLASSIFICAÇÃO DE VARIÁVEIS.....	19
2.4 ERROS EM MEDIDAS DE PROCESSO	21
2.4.1 ERROS ALEATÓRIOS	22
2.4.2 ERROS GROSSEIROS	23
2.4.3 DETECÇÃO DE ERROS GROSSEIROS.....	25
2.5 FUNÇÕES ROBUSTAS	29
3 METODOLOGIA	32
3.1 AVALIAÇÃO DA PREDIÇÃO DO MÉTODO SQP	32
3.2 RECONCiliaÇÃO DE DADOS APLICANDO FUNÇÕES ROBUSTAS	39
3.3 AVALIAÇÃO DE DESEMPENHO DAS FUNÇÕES.....	50
4 RESULTADOS E DISCUSSÃO	51
4.1 AVALIAÇÃO DA PREDIÇÃO DO MÉTODO SQP	51
4.2 RECONCiliaÇÃO DE DADOS APLICANDO FUNÇÕES ROBUSTAS	57
4.3 AVALIAÇÃO DE DESEMPENHO DAS FUNÇÕES ROBUSTAS NA PRESENÇA DE ERROS GROSSEIROS NO ESTUDO DE CASO 6	72
5 CONCLUSÃO	80
6 REFERÊNCIAS.....	82

APÊNDICES

APÊNDICE A - TÉCNICA DE PROGRAMAÇÃO QUADRÁTICA SEQUENCIAL (SQP)

APÊNDICE B - MÉTODO DE CONVERGÊNCIA DE GAUSS NEWTON

APÊNDICE C - AVALIAÇÃO DA PREDIÇÃO DO MÉTODO SQP

APÊNDICE D - DADOS RECONCILIADOS COM AS FUNÇÕES ROBUSTAS

LISTA DE ABREVIATURAS E SÍMBOLOS

SIGLAS	DESCRIÇÃO
<i>AIC</i>	Critério de informação Akaike
<i>ALARM</i>	Alamgir M-estimador redescendente
<i>CSTR</i>	<i>Continuous stirred tank reactor</i>
<i>DEG</i>	Detecção de erro grosseiro
<i>EG</i>	Erro grosseiro
<i>EMSO</i>	Ambiente para modelagem simulação e otimização
<i>FIFO</i>	Primeiro que entra primeiro que sai
<i>GLR</i>	Máxima verossimilhança generalizada
<i>GT</i>	Teste global
<i>MHE</i>	Horizonte móvel avançado
<i>MIMT</i>	Teste de medida iterativo modificado
<i>NT</i>	Teste nodal
<i>PCA</i>	Análise de componentes principais
<i>QWLS</i>	Matriz quadrada quasi-ponderada
<i>SQP</i>	Programação quadrática sequencial
<i>RDTR</i>	Reconciliação de dados em tempo real
<i>RTO</i>	Otimização em tempo real
<i>WLS</i>	<i>Weighted least squares</i>

SÍMBOLOS	DESCRIÇÃO [UNIDADE]
<i>CAT</i>	Massa de catalisador no reator [kg]
K_d	Constante de desativação do sítio catalítico [h^{-1}]
K_p	Constante cinética de reação [$\text{kg}^2/(\text{kg}^2 \cdot \text{h})$]

<i>L</i>	Nível do reator [%]
<i>m_{cat}</i>	Vazão de alimentação de catalisador [kg/h]
<i>m_e</i>	Vazão de alimentação propeno fresco (contaminado por propano) [kg/h]
<i>m_{PEEB}</i>	Vazão de alimentação de <i>PEEB</i> [kg/h]
<i>m_{Pol}</i>	Vazão de polipropileno produzido [kg/h]
<i>m_{Purga}</i>	Vazão de purga [kg/h]
<i>m_{TEA}</i>	Vazão de alimentação <i>TEA</i> [kg/h]
<i>Pa</i>	Massa de propano no reator [kg]
<i>Pe</i>	Massa de propeno no reator [kg]
<i>PEEB</i>	Massa de para-etoxietilbenzoato [kg]
<i>POL</i>	Massa de polímero no reator [kg]
<i>TEA</i>	Massa de Trietil alumínio no reator [kg]
<i>α</i>	Fator de recirculação de TEA [adimensional]
<i>w_a</i>	Fração mássica de propano na corrente de propeno [kg/kg]
<i>ρ_a</i>	Densidade do propano [kg/L]
<i>ρ_e</i>	Densidade do propeno [kg/L]
<i>ρ_{pol}</i>	Densidade do polímero [kg/L]

1 INTRODUÇÃO

Em qualquer planta química moderna, processo petroquímico ou refinaria, existem centenas ou mesmo milhares de variáveis, tais como vazões, temperaturas, pressões, níveis e composições, que são rotineiramente medidos e gravados automaticamente com a finalidade de controlar os processos, otimizar *on-line*, ou pelo simples processo de avaliação econômica. Os computadores modernos e sistemas de aquisição de dados facilitam a coleta e o processamento de um grande volume de dados, com uma frequência da ordem de minutos ou mesmo de segundos.

Em virtude dos dados serem gerados por instrumentação e por observadores muitas dessas medidas podem vir a se tornarem não exatas, ou seja, não refletem exatamente os dados reais do processo. Além disso, a falta de algumas medições, como, por exemplo, vazão das correntes ou dados de projeto dos equipamentos, dificulta, e pode até inviabilizar a aplicação de metodologias para otimização de processos químicos. Surge então a necessidade do uso de outros métodos para avaliar o balanço mássico de uma unidade industrial.

A estratégia conhecida para redução dessa imprecisão em medidas é chamada de condicionamento de dados, na qual está inclusa a reconciliação de dados. Essa estratégia permite que nas medições estejam todas as informações relevantes do processo, resultando em uma maior precisão e confiabilidade dos dados.

Em termos matemáticos, a reconciliação de dados consiste em um problema de otimização cujo objetivo é a melhoria da qualidade dos dados que representam processos experimentais ou industriais (PLÁCIDO, 1995). Os resultados dessa técnica quando aplicada corretamente, podem ser bastante expressivos, principalmente na otimização de processos, e no diagnóstico de falhas em plantas.

Entretanto, em termos estatísticos, a técnica de reconciliação de dados fundamenta-se em modelos estatísticos de distribuição; na regressão de dados e estimação de parâmetros, onde uma coerente escolha da função mérito e, em particular, na definição da função robusta para reconciliação, podem contribuir para a melhoria da qualidade do ajuste.

No que se refere às famílias de funções robustas, merece destaque o estudo pioneiro realizado por Huber (1981), através da estatística robusta utilizada para encontrar solução em problemas que fugiam à distribuição gaussiana normal. Essas funções possuem considerável distribuição independente e produzem resultados sem desvios na presença da derivada dos

dados em uma distribuição ideal e são insensíveis aos desvios da idealidade (ALBUQUERQUE E BIEGLER, 1995).

As funções robustas possuem habilidade de predição, em particular, uma robustez em tratar dados com erros grosseiros, analisadas em alguns trabalhos citados ao longo desta dissertação.

Nesse trabalho serão apresentados resultados comparativos das eficiências dos estimadores de processos químicos no estado estacionário em sistemas lineares e não lineares com indicativo de erros através da aplicação das funções robustas *Cauchy*, *Fair*, *Normal Contaminada*, *Logística*, *Alarm* e *New Target*. O critério de avaliação idealizado para definir a escolha de cada função foi realizado através do número de iterações e do erro médio relativo. O desafio da reconciliação com essas funções robustas foi explorá-las em sistemas não lineares corrompidos com erros ou não, e poder avaliar as funções *New Target* e a *Alarm*, principalmente por apresentarem bom desempenho tanto em problemas de reconciliação de dados contaminados com erros grosseiros em sistemas lineares, quanto em problemas direcionados à detecção de *outliers*.

Esta dissertação está subdividida em cinco capítulos incluindo esta introdução. O segundo capítulo compreende uma concisa revisão sobre a reconciliação de dados (teoria e aplicações), tipos de erros em medidas de processo (aleatórios e grosseiros) e funções robustas, o qual, auxiliará na definição da metodologia empregada no terceiro capítulo. No quarto capítulo são apresentados os estudos de casos utilizados, bem como, os resultados obtidos e as discussões pertinentes. As principais conclusões e as sugestões para trabalhos futuros estão resumidas no quinto capítulo.

Com relação aos estudos de casos, procurou-se avaliar a capacidade de predição das funções robustas de diferentes famílias com relação aos erros sistemáticos das medidas. O primeiro caso estudado foi referente a um sistema linear de equações com três unidades operacionais onde todas as variáveis são medidas. O segundo caso demonstrou a importância das reações químicas com base na reação de cloração do metano. O terceiro caso foi um sistema não linear com balanço energético. O quarto caso representou um trocador de calor com válvula em *bypass* em um sistema linear de equações com erro grosseiro. O quinto caso é uma modificação do caso anterior, após inserção do balanço energético transformando-o em um sistema não linear de equações. O sexto tratou de um caso real de um processo de

polimerização, com base na descrição do modelo fenomenológico da produção de polipropileno em massa via catálise de Ziegler-Natta em um reator CSTR, com reciclo e purga.

O que pode ser verificado com relação às funções robustas, é que mesmo nas diferentes famílias estudadas, todas tiveram desempenho favorável na reconciliação dos dados. Entretanto, as funções *Alarm* e *New Target* que possuíam pontos de corte, apresentaram critérios de desempenho diferenciados das demais, tanto em relação ao erro médio relativo quanto ao número de iterações das funções objetivo. E dentre as funções da família dos M-estimadores a *Normal Contaminada* foi a que mais se ajustou às medidas corrompidas com erros.

1.1 OBJETIVO GERAL

O objetivo deste trabalho é avaliar a eficiência do uso de funções robustas no processo de reconciliação de dados de processos químicos e contribuir assim para o ajuste ótimo de medidas e estimativas (valores mapeados).

1.1.1 OBJETIVOS ESPECÍFICOS

- 1) Avaliar as funções robustas em medidas contendo erro e sem erro;
- 2) Contribuir com o desenvolvimento das funções robustas aplicando-as em sistemas não lineares;
- 3) Avaliar a predição método numérico (*SQP*) utilizado;
- 4) Comparar a eficiência das funções robustas em dois estudos de caso teóricos e um caso real.

2 REVISÃO DA LITERATURA

Este capítulo, dividido em 5 (cinco) seções, apresenta uma revisão bibliográfica referente ao estudo. A primeira define a técnica de reconciliação de dados através de um apanhado histórico. A segunda apresenta os métodos para a reconciliação de dados bem como os sistemas utilizados nesse estudo. A terceira representa a classificação de variáveis de um processo industrial. A quarta seção versa sobre os erros em medidas de processo e testes para detecção dos erros grosseiros. A quinta seção mostra a utilização de estimadores robustos para aplicação da reconciliação de dados.

2.1 RECONCILIAÇÃO DE DADOS

Reconciliação de dados possui diversas definições na literatura, a saber:

- (a) Segundo Kuehn e Davidson (1961), precursores da técnica, a reconciliação de dados é uma ferramenta que, dentre uma gama de aplicações, permite o ajuste ótimo de medidas e estimativas com base em redundância espacial e no modelo, geralmente balanços de massa e/ou energia.
- (b) É um procedimento de ajuste dos valores medidos de forma que o conjunto reconciliado obedeça às leis de conservação e demais restrições físicas impostas ao sistema (PRATA, 2010).
- (c) É um procedimento clássico que se baseia em acrescentar às leituras redundantes do processo a informação considerada perfeita, proveniente dos balanços de conservação de massa e energia, podendo ainda contemplar modelos cinéticos, termodinâmicos ou correlações em geral (BARBOSA, 2003).
- (d) É uma técnica usada para estimar os valores verdadeiros das variáveis medidas de um processo que satisfaçam ao modelo matemático que o representa; visa à redução da variância das variáveis do processo, aumentando a precisão e fechando a equação de balanço (MAGALHÃES, 2013).

A sua implementação leva em consideração um modelo estatístico para cada sistema de medição, por observar que plantas industriais possuem grandes quantidades de equipamentos responsáveis por essas medidas, causando assim redundância nas informações.

Para Benqlilou (2004), o sistema de reconciliação de dados permite:

- ▶ Aumentar a confiança na medição;
- ▶ Obter o valor mais provável para as variáveis não medidas (também conhecido como problema de cooptação) incluindo parâmetros de modelo;
- ▶ Detectar falhas em instrumentos chamados de desvios sistemáticos (*bias*) e no processo (vazamentos);
- ▶ Determinar a localização ótima de instrumentos de medição;
- ▶ Obter uma melhor caracterização do estado atual do processo, permitindo que se possa operar mais próximo da especificação;
- ▶ Estimar a eficiência dos equipamentos;
- ▶ Reduzir erros de modelagem;
- ▶ Justificar tarefas de manutenção (como por exemplo, calibração de instrumentos, limpeza de equipamentos, melhorias no processo, entre outras...).

A reconciliação de dados faz parte de uma abordagem de medições derivada do condicionamento de dados (FARIAS, 2009). A Figura 2.1 mostra a estratégia de condicionamento de dados e suas aplicações:

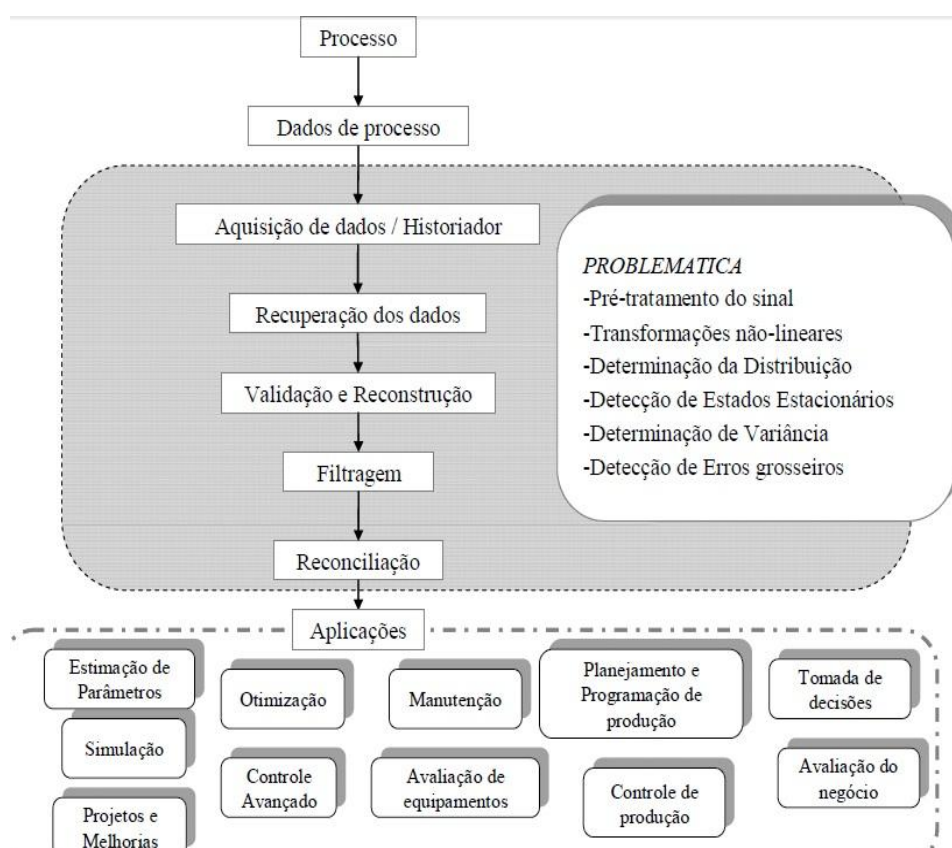


Figura 2.1 Condicionamento de dados e suas aplicações
Fonte: FARIAS, 2009.

A Figura 2.2 apresenta os conjuntos de dados de processos. Pode-se observar uma analogia com um “alvo” onde quanto mais próximo do centro do alvo esses dados estão, mais precisos e exatos eles serão.

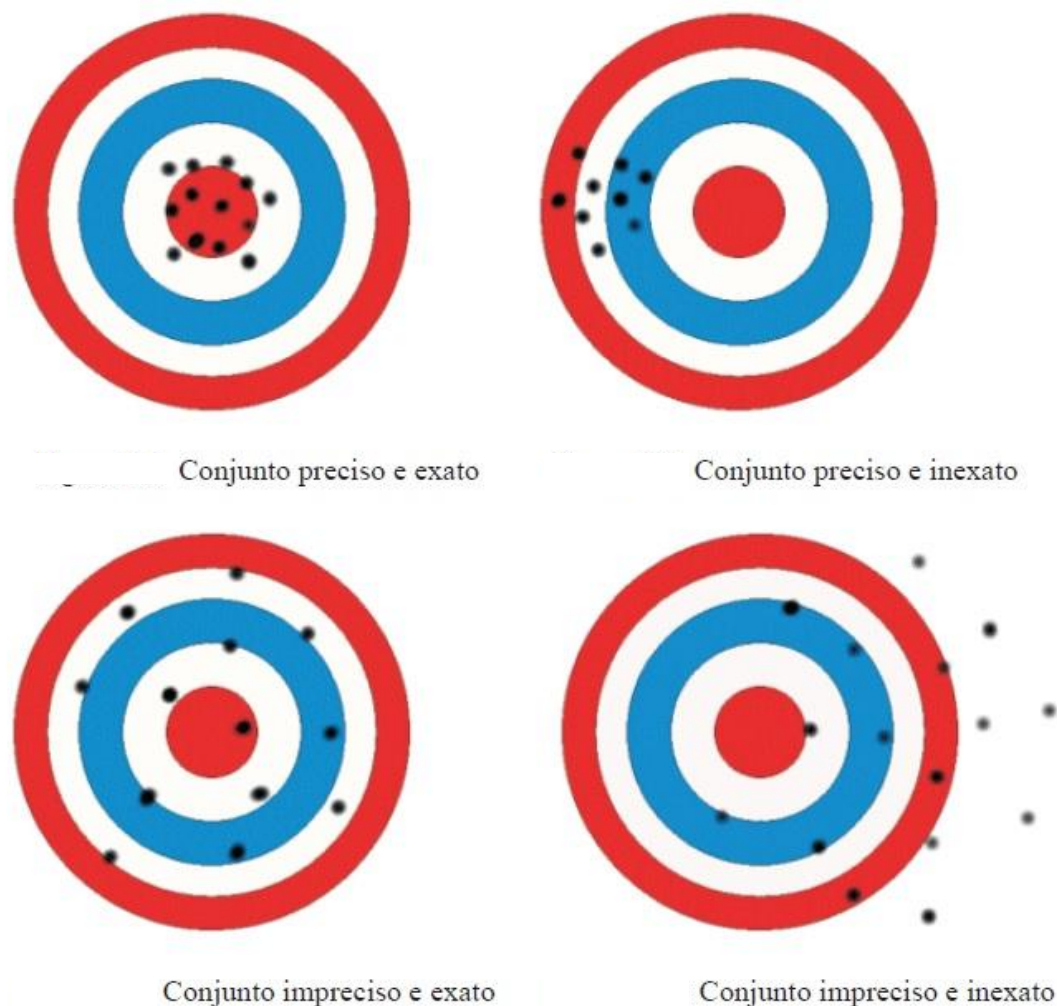


Figura 2.2 Características de conjuntos de dados
Fonte: GUINHO, 2003.

Com uma ferramenta desse tipo, abre-se a perspectiva real de implementar um sistema de gerenciamento de plantas químicas que produza melhores resultados, com um maior grau de confiança.

2.1.1 HISTÓRICO DA RECONCILIAÇÃO DE DADOS

Desde 1961, existem estudos sobre reconciliação de dados. Ao longo do tempo, eles foram aprimorados com a contribuição de diversos autores. Nesse histórico serão abordados apenas os trabalhos referentes aos métodos de resolução dos problemas de ajuste em medidas que é o foco principal desta dissertação.

A primeira solução proposta para o problema de reconciliação de dados em sistemas lineares foi apresentada por Kuehn e Davidson (1961), em cujo trabalho foi apresentada a solução para um sistema com todas as variáveis medidas.

Ripps (1965) ajustou os dados experimentais com base no fechamento dos balanços de massa e de energia em sistema estacionário, utilizando o método dos mínimos quadrados ponderados, que possuía balanços de massa nodais como condições, e como resolução do sistema aplicou os multiplicadores de Lagrange. Observou que quando há erros grosseiros em medidas, eles devem ser eliminados através de ajustes, através do chamado efeito da remoção de várias medidas, entretanto essa remoção deve ser feita uma de cada vez.

Vaclávek (1969a) apresentou um estudo com base em ajuste de sistemas complexos, entretanto, utilizou, assim como Ripps (1965), o método dos mínimos quadrados, e a solução através dos multiplicadores de Lagrange. Projetou o sistema de forma que a combinação dos dados medidos fosse coerente com as equações de balanço. Considerou somente o balanço de uma única quantidade, através do balanço total de massa ou o de apenas um dos componentes.

Vaclávek(1969b), em um estudo posterior, aprimorou os dados medidos através da derivação de um algoritmo que classificou as correntes medidas e as correntes não medidas, de acordo com as propriedades topológicas no esquema de balanço. Dispôs as medidas em dois grupos, o primeiro com medidas que podiam ter seu valor obtido através de outras correntes medidas sem necessidade de medição direta, e outro que somente teria valores através de medidas diretas. Definiu ainda o que chamou de esquema de balanço reduzido, que é a simplificação do balanço original do sistema.

Umedaet *al.* (1971) estudaram o ajuste de dados e a estimação de parâmetros em plantas químicas, com base no complexo método de Box para minimização da função objetivo, resolvendo o problema de otimização com várias restrições, que foram as expressões matemáticas das relações entre as variáveis do processo em sistema estacionário, aplicado em processos envolvendo dois reatores de tanque agitado contínuos (CSTR) e duas colunas de destilação com malha de reciclo.

Nogita (1972), assim com Ripps (1965) e Vaclavek (1969a,1969b), aplicou o método dos mínimos quadrados ponderados em estado estacionário na verificação de erros grosseiros. Foi apresentado um algoritmo para identificação dos erros grosseiros pela eliminação sequencial de valores medidos, onde avaliou o efeito da remoção de algumas variáveis na função-teste proposta. Considerou ainda que todas as variáveis eram medidas diretamente.

Murthy (1973,1974) contribuiu com os trabalhos anteriores ao utilizar o mesmo método para solução do problema de ajuste, entretanto, no primeiro trabalho utilizou um procedimento simples de ajustes das vazões molares de várias espécies e não exigiu qualquer conhecimento das reações químicas que ocorriam no reator. Já no segundo trabalho, com as reações já conhecidas e com restrições adicionais, como a irreversibilidade e inércia em algumas reações, foi formulado um problema de ajuste de dados. Se as reações independentes fossem iguais à quantidade máxima possível, ou seja, o mesmo que ser iguais ao número de espécies químicas menos a matriz característica dos elementos, e ainda fossem reversíveis, as correções seriam idênticas às que foram obtidas no primeiro estudo.

Almasy e Sztano (1975) basearam-se em uma estatística-teste para verificação de erros grosseiros usando a distribuição chi-quadrado que foi aplicada ao conjunto de dados como um todo, com matriz de dados conhecida, encontrando assim a fonte do erro grosseiro. Quando a matriz de dispersão do erro não era conhecida sugeriram aplicar o algoritmo para verificação de dados com base na matriz dispersão dos defeitos do modelo, sendo, portanto, considerado não existir uma matriz exata de dispersão dos erros, e sim uma aproximação.

Hlavacék (1977) relatou que os ajustes dos balanços multicomponentes eram mais complicados por se tratarem de restrições não-lineares. Sugeriu cinco modelos de solução: 1) substituição algébrica direta de variáveis das restrições na função objetivo; 2) minimização via multiplicadores de Lagrange; 3) solução direta por programação não-linear; 4) minimização por programação linear, e 5) critério minimax de Chebyshev.

Romagnoli e Stephanopoulos (1981), assim como Ripps (1965), estudaram a retificação de dados medidos contendo erros grosseiros. Apresentaram um método para identificar a fonte de erros através da eliminação seriada de uma ou mais medidas de cada vez. Utilizaram o método dos mínimos quadrados para a obtenção dos ajustes, tendo como restrições os balanços de massa nodais.

Mah (1981) apresentou uma revisão sobre estudos em diversas áreas relacionadas ao ajuste de dados de processo, envolvendo a reconciliação de dados, a classificação de

variáveis, a detecção e identificação de erros grosseiros e a definição de pontos de tomada de medidas que formam uma base para o projeto e análise de sistemas de monitoramento.

Stanley e Mah (1981a) apresentaram dois algoritmos para determinação das observabilidades local e global, da redundância de variáveis individuais e medições de processos baseadas em técnicas de análise teórico-gráfica.

Stanley e Mah (1981b) desenvolveram uma teoria sobre observabilidade e redundância aplicável a sistemas no estado estacionário, através de analogias com sistemas dinâmicos. Demonstraram a importância desses conceitos na predição do desempenho qualitativo de estimadores, como os mínimos quadrados ponderados e outros. Para restrições e medições lineares, as condições para observabilidade local também valeram para a observabilidade global, e o sistema foi subdividido em subsistemas (redundante, estritamente observável e não observável).

Keller (1992) propôs um algoritmo analítico para estimar a matriz de covariância dos erros de medida, que foi empregada na reconciliação de dados de processo, baseando-se nos resíduos de restrições lineares, calculados a partir dos dados disponíveis.

Bagajewicz (1996) abordou o problema da distribuição de probabilidade dos erros na reconciliação de dados. No caso de leituras como as de vazão, que implicariam em transformações não lineares sobre uma medida primária, mesmo que esta tivesse uma distribuição normal, a leitura final a distorceria e a formulação “clássica” tornar-se-ia ineficaz para captar a natureza dos erros nos procedimentos de reconciliação de dados.

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

Barbosa (2003) abordou a reconciliação de dados linear em estado estacionário com foco na melhoria da qualidade de informação sobre vazões mássicas que entravam e saíam da seção reator/regenerador do processo de craqueamento catalítico fluido (FCC), através do desenvolvimento de um *software* criado para reconciliar dados de processos químicos e petroquímicos. O sistema de testes de processos distribuído em rede se mostrou bastante útil nos testes de reconciliação de dados e é uma solução versátil para simular integração de processos em tempo real.

Pollo (2004) realizou um estudo no setor de utilidades de uma indústria petroquímica, especificamente em uma unidade de tratamento de água, visando minimizar a geração de efluentes e a redução dos custos envolvidos no processo. Foi aplicada a reconciliação de dados em um sistema hídrico, através do método de projeção matricial onde há o ajuste dos valores medidos e faz uma estimativa para aqueles valores não medidos.

Oliveira Jr. (2006), em seu trabalho de estimação de parâmetros representado por modelos de processos formados por um conjunto de equações algébrico-diferenciais, mostrou que os modelos de polimerização de propileno e o balanço hídrico de uma indústria de fertilizantes poderiam ser ajustados a partir de dados de processo, sem que nenhum dado experimental fosse gerado em laboratório. Foi desenvolvido um código de classificação de variáveis utilizando o método de Euler para o cálculo das derivadas. A estratégia computacional utilizada foi o método de Gauss-Newton para minimizar a função objetivo (Noronha, *et al.*, 1993). A função foi definida na forma da máxima verossimilhança.

Marques (2006) aplicou a reconciliação de dados na identificação e caracterização de balanços hídricos em plantas industriais. Também foi utilizado nesse estudo o pacote computacional desenvolvido por Noronha *et al.* (1993), como método de Gauss-Newton para minimizar a função objetivo que foi definida na forma da máxima verossimilhança onde o parâmetro de ponderação é a variância de cada medida. O objetivo do estudo foi efetuar uma análise do projeto em relação aos sensores previstos, avaliando a adequação com relação inicialmente à observabilidade e, em seguida, com relação à redundância.

Feldman (2007) aplicou a reconciliação de dados em tempo real para monitoramento e detecção de falhas no terminal de transporte e armazenamento de derivados de petróleo de Ribeirão Preto. Dada característica não linear do processo, aplicou o método de Euler explícito. A implementação da técnica em tempo real foi realizada utilizando a abordagem de janela de tempo, também chamada de janela móvel, que consiste a amostragem das variáveis de um processo em períodos regulares e no armazenamento desses dados na forma de uma fila, também conhecida por FIFO (*first in, first out*). Foi desenvolvido um software chamado RDTR (acrônimo de reconciliação de dados em tempo real). No problema de reconciliação foi utilizado o pacote computacional MAXIMA, baseado no algoritmo apresentado por Noronha *et al.* (1993), e implementado na linguagem de programação FORTRAN.

Ramis (2009) estudou a reconciliação dinâmica não linear, utilizando programação sequencial para resolver o problema da reconciliação de dados utilizando simuladores comerciais para representar o processo, considerando um modelo dinâmico em malha

fechada. Para o desenvolvimento do estudo, o autor aplicou três problemas de reconciliação de dados dinâmica não linear. Reconciliação de dados em um sistema modelo composto por uma linha com dois medidores de vazão mássica. Reconciliação de dados em um sistema modelo, composto por um separador bifásico, com medição de vazão mássica nas correntes de entrada e saída e medição de nível. Reconciliação de dados em um sistema real de condicionamento de gás combustível, para consumo em turbinas para geração de energia.

Poulin *et al.* (2010) desenvolveram critérios para avaliar o impacto de perturbações estacionárias em medidas de qualidade e, portanto, propor aplicabilidade de índices no método de reconciliação de dados para estado estacionário em plantas com perturbações. Um critério genérico com base na variância generalizada foi dado e outro critério foi proposto com base no desempenho da planta através de indicadores que reduzem a variância em direções de estado espacial. Os diversos índices de aplicabilidade foram ilustrados através de um exemplo de balanço material para um circuito de separação mineral simulado e a abordagem de estado estacionário foi comparado a um método ideal de reconciliação de dados em estado estacionário.

Prata *et al.* (2010) aplicou a reconciliação de dados em um reator de polimerização. Desenvolveu e implementou uma estratégia para monitorar em tempo real as propriedades finais e a produção de polipropileno em massa via catálise de Ziegler- Natta, em um reator tanque continuamente agitado com purga e reciclo. Através da reconciliação de dados dinâmica e não linear com estimação simultânea dos estados e parâmetros de interesse. O estudo foi realizado em conjunto com o pacote computacional ESTIMA (Noronha *et al.* 1993) que utiliza o método de Gauss-Newton para minimizar a função objetivo. Foi aplicada uma estratégia de solução sequencial, onde o sistema de equações algébrico-diferenciais foi resolvido por integração numérica através do código DASSL, sem necessidade de discretizar o sistema, e escolhendo a precisão de cada variável.

Jesus (2011) otimizou um processo industrial de produção de etileno em tempo real, considerando as dificuldades inerentes a aplicações industriais. Foram desenvolvidos modelos matemáticos para a planta de produção de etileno de uma unidade industrial situada em Duque de Caxias, Rio de Janeiro. Como procedimento de otimização em tempo real foi utilizada a estratégia estocástica de Enxame de Partículas. Para a reconciliação de dados foram utilizados dois modelos, sendo que o primeiro abrangeu um único forno, com o objetivo de identificar aspectos dinâmicos, e o forno de pirólise em operação, com o objetivo de otimizar o processo em tempo real.

Sarabia *et al.* (2012) apresentaram uma proposta de melhoria com base em dados industriais no funcionamento do abastecimento de redes de hidrogênio em uma refinaria de petróleo da Petronor do grupo espanhol Repsol. Os autores projetaram uma estratégia visando uma otimização integrada baseada em quadros com os produtos disponíveis para as instalações de consumo. Programaram essa otimização com linguagem em GAMS e abordagem sequencial para tempos estabelecidos. Com um grande sistema de distribuição, a solução do problema foi através de um sistema não linear que visou à redução do custo-utilidade da rede de hidrogênio.

Garcia (2013) utilizou através de uma modelagem simplificada do processo em escala industrial de uma coluna despropenizadora em uma refinaria em Mauá, SP. Para a plataforma de modelagem foi escolhido o simulador EMSO (Soares e Secchi, 2003), que é um simulador orientado a equações com características adequadas para a implementação da otimização em tempo real (RTO). Foram simulados dados e posteriormente reconciliados para validar três métodos escolhidos: o teste de variância de janela de tempo, o teste F modificado e o método de *wavelet*.

Maradiaga *et al.* (2013) reconciliaram dados em estado estacionário para sistemas de refrigeração por absorção analisando os graus de liberdade e tratamento de erros grosseiros através do teste MIMT que é a reconciliação de dados clássica com a estratégia de detecção de erros utilizando mínimos quadrados como solução do problema. Com linguagem de programação em Matlab[®], obteve melhoria nos balanços de massa e energia através do modelo de absorção por resfriamento.

Chen *et al.* (2013) através do balanço de massa e energia visou construir um modelo confiável para o monitoramento de processos aplicando um estimador robusto para a reconciliação de dados, que é também um critério do erro quadrático médio tradicional. O valor ideal da densidade *Kernel*¹, e foi escolhido através da minimização da informação do critério de *Akaike*² (AIC).

Dois casos foram estudados: (i) o primeiro utilizando os benefícios de um estimador robusto comparando com vários outros como função *Fair*, estimador redescendente, matriz quadrada quasi-ponderada (QWLS) e outros; (ii) o segundo, incluiu variáveis com valores não

¹Densidade Kernel: Medida onde o tamanho dos núcleos utilizados nas estimativas é variado dependendo da localização ou das amostras ou a localização do ponto de medição. É uma técnica eficaz quando o espaço de amostragem é multidimensional.

²Critério de Akaike (AIC): É uma medida da qualidade relativa de um modelo estatístico para um determinado conjunto de dados, fornecendo assim, um meio para seleção do modelo.

medidos e erros aleatórios com distribuição não Gaussiana. O resultado obtido mostrou que o estimador chamado de *correntropy* apresentou melhor estimação quando comparado aos outros.

Ghosh *et al.* (2014) propuseram um método baseado em otimização estocástica para identificar um subconjunto ótimo de medidas de monitoramento para variáveis de processo. Esse monitoramento efetivo utilizando o PCA foi baseado no problema de referência “Tennessee Eastman Challenge³”, e explorou a importância da classificação de variáveis como determinante no monitoramento de desempenho de uma planta. Ao utilizar o PCA (método baseado em análise de componentes principais) a classificação incidiria apenas nas variáveis mais relevantes, para isso a otimização estocástica foi escolhida para essa seleção.

Jiang *et al.* (2014) reconciliaram dados do monitoramento do desempenho *on-line* de turbinas a vapor em unidade geradora de energia. O objetivo do trabalho foi reduzir a incerteza nas medições dos fluxos e nas taxas de vapor das turbinas. Aplicou o teste global para garantir que erros grosseiros não afetassem as medidas do processo para aplicação mais eficaz da reconciliação de dados. Observou ainda o impacto da reconciliação de dados sobre os coeficientes de sensibilidade da taxa de calor em relação aos diferentes parâmetros medidos.

Alguns autores aplicaram a reconciliação de dados para analisar perdas operacionais em sistemas de vapor, Mavromatis e Kokossis (1998) na busca por uma metodologia para otimização de redes de turbinas de vapor a partir de conjunto de regras lógicas baseadas nas relações de Willans, relacionou as vazões de vapor extraído com a sua potência gerada. Já Hirata *et al.* (2003) buscou o aproveitamento em duas plantas petroquímicas que se fundiram em 1994 para interligar o sistema de vapor com o de energia elétrica. Onde utilizaram uma interface gráfica, e *solver* OSL via programação matemática (GAMS). Incluíram parâmetros que envolviam carga de produção, temperatura ambiente, *status* de parada ou partida de equipamentos, resgataram dados do histórico da planta.

³Tennessee Eastman Challenge: Proposto por Downs e Vogel de Fortran para Matlab®. É um código escrito exclusivamente em comandos do Matlab® e é fornecida uma interface através de uma aplicação na função S do Simulink®.

2.2 MÉTODOS PARA A RECONCILIAÇÃO DE DADOS

O conceito de reconciliação de dados foi inicialmente proposto para calcular o balanço de massa para processos em estado estacionário (Kuehn e Davidson, 1961; Mah *et al.*, 1976; Hodouin e Everell, 1980; Crowe *et al.*, 1983; Crowe, 1986). O interesse de aplicar a técnica de reconciliação de dados em plantas químicas e metalúrgicas começou na década de 1980. Gerentes de uma planta industrial realizaram a reconciliação de dados e observaram que poderiam transformar os dados brutos em informações mais consistentes e confiáveis, e que isso seria fundamental para a operação e gerenciamento da planta.

Atualmente, a reconciliação de dados em processos em estado estacionário é uma técnica amadurecida, e têm sido muito bem aplicada em química, bioquímica, petroquímica, celulose e papel e indústrias de processamento de minerais. O estado estacionário é uma suposição idealística para plantas reais, onde as condições do processo são continuamente submetidas a variações (BAI e THIBAUT, 2010).

2.2.1 RECONCILIAÇÃO DE DADOS EM ESTADO ESTACIONÁRIO

A reconciliação de dados em estado estacionário é uma técnica usada para melhorar a exatidão e a precisão das medidas reduzindo o impacto de erros aleatórios medidos como também de erros grosseiros em dados de processo em estado estacionário. A concepção de reconciliação de dados foi inicialmente proposta por Kuehn e Davidson (1961), e depois deles, foi exaustivamente estudada por Mah *et al.* (1976), Stanley e Mah (1977), Hodouin e Everell (1980), Romagnoli e Stephanopoulos (1981), Crowe *et al.* (1983), Jordache *et al.* (1985), Tamhane e Mah (1985), Serth e Heenan (1986), Crowe (1983, 1986, 1989), Narashimhan and Mah (1987), Rosenberg *et al.* (1987), Tjoa e Biegler (1991), Rollins e Davis (1992), Tong e Crowe (1995) e Sanchez e Romagnoli (1996). Para processos em estado estacionário, em geral, os valores reconciliados ou estimados das variáveis do processo são obtidos minimizando-se uma restrição, em geral mínimos quadrados ponderados, sujeito a restrições de igualdade e desigualdade, representado pela Equação (2.1) (BAI e THIBAUT, 2010):

$$\begin{aligned} \text{Min } J(\hat{\mathbf{y}}, \hat{\mathbf{z}}) &= (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) \\ \text{sujeito às restrições: } \mathbf{f}(\hat{\mathbf{y}}, \hat{\mathbf{z}}) &= \mathbf{0} \\ \mathbf{g}(\hat{\mathbf{y}}, \hat{\mathbf{z}}) &\geq \mathbf{0} \end{aligned} \quad (2.1)$$

Sendo $\mathbf{y}$ um vetor de variáveis medidas no processo, tendo como elementos valores médios de medidas brutas ao longo de um período de tempo, por exemplo, período de uma hora, ou quaisquer medidas em qualquer tempo de amostragem para os processos de estado estacionário. $\hat{\mathbf{y}}$ é um vetor de valores reconciliados para as variáveis do processo, $\hat{\mathbf{z}}$ é um vetor de estimativas de variáveis não medidas do processo ou parâmetros do modelo, $\mathbf{V}$ é a matriz de covariância das medidas no processo de variáveis medidas, $\mathbf{f}$ é um vetor função em uma restrição com modelo de igualdade, e $\mathbf{g}$ é um vetor função do modelo de restrição de desigualdade incluindo limites simples inferiores e superiores para $\hat{\mathbf{y}}$ e $\hat{\mathbf{z}}$. A Equação (2.1) $(\mathbf{y} - \hat{\mathbf{y}})^T$ denota a transposta de $(\mathbf{y} - \hat{\mathbf{y}})$, e $\mathbf{V}^{-1}$ denota a matriz inversa de $\mathbf{V}$.

O exemplo mais simples de reconciliação de dados em estado estacionário é para reconciliação das taxas de fluxo em uma planta onde todos os fluxos são medidos. Essas taxas de fluxo devem satisfazer os balanços de massa em torno dessa unidade, e são obtidas através da Equação (2.2):

$$\begin{aligned} \text{Min } J(\hat{\mathbf{y}}) &= (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) \\ \text{sujeito à restrição } \mathbf{A}\hat{\mathbf{y}} &= \mathbf{0} \end{aligned} \quad (2.2)$$

sendo $\mathbf{A}$ a matriz de incidência da planta. Cada linha da matriz de incidência representa uma unidade particular, isto é um nó na planta, e cada coluna representa um dado fluxo, respectivamente. Cada elemento em $\mathbf{A}$ ou é 1, ou -1 ou 0 dependendo se a corrente do fluxo é um fluxo de entrada, um fluxo de saída ou que não está associada a esse nó, respectivamente.

Usando-se o método dos multiplicadores de *Lagrange*, o problema de otimização da Equação (2.2) pode ser transformado para resolução do problema de otimização representado pela Equação (2.3):

$$\text{Min } J(\hat{\mathbf{y}}) = (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) - 2\boldsymbol{\lambda}^T \mathbf{A}\hat{\mathbf{y}} \quad (2.3)$$

sendo $\boldsymbol{\lambda}$ é o vetor multiplicador de *Lagrange*. As condições necessárias de mínimo para a Equação (2.3) são:

$$\frac{\partial J}{\partial \hat{\mathbf{y}}} = -2 \mathbf{V}^{-1}(\mathbf{y} - \hat{\mathbf{y}}) - 2\mathbf{A}^T \boldsymbol{\lambda} = \mathbf{0} \quad (2.4)$$

$$\frac{\partial J}{\partial \boldsymbol{\lambda}} = -2\mathbf{A}\hat{\mathbf{y}} = \mathbf{0} \quad (2.5)$$

Multiplicando-se cada termo pela matriz $\mathbf{V}$ na Equação (2.4) permite-nos escrever:

$$\mathbf{y} - \hat{\mathbf{y}} + \mathbf{V}\mathbf{A}^T \boldsymbol{\lambda} = \mathbf{0} \quad (2.6)$$

Novamente, multiplicando-se cada termo pela matriz $\mathbf{A}$ na Equação (2.6) e aplicando-se $\mathbf{A}\hat{\mathbf{y}} = \mathbf{0}$ obtém-se:

$$\mathbf{A}\mathbf{y} + \mathbf{A}\mathbf{V}\mathbf{A}^T \boldsymbol{\lambda} = \mathbf{0} \quad (2.7)$$

Rearranjando-se a Equação (2.7) é possível obter-se o multiplicador de *Lagrange* a partir de:

$$\boldsymbol{\lambda} = -(\mathbf{A}\mathbf{V}\mathbf{A}^T)^{-1} \mathbf{A}\mathbf{y} \quad (2.8)$$

Colocando-se a Equação (2.8) em (2.6) obtém-se a solução final da reconciliação em estado estacionário com todas as correntes de fluxos medidos.

$$\hat{\mathbf{y}} = \mathbf{y} - \mathbf{V}\mathbf{A}^T((\mathbf{A}\mathbf{V}\mathbf{A}^T)^{-1} \mathbf{A}\mathbf{y}) \quad (2.9)$$

Segundo Pinto e Lage (2001), sistemas lineares podem ter quatro tipos de solução: (i) uma única solução não trivial, (ii) nenhuma solução, (iii) um número infinito de soluções e (iv) a solução trivial ($\mathbf{x}=\mathbf{0}$). Os sistemas de uma única solução são os mais usuais nas soluções das equações matemáticas de um dado modelo físico.

Os sistemas lineares podem ser resolvidos através de dois tipos de métodos: os métodos diretos e os métodos iterativos.

Os métodos diretos consistem em procedimentos, baseados na eliminação algébrica, que obtêm a solução exata em um número fixo de operações. Entre estes métodos, encontram-

se: (i) a eliminação Gaussiana, (ii) a eliminação de Gauss-Jordan, (iii) a inversão matricial, (iv) a decomposição LU e outros métodos derivados para matrizes de coeficientes de formas especiais, como o algoritmo de Thomas (PINTO E LAGE, 2001).

Os métodos iterativos geram uma solução aproximada para o sistema de equações utilizando um procedimento iterativo a partir de uma solução aproximada inicial (“chute”). Exemplos: (i) iteração de Jacobi, (ii) iteração de Gauss-Seidel, (iii) sobre-relaxação sucessiva (*successive over relaxation*, SOR), (iv) relaxação por linhas (*line successive over relaxation*, LSOR), (v) método implícito de direção alternada (*Alternating-Direction Implicit*, ADI), além de outros métodos específicos para matrizes de coeficientes com estruturas especiais, como o método fortemente implícito modificado (*Modified Strong Implicit Procedure*, MSIP) (PINTO E LAGE, 2001).

Narashiman (1999) explica que em sistemas lineares, é assumido que os erros aleatórios medidos seguem uma distribuição normal com média zero e variância-covariância Σ .

Para Bai e Thibault (2010), uma abordagem de estado estacionário em sistemas não lineares aplica-se à linearização sucessiva. A ideia reside no fato de que modelos não lineares $f(\hat{y}, \hat{z}) = \mathbf{0}$ podem ser linearizados utilizando a primeira ordem da série de *Taylor* em torno das variáveis estimadas. Os modelos linearizados são considerados como restrições de equações lineares, transformando o problema de reconciliação de dados não linear em um problema de reconciliação de dados linear. Em seguida, o problema de reconciliação de dados linear é resolvido, e os resultados de $\hat{y}$ e $\hat{z}$ são as novas estimativas para linearizar os modelos não lineares novamente. A reconciliação de dados linear é resolvida repetidamente até os desvios dessas estimativas de suas últimas iterações, venham a convergir dentro de algumas tolerâncias especificadas.

A abordagem da linearização sucessiva pode ser aplicada quando não há restrições de desigualdade impostas ao processo de reconciliação de dados. Com restrições de desigualdade, por exemplo, $g(\hat{y}, \hat{z}) \geq \mathbf{0}$, um problema de reconciliação de dados não linear tem que ser resolvido por meio de técnicas de programação não linear (NLP).

Ao contrário de sistemas de equações lineares, que admitem apenas uma única solução, sistemas não lineares de equações podem apresentar soluções diferentes para um mesmo conjunto de equações.

Um problema geral de reconciliação de dados em sistemas não lineares pode ser formulado como um problema de minimização com uma função de mínimos quadrados conforme a apresentada na Equação (2.10) abaixo:

$$\mathbf{Min}_{\mathbf{x},\mathbf{u}}(\mathbf{y} - \mathbf{x})^T \Sigma^{-1}(\mathbf{y} - \mathbf{x}) \quad (2.10)$$

$$\text{Sujeito a } \mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0} \quad (2.11)$$

$$\mathbf{g}(\mathbf{x}, \mathbf{u}) \leq \mathbf{0} \quad (2.12)$$

Sendo:

$\mathbf{f}$: m x 1 vetor de restrições de igualdade;

$\mathbf{g}$: q x 1 vetor de restrições de desigualdade;

Σ : n x n matriz de variância-covariância;

$\mathbf{u}$: p x 1 vetor de variáveis não medidas;

$\mathbf{x}$: n x 1 vetor de variáveis medidas;

$\mathbf{y}$: n x 1 vetor de valores medidos da variável $\mathbf{x}$.

As restrições de igualdade definidas pela Equação (2.11) incluem todas as relações de conservação material e energética e restrições de equilíbrio termodinâmico. Já as restrições de desigualdades da Equação (2.12) podem ser relacionadas aos limites superiores e inferiores das variáveis ou restrições de viabilidade complexas relacionadas à operação de equipamentos.

Para a resolução de sistemas não lineares operando em estado estacionário com restrições de desigualdade como as expostas na Equação (2.12), alguns métodos podem ser usados. Para esse estudo foram utilizados a Técnica Sequencial de Programação Quadrática – *Sequential Quadratic Programming Technique* (SQP) e o Gauss-Newton (Ver Apêndices A e B). Ambos possuem a similaridade de resolver funções objetivo de natureza quadrática através de estimativas de máxima verossimilhança (MLE).

Essas restrições de desigualdade em variáveis medidas e não medidas possuem a seguinte forma:

$$\mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max} \quad (2.13)$$

$$\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max} \quad (2.14)$$

2.3 CLASSIFICAÇÃO DE VARIÁVEIS

Uma das etapas da técnica de reconciliação de dados é a classificação de variáveis. Ela consiste no estudo do modelo de cada processo através da importância de cada parâmetro.

Os precursores do estudo sobre classificação de variáveis foram Vaclavék e Loucka (1976) que, a fim de objetivar as variáveis de um processo, fizeram a descrição delas, em redundantes e observáveis. Depois disso alguns autores publicaram trabalhos com o mesmo intuito como Stanley e Mah (1981), Crowe (1989).

De acordo com Bai e Thibault (2010), variáveis de processo descrevem as características de uma planta. O número de variáveis de processo em uma planta pode ser da ordem de centenas ou milhares dependendo da complexidade da planta. Entretanto, nem todas as variáveis de processo são medidas como devido por razões econômicas e técnicas. Muitas variáveis críticas são normalmente selecionadas e medidas, e são chamadas de variáveis medidas do processo. E outras variáveis que não são selecionadas são chamadas de variáveis não medidas de processo.

Uma variável redundante é aquela medida cujo valor pode ser calculado por outra variável medida através do modelo de processo em adição à sua própria medida. Uma variável não redundante é aquela medida que não pode ser calculada por outras, somente por suas próprias medições. Para Bagajewicz (2001), a medição de uma variável é não redundante se após a remoção desta medição a variável não é observável, isto é, não pode ser calculado usando uma equação de equilíbrio envolvendo as outras medições do sistema.

A redundância é, portanto, uma propriedade desejável em um sistema, porque, caso um instrumento de medida venha falhar, a sua variável pode ser estimada por outros balanços. Além disso, é importante salientar que se o número de diferentes balanços que podem ser utilizados aumenta, haverá formas adicionais para o cálculo dessa variável. Quando mais medições redundantes estiverem envolvidas no cálculo das variáveis, é dito que a importância de tal sistema e a sua precisão de estimativa aumentam.

Uma variável observável é aquela não medida que pode ser calculada através de variáveis medidas ligada ao modelo de processo. Uma variável não observável é aquela que não está disponível qualquer informação no processo.

Segundo Oliveira Jr. (2011), o problema de otimização com restrição para ajuste das medidas é resolvido em intervalo delimitado. Dessa forma, classificar variáveis de um

processo qualquer, depende da análise de dados sobre fluxogramas complexos e segue as seguintes observações:

- 1) O fluxograma será apresentado de forma que, os NÓS representem divisores, misturadores e operações (Ver Figura 2.3);

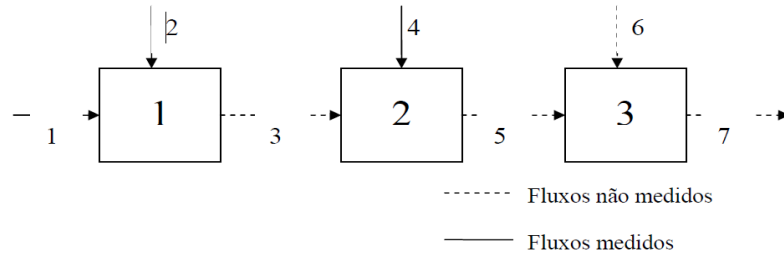


Figura 2.3 Fluxograma representativo

- 2) A identificabilidade deverá analisar todas as variáveis pertencentes às equações do processo;
 x_1 , x_2 e x_4 são variáveis medidas;
 x_3 , x_5 , x_6 e x_7 são variáveis não medidas.
- 3) Conceber equações para cada um dos NÓS do fluxograma, sempre relacionadas às variáveis em estudo;

$$x_1 + x_2 = x_3 \quad (2.15)$$

$$x_3 + x_4 = x_5 \quad (2.16)$$

$$x_5 + x_6 = x_7 \quad (2.17)$$

- 4) A topologia das equações do modelo seja representada por uma matriz de ocorrência (**A**) cujas linhas da matriz correspondam às equações e as colunas correspondam às variáveis de processo, tanto medidas como não medidas;

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Figura 2.4 Matriz representativa

5) E por fim, pelo algoritmo de classificação obtido da seguinte maneira:

- A coluna referente a variável medida x_n é zerada, de forma que $A_{in} = 0$ para todo i .

Repete-se para cada uma das variáveis medidas.

- Identificam-se as linhas de zero da matriz A . As linhas indicam equações redundantes que contém apenas variáveis medidas.

- Aplicam-se as técnicas de comparação para identificação dos conjuntos de n equações e n incógnitas (n linhas de A que apresentem as mesmas n ocorrências), partindo de $n=2$. Se realizada corretamente, as n variáveis identificadas são ditas observáveis, caso contrário, as variáveis remanescentes em A são ditas não observáveis.

Portanto, a eficiência da reconciliação de dados depende em muitos aspectos de uma classificação de variáveis correta, e de instrumentos para medição em perfeito estado, porque uma variável sendo caracterizada como medida, que tenha sido mal controlada, acarretará em sérios problemas à técnica de reconciliação de dados.

2.4 ERROS EM MEDIDAS DE PROCESSO

O uso de computadores não só permite que os dados obtidos cheguem a uma melhor frequência, mas também resultam na eliminação de erros presentes em medições manuais. Isto em si tem melhorado muito a precisão e validade dos dados do processo. O aumento da quantidade de informações, no entanto, deve ser explorada para melhorar ainda mais a precisão e a consistência dos dados do processo através de verificação e tratamento desses dados sistemáticos.

Medições de processo são inevitavelmente corrompidas por erros durante medidas, processamento e transmissão do sinal medido. O erro total de medição, que é a diferença entre o valor medido e o valor verdadeiro de uma variável, pode ser representado como a soma das contribuições de dois tipos de erros, erros aleatórios e erros grosseiros (BAI e THIBAUT, 2010).

2.4.1 ERROS ALEATÓRIOS

Ainda segundo Bai e Thibault (2010), observa-se geralmente que, se a medição de uma variável de processo é repetida sob condições idênticas, o mesmo valor de medida não é obtido. Isto é devido à presença de erros aleatórios nas medidas. Erros aleatórios não podem ser previstos nem explicados com precisão.

Erros aleatórios implicam que a magnitude e o sinal do erro de cada repetição não podem ser previstos, sendo a mesma caracterizada por uma variável aleatória a qual segue uma função de densidade probabilidade (FDP). Em condições normais, estes tipos de erros não podem ser removidos e não são desejáveis, porém, é possível aplicar técnicas de filtragem por equipamentos ou por *softwares*.

Assim, a relação entre o valor medido, o valor verdadeiro e o erro aleatório em uma variável medida é expresso na Equação (2.18);

$$y = x + \varepsilon \quad (2.18)$$

Sendo y o valor medido, x o valor verdadeiro e ε o erro aleatório. Erros aleatórios geralmente oscilam em torno de zero. Seu valor médio (ou esperança) pode ser dado pela Equação (2.19),

$$E(\varepsilon) = 0 \quad (2.19)$$

e sua variância

$$\text{var}(\varepsilon) = E[\varepsilon^2] = \sigma^2 \quad (2.20)$$

Sendo σ o desvio padrão do erro medido.

O desvio padrão é uma medida da precisão da medição. Quanto menor for o desvio padrão, mais precisa é a medição e maior a probabilidade de que o erro aleatório será próximo de zero.

Se os erros aleatórios nas medições de duas variáveis diferentes i e j forem considerados estatisticamente independentes, então eles devem possuir correlação nula, o que é representado pela Equação (2.21):

$$\text{cov}(\varepsilon_i, \varepsilon_j) = E[\varepsilon_i, \varepsilon_j] = 0 \quad (2.21)$$

2.4.2 ERROS GROSSEIROS

Normalmente erros grosseiros são associados a falhas de sensores. As falhas em instrumentos de medição estão associadas a: *bias*⁴, falha completa, derivações e degradação da precisão.

Os erros grosseiros são causados por eventos não aleatórios, tais como mau funcionamento de instrumentos (devido à instalação inadequada dos dispositivos de medição), descalibração, desgaste ou corrosão dos sensores, e incrustações. A natureza não aleatória desses erros implica que, em qualquer momento possuam certa magnitude e dão sinal que pode ser desconhecida. Assim, se a medição é repetida com o mesmo instrumento em condições idênticas, a contribuição de um erro grosseiro sistemático para um valor medido será sempre o mesmo.

Detectar e tratar os erros grosseiros tem sido o foco de diversos trabalhos sobre reconciliação de dados. Normalmente estes erros são classificados em dois tipos: erro grosseiro do tipo espúrio (*outlier*⁵) e erro grosseiro do tipo sistemático (*bias*).

⁴Bias: Valores consistentemente muito altos ou baixos das medidas em relação aos valores reais. O valor médio do erro não é nulo. Podem derivar da instalação ou calibração incorreta dos instrumentos de medidas (LIEBMAN *et al.* 1992, McBRAYER e EDGARD, 1995).

⁵Outliers: Medidas obtidas a partir de algum comportamento anormal, resultado de “picos” de processo ou distúrbios não medidos. É uma medida que não segue qualquer distribuição estatística representativa do conjunto dos dados (CHEN e ROMANGNOLI, 1998).

Em processos industriais, deseja-se detectar os efeitos dos erros grosseiros para acionar procedimentos de manutenção e manter as informações do processo sempre em níveis confiáveis.

Se existe erro grosseiro em um valor medido, ele pode ser mensurado através da Equação (2.22):

$$y = x + \varepsilon + \delta \quad (2.22)$$

onde δ é a magnitude desse erro grosseiro.

Erros grosseiros afetam significativamente a precisão de qualquer aplicação industrial através dos dados de processo. Eles têm que ser detectados e removidos. A reconciliação de dados é a técnica mais apropriada na maioria dos casos.

Alguns estudos recentes sobre erros grosseiros serão citados para fins de referencial teórico.

Maronna e Arcas (2009) desenvolveram o problema de reconciliação linear através de um modelo de regressão. Visando obter a máxima probabilidade de identificar os erros tipo (*outliers*) corretamente.

Sun *et al.* (2011) utilizaram dados históricos que ficavam restritos a banco de dados para melhorar a eficiência na detecção de erros grosseiros e reconciliação de dados. Incluíram uma nova estratégia com dois passos. O primeiro foi a utilização de informações através de um programa que usa técnicas mistas antes de detectar o erro grosseiro. O segundo passo foi estimar todos os erros grosseiros detectados e ajustar os dados do processo com os balanços material, e de energia e outras restrições.

Gonzalez *et al.* (2012) desenvolveram um algoritmo não linear para detecção e estimativa sistemática de erros grosseiros em balanço de massa e energia nos dados medidos de areias betuminosas. Através de uma detecção de erro grosseiro em tempo real utilizando o método *on-line* Bayesiano Dinâmico com um tipo de comutação com o Filtro de Kalman.

Nicholson *et al.* (2014) apresentaram a integração direta de modelos dinâmicos não lineares em um problema de otimização bem definido. Utilizando uma estimação por horizonte móvel avançado, que demonstrou um bom desempenho quando o ruído de medição segue uma distribuição gaussiana. O MHE considera medidas que estão “contaminadas” (corrompidas) com grandes erros. Aplicaram dois estimadores robustos, função *Huber's fair* e *Hampel's* redescendente em dois estudos de casos dinâmicos, (i) reator CSTR; (ii) coluna de

estilação. Obtiveram melhores resultados utilizando o estimador redescendente principalmente quando existiam grandes erros grosseiros nas medidas.

2.4.3 DETECÇÃO DE ERROS GROSSEIROS

Segundo Bai e Thibault (2010), várias técnicas podem ser usadas para detectar e eliminar os dois tipos de erros grosseiros. Qualquer estratégia para detecção de erros grosseiros deve preferencialmente possuir as seguintes habilidades:

- ▶ Habilidade de detectar a presença de um ou mais erros grosseiros nos dados **(problema de detecção)**;
- ▶ Habilidade de identificar o tipo e localização do erro grosseiro **(problema de identificação)**;
- ▶ Habilidade de localizar e identificar múltiplos erros grosseiros aqueles que podem estar presentes simultaneamente nos dados **(problema de identificação de múltiplos erros grosseiros)**;
- ▶ Habilidade de estimar a magnitude do erro grosseiro **(problema de estimação)**.

As técnicas de detecção de erros grosseiros utilizam-se da aplicação de testes estatísticos no conjunto de dados, onde se determina se a medida segue uma distribuição com média zero. Dessa forma, é aplicado um teste de hipótese na média dos dados, sendo:

H_0 : Hipótese nula: $\mu=0$

H_1 : Hipótese alternativa: $\mu \neq 0$

que o teste segue a Equação (2.23):

$$t = \frac{\hat{E}(y_i - x_i)}{\hat{\sigma}} \quad (2.23)$$

Este teste é a base para todos os procedimentos de detecção de erros grosseiros (DEG) clássicos. Se uma função densidade de probabilidade pode ser considerada para t , grandes valores de t descreverão situações menos esperadas e darão a prova de que H_1 seja verdadeira, o que prova que existam erros grosseiros.

O primeiro caso de detecção de erros grosseiros foi apresentado por Reilly e Carpani (1963). Foram apresentados dois dos mais populares testes de detecção de erros grosseiros, o teste global (GT) e o teste nodal (NT), baseados nos resíduos do modelo de restrição.

Teste Global – determina se todo o conjunto de dados segue a distribuição normal comparando a variância dos resíduos das restrições com a variância esperada.

Teste Nodal – avalia o quão bem as medições satisfazem cada uma das restrições utilizando os resíduos destas para detectar a existência de erros grosseiros.

A partir desses, surgiram outros testes buscando corrigir algumas desvantagens que os primeiros apresentavam. Almasy e Sztano (1975) estudaram, para sistemas com restrições lineares com todas as vazões medidas, o teste da máxima potência para detecção de um único erro grosseiro nas medidas quando a variância era conhecida. Este teste foi apresentado como tendo a maior probabilidade de detectar corretamente o erro grosseiro quando somente um desse está presente nas medidas, visto que, quando são realizados testes múltiplos, o intervalo de confiança é estendido aumentando assim a região de rejeição do teste estatístico.

Mah e Tamhane (1982) propuseram o teste da medição, um dos métodos mais populares para DEG. Baseia-se no teste de hipótese de *Neyman-Pearson*, usando os resíduos entre as medições e os seus valores estimados durante a reconciliação de dados. O teste da medição trata da principal desvantagem dos testes GT e NT, requer que os dados sejam reconciliados antes que seja feita a detecção de erros.

Madron (1985) aplicou o conceito de credibilidade da medida pela atribuição de um valor máximo de magnitude do erro grosseiro. Narasimhan e Mah (1987) utilizaram o conceito de razão da máxima verossimilhança generalizada (*generalized likelihood ratio test* – GLR). Apresentaram um modelo para medição e um para restrição, que avaliava a presença de diferentes tipos de erro grosseiro. Harikumar e Narasimhan (1993) estudaram a imposição de limites para valores estimados visando a adequação das falhas dos testes GT e NT. Tong e Crowe (1995) propuseram o teste baseado em Análise de Componentes Principais (PCA).

Um exemplo clássico da aplicação na estratégia de detecção de erros grosseiros foi apresentado por Mah (1990) e também por Knopf (2012) onde um fluxograma operacional de três unidades mostrado na Figura 2.5, foi estudado para detecção de erros grosseiros.

Esse sistema composto por misturador, reator e vaso separador possui 5 (cinco) variáveis sendo todas medidas com 1(uma) corrente sendo de reciclo. Para análise dessas correntes foi considerado pelos autores que o balanço energético devia ser desconsiderado e apenas o balanço mássico seria relevante para solução do sistema.

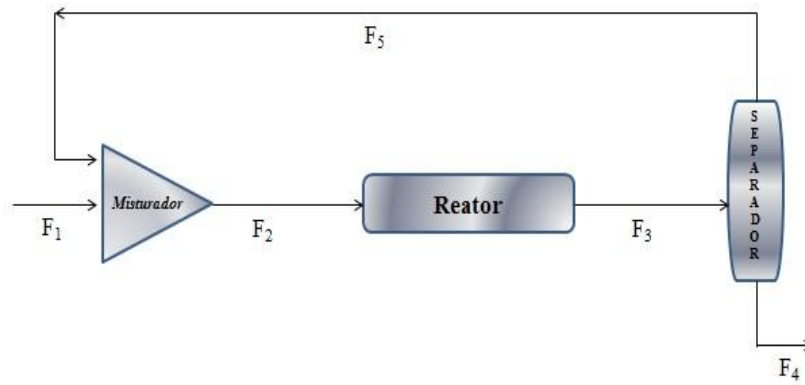


Figura 2.5 Fluxograma operacional de três unidades

A Tabela 2.1 mostra os dados de operação dessa unidade e também o desvio padrão de cada medida.

Tabela 2.1 Dados medidos da unidade

	Corrente 1	Corrente 2	Corrente 3	Corrente 4	Corrente 5
Fluxo medido	99.3	123.2	125.1	100.5	23.9
Σ medido	2.0	2.0	2.0	2.0	2.0

O estudo se baseou na aplicação do teste global para verificar se nesse sistema erros grosseiros eram presentes.

Considerando que o problema de reconciliação de dados é dado da seguinte forma:

$$\text{Minimizando } f = \sum_{i, \text{variáveis medidas}} \left(\frac{(x_i^+ - x_i)}{\sigma_i} \right)^2 \quad (2.24)$$

Sujeito a $h_k(x_i) = 0 \quad k = 1, \dots, K$.

A função objetivo da soma dos termos exposta na Equação (2.25):

$$\left(\frac{(x_i^+ - x_i)}{\sigma_i} \right)^2 \quad (2.25)$$

Onde x^+ são os dados medidos, x_i são os dados reconciliados e σ o desvio padrão das medidas.

Onde a Equação (2.26) é também uma probabilidade de distribuição qui-quadrado, $\chi^2(V)$, com grau de liberdade igual ao grau de redundância (DoR) no problema de reconciliação de dados.

De forma que é apresentada a Equação (2.26) como sendo o teste global de erros grosseiros:

$$GT = \sum \left(\frac{(x_i^+ - x_i)}{\sigma_i} \right)^2 < \chi^2_{(1-\alpha)}(V) \quad (2.26)$$

Sendo $\chi^2_{(1-\alpha)}(V)$ é o valor do limite superior da distribuição qui-quadrado onde erros grosseiros não são esperados e α é o nível de significância, que é geralmente dado por $\alpha=5\%$ (0.05).

Isto posto, há a necessidade de testes de hipótese e uma discussão de erros de primeiro e de segundo tipos. Entretanto, nessa proposta apenas avalia se o teste global (GT) é menor que $\chi^2_{(1-0.05)}(V)$, ocorrendo isso, não são esperados erros grosseiros no conjunto de dados do sistema. Busca-se o valor de:

$$\chi^2_{(1-0.05)}(V) = \text{CHI2INV}(\text{nível de significância}, \text{DoR}) = \text{CHI2INV}(0.05, \text{DoR}).$$

Retornando ao exemplo das três unidades operacionais tem-se o valor do teste global:

$$GT = \sum \left(\frac{(x_i^+ - x_i)}{\sigma_i} \right)^2 = 0.6466$$

Com grau de redundância (DoR) = o número de linhas da matriz de incidência A (dada pelo balanço material do processo) = 3.

Obtém-se o seguinte valor de $\chi^2_{(1-\alpha)}(V) = \text{CHI2INV}(0.95, 3) = 7.8147$, não sendo portanto, encontrados erros grosseiros.

O teste global é bem simples, mas exprime a primeira e fundamental estratégia de erros grosseiros em medida, que é checar a existência deles.

2.5 FUNÇÕES ROBUSTAS

Vários estudos foram realizados visando diminuir ou eliminar a influência dos erros grosseiros em dados medidos. O uso de funções robustas tem papel principal nesse intento porque na abordagem tradicional da reconciliação de dados, os erros de medição seguem uma distribuição normal com média zero, variância conhecida e as mesmas inferências estatísticas (FARIAS, 2006). Entretanto, quando existem erros grosseiros essas inferências estatísticas são inválidas e prejudicam todo o conjunto de dados, em virtude disso, a estimação desses dados através da chamada estatística robusta (Huber, 1964); (Huber, 1981) vem contribuir para sua aplicação na distribuição ideal.

Tjoa e Biegler (1991) buscaram uma junção do método clássico da reconciliação de dados e da identificação de erros grosseiros, de forma que o desvio em dados estimados fosse resolvido com o princípio da máxima verossimilhança através da Distribuição Gaussiana Contaminada onde os valores obtidos com erros tanto aleatórios quanto grosseiros eram considerados. Esse estudo buscou a minimização da função objetivo através de resíduos ponderados ajustados automaticamente por meio do método de Programação Quadrática Sucessiva (SQP).

Johnston e Kramer (1995) também aplicando o princípio da máxima verossimilhança para maximização da função de distribuição de probabilidades, entretanto o diferencial do estudo foi o uso desse princípio com a regressão robusta para tratar a sensibilidade do estimador aos erros grosseiros contidos nos dados. Através da função robusta Lorentziana e a distribuição gaussiana bivariada, obtiveram resultados satisfatórios no tratamento dos erros grosseiros.

Arora e Biegler (2001) formularam no problema da reconciliação de dados clássica a solução dos erros tanto grosseiros quanto sistemáticos através do uso de estimadores robustos como o *Hampel* de três partes, o M-estimador função *Fair*, aliados ao método de “Brauch and Bound” para resolução com estratégia automatizada. Aplicaram o método em dois conjuntos, um com medidas com falhas, e outro com medidas sem falhas. Concluíram que o estimador de três partes de *Hampel* foi o que apresentou maior robustez e com um ponto de corte que contribuiu com a reconciliação simultaneamente sem a necessidade do uso de análise exploratória.

Ozyurt e Pike (2004) estudaram a falta de robustez em alguns estimadores que prejudicava na remoção dos erros grosseiros em medidas de processo. Detectaram os erros grosseiros através de métodos estatísticos baseados em distribuição combinada com auxílio de método de Monte Carlo na obtenção das constantes e das eficiências relativas dos estimadores robustos. Para correção da pouca precisão que a função *Fair* apresentava nos ajustes da eficiência assintótica foi aplicado o que foi chamado de função da influência do estimador robusto com ponto de máximo, ponto de mínimo ou ponto de inflexão das derivadas primeira e segunda ordem.

Wongrat *et al.* (2005) utilizaram o algoritmo genético na geração de parâmetros para a reconciliação de dados não linear. Com a aplicação da robustez dos M-estimadores junto aos algoritmos genéticos apresentaram o algoritmo genético modificado (mGA) associado ao estimador redescendente de três partes de *Hampel* (HTPRE) e o ajuste das constantes do estimador foi realizado através da minimização do critério de informação de *Akaike*.

Mei *et al* (2007) avaliaram o efeito negativo dos erros grosseiros e sistemáticos em medidas para posterior aplicação da reconciliação de dados. Estudaram os desvios nas medidas através de perturbações nos processos e verificando o reflexo delas nos valores reconciliados. Assim como Arora e Biegler (2001) e Wongrat *et al.* (2005) aplicaram estimadores robustos para obtenção da detecção e eliminação dos erros que “contaminavam” as medidas.

Outros autores buscaram implementar as funções robustas no problema de estimação e reconciliação de dados. Zhang *et al.* (1995) aplicaram a Normal Contaminada em dados reais e estado estacionário assim como Bourouis *et al.* (1998) em unidade de dessalinização por separação flash com restrições não lineares, Schladt e Hu (2007) e Lid e Skogestad (2008b). Já Chen *et al.* (1998b) estudaram os M-estimadores Lorenziana e Fair em uma planta de ácido sulfúrico.

Jin *et al.* (2012) implementaram um novo estimador robusto chamado de *New Target* para reconciliação de dados robustos. Nesse estudo alguns exemplos da literatura foram comparados com a *New Target*, que as funções de Cauchy e Huber da família dos M-estimadores. A função proposta buscou detectar pequenos erros grosseiros nas medidas já que as outras funções estudadas não eram possuíam essa característica. Para a estimação robusta utilizou a função de influência $(I(r) = (d\rho(r))/dr)$ onde ρ representava a função robusta, o que permitia a comparação quanto a robustez dos estimadores. Como critério de desempenho foi aplicado a redução total de erro (TER) e a soma dos erros quadrados (SSE). Os autores

concluíram que a reconciliação de dados utilizando a *New Target* obteve dados mais precisos principalmente quando existia um ou dois erros grosseiros nas medidas, o que representou maior robustez entre as funções.

Alamgir *et al.* (2013) propôs um M-estimador da família dos redescendentes chamado *Alamgir Redescending M-estimator* (ALARM) baseado na modificação da tangente hiperbólica como função peso. Essa função foi concebida com a ideia de que a função de mínimos quadrados é sensível na presença de valores espúrios (*outliers*) o que afeta severamente o conjunto de dados. Os M-estimadores redescendentes são projetados para que *outliers* extremos sejam totalmente rejeitados nas medidas. A ALARM foi comparada a outros estimadores como Tukey Biweight (1960), Andrew Sine (1972), Hampel de três partes (1997), Huber(1964) e OLS (1960) com relação a robustez e eficiência, já que segundo os autores, o aumento da robustez dos outros estimadores para eliminação dos *outliers* resultava em diminuição da eficiência dos mesmos, o que não aconteceu com a ALARM.

3 METODOLOGIA

Neste capítulo serão explanadas as técnicas utilizadas para reconciliação de dados e avaliação das funções robustas. De início, foram selecionados 3 (três) estudos de casos visando avaliar a capacidade de predição do método de otimização SQP. Neste sentido, utilizou-se como método de referência, o método de Gauss-Newton. Em seguida, serão apresentadas as funções robustas e as características que definem sua robustez. Os índices de desempenho empregado no presente trabalho serão estabelecidos e, por fim, 3 (três) estudos de casos serão apresentados para avaliação das funções robustas.

Todos os casos utilizados versaram a respeito de medidas em sistemas químicos que possuíam como variáveis temperaturas, pressões, vazões e composições das espécies. Esse estudo foi realizado em um desktop com configuração Intel Core e 3GB de RAM.

3.1 AVALIAÇÃO DA PREDIÇÃO DO MÉTODO SQP

Durante o desenvolvimento do trabalho, optou-se por utilizar o método de otimização com base na Técnica de Programação Sequencial Quadrática (SQP). Contudo, uma prévia avaliação foi realizada para quantificar a capacidade de predição das suas estimativas. Para isso, foi selecionado como “*benchmark*” o método iterativo de Gauss-Newton descrito por Noronha *et al.* (1993) (Ver Apêndice A e B).

Três estudos de casos foram utilizados para esta avaliação se com tolerância especificada para os estudos de 10^{-6} .

O algoritmo do método SQP é apresentado na Figura 3.1.

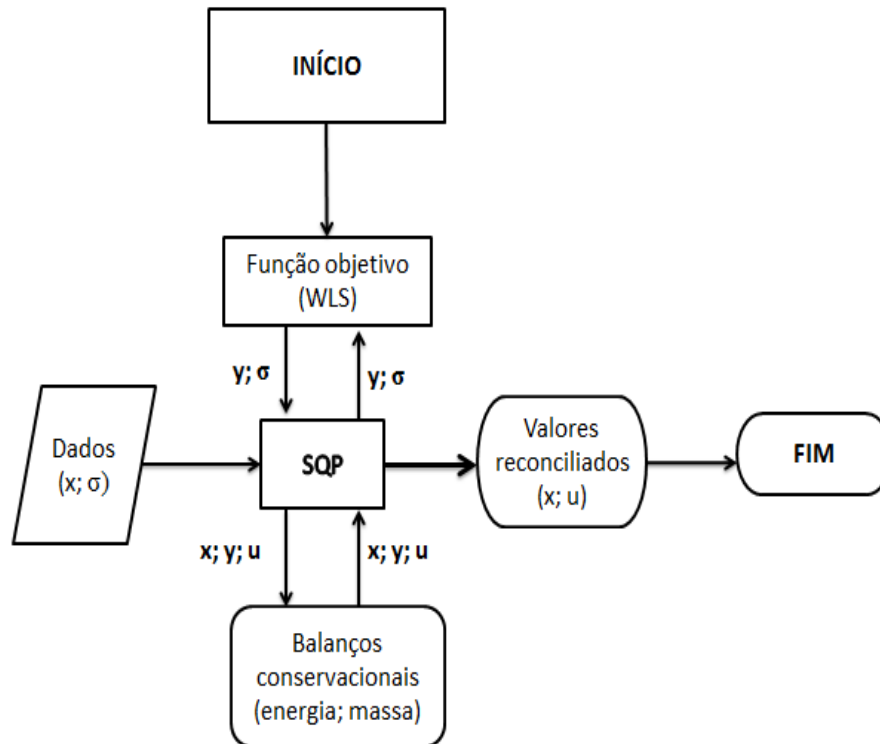


Figura 3.1 Fluxograma do método SQP

ESTUDO DE CASO 1: RECONCILIAÇÃO DE DADOS EM MÚLTIPLA UNIDADE (KNOPE, 2012)

Na Figura 2.5 foi apresentado um fluxograma de operação de três unidades. O sistema possui cinco correntes onde todas as variáveis são medidas, sendo que uma corrente é de reciclo. As unidades operacionais são representadas por um misturador, reator e um vaso separador.

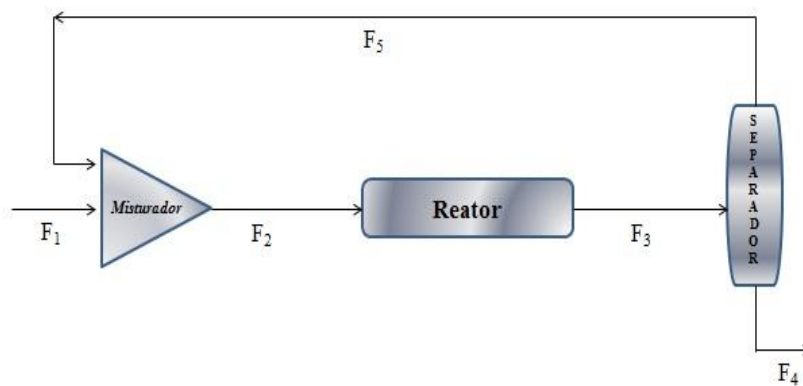


Figura 2.5 Fluxograma operacional de três unidades

A Tabela 2.1 mostrou os dados de operação dessa unidade e também o desvio padrão de cada medida.

Tabela 2.1 Dados medidos da unidade

	Corrente 1	Corrente 2	Corrente 3	Corrente 4	Corrente 5
Fluxo medido	99.3	123.2	125.1	100.5	23.9
Σ medido	2.0	2.0	2.0	2.0	2.0

ESTUDO DE CASO 2: CLORAÇÃO DO METANO RESTRIÇÕES LINEARES (MADRON *et al.*, 2007), (KNOPE, 2012).

Na Figura 3.2 está representado o fluxograma das espécies envolvidas no processo de cloração do metano com entrada e saída.



Figura 3.2 Espécies observadas no reator de cloração

Segundo Solomons (2000), se há mistura entre metano e cloro (ambas as substâncias gasosas à temperatura ambiente), diante de um aquecimento ou irradiação com luz, a reação deverá ocorrer vigorosamente. No início, os únicos compostos que estão presentes nessa mistura são cloro e metano com produção de clorometano e de cloreto de hidrogênio, reação demonstrada na Figura 3.3.

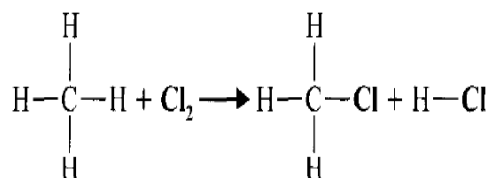


Figura 3.3 Reação de cloração do metano

Fonte: Solomons, 2000.

Contudo, à medida que a reação progride, a concentração de clorometano aumenta na mistura e uma segunda reação de substituição começa a ocorrer. O clorometano reage com o cloro para produzir o diclorometano segundo a Figura 3.4.

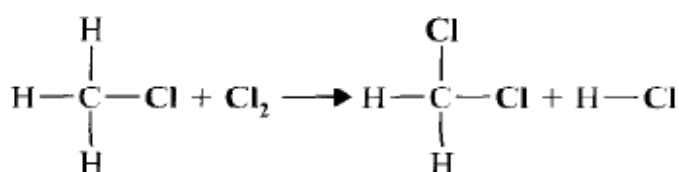


Figura 3.4 Reação de substituição para formação de diclorometano

Fonte: Solomons, 2000.

O diclorometano produzido pode então reagir para formar o triclorometano. E o triclorometano, à medida que se acumula na mistura, pode reagir com o cloro para produzir o tetraclorometano. Cada vez que ocorre uma substituição do -H pelo -Cl, uma molécula de H-Cl é produzida.

Neste problema o balanço material das espécies no sistema reativo foi formulado usando extensão das reações. Após determinar quantas reações independentes existiam obteve-se as extensões das reações (uma para cada reação independente).

O número de reações independentes = número de espécies - rank da matriz atômica (consiste no número do balanço por elemento). Neste estudo de caso tem-se 7 espécies e 3 participam do balanço por elementos (C, H, Cl) = 4 reações independentes para cada espécie formada. São representadas pelas reações (3.1) a (3.4):



Sendo as extensões das reações para cada uma das quatro reações (ξ_1 , ξ_2 , ξ_3 , ξ_4) baseadas nos moles da reação de cloração, o balanço mássico das espécies é dado a seguir:

$$Cl_2: N_{Cl_2}^{Reator_E} - (\xi_1) - (\xi_2) - (\xi_3) - (\xi_4) N_{Cl_2}^{Reator_S} = 0 \quad (3.5)$$

$$CH_4: N_{CH_4}^{Reator_E} - (\xi_1) - N_{CH_4}^{Reator_S} = 0 \quad (3.6)$$

$$CH_3Cl: N_{CH_3Cl}^{Reator_E} + (\xi_1) - (\xi_2) - N_{CH_3Cl}^{Reator_S} = 0 \quad (3.7)$$

$$CH_2Cl_2: N_{CH_2Cl_2}^{Reator_E} + (\xi_2) - (\xi_3) - N_{CH_2Cl_2}^{Reator_S} = 0 \quad (3.8)$$

$$CHCl_3: N_{CHCl_3}^{Reator_E} + (\xi_3) - (\xi_4) - N_{CHCl_3}^{Reator_S} = 0 \quad (3.9)$$

$$CCl_4: N_{CCl_4}^{Reator_E} + (\xi_4) - N_{CCl_4}^{Reator_S} = 0 \quad (3.10)$$

$$HCl: N_{HCl}^{Reator_E} + (\xi_1) + (\xi_2) + (\xi_3) + (\xi_4) - N_{HCl}^{Reator_S} = 0 \quad (3.11)$$

Os valores referentes às extensões são: $\xi_1 = 25.75747$; $\xi_2 = 29.69288$; $\xi_3 = 4.874971$ e $\xi_4 = 0.390384$.

Os dados medidos das espécies no reator são apresentados na Tabela 3.1:

Tabela 3.1 Dados de fluxo molar medido e desvio padrão

	Fluxo de entrada	Σ	Fluxo de saída	σ
	medido		medido	
Fluxo	100	Fixo		
Cl₂	62,1	3,105	0	Fixo
CH₄	26,9	1,345	1,57	0,0785
CH₃Cl	10,8	0,54	6,83	0,3415
CH₂Cl₂	0,78	0,039	24,8	1,24
CHCl₃	0,38	0,019	4,82	0,241
CCl₄	0	Fixo	0,39	0,0195
HCl	0	Fixo	59,3	2,965

ESTUDO DE CASO 3: TROCADOR DE CALOR ADIABÁTICO COM BALANÇO DE ENERGIA NÃO LINEAR

A Figura 3.5 é referente ao trocador de calor adiabático com balanço de energia não linear e assumindo que não há vazamento.

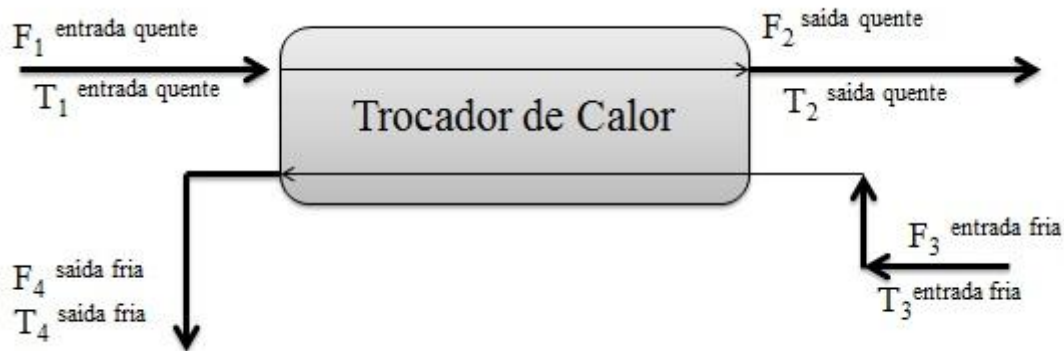


Figura 3.5 Fluxograma do trocador de calor

Assumem-se as seguintes capacidades caloríficas:

$$\begin{aligned}\hat{C}_{p,h(\text{correntes_quentes})} &= 5 + (0,1)T \\ \hat{C}_{p,c(\text{correntes_frias})} &= 3 + (0,2)T\end{aligned}\quad (3.12)$$

Essas capacidades caloríficas são advindas da Equação (3.13) onde:

$$\hat{C}_p = a + bT \quad (3.13)$$

Sendo $\hat{C}_p$ a capacidade calorífica de uma substância a pressão constante, com unidade de kJ/kg.C, T é a temperatura em Celsius, e a e b são os parâmetros dependentes das espécies.

Pode-se estimar a variação da entalpia de acordo com a Equação (3.14):

$$\Delta \hat{h} = \hat{h}_{(T_2,P)} - \hat{h}_{(T_1,P)} = \int_{T_1}^{T_2} \hat{C}_p dT = a(T_2 - T_1) + \frac{b}{2}(T_2^2 - T_1^2) \quad (3.14)$$

Integrando a Equação (3.14) com as condições de entrada e saída do sistema em cada corrente chegam-se as Equações (3.15) e (3.16).

$$\Delta \hat{h}_{quente} = \int_{T_1}^{T_2} \hat{C}_{p,h} dT = \int_{T_{-Qe}}^{T_{-Qs}} (5 + 0,1T) = \left[5T + \frac{0,1}{2} T^2 \right]_{T_{-Qe}}^{T_{-Qs}} \quad (3.15)$$

$$\Delta \hat{h}_{frio} = \int_{T_1}^{T_2} \hat{C}_{p,f} dT = \int_{T_{-fe}}^{T_{-fs}} (3 + 0,2T) = \left[3T + \frac{0,2}{2} T^2 \right]_{T_{-fe}}^{T_{-fs}} \quad (3.16)$$

Os dados referentes ao caso são apresentados na Tabela 3.2.

Tabela 3.2 Dados medidos e desvio padrão das medidas

Correntes	Quente (Entrada)	Quente (Saída)	Fria (Entrada)	Fria (Saída)
Taxa (kg/s)	10		20	
Temperatura (°C)	90	51	20	43
Σ fluxo	0.2		0.4	
Σ temperatura	1	1	1	1

3.2 RECONCILIAÇÃO DE DADOS APLICANDO FUNÇÕES ROBUSTAS

Reconciliar dados utilizando funções robustas é apresentar uma modificação da função mínimos quadrados ponderados para implementação de um estimador com função objetivo de verossimilhança sem restrições (Huber, 1960) conforme a Equação (3.17):

$$\min \sum_i \rho \left(\frac{y_i - x_i}{\sigma_i} \right) \quad (3.17)$$

sendo ρ a função robusta, y_i os valores medidos, x_i os valores reconciliados e σ_i o desvio padrão. E o erro padrão é expresso como $\varepsilon_i = (y_i - x_i) / \sigma_i$.

Dentre as principais famílias de funções robustas estão os L-estimadores que possuem combinações lineares de observações de ordem, os M-estimadores baseados na máxima verossimilhança e os R-estimadores com ranqueamento dos resíduos. As funções robustas estudadas nesse trabalho *Cauchy*, *Fair*, *Normal Contaminada*, *New Target* e *Logistica* pertencem aos M-estimadores enquanto a *Alarm* é um M-estimador redescendente.

Os M-estimadores robustos são aqueles que produzem resultados imparciais e são insensíveis com relação à distribuição estatística ideal. M-estimadores visam a redução do tempo computacional através da reconciliação de dados e estimação de parâmetros simultâneos porque os erros grosseiros são facilmente identificados por análise exploratória dos dados (CHENet al., 2013).

As funções ρ apresentadas nesse estudo são as seguintes:

Mínimos quadrados ponderados (WLS)

$$\frac{1}{2} \varepsilon_1^2 \quad (3.18)$$

Cauchy

$$c_c^2 \ln \left(1 + \frac{\varepsilon_i^2}{c_c^2} \right) \quad (3.19)$$

Fair

$$2c_f^2 \left[\frac{|\varepsilon_i|}{c_f} - \ln \left(1 + \frac{|\varepsilon_i|}{c_f} \right) \right] \quad (3.20)$$

Normal Contaminada

$$-\ln \left\{ (1 - p_{CN}) \exp \left(-\frac{\varepsilon_i^2}{2} \right) + \frac{p_{CN}}{b_{CN}} \exp \left(-\frac{\varepsilon_i^2}{2b_{CN}^2} \right) \right\} \quad (3.21)$$

Logística

$$2 \ln \left(1 + \exp \left(\frac{\varepsilon_i}{c_{Lo}} \right) \right) - \left(\frac{\varepsilon_i}{c_{Lo}} \right) \quad (3.22)$$

New Target

$$\begin{cases} \frac{c^2}{6} \left[1 - \left[1 - a \frac{\varepsilon^2}{c^2} \right]^3 \right] & \text{if } |\varepsilon| \leq c \\ A \ln \frac{\varepsilon^2}{c^2} + B, & \text{if } |\varepsilon| \geq c \end{cases} \quad (3.23)$$

Alarm Redescendente

$$\begin{cases} \left[\frac{2c^2}{3} \left[1 - \frac{2(1 + 3e^{\left(\frac{\varepsilon}{c}\right)^2})}{(1 + e^{\left(\frac{\varepsilon}{c}\right)^2})^3} \right] \right], & |\varepsilon| \leq c \\ \left[\frac{2c^2}{3} \left[1 - \frac{2(1 + 3e)}{(1 + e)^3} \right] \right], & |\varepsilon| \geq c \end{cases} \quad (3.24)$$

Tabela 3.3 Valor das constantes para diferentes funções com valores de eficiência de 95%

Função ρ	Valor das constantes
Cauchy	$c_c=2.3849$
Fair	$c_f=1.3998$
Normal Contaminada	$b_{cn}=10, p_{cn}=0.235$
Logística	$c_{Lo}=0.602$
New Target	$c=3, a=0.65$
Alarm	$c=3$

Os valores das constantes apresentados na Tabela 3.3, foram extraídos dos trabalhos de Ozyurt and Pike (2004), Jin *et al.* (2012) e Alamgir *et al.* (2013) e a Figura 3.6 apresenta o desempenho das funções em relação ao erro para fins de análise das constantes de ajustes de cada função.

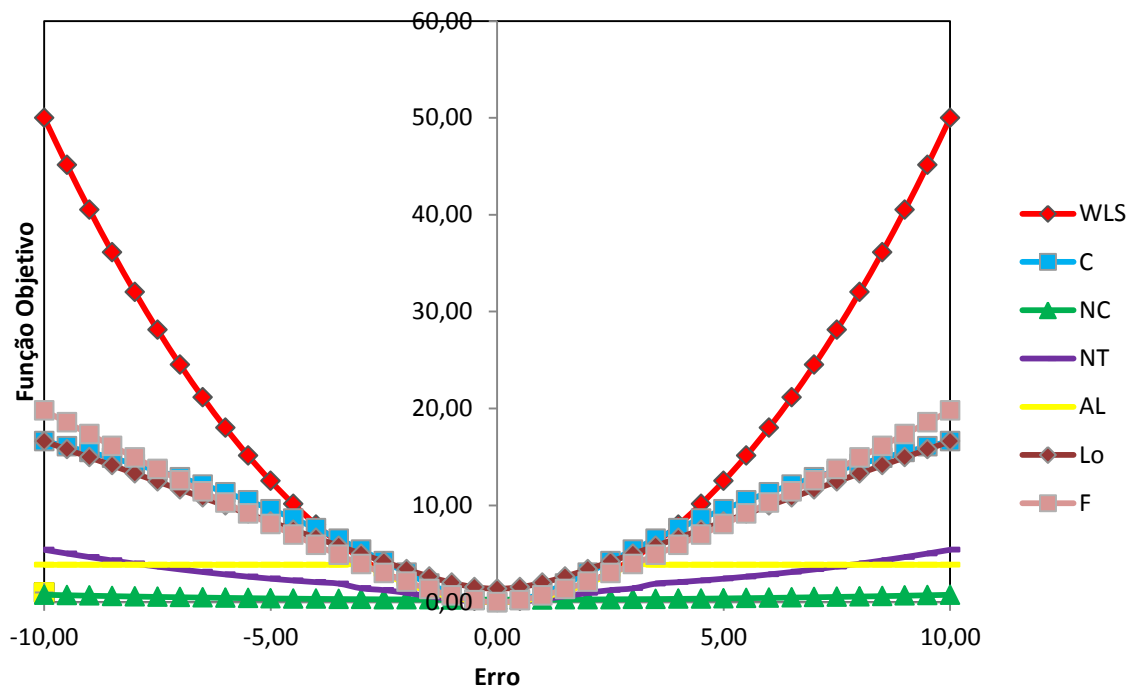


Figura 3.6 Comportamento das funções robustas em relação ao erro

Na Figura 3.6 pode-se evidenciar que todas possuem um único mínimo quando o erro aproxima-se de zero, configurando-se como uma função puramente convexa e que lhe atribuem consideráveis propriedades para a otimização. Além disso, as funções *Cauchy*, *Fair* e *Logística* possuem desempenho similar em relação à projeção da função, legitimando as conclusões apresentadas por Ozyurt and Pike (2004).

As funções *Normal Contaminada*, *New Target* e *Alarm*, devido a sua forma funcional, conseguem atenuar o valor da função em relação a grandes variações de erro. Este comportamento pode também ser observado ao se analisar Função Influência, a qual representa a sensibilidade da função em relação ao erro ou a contaminações referentes a erros grosseiros e *outliers* (ZHANG *et al.*, 2010) e (PRATA *et al.*, 2010). O comportamento idealizado para a Função de Influência é de, com o aumento do erro, convergir a um valor pequeno e constante, tornando-se assim indiferente a contaminações por grandes erros grosseiros.

A Figura 3.7 mostra que em algumas funções, a destacar *Normal Contaminada*, *New Target* e *Alarm*, uma pequena região de máximo para Função de Influência é definida que, em seguida, alcança um valor próximo de zero. Este comportamento sugere que erros acima do limite de máximo podem ser desprezados, o que motivou ao desenvolvimento de estratégias de simultânea reconciliação de dados e detecção de erros grosseiros (ARORA e BIEGLER, 2001) e (OZYURT e PIKE, 2004).

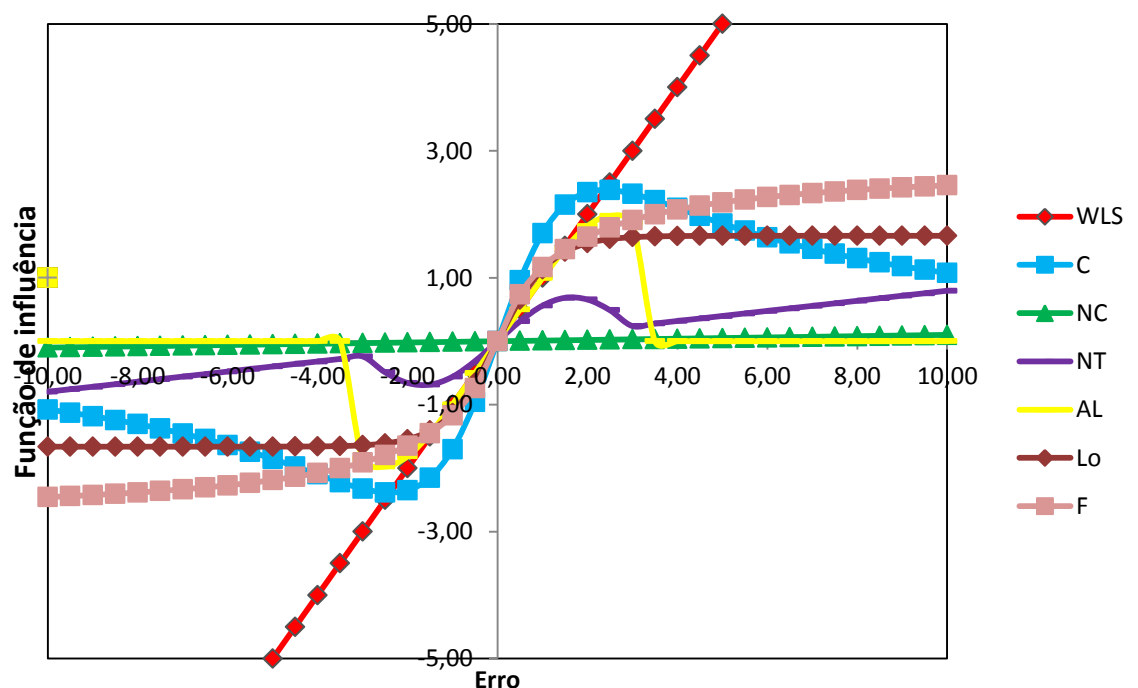


Figura 3.7 Comportamento das funções de influência em relação ao erro

Além disso, funções como essas possuem constantes como uma função de eficiência assintótica. Entretanto, essas variações assintóticas “apresentam apenas indicações brutas para as verdadeiras variâncias” no tamanho da amostra finita (HAMPEL, 1985). Portanto, a

aproximação finita no tamanho das amostras e consequentemente a sua relativa eficiência é obtida através de simulações de Monte Carlo com corridas variando de 28 a 2.000 simulações.

O ambiente computacional foi inicialmente utilizado através da função de mínimos quadrados com restrições não lineares a fim de encontrar o mínimo de uma função multivariável sujeito as suas restrições, fossem elas de igualdades ou desigualdades. E posteriormente testadas as funções robustas já mencionadas.

Três estudos de casos foram escolhidos para analisar a eficiência das funções robustas, eles estão apresentados a seguir.

ESTUDO DE CASO 4: IDENTIFICAÇÃO E DETECÇÃO DE ERROS GROSSEIROS: SISTEMA LINEAR

A Figura 3.8 trata de um sistema composto por 6 (seis) correntes todas medidas e contendo um trocador de calor com válvula *bypass*, misturador e separador. Foi estudado inicialmente por Narashiman e Jordache (2000) e recentemente por Knopf (2012). Neste caso, uma das correntes possui erros grosseiros (a medida da taxa da corrente F_2 contém 4 *bias* positivos), os dados de operação são mostrados na Tabela 3.4.

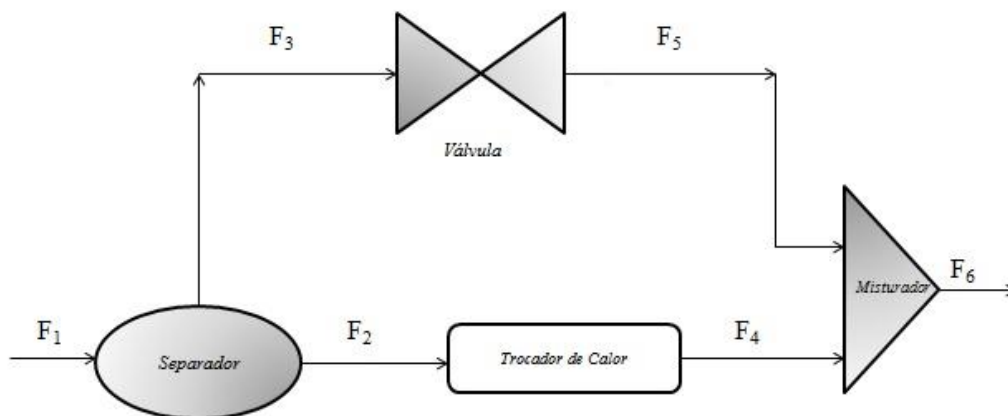


Figura 3.8 Trocador de calor com válvula de desvio (bypass)

Na Tabela 3.4 constam os dados referentes ao trocador de calor apresentado na Figura 3.8. Nela se encontram os valores verdadeiros das taxas que são aqueles atribuídos ao projeto do equipamento e também o valor medido da taxa que são os aferidos por instrumentos através de observadores

Tabela 3.4 Dados do trocador de calor com bypass

Número das Correntes	Valor verdadeiro da Taxa	Valor medido da Taxa	Σ Taxa
1	100	101.9	1
2	64	68.45	1
3	36	34,65	1
4	64	64,20	1
5	36	36,64	1
6	100	98,88	1

ESTUDO DE CASO 5: TROCADOR DE CALOR SISTEMA NÃO LINEAR

A Figura 3.9 acrescenta ao mesmo trocador de calor da Figura 3.8 medidas referentes às correntes de temperaturas, e o *bias* que existia antes na corrente F_2 foi removido. Entretanto a corrente F_7 que faz parte do balanço energético contém 5 *bias* positivos. Os valores verdadeiros e medidos são apresentados na Tabela 3.5.

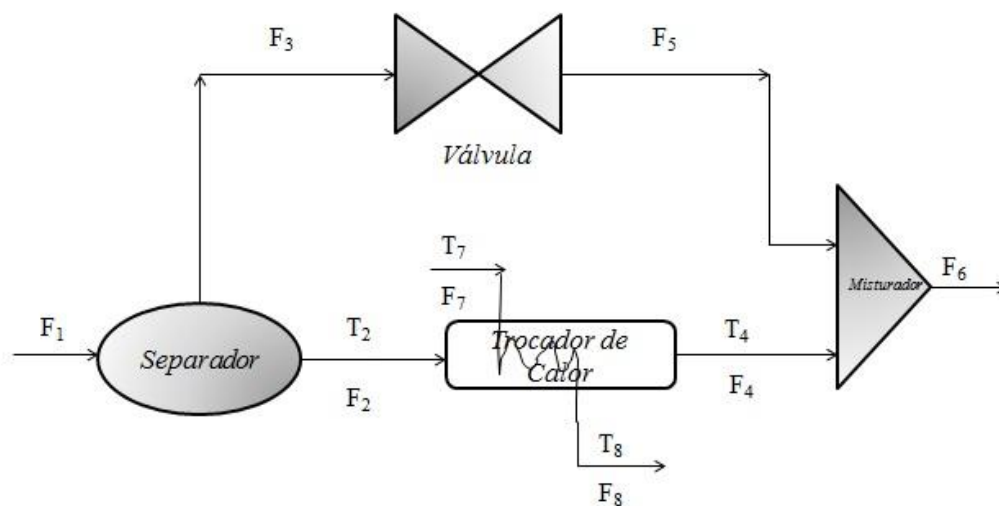


Figura 3.9 Trocador de calor com válvula bypass e balanço de energia

Tabela 3.5 Dados do trocador de calor com válvula bypass

Número de correntes	Valor verdadeiro da taxa	Valor verdadeiro da temperatura	Valor medido	Σ Taxa ou Σ Temperatura
1	100		101,91	1
2	64		64,45	1
3	36		34,65	1
4	64		64,20	1
5	36		36,44	1
6	100		98,88	1
7	140,6		140,0	1
2		90	90	1
4		51	51	1
7		20	25	1
8		43	43	1

Na Tabela 3.5 é possível observar os dados dos balanços energético e mássico, com os valores verdadeiros das taxas e das temperaturas que são os dados de projeto do equipamento e os valores medidos que são os aferidos pelos observadores através de instrumentos.

ESTUDO DE CASO 6: O PROCESSO DE POLIMERIZAÇÃO

Nesse estudo de caso são apresentados dados reais de um processo de polipropileno. Esse processo foi estudado por Oliveira Jr. (2006) e Prata (2009). O primeiro autor estimou os parâmetros da unidade industrial aplicando a técnica de reconciliação de dados em estado estacionário juntamente com a etapa de classificação de variáveis. Já o segundo focou na reconciliação de dados dinâmicos através do modelo do processo.

O sistema simplificado é mostrado na Figura 3.10.

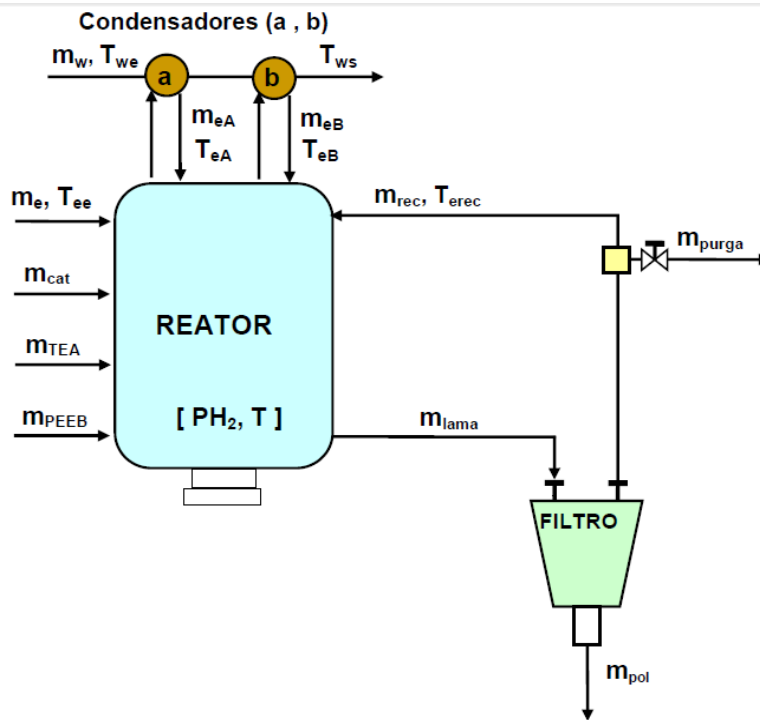


Figura 3.10 Processo de polimerização simplificado
Fonte: Oliveira Jr., 2006.

Para esse estudo, será apresentada uma simplificação do processo para avaliação da aplicação das funções robustas em processos industriais com dados reais.

Nos dados utilizados nessa dissertação buscou-se um intervalo de tempo a ser estudado com características estacionárias, todas as variáveis foram medidas em um intervalo maior do que o demonstrado nesse estudo. O que se pretendia era apresentar a comparação das funções robustas em um modelo estacionário de um processo industrial. Foi estudado o balanço mássico em relação ao reator de polimerização, o balanço energético não fez parte do escopo desse trabalho.

De acordo com Prata (2010), a polimerização do polipropileno em massa é um processo em lama no qual o diluente é o próprio monômero. São utilizados reatores CSTR em combinação com catalisadores Ziegler-Natta de quarta geração ($\text{TiCl}_4/\text{MgCl}_2 + \text{PEEB} + \text{TEA}$), onde PEEB é o doador de elétrons e o TEA é co-catalisador.

Cita ainda que, o monômero propeno (contaminado por propano) é alimentado na fase líquida. O catalisador Ziegler-Natta, o co-catalisador (TEA), o doador de elétrons (PEEB) e outros agentes são introduzidos mediante uma corrente de alimentação em fase líquida. A corrente de saída do reator é composta basicamente de propano, propeno e polipropileno.

Nesse estudo, o modelo matemático utilizado bem como hipóteses simplificadoras foi proposto por Oliveira Jr. (2006), e possui seis componentes: propano, propeno, polipropileno, TEA, PEEB e catalisador.

Desses, foram escolhidas as variáveis medidas e as variáveis não medidas a serem estimadas com as funções robustas são elas:

Variáveis medidas: m_e , m_{CAT} , m_{TEA} , m_{PEEB} , m_{Pol} e L ;

Variáveis não medidas: m_{purga} , Pe , Pa , Pol , CAT , TEA e $PEEB$.

O modelo utilizado nessa dissertação baseou-se nas equações a seguir, bem como nas variâncias medidas apresentadas na Tabela 3.6 e nos parâmetros apresentados na Tabela 3.7.

Balanço de massa para o Propano:

$$m_e \cdot w_a - \left(\frac{Pa}{Pa + Pe} \right) \cdot m_{purga} = 0 \quad (3.25)$$

Balanço de massa para o Propeno:

$$m_e \cdot (1 - w_a) - R_{pol} - \left(\frac{Pe}{Pa + Pe} \right) \cdot m_{purga} = 0 \quad (3.26)$$

Balanço de massa para o Polipropileno:

$$R_{pol} - m_{pol} = 0 \quad (3.27)$$

Cálculo da taxa de polimerização (produção de polipropileno):

$$R_{pol} = Kp \cdot \left(\frac{CAT \cdot Pe}{M} \right) \quad (3.28)$$

Balanço de massa para o Catalisador:

$$m_{CAT} - \left(\frac{CAT}{POL} \right) \cdot m_{pol} - Kd \cdot CAT = 0 \quad (3.29)$$

Balanço de massa para o PEEB:

$$m_{PEEB} - \left(\frac{PEEB}{Pol} \right) \cdot m_{pol} = 0 \quad (3.30)$$

Balço de massa para o TEA:

$$m_{TEA} - \alpha \cdot \left(\frac{TEA}{Pe + Pa} \right) \cdot m_{purga} - (1 - \alpha) \cdot \left(\frac{TEA}{Pol} \right) \cdot m_{pol} = 0 \quad (3.31)$$

Restriço de volume:

$$V = \left(\frac{Pa}{\rho a} + \frac{Pe}{\rho e} + \frac{Pol}{\rho pol} \right) \quad (3.32)$$

Cálculo da densidade do propeno:

$$\rho_e = 1.10^{-3} \cdot [5.099.10^2 + 1.7489.T - 5.840.10^{-3}.T^2] \quad (3.33)$$

Cálculo da densidade do propano:

$$\rho_a = 1.10^{-3} \cdot [5.099.10^2 + 1.7489.T - 5.840.10^{-3}.T^2] \quad (3.34)$$

Cálculo da densidade do polipropileno (polímero):

$$\rho_{pol} = 1.10^{-3} \cdot [6.802.10^2 + 1.8190.T - 3.575.10^{-3}.T^2] \quad (3.35)$$

Cálculo do volume:

$$V(L) = V = 2,300^4 + \left[(L - 16,5\%) \cdot 1,544^3 \right] \quad (3.36)$$

Tabela 3.6 Variâncias para as variáveis medidas do processo de polimerização

Variável	Variância
m_e	$3,600 \cdot 10^5$
m_{CAT}	$2,500 \cdot 10^{-3}$
m_{TEA}	$2,500 \cdot 10^{-3}$
m_{PEEB}	$2,500 \cdot 10^{-3}$
m_{Pol}	$3,600 \cdot 10^5$
L	$5,000 \cdot 10^{-1}$

Tabela 3.7 Valores dos parâmetros fixos do modelo

Parâmetros	Descrição	Valores
T	Temperatura controlada de operação do reator	341K
α	Fator de reciclo de TEA	0.0002
K_d	Constante de iniciação	2
K_p	Constante cinética de propagação	5.5
w_a	Fração mássica de propano	0,004kg/kg

Os valores dos parâmetros das Tabelas 3.6 e 3.7 foram apresentados por Prata (2010), com base em trabalho anterior de Schwaab e Pinto (2007).

Através desse modelo foi realizada a reconciliação de dados das variáveis medidas e estimação das variáveis não medidas que são observáveis com a aplicação das funções robustas escolhidas.

3.3 AVALIAÇÃO DE DESEMPENHO DAS FUNÇÕES

Para a definição dos critérios de desempenho dos estudos de casos extraídos da literatura, foram considerados aspectos como: convergência, número de avaliações da função objetivo e erro médio relativo. Os dois primeiros aspectos indicam se a função pode ser utilizada para aplicações em tempo real, já o último aspecto refere-se a habilidade da função se aproximar dos valores reais.

O critério do erro médio relativo é exposto na Equação (3.37):

$$EMR = \frac{(x^t - x)}{\sigma} \quad (3.37)$$

Sendo x os valores verdadeiros e σ o desvio padrão da amostra.

Aliados a esses critérios, com relação aos três últimos estudos de casos, o desempenho dessas funções foi primeiramente comparado aos ajustes obtidos nas medidas sem erro, e posteriormente aplicados os índices que se baseiam em medidas contaminadas com erro.

Os índices estudados nessa dissertação foram apresentados por Zhang *et al.* (2010):

Erro Relativo Medido:

$$ERM_i = \frac{|x_i - x_i^m|}{x_i} \quad (3.38)$$

Erro Relativo Reconciliado:

$$ERR_i = \frac{|x_i - x_i^r|}{x_i} \quad (3.39)$$

Redução do Erro Relativo:

$$RER = \frac{\sum_i (ERM_i - ERR_i)}{\sum_i ERM_i} \cdot 100\% \quad (3.40)$$

Sendo x_i o valor verdadeiro, x_i^r o valor reconciliado, e x_i^m o valor medido.

4 RESULTADOS E DISCUSSÃO

Neste capítulo serão discutidos os resultados referentes à reconciliação de dados com funções robustas. Primeiramente será mostrada a avaliação da predição do método SQP para a reconciliação (*solvers*), em seguida os estudos de casos com as funções robustas e por fim a avaliação destas na presença erros.

Com relação a avaliação da predição do método SQP foram escolhidos os estudos de caso 1, 2 e 3. Para a avaliação das funções robustas contendo ou não erros foram selecionados os estudos de caso 4, 5 e 6.

4.1 AVALIAÇÃO DA PREDIÇÃO DO MÉTODO SQP

Serão apresentados os casos estudados para fins de avaliação dos métodos com base nos processos. As Figuras 2.5, 3.2 e 3.5 representam os fluxogramas; as Tabelas 2.1, 3.1 e 3.2 referem-se aos dados medidos e desvios padrões e as Tabelas 4.1 a 4.6 apresentam os dados obtidos aplicando os métodos de Gauss-Newton e o SQP.

A função objetivo considerada para este estudo foi a Mínimos Quadrados Ponderados (pelo desvio padrão) devido, principalmente, a sua simplicidade. Neste caso, apenas a influência do método pode ser considerada.

ESTUDO DE CASO 1: RECONCILIAÇÃO DE DADOS EM MÚLTIPLA UNIDADE (KNOPE, 2012)

Na Figura 2.5 foi apresentado um fluxograma de operação de três unidades visando verificar os solucionadores e escolher o que representasse resultados mais eficientes mesmo com as restrições do sistema. O sistema possui cinco correntes, sendo que uma corrente é de reciclo. As unidades operacionais são representadas por um misturador, reator e um vaso separador.

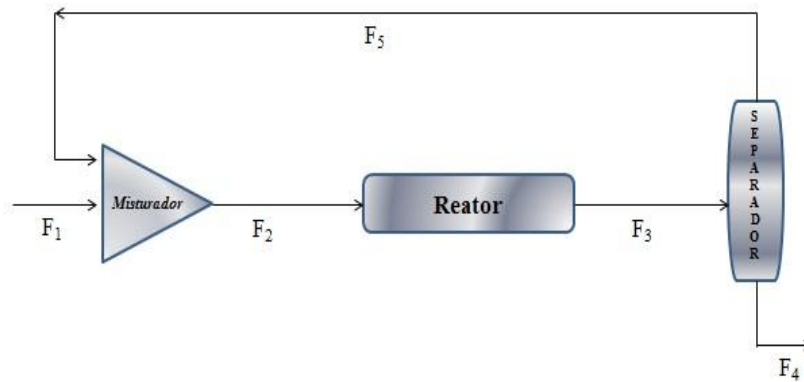


Figura 2.5 Fluxograma operacional de três unidades

A Tabela 2.1 mostrou os dados de operação dessa unidade e também o desvio padrão de cada medida.

Tabela 2.1 Dados medidos da unidade

	Corrente 1	Corrente 2	Corrente 3	Corrente 4	Corrente 5
Fluxo medido	99.3	123.2	125.1	100.5	23.9
Σ medido	2.0	2.0	2.0	2.0	2.0

Uma análise dos resultados reportados nas Tabelas 4.1 e 4.2 indicam que ambos chegaram as mesmas estimativas, contudo, o método SQP apresentou um número de iterações menor mas não destoante do método de Gauss-Newton.

Tabela 4.1 Análise do método de Gauss-Newton para o estudo de caso 1

Método de Gauss-Newton					
Estudo de caso 1					
Medido	99,30	123,20	125,10	100,50	23,90
Estimado	99,99	124,06	124,06	99,99	24,08
Iterações	3				

Tabela 4.2 Análise do método SQP para o estudo de caso 1

Método SQP					
Estudo de caso 1					
Medido	99,30	123,20	125,10	100,50	23,90
Estimado	99,99	124,06	124,06	99,99	24,08
Iterações	2				

ESTUDO DE CASO 2: CLORAÇÃO DO METANO RESTRIÇÕES LINEARES (MADRON *et al.*, 2007), (KNOPE, 2012).

Na Figura 3.2 está representado o fluxograma das espécies envolvidas no processo de cloração do metano com entrada e saída.



Figura 3.2 Espécies observadas no reator de cloração

Os dados medidos das espécies no reator são apresentados na Tabela 3.1:

Tabela 3.1 Dados de fluxo molar medido e desvio padrão

	Fluxo de entrada	Σ	Fluxo de saída	Σ
	medido		medido	
Fluxo	100	Fixo		
Cl₂	62,1	3,105	0	Fixo
CH₄	26,9	1,345	1,57	0,0785
CH₃Cl	10,8	0,54	6,83	0,3415
CH₂Cl₂	0,78	0,039	24,8	1,24
CHCl₃	0,38	0,019	4,82	0,241
CCl₄	0	Fixo	0,39	0,0195
HCl	0	Fixo	59,3	2,965

Neste caso, as Tabelas 4.3 e 4.4 indicam que o método de Gauss-Newton apresentou um número substancialmente menor de iterações quando comparado ao SQP.

Tabela 4.3 Análise do método de Gauss Newton para o estudo de caso 2

Método Gauss-Newton											
	Estudo de caso 2										
Medido	62,1	26,9	10,8	0,78	0,38	1,57	6,83	24,8	4,82	0,39	59,3
Reconciliado	60,33	27,08	10,75	0,77	0,37	1,56	6,86	25,19	4,95	0,40	60,33
Iterações	3										

Tabela 4.4 Análise do método SQP para o estudo de caso 2

Método SQP											
	Estudo de caso 2										
Medido	62,1	26,9	10,8	0,78	0,38	1,57	6,83	24,8	4,82	0,39	59,3
Reconciliado	59,93	26,95	10,76	0,78	0,38	1,57	6,84	25,2	4,84	0,39	59,93
Iterações	17										

Em termos dos valores reconciliados, os resultados obtidos foram diferentes, mas não destoantes dos valores medidos. Acredita-se que esta diferença observada no número de iterações possa ser atribuída as grandezas das correntes, favorecendo o método de Gauss-Newton.

Entretanto, Gauss-Newton apresenta uma dificuldade com relação a convergências, pois, de acordo com estudos o método pode falhar, ou mesmo deixar de gerar uma sequência infinita sendo necessário impor condições ao sistema (Allende e Gonçalves, 2011), (Lima, 2013), (Argyros e Magreñan, 2014).

ESTUDO DE CASO 3: TROCADOR DE CALOR ADIABÁTICO COM BALANÇO DE ENERGIA NÃO LINEAR

A Figura 3.5 apresentada no capítulo anterior trata-se do trocador de calor adiabático com balanço de energia não linear onde se assume não haver vazamento.

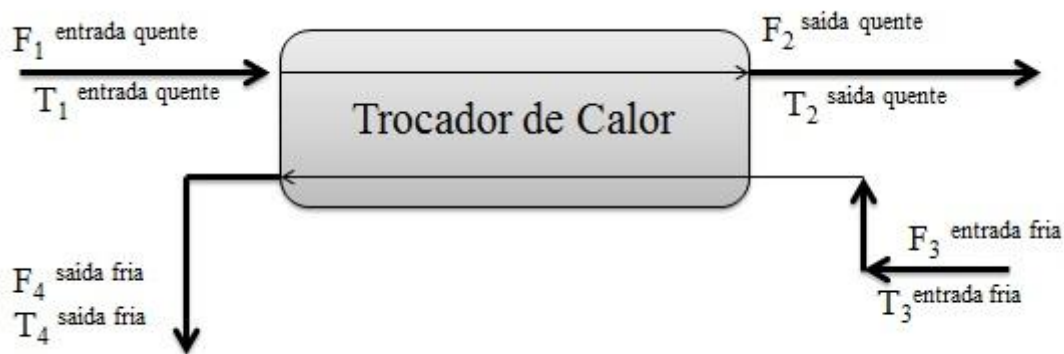


Figura 3.5 Fluxograma do trocador de calor

Tabela 3.2 Dados medidos e desvio padrão das medidas

Correntes	Quente (Entrada)	Quente (Saída)	Fria (Entrada)	Fria (Saída)
Taxa (kg/s)	10		20	
Temperatura (°C)	90	51	20	43
Σ fluxo	0.2		0.4	
Σ temperatura	1	1	1	1

Os resultados reportados na Tabela 4.5 e 4.6 mostram que o método de SQP apresentou um número menor de iterações. Esta redução pode ser atribuída a habilidade do método SQP em lidar com não linearidades se comparada ao método de Gauss-Newton.

Tabela 4.5 Análise do método de Gauss Newton para o estudo de caso 3

Método de Gauss-Newton								
Estudo de caso 3								
Medido	10	10	20	20	90	51	20	43
Estimado	9,86	9,86	20,24	20,24	89,60	51,29	19,59	43,70
Iterações	11							

Tabela 4.6 Análise do método SQP para o estudo de caso 3

Método SQP								
Estudo de caso 3								
Medido	10	10	20	20	90	51	20	43
Estimado	9,9359	9,9359	20,1266	20,1266	89,5167	51,351	19,5157	43,8257
Iterações	9							

A vantagem do SQP em relação ao Gauss-Newton se estende também a relação com o tratamento das restrições, o SQP responde melhor a função objetivo de mínimos quadrados, apresentando assim maior robustez numérica. O Gauss-Newton precisa ser modificado resultando em diversas variações do método como, por exemplo, o método inexato de Gauss-Newton, método quase inexato de Gauss-Newton, método de Gauss-Newton modificado inexato.

4.2 RECONCILIAÇÃO DE DADOS APLICANDO FUNÇÕES ROBUSTAS

A seguir serão apresentados os resultados referentes à aplicação da técnica de reconciliação de dados utilizando as funções robustas *Mínimos Quadrados Ponderados* (WLS), *Cauchy* (C), *Fair* (F), *Normal Contaminada* (NC), *New Target* (NT), *Alarm* (AL) e *Logística* (Lo).

Nesse tópico serão apresentados outros 3 (três) estudos de casos, dois extraídos da literatura onde as funções robustas são testadas em um sistema linear e o outro em sistema não-linear, e um com dados reais de um sistema não linear que é o processo de polimerização.

ESTUDO DE CASO 4: IDENTIFICAÇÃO E DETECÇÃO DE ERROS GROSSEIROS: SISTEMA LINEAR

Esse estudo de caso possui uma particularidade que contribui para a estimação com funções robustas que é a identificação de erros. Sabe-se que na corrente F_2 (Ver Figura 3.8) contém 4 *bias* positivos que são erros do tipo espúrio e se espalham pelo conjunto de dados podendo corromper as medidas (Ver Tabela 3.4).

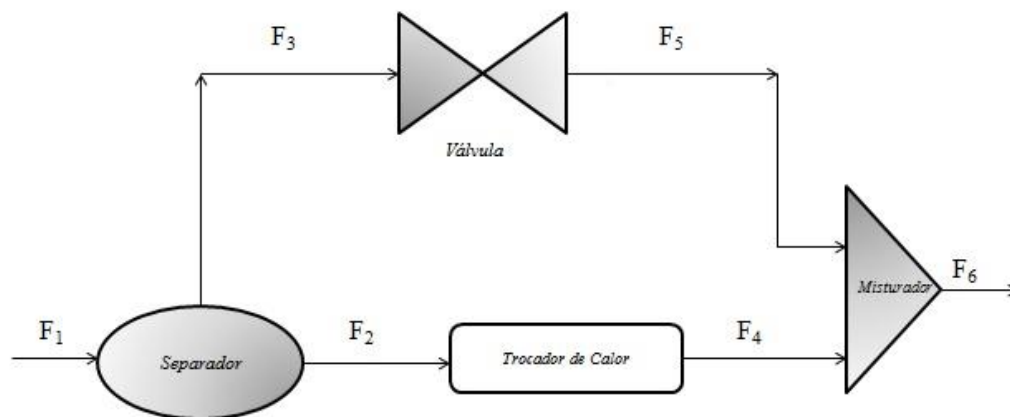


Figura 3.8 Trocador de calor com válvula de desvio (bypass)

Tabela 3.4 Dados do trocador de calor com bypass

Número das Correntes	Valor verdadeiro da Taxa	Valor medido da Taxa	Σ Taxa
1	100	101.9	1
2	64	68.45	1
3	36	34,65	1
4	64	64,20	1
5	36	36,64	1
6	100	98,88	1

Tabela 4.7 Dados reconciliados utilizando funções robustas: DEG linear

FUNÇÕES ROBUSTAS									
Corr	V	M	WLS	C	NC	NT	AL	Lo	F
1	100	101.9	100.8867	100.7421	100.8983	100.1536	100.1591	100.8149	100.8749
2	64	68.45	65.8333	65.3606	65.8630	64.5579	64.4618	65.4719	65.6306
3	36	34.65	35.0533	35.3814	35.0354	35.5957	35.6973	35.3430	35.2443
4	64	64.20	65.8333	65.3606	65.8630	64.5579	64.4618	65.4719	65.6306
5	36	36.44	35.0533	35.3814	35.0354	35.5957	35.6973	35.3430	35.2443
6	100	98.88	100.8867	100.7421	100.8983	100.1536	100.1591	100.8149	100.8749

Corr: Correntes; V: Valor verdadeiro; M: Valor medido.

Tabela 4.8 Resultados das iterações

Resultados das Iterações							
Funções	WLS	C	NC	NT	AL	Lo	F
Iterações	14	9	6	14	14	8	11

Observa-se na Tabela 4.8 que as funções *Mínimos Quadrados Ponderados*, *New Target*, *Alarm* e *Fair* obtiveram um maior número de iterações. Em se tratando de *Fair* sabe-se que ela é limitada a um valor constante e quando aplicada a um conjunto de dados com erro grande a estimação cresce ilimitadamente, o mesmo acontece com funções que possuem distribuição normal. Farias (2006) explica que quanto mais a c_F (constante de *Fair*) aumenta, sua função se aproxima da distribuição normal e o ajuste aumenta mais rapidamente em

função do erro. Como a c_F utilizada nesse estudo é a mesma aplicada por Ozyurt e Pike (2004), entende-se que o comportamento da *Fair* foi similar ao obtido no citado estudo.

A função *Normal Contaminada* obteve o menor número de iterações superando inclusive o estimador redescendente que deveria diante de somente um erro grosseiro ter menos iterações. Essa função e a de *Cauchy* aproximam-se de zero para valores de erros grandes, sendo que *Cauchy* mais lentamente. E a medida que o tamanho do erro grosseiro aumenta esses dois estimadores são os que apresentam melhores resultados (FARIAS, 2006).

Já com relação a capacidade de reconciliação das funções robustas na presença de erros grosseiros na Tabela 4.7, pode-se inferir que todas as funções aproximaram-se dos valores verdadeiros.

Com relação ao critério de desempenho baseado nas medidas contaminadas com erros a Figura 4.1 permite verificar como foi o comportamento de cada função nesse estudo de caso.

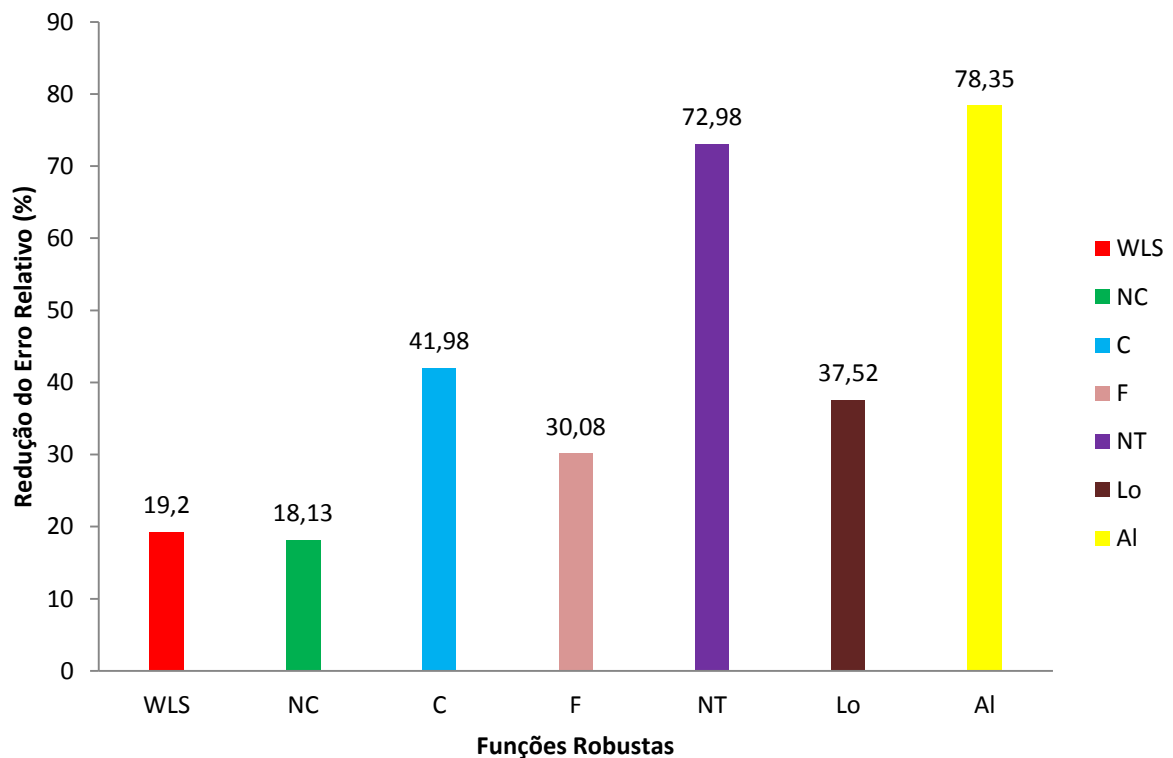


Figura 4.1 Análise do desempenho das funções robustas em função da redução do erro relativo para o Caso 4.

Conforme apresentado na Figura 4.1 a função que obteve uma maior redução do erro relativo foi a *Alarm* seguida da *New Target*. O que pode entender comparando as Tabelas 4.7

e 4.8 com a Figura 4.1 é que mesmo com um número de convergência superior, essas funções conseguem reduzir o impacto da contaminação da medida dada a sua robustez e reconciliar os dados. A robustez a que se refere é com relação a insensibilidade diante de erros grosseiros, que permite a obtenção de dados mais precisos e confiáveis.

ESTUDO DE CASO 5: IDENTIFICAÇÃO E DETECÇÃO DE ERROS GROSSEIROS: SISTEMA NÃO LINEAR

Este estudo de caso tem a particularidade de ser não linear e a temperatura da corrente 7 representada pela corrente F_{10} , possuir 5 *bias* positivos conforme mostrado na Figura 3.9 e os dados da Tabela 3.5.

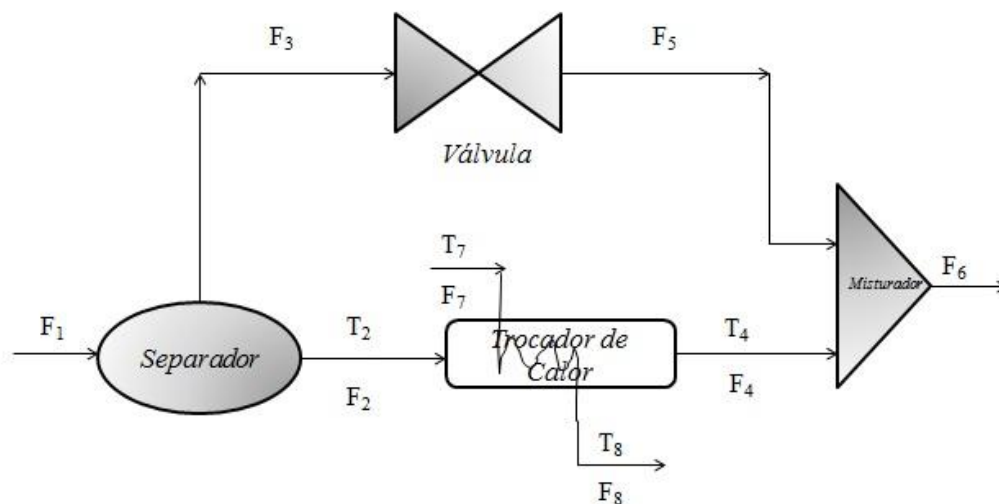


Figura 3.9 Trocador de calor com válvula bypass e balanço de energia

Tabela 3.5 Dados do trocador de calor com válvula bypass

Número de correntes	Valor verdadeiro da taxa	Valor verdadeiro da temperatura	Valor medido	Σ Taxa ou Σ Temperatura
1	100		101,91	1
2	64		64,45	1
3	36		34,65	1
4	64		64,20	1
5	36		36,44	1
6	100		98,88	1
7	140,6		140,0	1
2		90	90	1
4		51	51	1
7		20	25	1
8		43	43	1

Os resultados obtidos na reconciliação de dados aplicando funções robustas no estudo de caso 5 são mostrados na Tabela 4.9. E a comparação do número de iteração de cada função é apresentada na Tabela 4.10.

Tabela 4.9 Dados reconciliados utilizando funções robustas: DEG não linear

FUNÇÕES ROBUSTAS									
Corr	V	M	WLS	NC	C	F	NT	Lo	AI
1	100	101.9	100.14	100.14	100.01	100.04	99.96	100.00	100.06
2	64	64.45	64.34	64.34	64.28	64.28	64.33	64.28	64.30
3	36	36.65	35.79	35.80	35.72	35.76	35.63	35.72	35.75
4	64	64.20	64.34	64.34	64.28	64.28	64.33	64.28	64.30
5	36	36.44	35.79	35.80	35.72	35.76	35.63	35.72	35.75
6	100	98.88	100.14	100.14	100.01	100.04	99.96	100.00	100.06
7	140.6	140.0	140.21	140.21	140.14	140.11	140.04	140.14	140.18
8	90	90	89.05	89.05	89.31	89.33	89.79	89.33	89.18
9	51	51	51.69	51.69	51.47	51.43	51.14	51.46	51.58
10	20	25	23.84	23.84	24.12	24.08	24.75	24.13	23.97
11	43	43	44.77	44.77	45.16	45.17	45.98	45.18	44.95

Corr: Correntes; V: Valor verdadeiro; M: Valor medido.

Tabela 4.10 Resultado das iterações: Não linear

Resultados das Iterações							
Função	WLS	NC	C	F	NT	Lo	AI
Iterações	11	11	10	14	20	14	12

Neste estudo de caso, de acordo com a Tabela 4.10, apenas a *New Target* obteve o número de iterações superior aos demais. A função *Cauchy* apresentou a menor quantidade de iterações. Este comportamento também foi observado no caso 4. Mais uma vez as funções robustas apresentaram resultados coerentes na presença de desvios. A temperatura da corrente T_7 representada na Tabela 4.9 pela corrente 10, possuía 5 *bias* positivos e as funções demonstraram isso na reconciliação com base nos valores obtidos, a exceção foi a função *New Target* que além de apresentar um maior número de iteração foi sensível ao erro do conjunto de dados resultando em um valor equivalente ao medido.

Com relação ao critério de desempenho baseado nas medidas contaminadas com erros a Figura 4.2 permite verificar como foi o comportamento de cada função nesse estudo de caso.

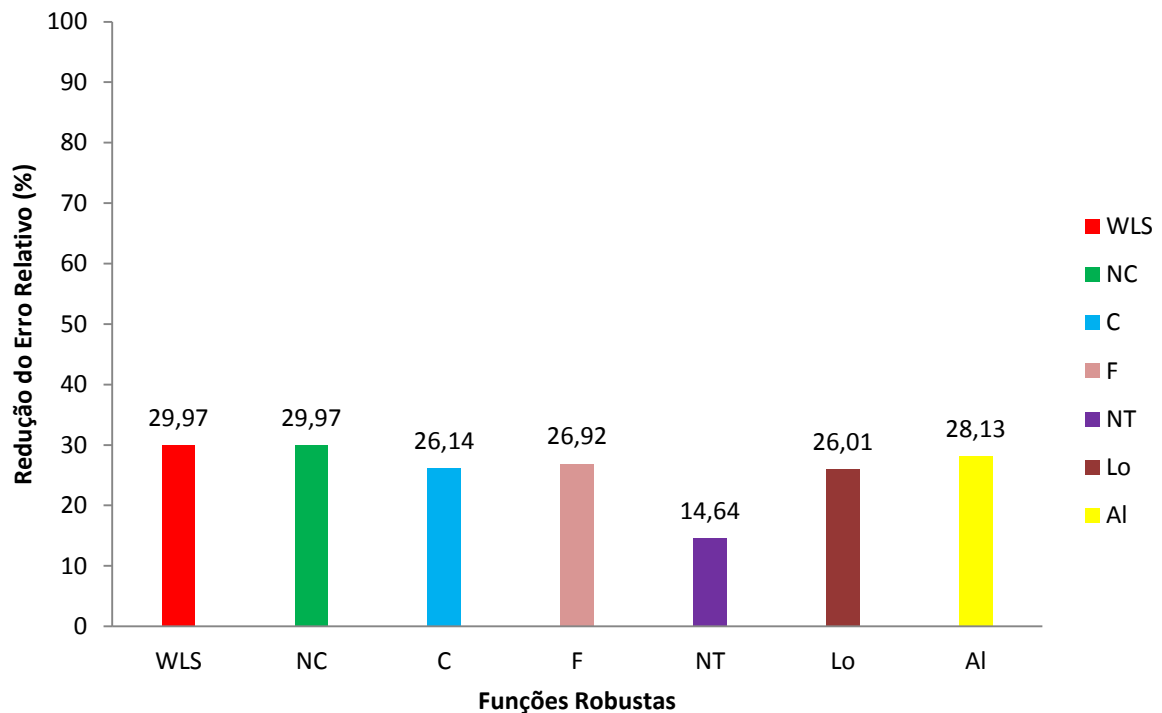


Figura 4.2 Avaliação do desempenho das funções robustas em relação a redução do erro relativo para o Caso 5

Observa-se na Figura 4.2 que a função *New Target* apresentou um valor inferior a 15 % para a redução do erro relativo e muito embora o valor reconciliado tenha se aproximado do valor medido, houve a atenuação do erro grosseiro. As demais funções apresentaram valores próximos de 30% para a redução do erro relativo, Zhang *et al.* (2010) em seu estudo realizaram testes com a magnitude de erro próximas as utilizadas neste teste e também obtiveram uma redução do erro entre 26 % as 35 %, considerando-as satisfatórias.

O que pode ser observado nesse estudo de caso é que muito embora algumas funções tenham características matemáticas diferentes, como por exemplo, a *New Target* e *Alarm* que possuem pontos de corte em sua formulação e por isso apresentam maior robustez em seus resultados, as funções robustas aqui estudadas da mais simples a mais complexa apresentaram resultados satisfatórios em um sistema não linear corrompido com erro em uma corrente. Isso corrobora com a literatura que versa que diante da presença de erro em somente uma corrente medida, todas as funções são equiparadas.

Entretanto, é importante salientar nesse estudo de caso que a temperatura da corrente 7 possuía 5 desvios da medida verdadeira, e nesse caso, apenas algumas funções conseguiriam demonstrar bom desempenho, o que não ocorreu nesse estudo. Diante disso, aventou-se a possibilidade de testar as funções em um sistema não linear com dados reais visando verificar o desempenho das funções.

ESTUDO DE CASO 6: O PROCESSO DE POLIMERIZAÇÃO

Conforme apresentado no capítulo anterior esse estudo de caso é referente aos reais de um processo de polipropileno. Estudado por Oliveira Jr. (2006) e Prata (2009) onde houve a estimação dos parâmetros da unidade industrial aplicando a técnica de reconciliação de dados e a reconciliação de dados aplicada em um modelo dinâmico do processo.

O sistema simplificado é mostrado na Figura 3.10.

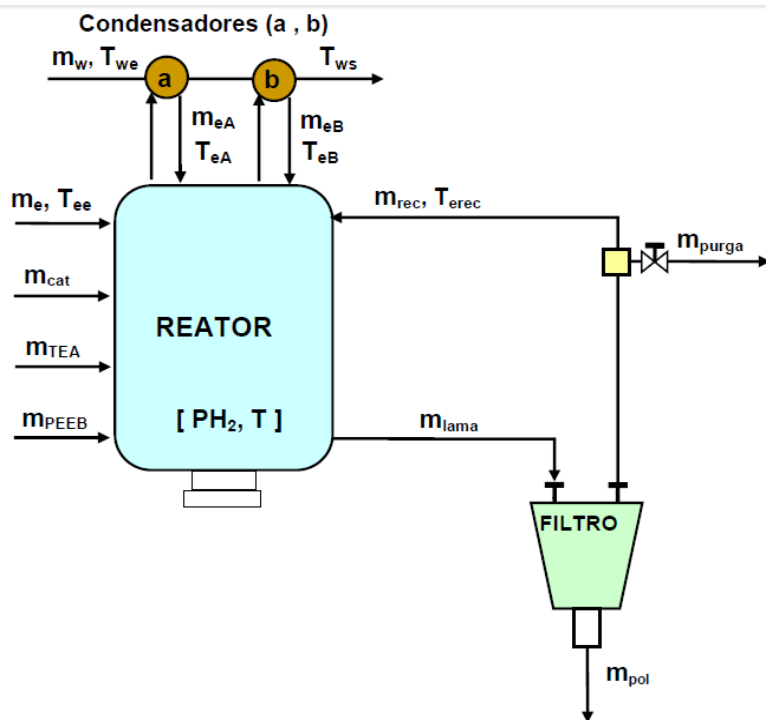


Figura 3.10 Processo de polimerização simplificado
Fonte: Oliveira Jr., 2006.

Nesse estudo de caso, foram reconciliados dados referentes a um ciclo de 2 horas de produção.

Cabe explicar que havia confiabilidade nesses dados medidos por conta das condições de controle do processo, e, portanto, além de erros aleatórios que eventualmente estão contidos nas medições considera-se que não existiam erros grosseiros.

As Figuras 4.3 a 4.9 são referente aos dados reconciliados com as funções robustas para a variável vazão de propeno (m_e), as Figuras 4.10 a 4.16 para a variável vazão do polímero (m_{pol}) e as Figuras 4.17 a 4.23 para a variável nível do reator (L).

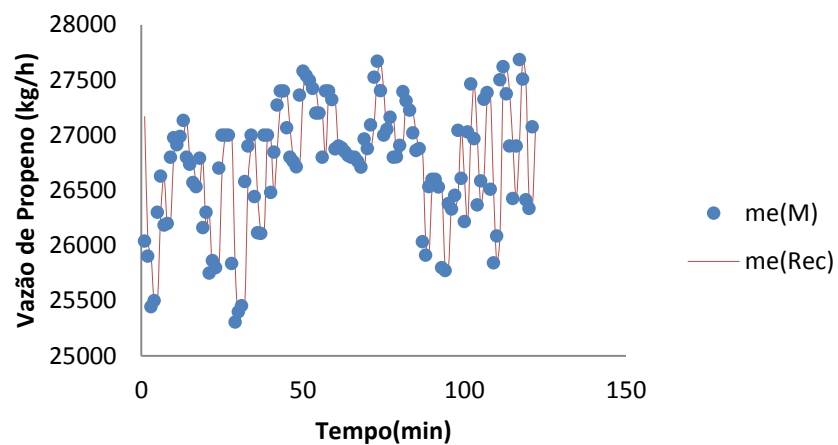


Figura 4.3 Variação da vazão medida de propeno e reconciliada através da função WLS

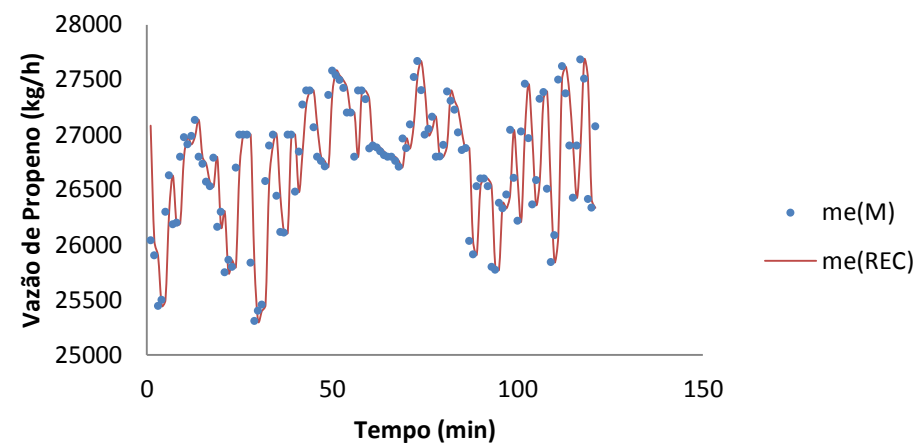


Figura 4.4 Variação da vazão medida de propeno e reconciliada através da função CAUCHY

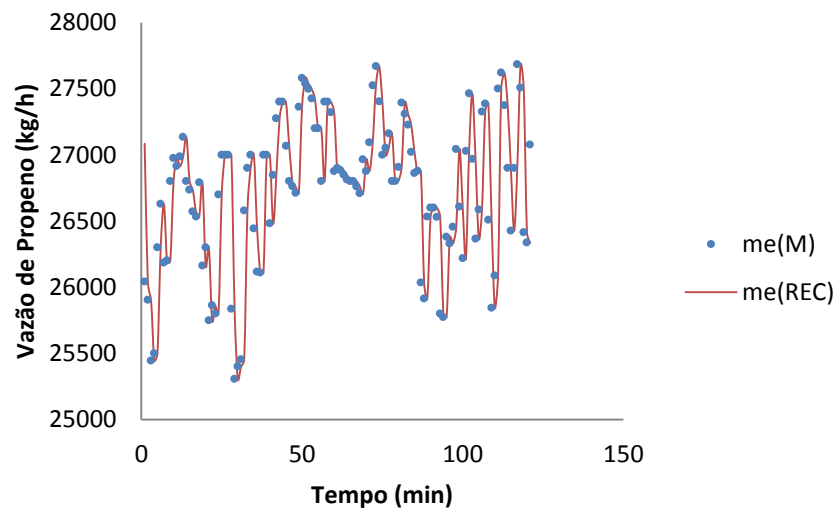


Figura 4.5 Variação da vazão medida de propeno e reconciliada através da função FAIR

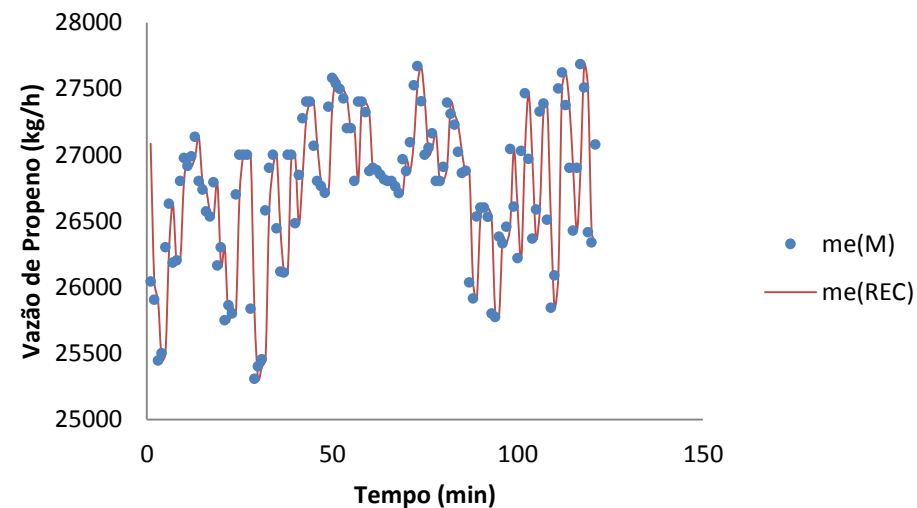


Figura 4.6 Variação da vazão medida de propeno e reconciliada através da função LOGÍSTICA

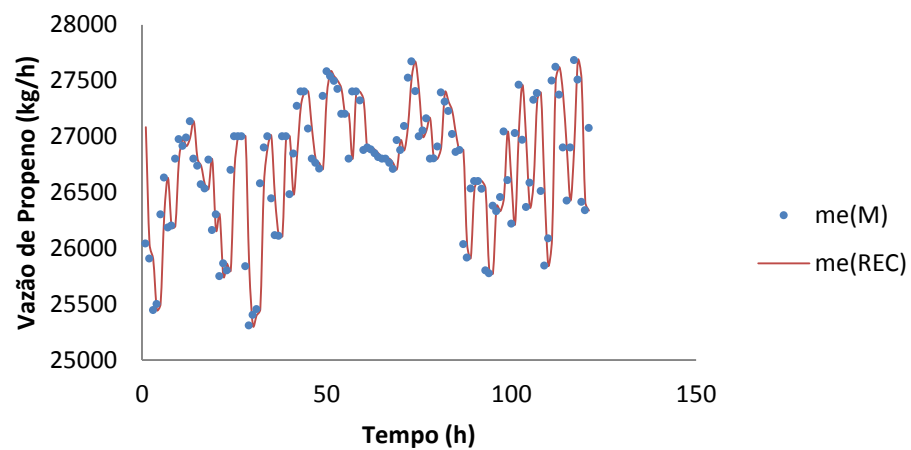


Figura 4.7 Variação da vazão medida de propeno e reconciliada através da função ALARM

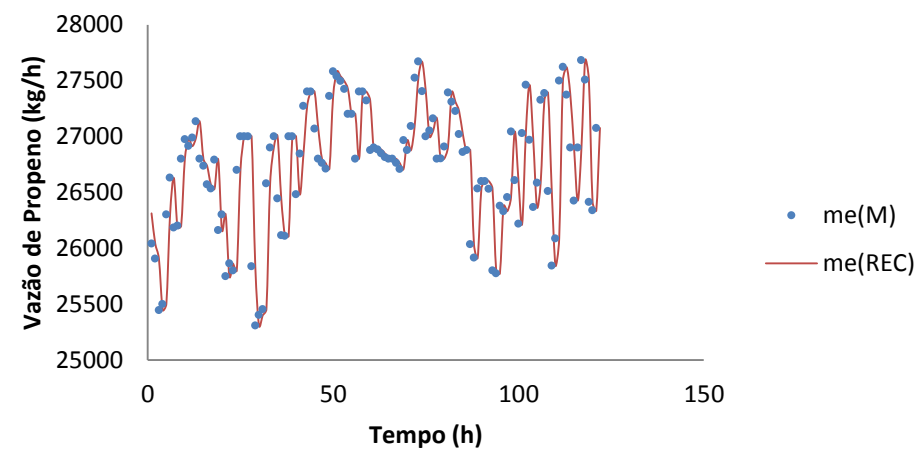


Figura 4.8 Variação da vazão medida de propeno e reconciliada através da função NEW TARGET

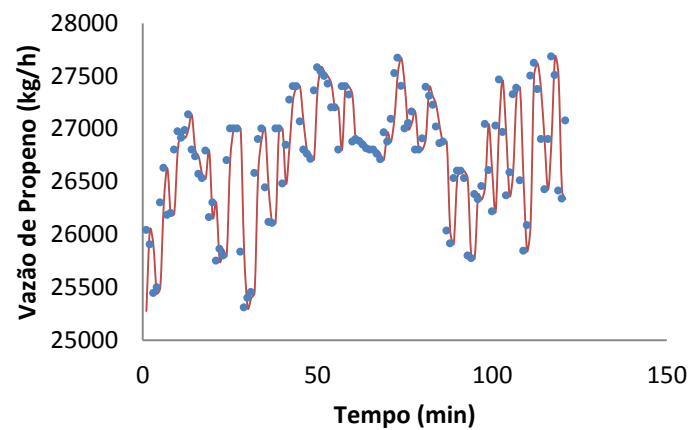


Figura 4.9 Variação da vazão medida de propeno e reconciliada através da função NORMAL CONTAMINADA

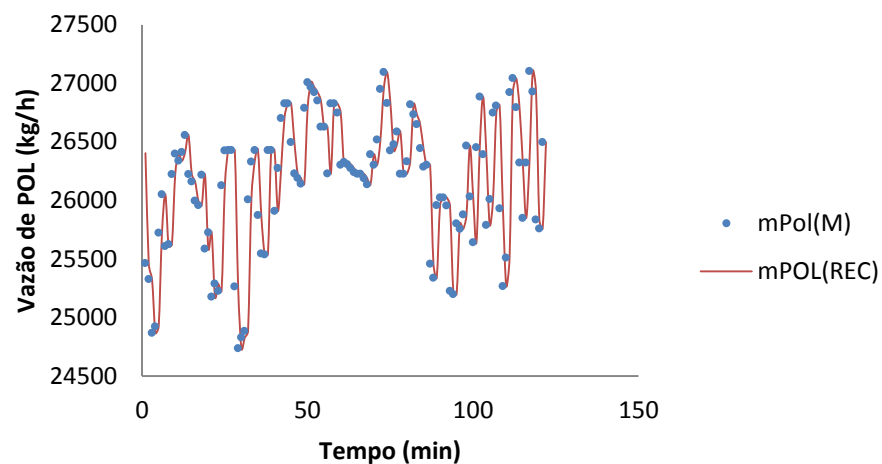


Figura 4.10 Variação da vazão medida de polímero e reconciliada através da função WLS

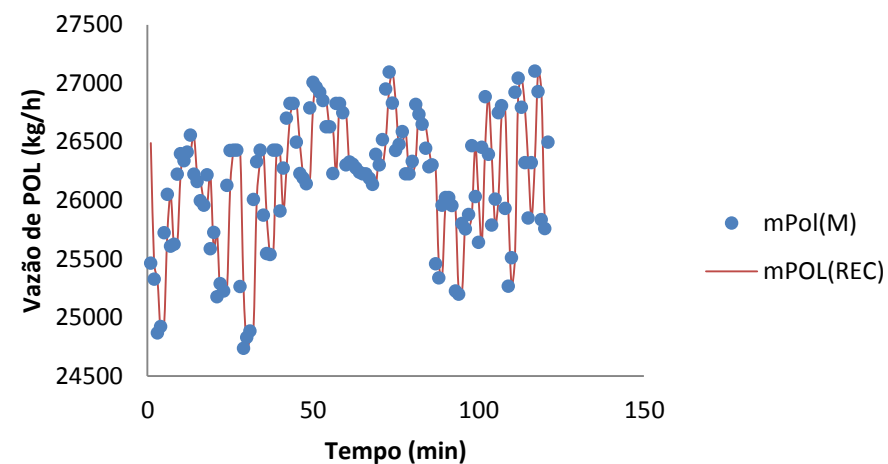


Figura 4.11 Variação da vazão medida de polímero e reconciliada através da função ALARM

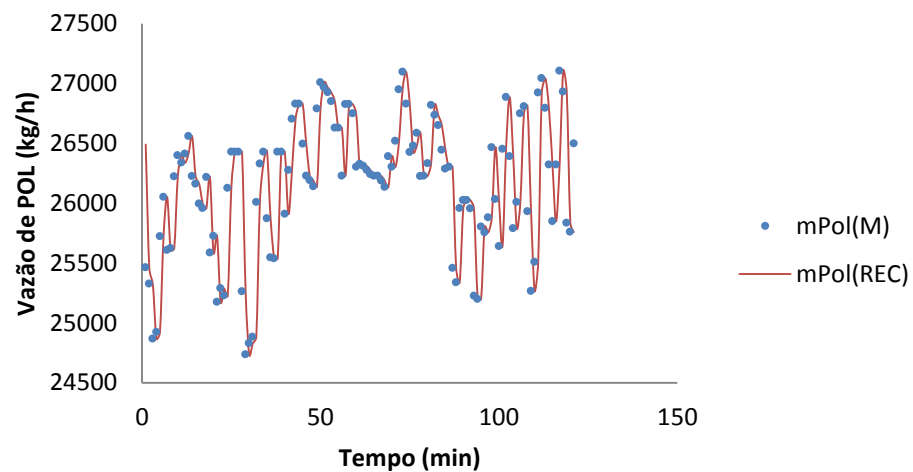


Figura 4.12 Variação da vazão medida de polímero e reconciliada através da função CAUCHY

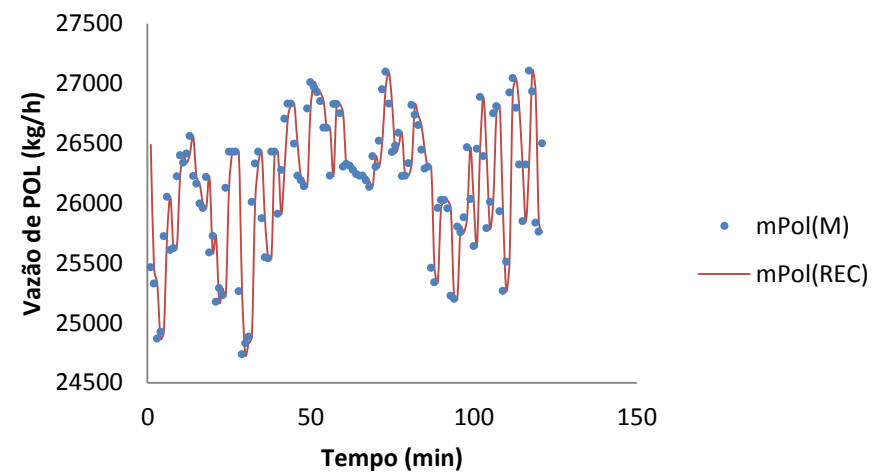


Figura 4.13 Variação da vazão medida de polímero e reconciliada através da função FAIR

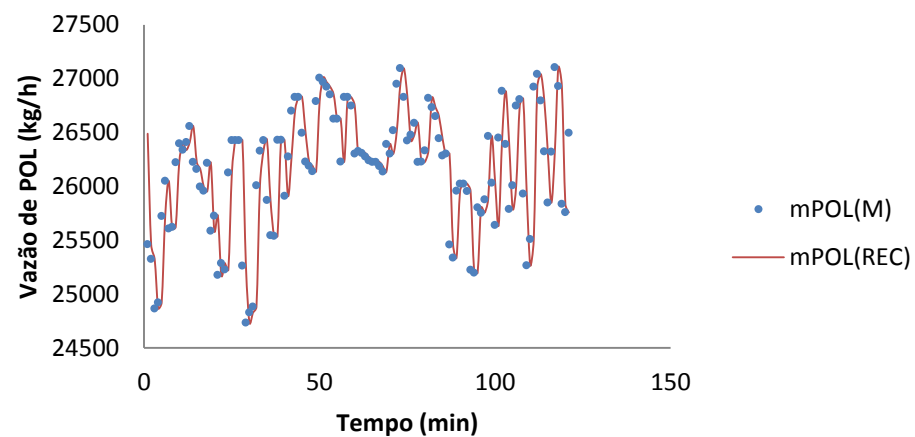


Figura 4.14 Variação da vazão medida de polímero e reconciliada através da função LOGÍSTICA

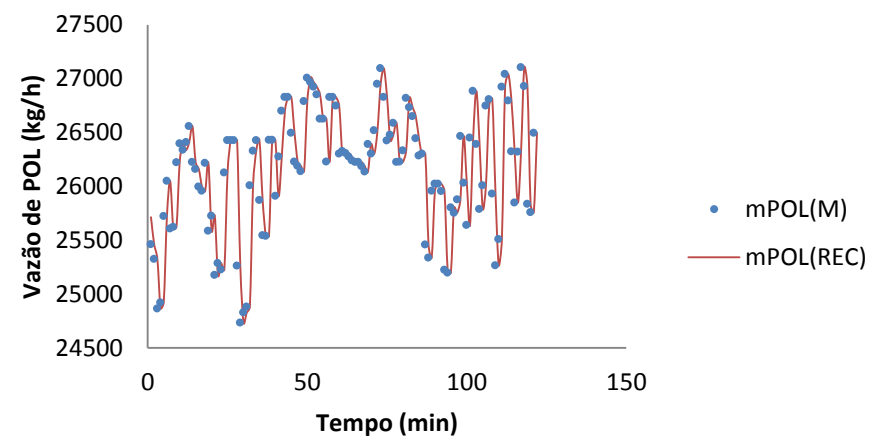


Figura 4.15 Variação da vazão medida de polímero e reconciliada através da função NEW TARGET

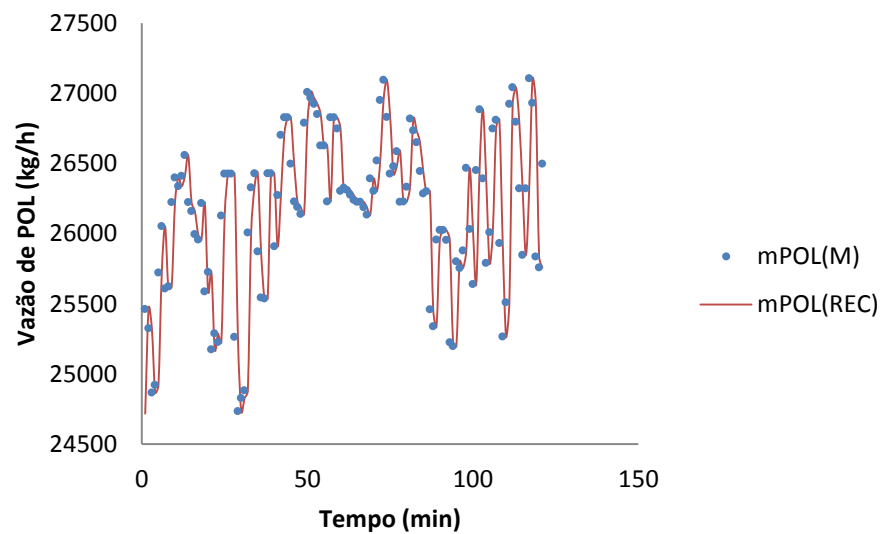


Figura 4.16 Variação da vazão medida de polímero e reconciliada através da função NORMAL CONTAMINADA

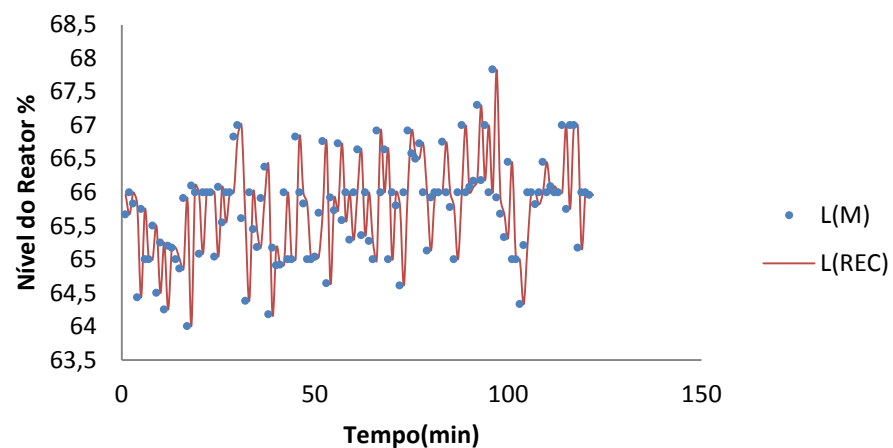


Figura 4.17 Variação do nível medido do reator e reconciliado através da função WLS

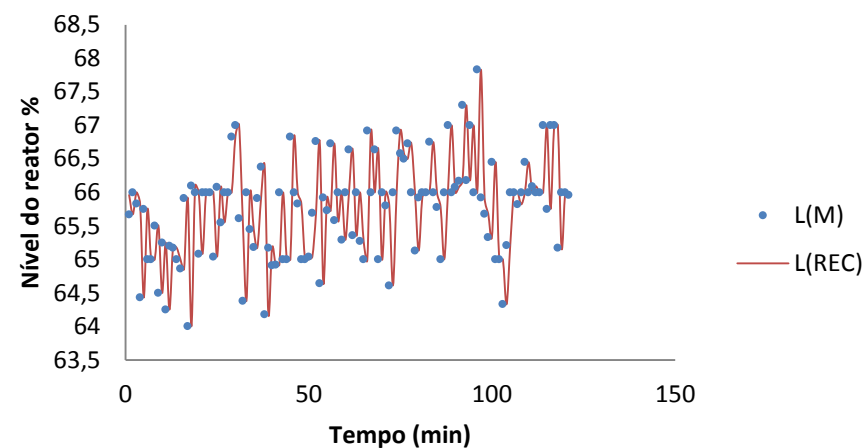


Figura 4.18 Variação do nível medido do reator e reconciliado através da função CAUCHY

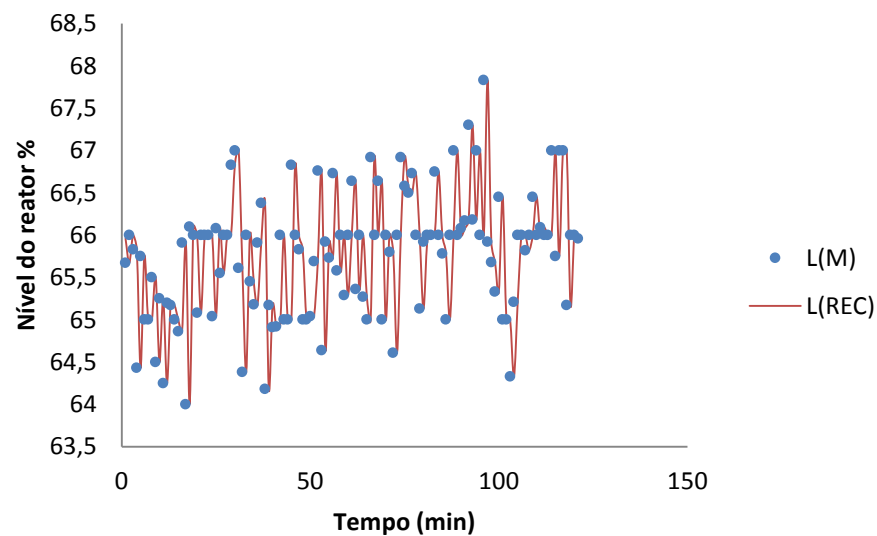


Figura 4.19 Variação do nível medido do reator e reconciliado através da função FAIR

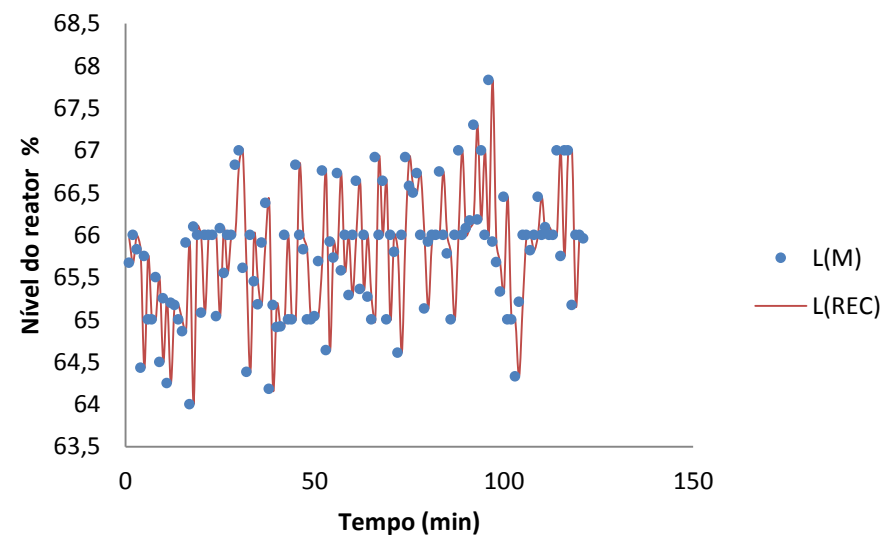


Figura 4.20 Variação do nível medido do reator e reconciliado através da função LOGÍSTICA

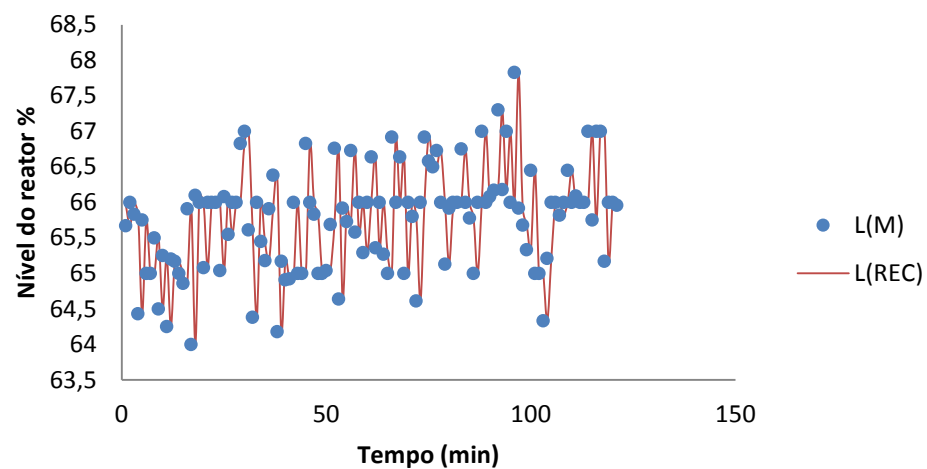


Figura 4.21 Variação do nível medido do reator e reconciliado através da função ALARM

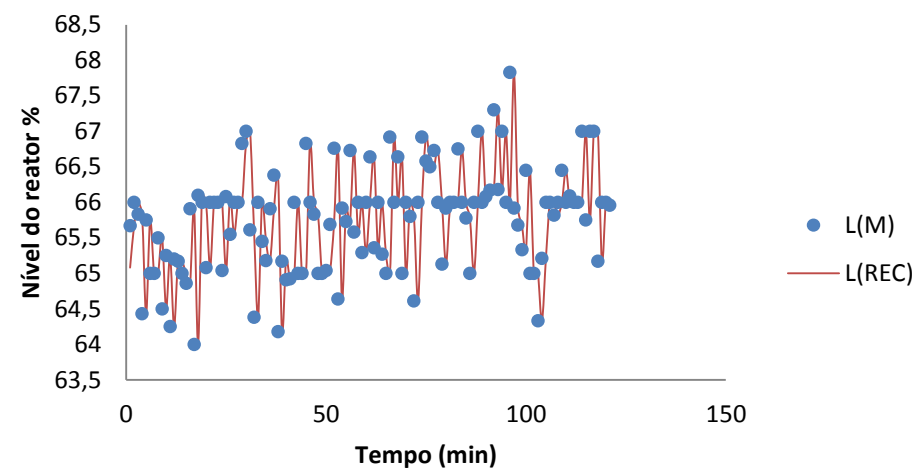


Figura 4.22 Variação do nível medido do reator e reconciliado através da função NEW TARGET

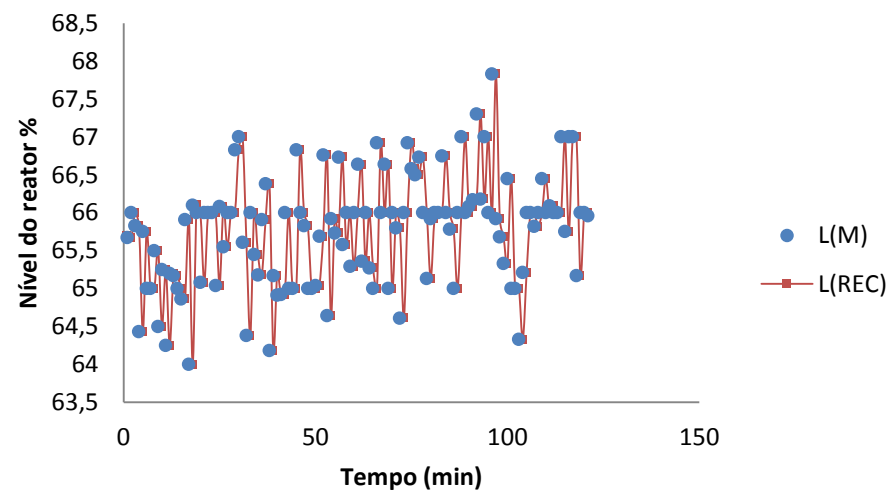


Figura 4.23 Variação do nível medido do reator e reconciliado através da função NORMAL CONTAMINADA

Com base nas Figuras 4.3 a 4.23 sobre as vazões medidas e reconciliadas com as funções robustas de propeno, polipropileno e nível do reator é possível observar que todas as funções robustas obtiveram desempenho similar favorável na reconciliação com os dados reais, e ficou demonstrado que não havia nenhum ajuste com desvio sistemático nas 3 (três) variáveis apresentadas.

Oliveira Jr.(2006) explica que a vazão de propeno fresco no processo de polimerização é observada através de um controlador de fluxo ajustado pelo indicador de nível do reator sendo seus dados medidos confiáveis, o que para o entendimento desse estudo contribuiu na reconciliação com as funções robustas. Para essa variável as funções robustas *New Target* Figura 4.8 e *Normal Contaminada* na Figura 4.9 apresentaram uma melhor precisão nos ajustes dos dados.

Com relação a vazão de polímero das Figuras 4.10 a 4.16 todas as funções obtiveram resultado esperado onde os dados reconciliados foram de acordo aos dados medidos, as funções *New Target* Figura 4.11 e *Normal Contaminada* Figura 4.12 foram as que mais se aproximaram nesse aspecto. As oscilações inicial e final nos valores que podem ser percebidas nos gráficos e nos dados reconciliados dessa variável devem-se a manipulação durante o processo de outras variáveis que compõem o balanço energético e que não fizeram parte do escopo desse estudo.

A variável de entrada nível do reator é extremamente importante dada às diversas espécies que participam do processo de polimerização. Como pode ser visto nas Figuras 4.17 a 4.23 mais uma vez todas as funções robustas obtiveram desempenho satisfatório, mesmo quando houve oscilações referente à alimentação de propeno fresco para o controle do processo.

Farias (2006) explica que quando são considerados apenas erros aleatórios presentes nas medidas, as funções objetivo robustas realizarão ajustes semelhantes ao realizado pela função objetivo convencional, nessa dissertação representada pela função de Mínimos Quadrados Ponderados, todas elas com um intervalo restrito ao redor da média. O que contribui com os resultados obtidos nesse estudo de caso com dados reais de processo.

Em relação às variáveis estimadas através das funções robustas, observou-se que a função *Alarm* apresentou erro de convergência nas variáveis **m_{PEEB}** que era variável medida e **PEEB** que era variável não medida porém observável, resultando em dados negativos bem como a estimação sendo interrompida prematuramente. Diante disso, fez-necessário a inclusão do limite inferior para obter um resultado favorável na reconciliação.

4.3 AVALIAÇÃO DE DESEMPENHO DAS FUNÇÕES ROBUSTAS NA PRESENÇA DE ERROS GROSSEIROS NO ESTUDO DE CASO 6

Para o Estudo de Caso 6 do processo de polimerização, foram acrescentados ruídos sistemáticos (*bias*) de 3σ e 5σ na variável m_{CAT} . A escolha dessa variável é devido ela ser a vazão de alimentação do catalisador Ziegler Natta, de extrema importância para o processo. As Figuras 4.24 a 4.30 representam a contaminação da medida da vazão do catalisador com 3σ positivos e as Figuras 4.31 a 4.37 a contaminação com 5σ positivos.

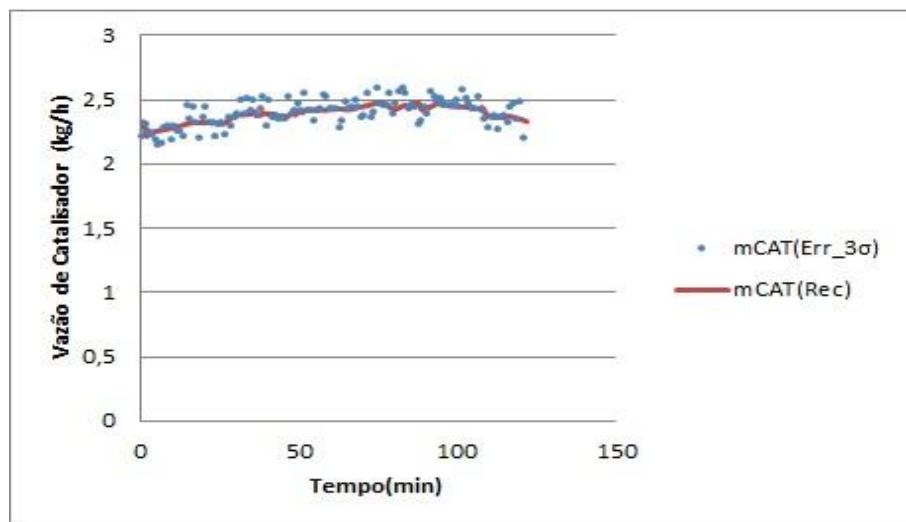


Figura 4.24 Variação da vazão medida contaminada com erro e reconciliada com WLS

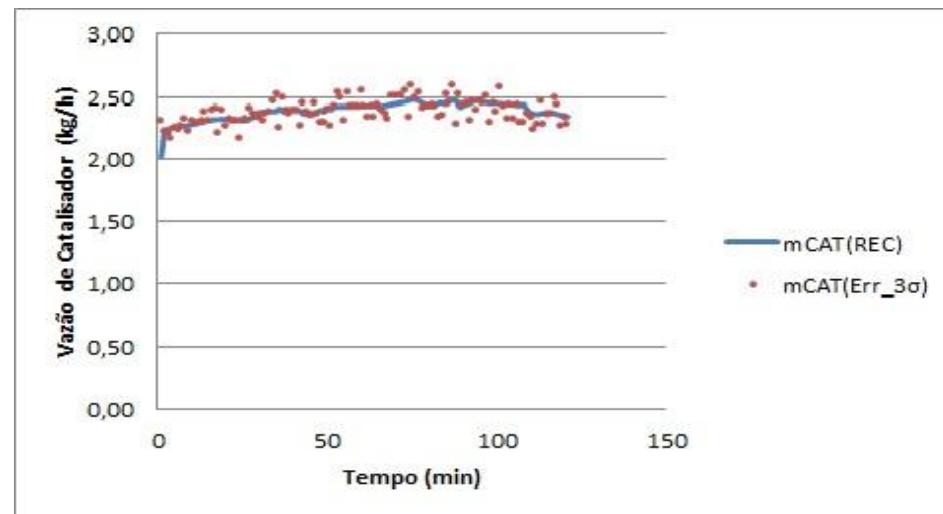


Figura 4.25 Variação da vazão medida contaminada com erro e reconciliada com ALARM

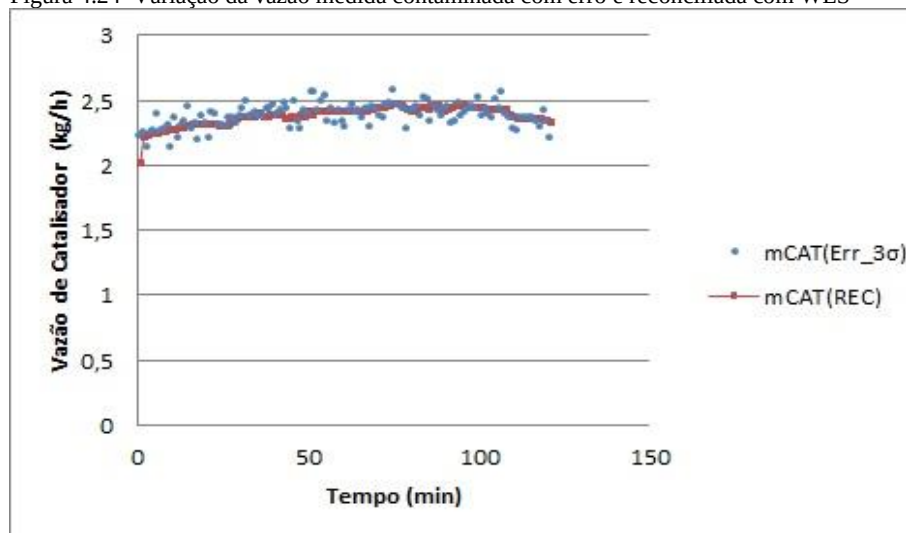


Figura 4.26 Variação da vazão medida contaminada com erro e reconciliada com CAUCHY

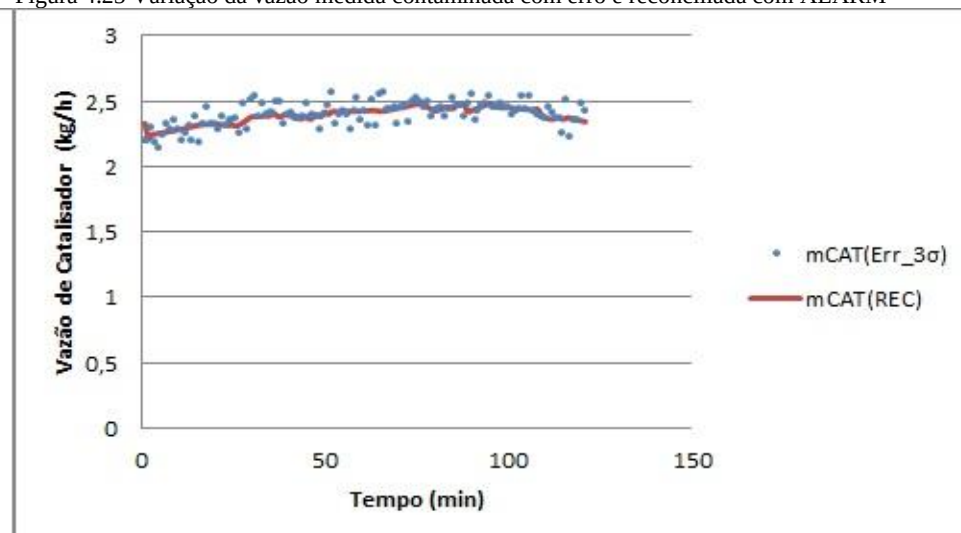


Figura 4.27 Variação da vazão medida contaminada com erro e reconciliada com FAIR

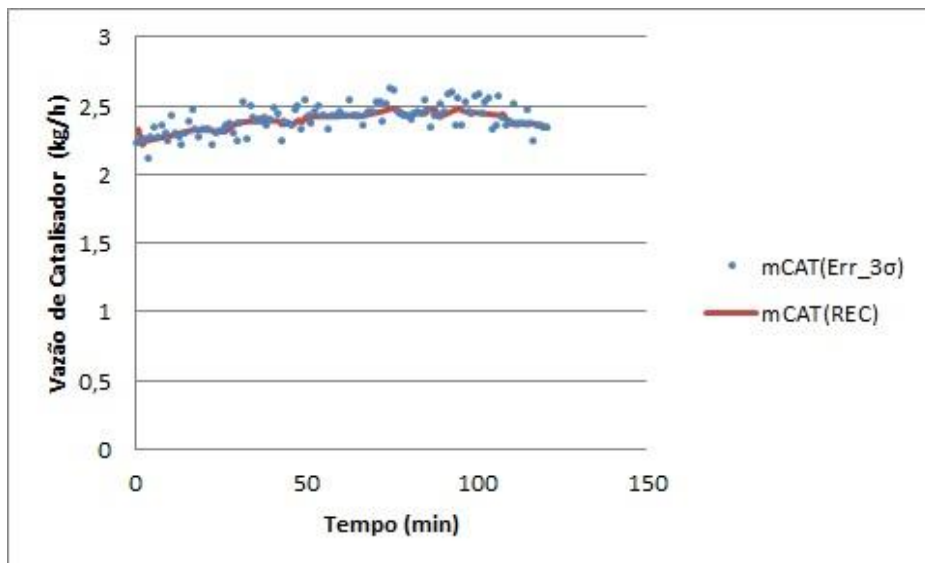


Figura 4.28 Variação da vazão medida contaminada com erro e reconciliada com LOGÍSTICA

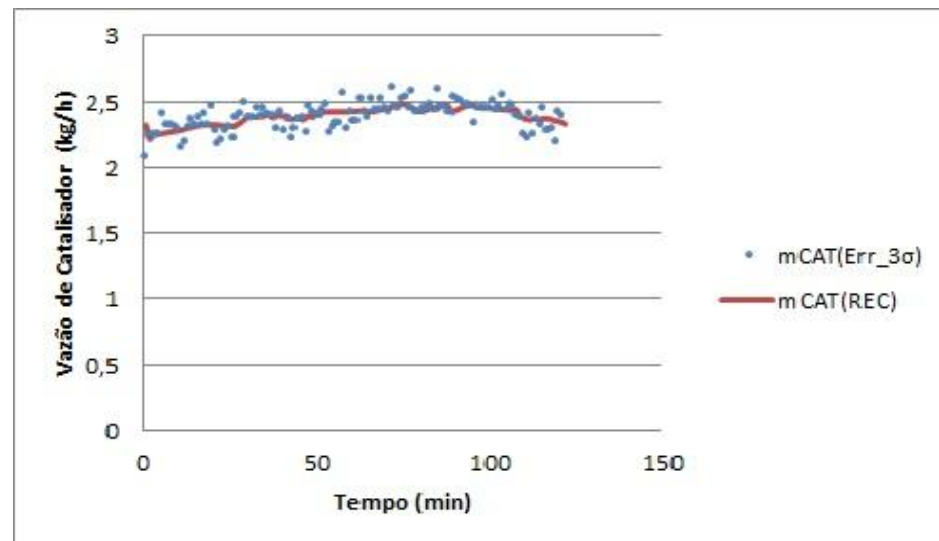


Figura 4.29 Variação da vazão medida contaminada com erro e reconciliada com NEW TARGET

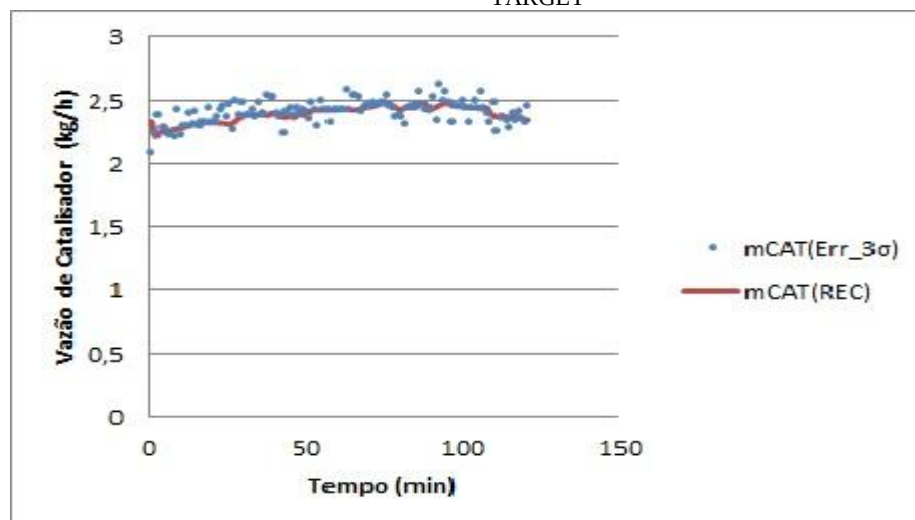


Figura 4.30 Variação da vazão medida contaminada com erro e reconciliada com NORMAL CONTAMINADA

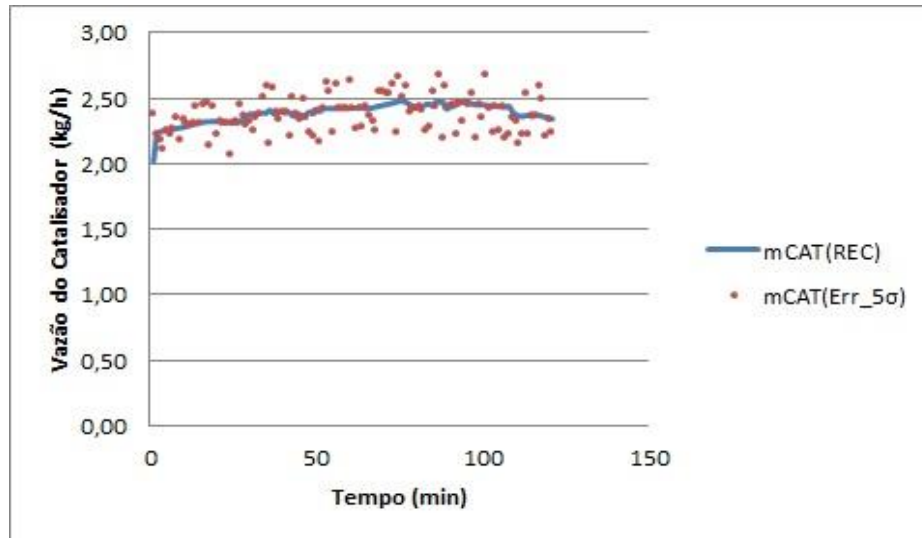


Figura 4.31 Variação da vazão medida contaminada com erro e reconciliada com ALARM

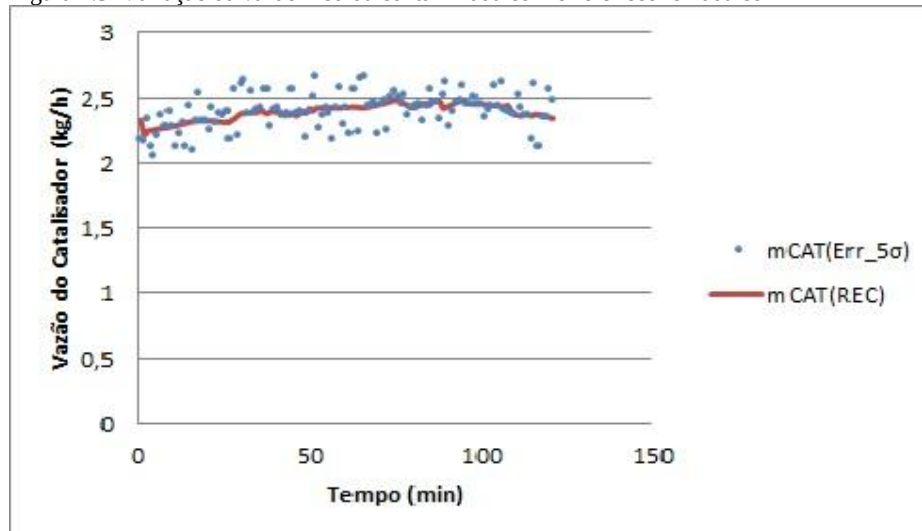


Figura 4.33 Variação da vazão medida contaminada com erro e reconciliada com FAIR

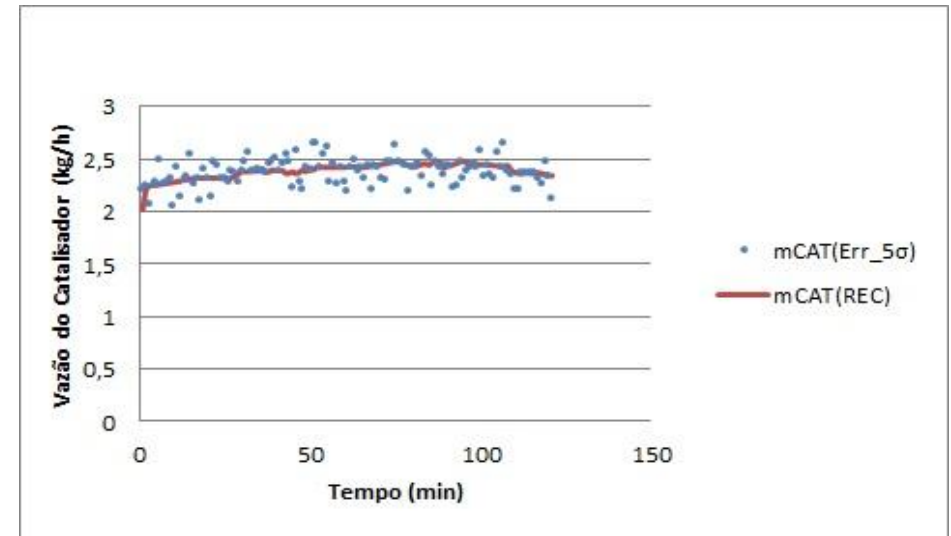


Figura 4.32 Variação da vazão medida contaminada com erro e reconciliada com CAUCHY

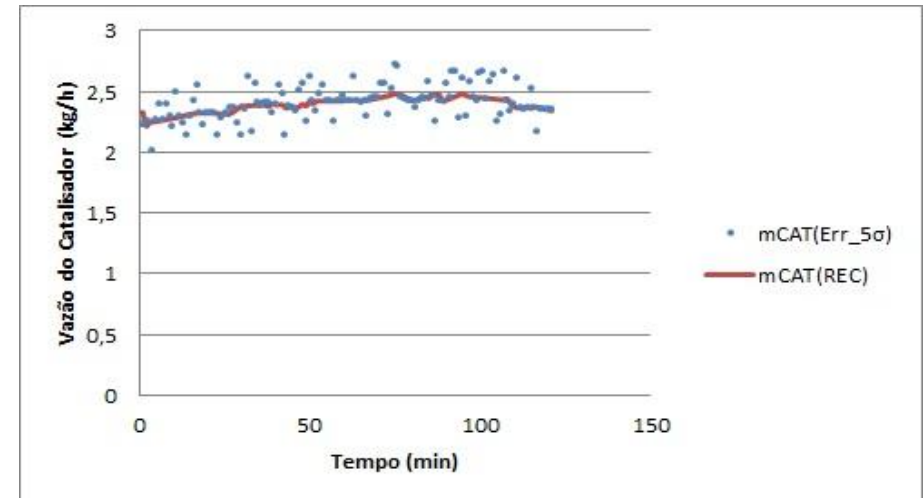


Figura 4.34 Variação da vazão medida contaminada com erro e reconciliada com LOGÍSTICA

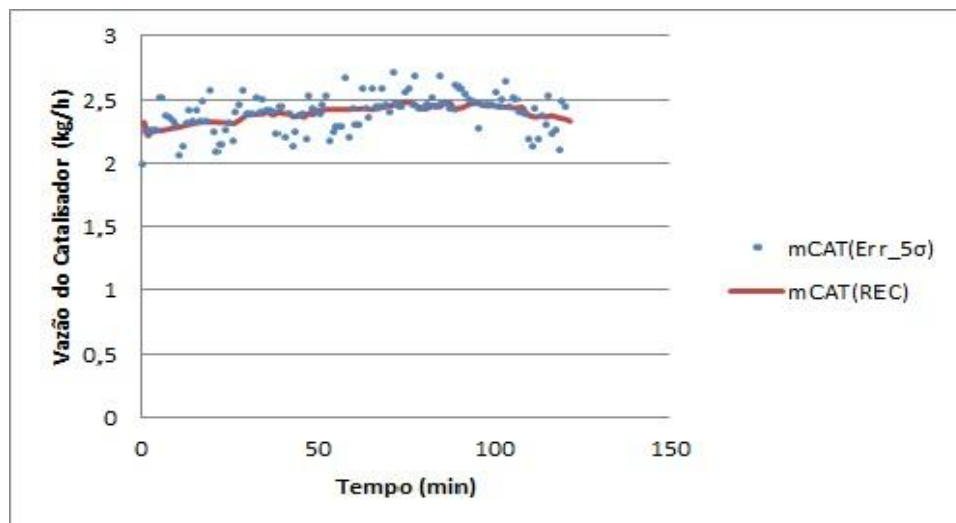


Figura 4.35 Variação da vazão medida contaminada com erro e reconciliada com NEW TARGET

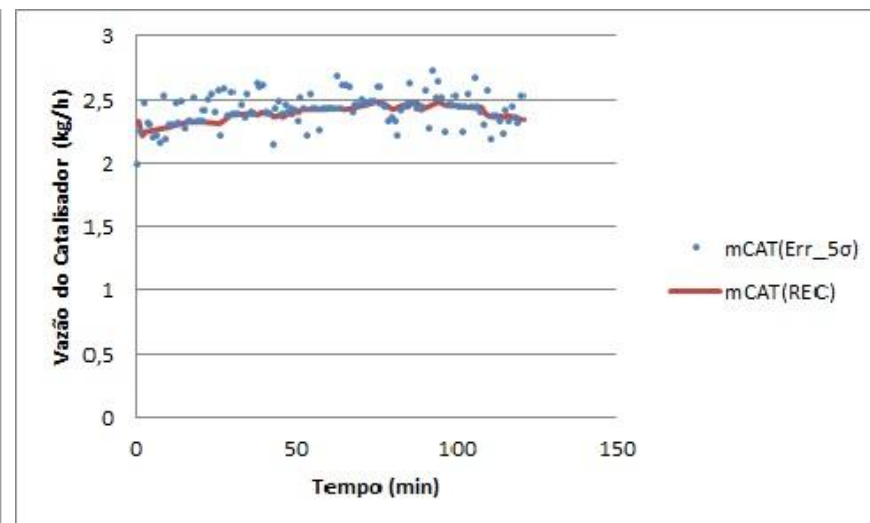


Figura 4.36 Variação da vazão medida contaminada com erro e reconciliada com NORMALCONTAMINADA

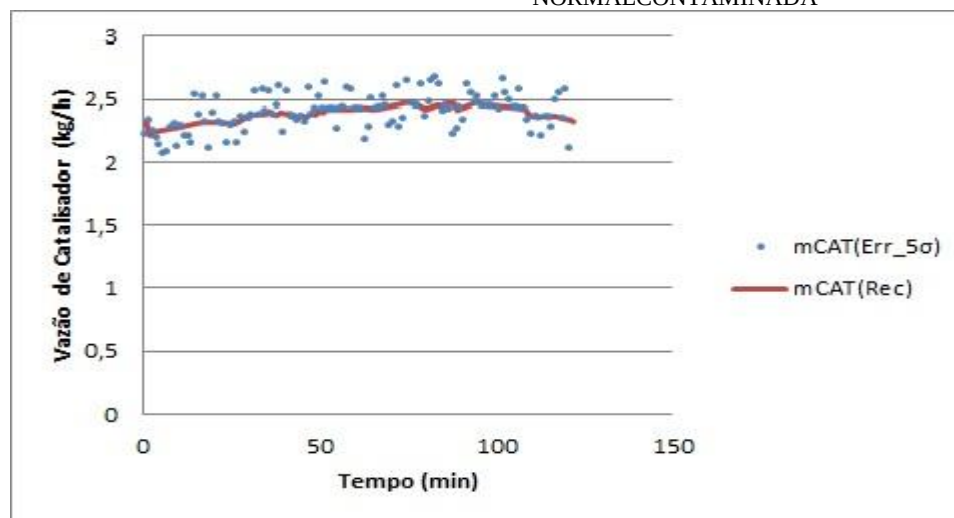


Figura 4.37 Variação da vazão medida contaminada com erro e reconciliada com WLS

Em relação a geração de ruídos na variável m_{CAT} pôde-se observar que algumas funções foram insensíveis a presença de *bias*.

Para fins de análise de desempenho das funções robustas aplicou o critério de redução do erro relativo (Zhang *et al.* 2010) sendo gerado um gráfico com o percentual de cada uma das funções.

A Figura 4.38 representa o *bias* de $3*\sigma$, ou seja, três vezes o desvio padrão da medida que para a variável m_{CAT} é 0.05.

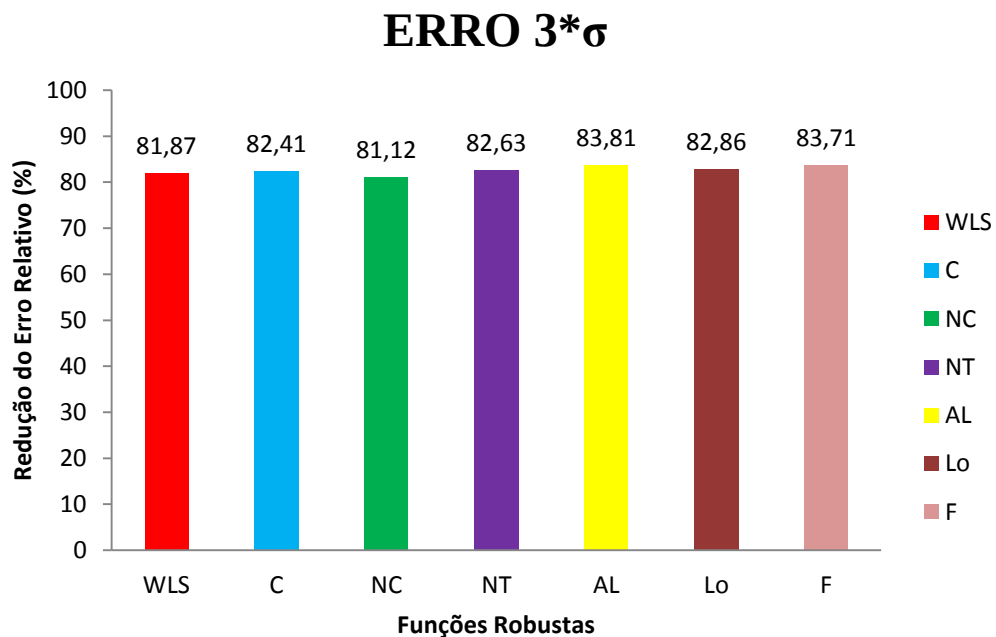


Figura 4.38 Análise do desempenho das funções robustas em relação a redução do erro relativo.

É possível observar que as funções robustas *Alarm* e *Fair* obtiveram desempenhos com relação a redução do erro relativo acima mas não destoante das outras funções. Entretanto, em se tratando de dados reais e com um desvio padrão pequeno como é o caso da variável m_{CAT} , esse resultado é importante, pois, mostra que, conforme foi observado na Figura 3.7, a função de influência quando a *Alarm* e a *Fair* estão na margem do erro de três vezes o desvio da medida elas possuem comportamentos semelhantes.

Com relação ao gráfico da redução do erro relativo para o *bias* de $5*\sigma$ onde as funções robustas obtiveram uma faixa de reconciliação em relação aos dados com um erro maior e, portanto, mais significativo para a medida, é possível observar na Figura 4.39 que mais uma vez as funções tiveram comportamento semelhante, contudo a função *Alarm* apresentou uma maior redução do erro diante das outras funções.

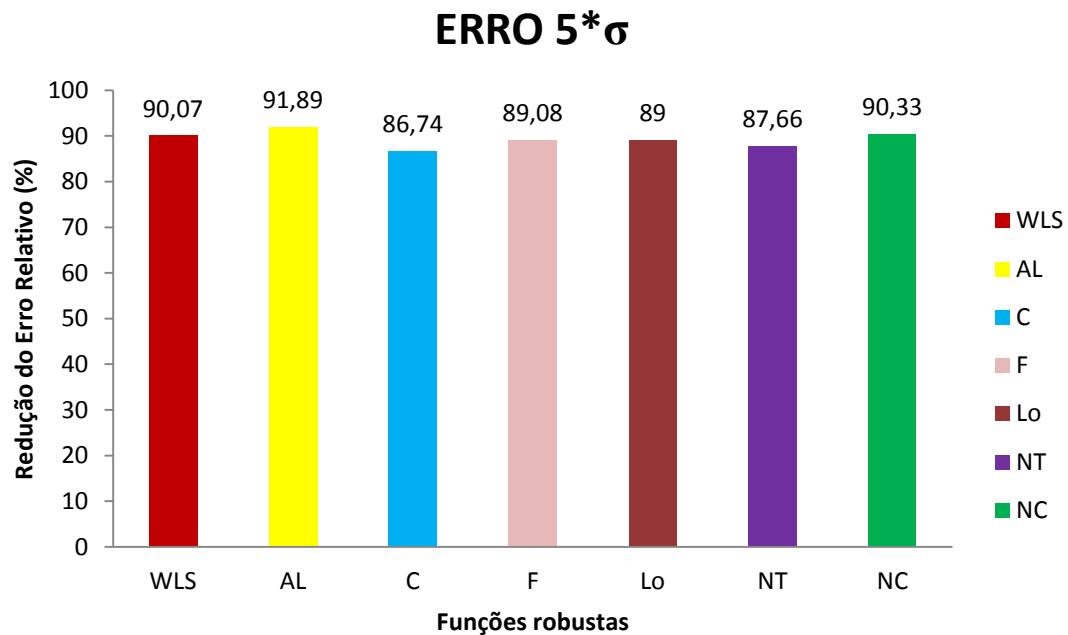


Figura 4.39 Análise do desempenho das funções robustas em relação a redução do erro relativo

Com base no desempenho dessas funções robustas, segue a lógica de que para um erro grosseiro de até três vezes desvios padrões, elas apresentam o mesmo desempenho inclusive similar a função de *Mínimos Quadrados Ponderados* com redução de erro em aproximadamente 90%. Mas com o aumento da magnitude do erro grosseiro, elas apresentam comportamento diferente. Entretanto, para esse estudo de caso utilizando dados reais a função normal tradicional também teve comportamento favorável ficando muito próxima a *Normal Contaminada* com redução de erro também em torno de 90%. O bom desempenho da função *Normal Contaminada* também pode ser visto nos trabalhos de Farias (2006) e Ozyurt e Pike (2004) no uso de estudos de casos da literatura em sistemas lineares.

Contudo, torna-se pertinente observar o que a quantidade de iterações dessas funções robustas para um erro dessa magnitude pode agregar a esse estudo, o que é mostrado na Figura 4.40.

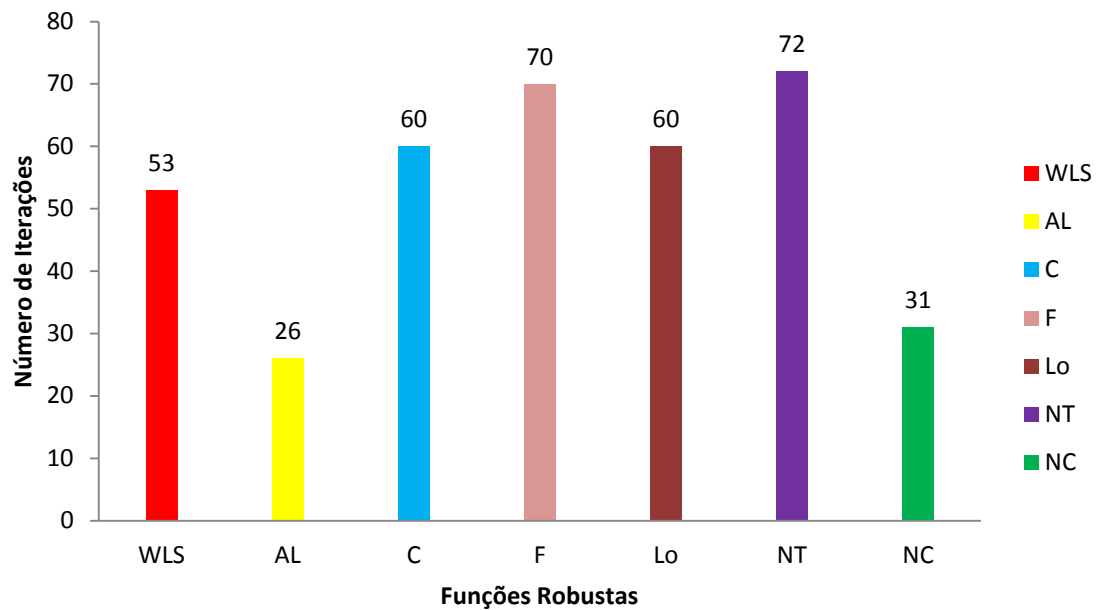


Figura 4.40 Número de iterações em relação a reconciliação de dados com as funções robustas

Dessa forma, muito embora as funções tenham tido desempenho semelhantes na redução do erro relativo, a *Alarm* destoa das demais com relação ao número de iterações na reconciliação dos dados corrompidos com erros. Ela apresenta um número reduzido de iterações principalmente com relação a *New Target* que muito embora tenha concepção matemática semelhante a *Alarm*, apresentou quantidade maior que duas vezes do número de iterações.

Torna-se, portanto claro que função *Alarm redescendente* apresenta bons resultados tanto com pequenos desvios quanto com grandes desvios, e também com relação aos critérios de desempenhos baseado no erro médio relativo e no número de iterações com relação a função objetivo.

5 CONCLUSÃO

O que se conclui desse estudo é que o uso das funções robustas na reconciliação de dados, seja ela em estudos da literatura ou em dados reais de uma planta industrial, é de grande valia, sobretudo por não necessitar de rotinas computacionais extensas e de difícil compreensão.

Com relação a reconciliação de dados em sistemas lineares e não lineares, as funções robustas apresentaram grau considerável de eficiência, com um pequeno número de iterações e resultados próximos aos valores verdadeiros nos estudos de casos extraídos da literatura. A exceção foi o exemplo do sistema linear da cloração do metano onde as funções robustas *Normal Contaminada* e *Alarm* apresentaram problemas de convergência sendo necessário atribuir limites a suas estimações.

Na geração do ruído nas correntes indica-se que para pequenos erros, as funções robustas apresentam resultados similares e a sua escolha deve ser guiada por outros critérios como convergência, número de avaliações da função objetivo, critério de redução do erro relativo. Contudo, com o aumento do erro e do número de correntes, para os estudos de casos extraídos da literatura, as funções *New Target* e *Alarm* apresentaram vantagens em relação às demais funções. No estudo de caso do processo de polimerização, onde foi testado um critério de desempenho diferenciado, pode-se inferir que as funções robustas *Normal Contaminada* e *Alarm* mesmo com problemas de convergência inicialmente na variável m_{PEEB} apresentaram melhor eficiência comparada as demais.

Com relação ao uso de dados reais de planta industrial é importante analisar a escolha das variáveis a serem estudadas com relação aos erros grosseiros contidos ou gerados nas medidas. Nessa dissertação a variável escolhida para geração do erro grosseiro m_{CAT} possuía um desvio pequeno em relação a sua vazão medida e por conta disso não foi possível obter melhores critérios de desempenho das funções robustas estudadas.

SUGESTÕES PARA TRABALHOS FUTUROS

- 1) Realizar a reconciliação de dados utilizando funções robustas com todas as etapas que competem a essa técnica em ambiente industrial. O que se pretende é verificar totalmente os impactos dos ajustes nos custos das indústrias;
- 2) Estudar as funções robustas em modelos dinâmicos visando comparar a eficiência das mesmas em função de cada batelada;
- 3) Implementar um teste de detecção de erro robusto que busque contribuir com as estimações robustas.

6 REFERÊNCIAS

ABUL-EL-ZEET, Z.H., BECERRA, V.M., ROBERTS, P.D., Combined Bias And Outlier Identification In Dynamic Data Reconciliation, **Computers and Chemical Engineering**, v.26, p. 921-935, 2002.

ALAMGIR, A. A., KHAN, S. A., KHAN, D. M., KHALIL, U. A new efficient redescending M-estimator: Alamgir Redescending M-estimator. **Research Journal of Recent Sciences**, v. 2(8), pp. 79-91, 2013.

ALBUQUERQUE, J.S., BIEGLER, L.T., Decomposition algorithms for on-line estimation with nonlinear models, **Computers and Chemical Engineering**, v.19,p. 1031-1039, 1995.

ALBUQUERQUE, J., BIEGLER, L. Data reconciliation and gross error detection for dynamic systems. **AIChE Journal**, n. 42, p. 2841-2856, 1996.

ALICI, S., EDGAR, T.F., Nonlinear Dynamic Data Reconciliation via Process Simulation Software and Model Identification Tools, **Industrial and Engineering Chemistry Research**, v. 41, p. 3984-3992, 2002.

ALLENDE, G. B., GONÇALVES, M. L. N., Local convergence analysis of a proximal Gauss-Newton method under a majorant condition preprint. **Journal of Complexity**. v. 27, p. 111-125, 2011.

ALMASY, G. A., SZTANO, T. Checking and correction of measurements on the basis of linear system model, **Problems of Control and Information Theory**, v. 4, p. 57-69, 1975.

ARGYROS, I. K., MAGREÑAN, Á. A., Local convergence analysis of proximal Gauss-Newton method for penalized nonlinear least squares problems. **Applied Mathematics and Computation**, v. 241, p. 401-405, 2014.

ARORA, N., BIEGLER, L.T., Redescending estimators for data reconciliation and parameter estimation. **Computers and Chemical Engineering**. v.25, p.1585-1599, 2001.

ARORA, N., BIEGLER, L.T., Parameter Estimation for a Polymerization Reactor Model with a composite-step Trust-Region NLP Algorithm, **Industrial and Engineering Chemistry Research**, v. 43, p. 3616-3631, 2004.

BAI, S.; THIBAUT, J. **Dynamic data reconciliation: teory and practice**. Saarbrücker, Alemanha, 2010.

BARBOSA Jr., V.P., WOLF, M.R.M., MACIEL FO, R., Development of Data Reconciliation for Dynamic nonlinear system: Applcation The Polymerization Reactor, **Computers and Chemical Engineering**, v.24, p. 501-506, 2000.

BARBOSA, A. G. **Desenvolvimento de um software para reconciliação de dados de processos químicos e petroquímicos**. Dissertação de Mestrado. Universidade de Campinas. 2003.

BAGAJEWICZ, M. J. On the probability distribution and reconciliation of process plant data. **Computers Chemical Engineering**, v. 20, n. 67, p. 813–19, 1996.

BAGAJEWICZ, M. J., JIANG, Q. Integral approach to plant linear dynamic reconciliation. **AIChE Journal**, 43, 2546-2558, 1997.

BAGAJEWICZ, M. J.; JIANG, Q. Comparison of steady state and integral dynamic data reconciliation. **Computers Chemical Engineering**, v. 24, p. 2367–83, 2000.

BAGAJEWICZ, M. J. **Process plant instrumentation: Design and upgrade**. Technomic Publishing Company, Pennsylvania, 2001.

BENQLILOU C. Data reconciliation as framework for chemical processes optimization and Control, Thesis, Universitat Politècnica de Catalunya, 2004.

BINDER, T., BLANK, L., DAHMEN, W., MARQUARDT, W. On the regularization of dynamic data reconciliation problems. **Journal of Process Control**, n. 12, p. 557-567, 2002.

BOUROUIS, M., PI-BOULEAU, L., FLOQUET, P., Simulation and data validation in multistage flash desalination plants. **Desalination**, v. 115, p. 1-14.

CHEN, J., PENG, Y., MUNOZ, J. C. Correntropy estimator for data reconciliation. **Chemical Engineering Science**, n. 104, p. 1019-1027, 2013.

CROWE, C. M., CAMPOS, Y. A. G., HRYMAK, A. Reconciliation of process flow rates by matrix projection. Part I: Linear case. **AIChE Journal**, n. 29, p. 881-888, 1983.

CROWE, C. M. Reconciliation of process flow rates by matrix projection. Part II: Nonlinear case. **AIChE Journal**, n. 32, p. 616-623, 1986.

CROWE, C. M. Observability and redundancy of process data for steady state reconciliation. **Chemical Engineering Science**, n. 44, p. 2909–17, 1989.

DAROUACH, M., ZASADZINSKI, M. Data reconciliation in generalized linear dynamic systems. **AIChE Journal**, 37, 193-201, 1991.

FABER, R., LI, B., LI, P., WOZNY G. Data reconciliation for real-time optimization of an industrial coke-oven-gas purification process. **Simulation Modelling Practice and Theory**, v.14, p. 1121-1134, 2006.

FARIAS, A. C. Avaliação de estratégias para reconciliação de dados e detecção de erros grosseiros. Dissertação de mestrado. Universidade Federal do Rio Grande do Sul, 2009.

FELDMAN, R. N. **Reconciliação de dados em tempo real para monitoração e detecção de falhas em terminal de transporte e armazenamento de derivados de petróleo**. Dissertação de Mestrado. Universidade Federal do Rio de Janeiro, 2007.

FERREIRA, O.P, GONÇALVES, M. L. N., OLIVEIRA, P. R. Local convergence analysis of the Gauss-Newton method under a majorant condition. **J. Complexity**, nº 27(1), pp. 111-125, 2011.

GARCIA, L. M. P. **Metodologias de análise de dados para um sistema de otimização em tempo real**. Dissertação de mestrado. Escola Politécnica da Universidade de São Paulo, 2013.

GONZALEZ, R. Data reconciliation and optimal management of hydrogen networks in a petrol refinery. **Control Engineering Practice**, v. 20, p. 343-354, 2012.

GONZALEZ, R., HUANG, B., XU, F., ESPEJO, A. Dynamic Bayesian approach to gross error detection and compensation with application toward an oil sands process. **Chemical Engineering Science**, p.44-56, 2012.

GHOSH, K., RAMTEKE, M., SRINIVASAN, R., optimal variable selection for effective statistical process monitoring. **Computers and Chemical Engineering**, n. 60, p. 260-276, Singapore, 2014.

HAMPEL, F. R. The breakdown points of the mean combined with some rejection rules. **Technometrics**, v. 27, pp. 95-107, 1985.

HARIKUMAR, P., NARASHIMAN, S. A method to incorporate Bounds in Data Reconciliation and Gross Error Detection - II. Gross Error Detection Strategies, **Computer and Chemical Engineering**, v. 17, p. 1121-1128, 1993.

HLAVÁČEK, V. Analysis of a complex plant-steady state and transient behavior. **Computer and Chemical Engineering**, v. 1, pp. 75-80, 1977.

HODOUIN, D., EVERELL, M. D. A hierarchical procedure for adjustment and material balancing of mineral processes data. Int. **J. Miner. Process.**, n. 7, p. 91-116, 1980.

HIRATA, K., SAKAMOTO, H., O'YOUNG, L., CHEUNG, K. Y., HUI, C. W. Multi-site integration - an industrial case study. **Computers and Chemical Engineering**, v. 28, p. 139-148, 2003.

HUANG, M., PU, D., A trust-region SQP method without a penalty or a filter for nonlinear programming. **Journal of Computational and Applied Mathematics**, v. 281, pp. 107-119, 2015.

HUBER, P. J., **Robust statistics**. New York, John Wiley, 1981.

JESUS, N. J. C. **Otimização em tempo real em um processo industrial de produção de etileno**. Tese de doutorado. Universidade Federal do Rio de Janeiro, 2011.

JIN, S., HUNG, L. X., LIU, M. A New Target function for robust data reconciliation. **Industrial and Engineering Chemistry Research**, v. 51, pp. 10220-10224, 2012.

JIANG, X., LIU, P., ZHENG, L., Data reconciliation for steam turbine on-line performance monitoring. **Applied Thermal Engineering**, accepted manuscript, 2014.

JIANG, X., LIU, P., ZHENG, L., Data reconciliation and gross error detection for operational data in power plants. **Energy**, p. 1-10, 2014.

JOHNSTON, L. P. M., KRAMER, M. A. Probability density estimation using elliptical basis function, **AICHE**, v. 40, pp. 1639, 1994.

JORDACHE, C., MAH, R. S. H., TAMHANE, A. C. Performance studies of the measurement test for detecting of gross error in process data. **AICHE J.**, n. 31, p. 1187-1201, 1985.

KELLER, J. Y. Analitical estimator of measurement error variances in data reconciliation. **Computers Chemical Engineering**, v. 16, n. 3, p. 185–8, 1992.

KNOPE, F. C., **Modelling, Analysis and Optimization of Process and Energy Systems**, 1^a Ed., John Wiley and Sons Inc., 2012.

KONG, M., CHEN, B., LI, B., An Integral to Dynamic Data Rectification, **Computers and Chemical Engineering**, v.24, p. 749-753, 2000.

KUEHN, D. R., DAVIDSON, H. Computer control. II. Mathematics of control. **Chemical Engineering Progress**, n. 57, p. 44, 1961.

LID, T., SKOGESTAD, S. Data reconciliation and optimal of a catalytic Naphtha Reformer. **Journal of Process Control**, v. 18, pp. 320-331, 1998.

LIEBMAN, M. J., EDGAR, T. F., LASDSON, L. S. Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques. **Computers Chemical Engineering**, n.16, p. 963-968, 1992.

LIMA, S. C., **Convergência local do método de Gauss-Newton para sistemas de equações não lineares sob condição majorante**. Dissertação de mestrado. Universidade Federal do Piauí, 2013.

MADRON, F. A new approach to the identification of gross errors in chemical engineering measurements. **Chemical Engineering Science**, v. 40, n. 10, p. 1855–60, 1985.

MAGALHÃES, A. **Reconciliação de dados para estimar a vazão de água de alimentação de uma usina nuclear tipo PWR utilizando um algoritmo de otimização com inspiração quântica**. Dissertação de mestrado, 2013.

MAH, R. S. Design and analysis of process performance monitoring systems. **Engineering Foundation Conf. on Chemical Process Control II**, p. 525–40, 1981.

MAH, R. S. H., G. STANLEY, D. DOWNING, Reconciliation and rectification of process flow and inventory data. **Industry Chemical Process Des. Dev.**, n. 15, p. 175-183, 1976.

MARADIAGA, D. M., BRUNO, J. C. CORONAS, A., Steady-state data reconciliation for absorption refrigeration systems. **Applied Thermal Engineering**, n. 51, p. 1170-1180, Spain, 2013.

MARONNA, R.; ARCAS, J. Data reconciliation and gross error diagnosis based on regression. **Computers and Chemical Engineering**, p.65-71. 2009.

MARQUES, J. A. **Reconciliação de dados na identificação e caracterização de balanços hídricos em plantas industriais**. Dissertação de mestrado. Universidade Federal do Rio de Janeiro, 2006.

MARTINI, A., SORCE, A., TRAVERSO, A., MASSARDO, A. Data reconciliation for power systems monitoring: Application to a microturbine-based test rig. **Applied Energy**. Italy, 2013.

MAVROMATIS, S. P., KOKOSSIS, A. C. A logic model for the analysis and optimization of steam turbine network. **Computers in Industry**, v. 36, p. 165-179, 1998.

MCBRAYER, K.F., EDGAR, T.F., Bias Detection And Estimation In Dynamic Data Reconciliation, **Journal of Process Control** v. 5, p. 285-289, 1998.

MCBRAYER, K.F., SODERSTROM, T.A., EDGAR, T.F., YOUNG, R.E., The Application of Nonlinear Dynamic Data Reconciliation to Plant Data, **Computers and Chemical Engineering**, v. 22, p. 1907-1911, 1998.

MEI, C., SU, H., CHU, J. Detection of gross errors using mixed integer optimization approach in process industry, **Journal of Zhejiang University SCIENCE A**. n.8(6), p.904-909, 2007.

MURTHY, A. K. S. A least squares solution to mass balance around a chemical reaction. **Industrial Engineering and Chemical Process Design Development**, v. 12, n. 3, pp. 246-8, 1973.

MURTHY, A. K. S. Material balance around a chemical reaction II. **Industrial Engineering and Chemical Process Design Development**, v. 13, n. 4, pp. 347-9, 1974.

NARASIMHAM, S.; MAH, R. S. H. Generalized likelihood ratio method for gross error identification. **AIChE Journal**, v. 33, n. 8, p. 1514-21, 1987.

NARASIMHAN, S.; JORDACHE, C. **Data Reconciliation and Gross Error Detection and intelligent use of process data**. [S.l.]: Gulf Publishing Company, 2000.

NICHOLSON, B., NEGRETE, R. L., LORENZ, T. B. On-line state estimation of nonlinear dynamic systems with gross errors. **Computers and Chemical Engineering**. USA, 2014.

NOGITA, S. Statistical test and adjustment of process data. **Industrial Engineering and Chemical Process Design Development**, v. 11, pp. 197-200, 1972.

NORONHA, F. B., PINTO, J. C., MONTEIRO, J. L., LOBÃO, M. W., SANTOS, T. J. **Um Pacote Computacional para Estimação de Parâmetros e Projeto de Experimentos**, Relatório Técnico COPPE/UFRJ, Rio de Janeiro, 1993.

OLIVEIRA Jr, A. M. **Estimação de parâmetros em modelos de processo usando dados industriais e técnica de reconciliação de dados**. Tese de doutorado. PEQ/COPPE. Universidade Federal do Rio de Janeiro, 2006.

OLIVEIRA JUNIOR, A. M.; LIMA, E. L.; PINTO, J.C.; A identificação estrutural como proposta de classificação de variáveis para a reconciliação de dados industriais. **Scientia**

Plena, vol. 7, num. 11, 2011.

OZYURT, D.B., PIKE, R.W., Theory and practice of simultaneous data reconciliation and gross error detection for chemical process. **Computers and Chemical Engineering**, n.28, p.381-402, 2004.

PLÁCIDO, J. **Desenvolvimento e aplicação industrial da técnica de reconciliação dedados**. Dissertação (Tese de Mestrado)— POLI/USP, 1995.

POLLO, L. D. **Reaproveitamento de águas e efluentes inorgânicos de uma indústria petroquímica**. Dissertação de mestrado. Universidade Federal do Rio Grande do Sul, 2004.

POULIN, E., HODOUIN, D., LACHANCE, L. Impact of plant dynamics on the performance of steady-state data reconciliation. **Computers and Chemical Engineering**, n. 34, p. 354-360, Canada,2010.

PRATA, D. M. SCHWAAB, M., LIMA, E.L., PINTO, J. C., Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for na industrial polypropylene reactor. **Chemical Engineering Science**, n. 65 (17), pp. 4943-4954, 2010.

RAMAMURTHI, Y., SISTU, P. B., BEQUETTE, B. W. Control relevant dynamics data reconciliation and parameter estimation. **Computers Chemical Engineering**, n.17, p.41-59, 1993.

RAMIS, M. E. S. **Reconciliação de dados aplicada ao tratamento de gás combustível em unidades marítimas de produção de petróleo**. Dissertação de mestrado. Universidade Federal do Rio de Janeiro, 2009.

REILLY, P., CARPANI, R. Application of statistical theory of adjustment to material balances, In Proceedings of the 13th **Canadian Chemical Engineering Conference**, Montreal, Quebec, 1963.

RIPPS, D. L. Adjustment of experimental data. **Chem. Eng. Progress - Symp. Series**, n. 61, p. 8–13, 1965.

ROLLINS, D. K.; DAVIS, J. F. Unbiased estimation of gross errors in process measurements. **AIChE Journal**, v. 38, p. 563–72, 1992.

ROLLINS,D.K., DAVIS, J. F. Gross error detection when variance-covariance matrices are unknown, **AIChE Journal**, v. 39, pp. 1335-1341, 1993.

ROMAGNOLI, J. A. On data reconciliation: Constraints processing and treatment of bias. **Chemical Engineering. Science**, v. 38, n. 7, p. 1107–17, 1983.

ROMAGNOLI, J. A.; STEPHANOPOULOS, G. Rectification of process measurements data in the presence of gross errors. **Chemical Engineering. Science**, v. 36, n. 11, p. 1849–63, 1981.

ROSENBERG, J., MAH, R. S. H., JORDACHE, C. Evaluation of schemes for detecting and identification of gross errors in process data. **Ind. & Eng. Chem. Proc. Des. Dev.**, n. 26, p.555-564, 1987.

SANCHEZ, M., ROMAGNOLI, J. Use of orthogonal transformations in data classification-Reconciliation. **Computers Chemical Engineering**, n.20, p. 483-493, 1996.

SARABIA, D., PRADA, C., GOMES, E., GUTIERREZ, C., CRISTEA, S., SOLA, J. M. Data reconciliation and optimal management of hydrogen networks in a petrol. Refinery. **Control Engineering Practice**, v. 20, p.343-354, 2012.

SCHLADT, M., HU, B., Soft sensor based on nonlinear steady-state data reconciliation in the process industry. **Chemical Engineering and Processing**, v. 46, pp. 1107-1115.

SCHWAAB, M., BISCAIA Jr, E. C., MONTEIRO, J. L., PINTO, J. C. Nonlinear parameter estimation through particle swarm optimization. **Chemical Engineering Science**. Rio de Janeiro, 2008.

SERTH, R. W., HEENAN, W. A. Gross error detection and data reconciliation in steam-metering systems. **AIChE J.**, 30, 743-747, 1986.

SILVA, E. D. V., **Reconciliação robusta de dados com seleção de modelo simultânea aplicada ao cálculo da potência térmica de um reator nuclear tipo PWR**, p. 94-99, Tese de Doutorado, Engenharia Nuclear, UFRJ/COPPE, Rio de Janeiro, 2012.

SODERSTROM, T.A., EDGAR, T.F., RUSSO, L.P., YOUNG, R.E., Industrial application of a large-scale dynamic data reconciliation strategy, **Industrial and Engineering Chemistry Research.**, v. 39, p. 1683-1693, 2000.

SOLOMONS, G., FRYHLE, C. **Química Orgânica**. 7ª edição. LTC-Livros Técnicos e Científicos Editora S.A., 2001.

STANLEY, G. M., MAH, R. S. H. Estimation of flows and temperatures in process network. **AIChE J.**, n.23, p.642-650, 1977.

STANLEY, G.; MAH, R. S. H. Observability and redundancy in process data estimation. **Chemical Engineering Science**, v. 36, p. 259–83, 1981a.

STANLEY, G.; MAH, R. S. H. Observability and redundancy classification in process networks. **Chemical Engineering Science**, v. 36, p. 1941–54, 1981b.

SUN, S., HUANG, D., GONG, Y. Gross error detection and data reconciliation using historical data. **Procedia Engineering**, n.55-59. 2011.

TAMHANE, A. C., MAH, R. S.H. Data reconciliation and gross error detection in chemical process network. **Technometrics**, n. 27, p.409-422, 1985.

TJOA, I. B., BIEGLER, L. T. Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems. **Computers Chemical Engineering**, n.15, p. 679-690.

TONG, H., CROWE, C. M. Detection of gross errors in Data Reconciliation by principal component analysis, **AIChE Journal**, v. 41, p. 1712-1722, 1995.

UMEDA, T., NISHIO, M., KOMATSU, S. A method for plant data analysis and parameters estimation. **Industrial Engineering and Chemical Process Design Development**, v. 10, pp. 236-243, 1971.

VACHHANI, P., RENGASWAMY, R., VENKATASUBRAMANIAN, V., A Framework for Integrating Diagnostic Knowledge with Nonlinear Optimization for Data Reconciliation and Parameter Estimation in Dynamic Systems, **Chemical Engineering Science**, v. 56, p. 2133-2148, 2001.

VÁCLAVEK, V.; LOUCKA, M. Selection of measurements necessary to achieve multicomponent mass balances in chemical plants. **Chemical Engineering Science**, n. 31, p. 1199–205, 1976.

VASEBI, A., POULIN, E., HODOUIN, D. Dynamic data reconciliation based on node imbalance autocovariance functions. **Computers and Chemical Engineering**, n. 43, p. 81-90, Canada, 2012.

VASEBI, A., POULIN, E., HODOUIN, D. Dynamic data reconciliation in mineral and metallurgical plants. **Annual Reviews in Control**. Canada, 2012.

WONGRAT, W., SRINOPHAKUN, T., SRINOPHAKUN, P., Modified genetic algorithm for nonlinear data reconciliation. **Computers and Chemical Engineering**, n.29, p.1059-1067, 2005.

ZHANG, P., RONG, G., WANG, Y. A new method of redundancy analysis in data reconciliation and its application. **Computers and Chemical Engineering**, v. 25, pp. 941-949, 2001.

ZHANG, Z., SHAO, Z., CHEN, X., WANG, K., QIAN, J. Quasi-weighted least squares estimator for data reconciliation. **Computers and Chemical Engineering**, v. 34, p. 154-162, 2010.

APÊNDICE A - TÉCNICA DE PROGRAMAÇÃO QUADRÁTICA SEQUENCIAL (SQP)

Consiste em resolver um problema de otimização não-linear por resolução sucessiva através de programação quadrática em série. A cada iteração, uma aproximação quadrática do problema é obtida pela aproximação da função quadrática com sua função objetivo, e uma aproximação linear das restrições, ambas aplicando a expansão da série de Taylor em torno das estimativas reais (NARASHIMAN, 1999), (HUANG E PU, 2015).

Nas Equações (2.10) a (2.12) observou-se que a função objetivo já era quadrática e somente as restrições podiam ser linearizadas. O resultado da programação quadrática com a iteração $k+1$ deve ser formulado como:

$$\text{Min}_s (\nabla \phi)^T s + s^T B s \quad (A.1)$$

Sujeito a;

$$f_i(\hat{z}_k) + [\nabla f_i(\hat{z}_k)]^T s = 0 \quad i = 1 \dots m \quad (A.2)$$

$$g_j(\hat{z}_k) + [\nabla g_j(\hat{z}_k)]^T s \leq 0 \quad j = 1 \dots q \quad (A.3)$$

Sendo:

z o vetor das variáveis originais (x, u);

$s = z - \hat{z}_k$ é a busca de direção para a iteração $k+1$;

$\nabla \phi, \nabla f_i, \nabla g_j$ são, respectivamente, os gradientes (derivada que diz respeito a variável z) da função objetivo, restrição de igualdade i , e restrição de desigualdade j , todos avaliados para estimação da corrente $\hat{z}_k$.

B é a Hessiana (matriz da derivada de segunda ordem da função objetivo que diz respeito a variável z) avaliada para a estimação da corrente $\hat{z}_k$.

Na formulação da programação quadrática, todas as variáveis são incluídas na função objetivo. Comparando com a função objetivo da reconciliação de dados da Equação (2.10), podem-se deduzir as seguintes:

$$\nabla \phi = \begin{bmatrix} -2\Sigma^{-1}(y - \hat{x}_k) \\ 0 \end{bmatrix} \quad (A.4)$$

$$B = 2 \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} \quad (A.5)$$

Observa-se que a matriz Hessiana é constante e também singular quando há variáveis não medidas no processo.

A solução da programação quadrática dá a direção para a obtenção das estimativas. Uma busca unidimensional é realizada na direção s_k em cada iteração k , de modo que o novo valor para z na iteração seguinte é exposta na Equação (A.6):

$$\hat{z}_{k+1} = \hat{z}_k + \alpha_k s_k \quad (A.6)$$

Onde α_k é um parâmetro de comprimento entre 0 e 1. Esse parâmetro de comprimento é obtido pela minimização da função de penalidade (similar ao Lagrangiano). O procedimento é repetido usando as novas estimativas até a convergência dos dados.

Se a função objetivo é não linear e seus termos são de ordem superior à quadrática, então a linha de minimização deverá fornecer estimativa mais precisa e menos inviável. No caso da reconciliação de dados, no entanto, desde que a função objetivo seja quadrática, o comprimento das unidades dão as estimativas ótimas para satisfazerem a linearização aproximada das restrições.

Neste caso, a minimização deverá melhorar a viabilidade com respeito às restrições não lineares. Mesmo que isso não leve a uma redução do número de iterações necessárias para a convergência.

APÊNDICE B - MÉTODO DE CONVERGÊNCIA DE GAUSS-NEWTON

Segundo Argyrus e Magreñan (2014) e Allende e Gonçalves (2011) o método de Gauss-Newton e suas variações são os mais eficientes métodos conhecidos para resolver problemas de mínimos quadrados não lineares e de otimização de composição convexa.

Lima (2013) descreve que o algoritmo obtido pelo método de Gauss-Newton é um método para se obter soluções de problemas de mínimos quadrados não lineares, que é uma modificação do método de Newton para minimizar (ou maximizar) funções.

Com a seguinte função de mínimos quadrados não linear apresentada na Equação (B.1):

$$\min \|F(x)\|^2 \quad (B.1)$$

Allende e Gonçalves (2011) descrevem o método como se $F'(x)$ é injetiva e tem imagem fechada para todo $x \in D$, e encontra os pontos estacionários para a solução dos mínimos quadrados em sistemas não lineares sendo,

$$F'(x)^* F(x) = 0 \quad (B.2)$$

Ainda explicam que formalmente o método é descrito como dado $x_0 \in D$ defina

$$x_{k+1} = x_k + S_k, \quad F'(x_k)^* F'(x_k) S_k = -F'(x_k)^* F(x_k), \quad k = 0, 1, \dots \quad (B.3)$$

Ou,

$$x_{k+1} = x_k - [F'(x_k)^* F'(x_k)]^{-1} F'(x_k)^* F(x_k), \quad k = 0, 1, \dots \quad (B.4)$$

Sendo $F'(x)^* F'(x)$ invertível em $B(x_*, r)$ é dada a seguinte aplicação do método:

Seja $G_F: B(x_*, r) \rightarrow Y$, definida por:

$$G_F(x) = x - F' x^\dagger F(x) \quad (B.5)$$

Sendo $\mathbf{G}_F(\mathbf{x})$ a iteração de Gauss-Newton para qualquer $\mathbf{x} \in \mathbf{B}(\mathbf{x}_*, \mathbf{r})$, pode não pertencer a $\mathbf{B}(\mathbf{x}_*, \mathbf{r})$ ou até mesmo não pertencer ao domínio de $\mathbf{F}$. Nesta situação a boa-definição a ser garantida seria de apenas uma iteração (LIMA, 2013). Entretanto é necessário que a iteração de Gauss-Newton seja repetida indefinidamente até a convergência do método, para tanto diversas variações do método são encontradas na literatura como explicado por Argyros e Magreñán (2014), Allende e Gonçalves (2011), Ferreira *et al.* (2011) e Lima (2011).

APÊNDICE C - AVALIAÇÃO DA PREDIÇÃO DO MÉTODO SQP

ESTUDO DE CASO 4: IDENTIFICAÇÃO E DETECÇÃO DE ERROS GROSSEIROS: SISTEMA LINEAR

A Figura 3.8 trata de um sistema composto por 6 (seis) correntes todas medidas e contendo um trocador de calor com válvula *bypass*, misturador e separador. Foi estudado inicialmente por Narashiman e Jordache (2000) e mais recentemente por Knopf (2012). Neste caso, uma das correntes possui erros grosseiros (a medida da taxa da corrente F_2 contém 4 *bias* positivos), os dados de operação são mostrados na Tabela 3.4.

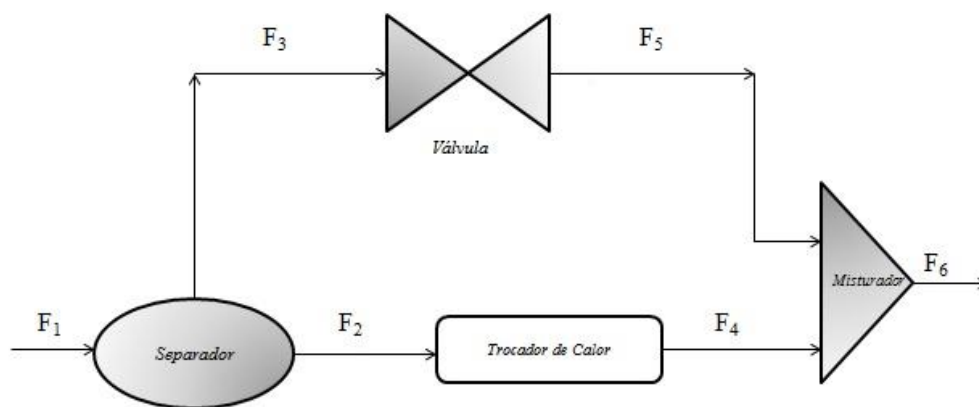


Figura 3.8 Trocador de calor com válvula de desvio (bypass)

Fonte: Knopf, 2012.

Tabela 3.4 Dados do trocador de calor com bypass

Número das Correntes	Valor verdadeiro da Taxa	Valor medido da Taxa	Σ Taxa
1	100	101.9	1
2	64	68.45	1
3	36	34,65	1
4	64	64,20	1
5	36	36,64	1

6	100	98,88	1
---	-----	-------	---

Nas Tabelas C.1 e C.2 em termos do número de iterações o método de SQP apresentou um valor menor, mas não muito destoante do valor obtido por Gauss Newton. Com relação aos valores reconciliados, não há diferenças entre os resultados o que indica que o erro na corrente da temperatura não afeta as propriedades de convergência do *solver*.

Tabela C.1 Análise do método de Gauss Newton para o estudo de caso 4

Método de Gauss-Newton						
Estudo de caso 4						
Correntes	1	2	3	4	5	6
Medido	101,91	68,45	34,65	64,2	36,44	98,88
Estimado	100,89	65,83	35,05	65,83	35,05	100,89
Iterações	3					

Tabela C.2 Análise do método SQP para o estudo de caso 4

Método SQP						
Estudo de caso 4						
Correntes	1	2	3	4	5	6
Medido	101,91	68,45	34,65	64,2	36,44	98,88
Estimado	100,89	65,83	35,05	65,83	35,05	100,89
Iterações	2					

ESTUDO DE CASO 5: TROCADOR DE CALOR SISTEMA NÃO LINEAR

A Figura 3.9 acrescenta ao mesmo trocador de calor da Figura 3.8 medidas referentes às correntes de temperaturas, e o *bias* que existia antes na corrente F_2 foi removido. Entretanto a corrente F_7 que faz parte do balanço energético contém 5 *bias* positivos. Os valores verdadeiros e medidos são apresentados na Tabela 3.5.

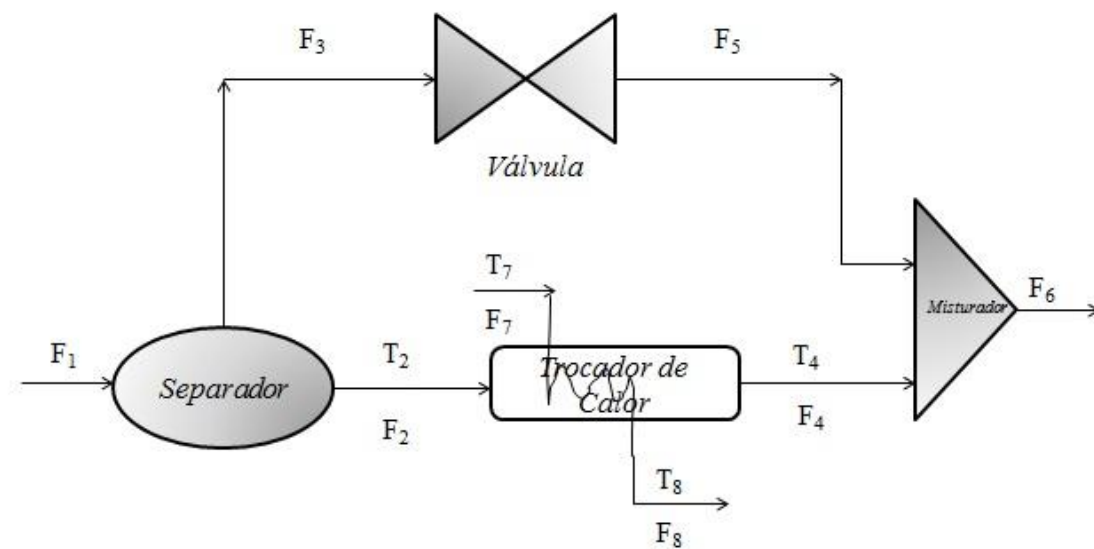


Figura 3.9 Trocador de calor com válvula *bypass* balanço de energia

Tabela 3.5 Dados do trocador de calor com válvula *bypass*

Número de correntes	Valor verdadeiro da taxa	Valor verdadeiro da temperatura	Valor medido	Σ Taxa ou Σ Temperatura
1	100		101,91	1
2	64		64,45	1
3	36		34,65	1
4	64		64,20	1
5	36		36,44	1
6	100		98,88	1

7	140,6	140,0	1
2	90	90	1
4	51	51	1
7	20	25	1
8	43	43	1

De modo similar ao caso 4, o método de SQP apresentou um valor menor de iterações (ver Tabela C.3 e C.4).

Para os valores reconciliados, não há diferenças entre os resultados o que indica que o erro na corrente (e até mesmo a mudança na corrente), não afetou as propriedades de convergência do *solver*.

Tabela C.3 Análise do método de Gauss Newton para o estudo de caso 5

Comparação de Métodos utilizando o WLS											
Método de Gauss-Newton											
Estudo de caso 5											
Correntes	1	2	3	4	5	6	7	2	4	7	8
Medido	101,91	64,45	34,65	64,2	36,44	98,88	140	90	51	25	43
Estimado	100,14	64,34	35,80	64,34	35,80	100,14	140,22	89,05	51,69	23,85	44,77
Iterações	8										

Tabela C.4 Análise do método SQP para o estudo de caso 5

Método SQP											
Estudo de caso 5											
Correntes	1	2	3	4	5	6	7	2	4	7	8
Medido	101,91	64,45	34,65	64,2	36,44	98,88	140	90	51	25	43
Estimado	100,14	64,34	35,80	64,34	35,80	100,14	140,22	89,05	51,69	23,85	44,77
Iterações	7										

Com base nos dados obtidos na Tabela C.1 e C.3 referem-se aos testes obtidos aplicando o método de Gauss-Newton, e Tabela C.2 e Tabela C.4 foram os testes obtidos aplicando a técnica de programação quadrática sequencial (SQP). Pode-se entender que o SQP foi o método que teve um desempenho mais favorável na resolução dos problemas. Os

algoritmos implementados demonstraram que em relação ao número de iterações o SQP foi mais eficiente, e ainda com relação ao tratamento das restrições dos estudos de caso.

APÊNDICE D - DADOS RECONCILIADOS COM AS FUNÇÕES ROBUSTAS

ESTUDO DE CASO 1: RECONCILIAÇÃO DE DADOS EM MÚLTIPLA UNIDADE (KNOPE, 2012)

Visando determinar os fluxos reconciliados desse sistema de operação com três unidades foram testadas as funções robustas. Em termos de comparação dos resultados todas as funções foram eficientes para esse sistema com restrição linear, pois, os dados obtidos foram coerentes com os medidos, com algumas funções inclusive resultando com valores idênticos (ver Tabela D.1).

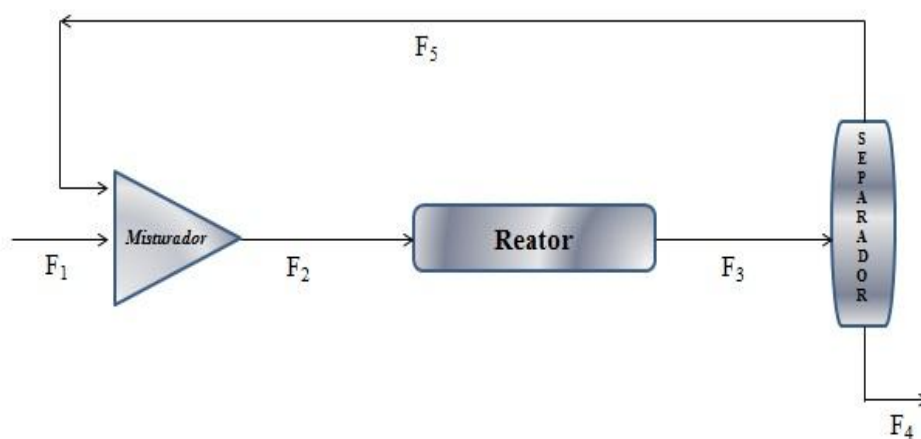


Figura 2.5 Fluxograma operacional de três unidades

Tabela 2.1 Dados medidos da unidade

	Corrente 1	Corrente 2	Corrente 3	Corrente 4	Corrente 5
Fluxo medido	99.3	123.2	125.1	100.5	23.9
Σ medido	2.0	2.0	2.0	2.0	2.0

Tabela D.1 Dados reconciliados utilizando funções robustas: Três unidades

FUNÇÕES ROBUSTAS								
Corr	M	WLS	C	NC	NT	AL	Lo	F
1	99.3	99.9875	99.9912	99.9875	99.9875	99.9875	99.9882	99.9961
2	123.2	124.0625	124.0611	124.0625	124.0625	124.0625	124.0535	124.0334
3	125.1	124.0625	124.0611	124.0625	124.0625	124.0625	124.0535	124.0334
4	100.5	99.9875	99.9912	99.9875	99.9875	99.9875	99.9882	99.9961
5	23.9	24.0750	24.0699	24.0750	24.0750	24.0750	24.0653	24.0372

Corr: Correntes; M:Dados medidos.

Com relação à avaliação de desempenho dessas funções tem-se o número de iteração (Ver Tabela D.2):

Tabela D.2 Número de iterações

Resultados das Iterações							
Funções	WLS	NC	C	F	NT	Lo	Al
	2	2	4	6	2	6	2

As funções *Mínimos Quadrados Ponderados* (WLS), *Normal Contaminada* (NC), *New Target* (NT) e *Alarm* (Al) obtiveram o menor número de iterações, fato este que comprova a literatura, pois, segundo alguns trabalhos, essas funções têm comportamentos semelhantes em sistemas sujeitos a restrições lineares Ozyurt e Pike (2004); Jin *et al.* (2012); Alamgir *et al.* (2013); Chen *et al.* (2013).

ESTUDO DE CASO 2: CLORAÇÃO DO METANO RESTRIÇÕES LINEARES (MADRON *et al.*, 2007), (KNOPF, 2012).

Na Figura 3.2 está representado o fluxograma das espécies envolvidas no processo de cloração do metano com entrada e saída.



Figura 3.2 Espécies observadas no reator de cloração

Tabela 3.1 Dados de fluxo molar medido e desvio padrão

	Fluxo de entrada	Σ	Fluxo de saída	σ
	medido		medido	
Fluxo	100	Fixo		
Cl₂	62,1	3,105	0	Fixo
CH₄	26,9	1,345	1,57	0,0785
CH₃Cl	10,8	0,54	6,83	0,3415
CH₂Cl₂	0,78	0,039	24,8	1,24
CHCl₃	0,38	0,019	4,82	0,241
CCl₄	0	Fixo	0,39	0,0195
HCl	0	Fixo	59,3	2,965

Tabela D.3 Dados reconciliados utilizando funções robustas: Cloração do metano

FUNÇÕES ROBUSTAS								
	M	WLS	C	NC	NT	AL	Lo	F
Fluxo	100							
(E-S)								
Cl₂	62,1	59.9309	59.8990	62.5392	59.9025	60.6351	59.8907	59.8227
CH₄	26,9	26.9528	26.9371	26.7510	26.9388	28.8996	26.9330	26.8985
CH₃Cl	10,8	10.7638	10.7645	10.8963	10.7644	10.8000	10.7647	10.7687
CH₂Cl₂	0,78	0.7796	0.7796	0.7811	0.7796	0.7800	0.7796	0.7797
CHCl₃	0,38	0.3798	0.3799	0.0000	0.3799	0.2672	0.3799	0.3799
CCl₄	0	-	-	-	-	-	-	-
HCl	0	-	-	-	-	-	-	-
Cl₂	0	-	-	-	-	-	-	-
CH₄	1,57	1.5698	1.5699	1.5705	1.5699	1.5700	1.5699	1.5700
CH₃Cl	6,83	6.8445	6.8442	6.7915	6.8442	6.8300	6.8441	6.8424
CH₂Cl₂	24,8	25.2264	25.2132	23.6534	25.2145	24.8001	25.2098	25.1855
CHCl₃	4,82	4.8450	4.8436	4.7527	4.8438	4.8200	4.8432	4.8332
CCl₄	0,39	0.3902	0.3902	1.6603	0.3902	0.7266	0.3902	0.390
HCl	59,3	59.9309	59.8990	62.5392	59.9025	60.6351	59.8907	59.8227

Com relação ao estudo de caso 2 referente a cloração do metano, algumas funções não tiveram o mesmo comportamento do estudo de caso 1. A função *Normal Contaminada* apresentou um erro de convergência na espécie CHCl₃ o que resultou em um valor negativo, sendo necessário a inserção do limite inferior.

A função *Alarm* também gerou um erro de convergência na mesma espécie, fez-se necessário então impor ao limite inferior o zero para que pudesse obter uma estimativa coerente. Isso demonstra uma falha de convergência na função.

Com relação ao número de iterações tem-se a Tabela D.4:

Tabela D.4 Resultados das iterações: Cloração do metano

Funções	Resultados das Iterações						
	WLS	NC	C	F	NT	Lo	Al
	20	17	29	26	20	23	30

Em virtude dos problemas de convergência das funções *Normal Contaminada* e *Alarm* o número de iterações destas não pode ser considerado, haja vista ter optado pela técnica de Programação Quadrática Sequencial (SQP) para não necessitar de modificar o algoritmo.

É importante salientar que o SQP de acordo com Huang e Pu (2015) em *trust-region* mesmo em poucos casos pode levar a pontos inviáveis e amortecer o efeito da função objetivo levando a uma lenta convergência. Para eliminar esse efeito os autores estudaram um filtro para programação não linear da função penalidade do SQP. Entretanto como o esboço deste estudo é testar as funções robustas sem a necessidade de modificação do método, entende-se que a função *Mínimos Quadrados Ponderados* e *New Target* obtiveram os melhores resultados com relação ao número de iterações.

ESTUDO DE CASO 3: TROCADOR DE CALOR ADIABÁTICO COM BALANÇO DE ENERGIA NÃO LINEAR

A Figura 3.5 é referente ao trocador de calor adiabático com balanço de energia não linear e assumindo que não há vazamento e a Tabela 3.2 são os dados medidos das variáveis.

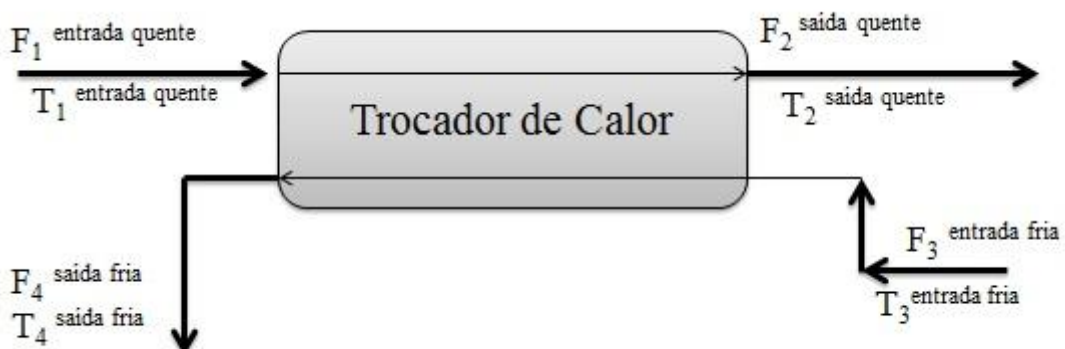


Figura 3.5 Fluxograma do trocador de calor

Tabela 3.2 Dados medidos e desvio padrão das medidas

Correntes	Quente	Quente	Fria	Fria
	(Entrada)	(Saída)	(Entrada)	(Saída)
Taxa (kg/s)	10		20	
Temperatura (°C)	90	51	20	43
Σ fluxo	0.2		0.4	
Σ temperatura	1	1	1	1

Na Tabela D.5 são apresentados os dados reconciliados aplicando as funções robustas.

Tabela D.5 Dados reconciliados utilizando funções robustas: Não linear

FUNÇÕES ROBUSTAS								
Fx	M	WLS	C	NC	NT	AL	Lo	F
1	10	9.9359	9.9395	9.9359	9.9391	9.9371	9.9405	9.9477
2	10	9.9359	9.9395	9.9359	9.9391	9.9371	9.9405	9.9477
3	20	20.1266	20.1196	20.1266	20.1203	20.1242	20.1176	20.1034
4	20	20.1266	20.1196	20.1266	20.1203	20.1242	20.1176	20.1034
5	90	89.5167	89.5341	89.5167	89.5327	89.5224	89.5392	89.5647
6	51	51.3510	51.3321	51.3511	51.3341	51.3447	51.3270	51.2912
7	20	19.5157	19.5331	19.5156	19.5318	19.5212	19.5382	19.5642
8	43	43.8257	43.8679	43.8255	43.8639	43.8395	43.8798	43.9528

Fx: Fluxo; M: Dados medidos.

Analisando-se um sistema não linear que é um dos objetivos desse estudo em virtude das funções apresentadas não terem sido testadas nesse tipo de sistema, demonstra a capacidade delas em obter dados bem próximos dos valores reais sem problemas de convergência e o mais importante, na presença de erros grosseiros.

De acordo com número de iterações:

Tabela D.6 Resultados das iterações: Não linear

Resultados das Iterações							
Funções	WLS	NC	C	F	NT	Lo	Al
	9	7	9	9	11	9	9

A expectativa em torno das iterações da função *Alarm* redescendente por sua característica de melhor convergência em sistemas lineares não foi superada, pois, a mesma apresentou quantidade de iterações iguais a *Mínimos Quadrados Ponderados*, *Cauchy*, *Fair* e *Logística* que são M-estimadores. Além da *New Target* também um M-estimador que obteve um número maior de iterações.