



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA QUÍMICA
ENG07053 - TRABALHO DE DIPLOMAÇÃO EM ENGENHARIA
QUÍMICA



Detecção de Erros Grosseiros e Reconciliação de Dados numa Central Petroquímica

Autor: João Pedro Chagas Zanetti

Orientador: Pedro Rafael Bolognese Fernandes

Porto Alegre, janeiro de 13.

Sumário

Sumário	ii
Agradecimentos	iii
Resumo	iv
Lista de Figuras	v
Lista de Tabelas	vi
Lista de Símbolos	vii
Lista de Abreviaturas e Siglas	viii
1 Introdução	1
2 Revisão Bibliográfica	3
2.1 Medições e Erros de Medição	3
2.2 Reconciliação de Dados	5
2.2.1 Reconciliação de Dados Clássica	6
2.2.2 Reconciliação de Dados Robusta	7
2.3 Detecção de Erros Grosseiros	9
2.3.1 Teste Global (GT)	9
2.3.2 Teste da Medição (MT)	10
3 Estudo de Caso	12
3.1 Descrição da Unidade Industrial	12
3.1.1 Determinação do Estado Estacionário da Planta	14
3.2 Apresentação do Problema	14
3.3 Metodologia	15
4 Resultados e Discussão	17
4.1 Teste Global para Detecção de Erros Grosseiros	17
4.2 Reconciliação de Dados Clássica	18
4.3 Reconciliação de Dados Robusta	19
4.3.1 Função Fair	19
4.3.2 Função Normal Contaminada	23
4.4 Teste de Medição para Detecção de Erros Grosseiros	27
5 Conclusões e Trabalhos Futuros	29
6 Referências Bibliográficas	31
Apêndices	32

Agradecimentos

À minha mãe, por sempre ter me incentivado, inspirado e ter feito seu melhor para contribuir ao meu crescimento.

Ao Time do Processo, onde realizei meu estágio, pelos ensinamentos e bons momentos, fundamentais à minha formação. Em especial ao Coordenador Cristiano Cardoso, pela apresentação do tema deste trabalho, que, depois de certa dor de cabeça no começo, acabou por fascinar-me. Também à amiga e colega Camila (ZS), pela enorme ajuda na busca dos *tags* de medição utilizados neste trabalho e compartilhamento dos anseios relativos ao fim da faculdade no fundão do *bus* na volta do Polo.

Ao meu orientador, Pedro, por ter aceitado esta “aventura” no mundo das medições. Além da ajuda nas questões técnicas, certamente sua paciência, bom humor e motivação foram fatores importantíssimos que contribuíram na execução deste trabalho.

A todos os amigos que tenho a felicidade de ter, presentes na alegria e na tristeza, que fizeram estes seis anos muito mais divertidos. Não seria justo, entretanto, não destacar aqueles que estiveram mais presentes neste último e conturbado semestre, em que tanto precisei de seu suporte: Arthur, Felipe (um amigo e, mais importante, teto em SP), Giulia*, Jaque, Mari “Bixo”*, Mari “Med”, e Vane*.

À minha família, em especial às tias Tita e Marília, que me “adotaram” após a mudança de minha mãe e nunca deixaram faltar duas importantes coisas que eu sempre tinha tido em casa: boa comida e bom vinho.

Aos Kura-Araxes, que há mais de seis mil anos já haviam domesticado a *Vitis vinifera* e tinham uma desenvolvida técnica de vinificação, contribuindo para minhas comemorações no fim de uma fase e começo de uma nova.

* – Colegas do grupo que fez o melhor (ou, pelo menos, mais divertido) projeto de vinícola jamais apresentado nas disciplinas de Projetos I e II. *Famosi perche siamo favolosi!*

Resumo

Este trabalho está focado na aplicação de diferentes técnicas de reconciliação de dados e detecção de erros grosseiros a um conjunto de dados reais de processo de uma unidade de produção de olefinas leves, localizada no Polo Petroquímico de Triunfo, RS. Foram aplicadas técnicas de reconciliação estacionárias às correntes de balanço global da unidade. Os resultados foram comparados aos fornecidos pelo *software* comercial de reconciliação de dados instalado na empresa. Um dos estimadores atingiu resultados satisfatórios para certos valores dos seus parâmetros, porém, a análise é prejudicada devido à simplicidade do problema proposto.

Lista de Figuras

Figura 2.1 – Classificação de variáveis (Retirado de Farias, 2009, p. 30).....	6
Figura 3.1 – Diferentes Blocos da unidade, com suas correntes envolvidas, e o volume de controle do problema (tracejado em laranja).....	12
Figura 4.1 – Ajustes percentuais aplicados às correntes de entrada (a) e saída (b).	18
Figura 4.2 – Variação dos ajustes aplicados às correntes de entrada em função do parâmetro c	20
Figura 4.3 - Variação dos ajustes aplicados às correntes de saída em função do parâmetro C	20
Figura 4.4 – Ajustes aplicados à corrente E1 em função do parâmetro c	21
Figura 4.5 – Massas totalizadas para a corrente S7, registro do Aspen AORA, medição original e valores reconciliados para diferentes c 's.	22
Figura 4.6 – Efeito do parâmetro p sobre os ajustes aplicados às correntes de entrada... ..	24
Figura 4.7 – Efeito do parâmetro b sobre os ajustes aplicados às correntes de entrada... ..	24
Figura 4.8 – Efeito do parâmetro p sobre os ajustes aplicados às correntes de saída.	24
Figura 4.9 – Efeito do parâmetro b sobre os ajustes aplicados às correntes de saída.	25
Figura 4.10 – Efeito dos parâmetros da função Normal contaminada sobre os ajustes à corrente E1.	26

Lista de Tabelas

Tabela 4.1 – Resumo das informações referentes aos GT's executados.....	17
Tabela 4.2 – Diferentes resultados totalizados para um período de 48h.....	18
Tabela 4.3 – Comparação das massas totalizadas das correntes suspeitas.....	22
Tabela 4.4 – Valores arbitrados para os parâmetros p e b utilizados.....	23
Tabela 4.5 – Diferentes ajustes aplicados à corrente S7.....	26
Tabela 4.6 – Maiores valores de Z_{ai} para o MT com base na RD Clássica.....	27
Tabela 4.7 – Maiores valores de Z_{ai} para o MT com base na RD <i>Fair</i> ($c=0,001$).	27

Lista de Símbolos

e	Erro total da medição
ε	Erro aleatório
δ	Erro grosseiro
m	Número de variáveis medidas
n	Número de restrições de igualdade do modelo
A	Matriz de incidência
W	Matriz diagonal de variâncias
H_0	Hipótese nula
H_1	Hipótese alternativa

Lista de Abreviaturas e Siglas

DEG	Detecção de Erros Grosseiros
EG	Erros Grosseiros
GT	Teste Global para Detecção de Erros Grosseiros
MT	Teste da Medição para Detecção de Erros Grosseiros
RD	Reconciliação de Dados
VC	Volume de Controle

1 Introdução

A monitoração das variáveis de processo numa indústria é de grande importância; operação, controle, otimizações e avaliações econômicas dependem de uma correta medição e interpretação de dados como temperaturas, vazões e níveis. A correta atuação das malhas de controle depende da precisão destes dados que, caso seja baixa, pode levar o processo a operar de forma pouco econômica, insegura ou até mesmo provocar um acidente.

Nos mercados de alta concorrência, como o de *commodities*, a liderança em custos é a principal estratégia adotada pelas empresas para aumentar sua competitividade. Inserida neste cenário está a indústria petroquímica, em que a minimização de perdas e de consumos e o aumento de rendimentos são áreas fundamentais, visando a maximizar os lucros. Para tanto, é imprescindível que as informações utilizadas na otimização dos processos e nas transferências de custódia (transações em que as posses de inventários são transferidas) apresentem elevada precisão, tanto o quanto possível.

Apesar do grande desenvolvimento na área de instrumentação, erros são inerentes às medições, tornando o condicionamento de dados de processo uma importante ferramenta de processo. É este contexto que faz com que a detecção de erros grosseiros (DEG) e a reconciliação de dados (RD) ganhem destaque. A primeira técnica trata de identificar medições que possivelmente apresentem um erro sistemático, já a segunda serve para ajustar os resultados das medições em face dos erros aleatórios, de forma a obedecer às restrições de modelos de processo.

Neste trabalho são aplicadas e avaliadas diferentes técnicas de DEG e RD a medições reais de fechamento de balanço material de uma unidade de petroquímicos básicos. Os resultados obtidos são comparados aos dados do *software* comercial de RD utilizado pela empresa. A comparação demonstra que, para certos valores de parâmetros das técnicas mais robustas, a RD foi satisfatória, corrigindo adequadamente mesmo medições com erros grosseiros (EG).

Esta monografia apresenta-se dividida em seis capítulos. O Capítulo 1 trata da introdução ao tema a ser abordado e a motivação para sua escolha. O Capítulo 2 apresenta uma revisão bibliográfica, que visa a fornecer os principais fundamentos

teóricos dos assuntos abordados neste estudo. É apresentado também um breve desenvolvimento das técnicas implementadas.

O Capítulo 3 é dedicado à apresentação da planta industrial onde foram coletados os dados a serem reconciliados e testados, contando com um descritivo sobre a determinação do estado estacionário na unidade. Aqui, também, são apresentados o problema proposto e a metodologia de implementação das técnicas estudadas.

No Capítulo 4 são mostrados e avaliados os resultados obtidos. Ele está dividido em 4 seções, apresentando os resultados para cada técnica. Finalmente, no Capítulo 5, são apresentadas as conclusões e sugestões para trabalhos futuros. As referências bibliográficas estão listadas alfabeticamente no Capítulo 6.

Os resultados completos das diferentes técnicas de reconciliação de dados são apresentados em tabelas como apêndices ao presente trabalho.

2 Revisão Bibliográfica

Este capítulo está dividido em três seções, sendo que a primeira aborda os aspectos gerais de medições e erros, a segunda trata da Reconciliação de Dados, e, a terceira, de Detecção de Erros Grosseiros. Nas duas últimas seções são apresentadas as técnicas utilizadas neste trabalho.

2.1 Medições e Erros de Medição

As medições são de extrema importância na indústria. Variáveis como vazões, níveis, composições, etc. são continuamente medidas e armazenadas. Entre os objetivos destas medições estão o controle e a avaliação econômica do processo e a otimização *online* (Narasimhan e Jordache, 2000).

Atualmente, a utilização de computadores e *softwares* possibilita a coleta automatizada de uma enorme quantidade de dados, obtidos com períodos de minutos ou mesmo segundos. Esta automatização possibilita um registro detalhado e preciso das diferentes variáveis de processo, além de eliminar a possibilidade de erros advindos da leitura manual das variáveis (Narasimhan e Jordache, 2000).

Imperfeições, entretanto, são inerentes às medições. Estas imperfeições dão origem a erros, o que faz com que o resultado de uma medição seja somente uma aproximação, ou estimativa, do valor real do mensurando (conforme definido pelo *Bureau International des Poids et Mesures – BIPM*, 2008). Nas indústrias, os dados oriundos das medições podem ser corrompidos em pelo menos uma das seguintes etapas: medição propriamente dita, processamento ou transmissão do sinal medido (Narasimhan e Jordache, 2000).

Tradicionalmente, o erro total em uma medição – isto é, a diferença entre o valor medido (y) e o valor real (x) – é visto como a soma de duas distintas contribuições: uma componente aleatória (ε) e uma componente sistemática (δ) desconhecida:

$$e = y - x = \varepsilon + \delta \quad (2.1)$$

O erro aleatório surge de variações temporais e espaciais imprevisíveis nas quantidades que influenciam certa medida (BIPM, 2008). O termo aleatório sugere que nem a magnitude, nem o sinal do erro podem ser preditos. Isto significa que mesmo que

uma medição seja repetida sob condições idênticas e utilizando o mesmo medidor, um valor diferente pode ser obtido como resultado. Este tipo de erro pode ser causado por diversos fatores, tais como: oscilação no fornecimento de eletricidade, mudança nas condições ambientes, ruído na conversão e transmissão do sinal. Não existe forma de “compensar” este erro, porém seu efeito pode ser minimizado ao aumentar-se o número de observações. Além disso, o erro aleatório geralmente corresponde às componentes de alta frequência do sinal medido e tem magnitude pequena, com exceção de picos ocasionais (Narasimhan e Jordache, 2000). Usualmente considera-se que este tipo de erro segue uma distribuição normal com média zero, em virtude do exposto pelo teorema do limite central.

Por outro lado, os erros sistemáticos, doravante denominados erros grosseiros (EG), são causados por eventos não aleatórios, tais como o funcionamento incorreto do medidor (devido à instalação ou escolha incorreta de instrumento), calibração errada (erro de zero), erosão, corrosão ou deposição (*fouling*) sobre os sensores. A natureza deste tipo de erro faz com que em qualquer instante de tempo, os mesmos tenham certa magnitude e sinal, que podem, no entanto, ser desconhecidos (Narasimhan e Jordache, 2000). Caso o EG tenha sua origem num efeito conhecido de uma quantidade influente no resultado da medição, este efeito pode ser quantificado e um fator de correção (*bias*) pode ser aplicado para reduzir, senão completamente eliminar, seu efeito (BIPM, 2008).

Diferentes aplicações exigem diferentes níveis de precisão nas medições. Segundo Peramanu e Wah (2011), medições utilizadas para fins de operação de plantas requerem exatidão de $\pm 5\%$, enquanto medições para fins de controle de processo devem ser aceitas com exatidão de $\pm 2\%$. Por outro lado, devido à sua natureza contratual, medições para operações de transferência de custódia podem exigir imprecisões de $\pm 0,1\%$, ou menores. A maioria dos medidores opera com faixas de incerteza conhecidas e especificadas dentro dos limites de capacidade nominal, tipicamente entre 25% e 95% da capacidade máxima do medidor (Peramanu e Wah, 2011). Este fato demonstra que a correta escolha do instrumento é primordial, pois vazões fora dos limites podem acarretar consigo erros grosseiros.

Fica evidenciado, assim, que erros em dados medidos podem trazer deterioração significativa no desempenho de uma planta. Pequenos erros podem levar a problemas no desempenho de sistemas de controle, enquanto erros maiores podem chegar a anular

ganhos obtidos através de otimizações de processo, levar a operação a um regime não econômico ou, ainda pior, inseguro (Narasimhan e Jordache, 2000). É, portanto, imprescindível reduzir ao máximo o efeito dos erros de medição.

2.2 Reconciliação de Dados

Muito tem sido feito na área de pesquisa e desenvolvimento com o objetivo de minimizar erros de medição. Existem hoje filtros de dados (digitais e analógicos), utilizados para atenuar ruídos de alta frequência, e até mesmo sensores inteligentes capazes de detectar a existência de problemas de *hardware* e de identificar se os dados medidos estão dentro de limites aceitáveis. Além disto, a validação de dados pode ser utilizadas para detectar EG's de grande magnitude, e técnicas mais sofisticadas como testes de controle de qualidade estatísticos podem identificar erros. Estas ferramentas são aplicadas a cada medição separadamente e, desta forma, não garantem consistência dos dados no que diz respeito à inter-relação das quantidades medidas, mas devem ser usadas como primeira abordagem na redução de erros aleatórios e EG's (Narasimhan e Jordache, 2000).

Explorar as relações e restrições de modelos de processo e a grande quantidade de dados medidos e armazenados em *softwares* historiadores pode ser de grande valia para aumentar a exatidão de dados, diminuindo o efeito de erros aleatórios nas medições. Desta forma surge a Reconciliação de Dados (RD), técnica que usa tais relações e restrições e obtém estimativas das variáveis de processo, ajustando as medições de forma a satisfazer as restrições e aumentar a exatidão de dados de processo (Peramanu e Wah, 2011).

Dois conceitos importantes no contexto de RD são os de observabilidade e redundância de variáveis, cujas definições são mostradas abaixo, conforme discutido por Farias (2009). Já a Figura 2.1 mostra um esquema de classificação de variáveis.

- Variável observável: é a variável que pode ser estimada a partir das variáveis medidas usando as equações do modelo de processo.
- Variável não observável: é a variável que não pode ser estimada a partir das variáveis medidas usando as equações do modelo.

- Variável redundante: é a variável que continua sendo observável mesmo após sua medição ser removida.
- Variável não redundante: é a variável que se torna não observável após sua medição ser removida.

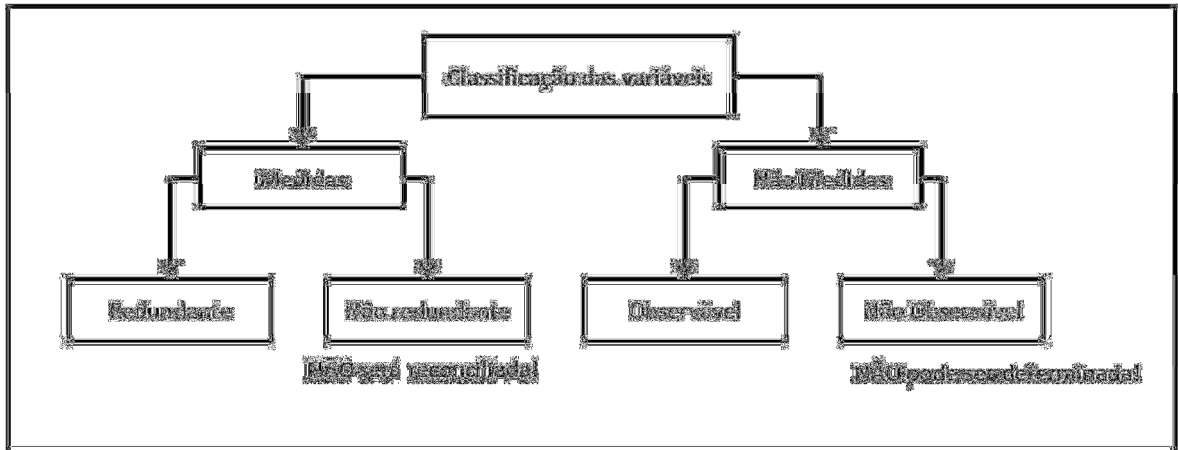


Figura 2.1 – Classificação de variáveis (Retirado de Farias, 2009, p. 30).

Um requisito para a RD ser efetiva é que as variáveis sejam redundantes. Ainda, quanto maior for a redundância de uma variável, mais precisa será a estimativa gerada (Farias, 2009).

2.2.1 Reconciliação de Dados Clássica

Conforme mencionado anteriormente, a reconciliação de dados aumenta a exatidão dos dados de processo ajustando os valores medidos de forma a satisfazer restrições de processo. O método clássico da RD, conhecido como *Mínimos Quadrados Ponderados*, foi proposto por Gauss (1895 *apud* Farias, 2009) ainda no século XVIII, porém o primeiro caso de aplicação industrial data da década de 1960 (Reilly *et al.*, 1963 *apud* Farias, 2009). O valor do ajuste feito é minimizado, dado o fato de os erros aleatórios presentes nas medições serem presumidamente pequenos.

A RD pode ser formulada pelo seguinte problema de otimização.

$$\min_{x_i} \sum_{i=1}^m \frac{(y_i - x_i)^2}{\sigma_i^2} \quad (2.2)$$

Sujeita à restrição

$$g_k(x_i) = 0, \quad k = 1, 2, \dots, n \quad (2.3)$$

A função objetivo (2.2) define a soma total ponderada dos quadrados dos ajustes feitos às medições, onde y_i é o valor da medição e x_i é a estimativa reconciliada, e os pesos $1/\sigma_i$ representam a exatidão de cada medição. A equação (2.3) define o conjunto de n restrições de igualdade do modelo do processo.

As restrições podem incluir desigualdades impostas pelos limites operacionais (Narasimhan e Jordache, 2000), porém o mais comum é que sejam simplesmente balanços materiais ou energéticos do modelo. Ainda, de acordo com estes mesmos autores, é preferível utilizar modelos em detrimento de equações empíricas, pois estas tem um grau de exatidão pouco confiável, podendo provocar reconciliações e detecções de erros grosseiros inexatas.

2.2.2 Reconciliação de Dados Robusta

O estimador clássico funciona como um distribuidor, transferindo a cada uma das medições uma parcela do “desbalanço”, ponderadas de acordo com suas respectivas incertezas. Este mecanismo é perfeitamente aplicável quando se aceita a premissa de que somente erros aleatórios – seguindo uma distribuição normal, com média zero – estão presentes. Se, no entanto, existem EG's e esta distribuição ideal não é atendida, a RD clássica é pouco eficiente, pois as inferências estatísticas em que está baseada não são válidas. Neste caso o EG é distribuído a todas as medições, gerando estimativas contaminadas, o que é conhecido como efeito *smearing* (Farias, 2009).

A RD robusta surge como alternativa para contornar este problema, utilizando outras funções objetivo, potencialmente mais robustas do ponto de vista estatístico, melhores candidatas para gerar estimativas representativas. Os estimadores robustos são concebidos de forma a gerar estimativas sem *bias* caso a distribuição ideal seja atendida, mas que, também, serão menos sensíveis, até certo grau, a desvios da idealidade (Albuquerque e Biegler, 1996). Alguns destes estimadores potencialmente robustos são apresentados a seguir.

2.2.2.1 Função *Fair*

A *Fair* é uma função de estimativa das medições dada por:

$$J_F = c^2 \left[\frac{|\varepsilon_i|}{c\sigma_i} - \ln \left(1 + \frac{|\varepsilon_i|}{c\sigma_i} \right) \right] \quad (2.4)$$

Onde c é um parâmetro de sintonia. À medida que o valor de c diminui, a função *fair* se torna menos sensível a desvios da idealidade, porém a eficiência computacional do estimador também é reduzida (Albuquerque e Biegler, 1996). O efeito prático de uma eficiência mais baixa é o maior tempo computacional exigido para que a otimização atinja a convergência. O equilíbrio do “binômio robustez-eficiência” pode se tornar um dilema, devendo-se analisar caso a caso a disponibilidade computacional.

2.2.2.1 Função Normal Contaminada

Assumindo-se que os erros grosseiros também seguem uma distribuição normal, porém deslocada (com média e desvio padrão desconhecidos), pode-se adotar outra abordagem para o desenvolvimento de uma função objetivo: uma combinação linear de duas funções de distribuições normais. Conforme descrito por Farias (2009), pode-se partir do princípio de que um percentual $(1-p)$ das observações das variáveis é exatamente descrito pela distribuição normal, G , e que o restante seja proveniente de outra distribuição, H . A distribuição resultante se torna

$$F = (1 - p)G + pH \quad (2.5)$$

A função objetivo a ser minimizada na hipótese de distribuição normal contaminada é a seguinte, a partir da função conjunta de probabilidade dos erros ε_i :

$$J_{CG}(\varepsilon) = - \sum_{i=1}^m \ln \left[(1-p) \exp \left(-\frac{\varepsilon_i^2}{\sigma_i^2} \right) + \frac{p}{b} \exp \left(-0,5 \frac{\varepsilon_i^2}{\sigma_i^2 b^2} \right) \right] \quad (2.6)$$

Onde p é a probabilidade de existência de EG's nas medições e b é a razão dos desvios padrão dos EG's e dos erros aleatórios (é uma expressão da magnitude dos EG's em relação aos erros aleatórios). Segundo Farias (2009), a distribuição normal contaminada é limitada para EG's de grandes magnitudes, porém esta restrição pode ser contornada através do ajuste dos parâmetros p e b , sendo que este ajuste é tratado como sintonia do

estimador. Quanto menores forem os valores de cada um dos parâmetros, mais próximo da distribuição normal será o comportamento do estimador.

2.3 Detecção de Erros Grosseiros

Erros grosseiros, conforme abordado na seção (2.1), podem ter origem em diversos fatores. A definição mais simples sobre seu comportamento é: “*Se a medição é repetida com o mesmo instrumento, sob as mesmas condições, a contribuição sistemática do erro grosseiro no valor medido será a mesma*” (Narasimhan e Jordache, 2000).

Uma RD executada em um conjunto de dados que apresente um ou mais EG's resultará em estimativas afetadas pelo espalhamento destes erros em todas as variáveis, fazendo com que os dados reconciliados não representem fielmente o estado do processo. O tratamento matemático de variáveis com erros grosseiros é denominado Detecção de Erros Grosseiros (DEG) (Farias, 2009). As técnicas de DEG clássicas são baseadas em testes estatísticos da hipótese de que as medições seguem uma distribuição com média zero.

A seguir são apresentadas as bases de dois testes para DEG: o Teste Global e o Teste da Medição. Ambos os testes são bastante simples e frequentemente utilizados em conjunto na literatura (Farias, 2009).

2.3.1 Teste Global (GT)

Este teste é bastante simples e se baseia no valor da função objetivo do problema de RD clássica, considerando simultaneamente a média de todas as medições. Neste teste, a hipótese nula (H_0) corresponde à inexistência de erros grosseiros, enquanto a hipótese alternativa (H_1) é a existência de EG's (Farias, 2009). A seguir serão definidas as variáveis utilizadas no GT.

O GT é baseado no vetor de resíduos das restrições, r . Sendo y o vetor das medições, A é a matriz de incidência e W é a matriz (diagonal) das variâncias das medições, r e sua variância, V , são definidos como:

$$r = Ay \quad (2.7)$$

$$V = AW A^T \quad (2.8)$$

Caso existam violações das restrições de processo, estas são refletidas em r . A variável de teste para DEG, que coincide com o cálculo da função objetivo da RD Clássica, é a seguinte:

$$\gamma = r^T V^{-1} r \quad (2.9)$$

Sob H_0 , γ segue uma distribuição qui-quadrado com grau de confiança α e ν graus de liberdade, onde $\nu = \text{posto}(A)$. Se $\gamma \geq \chi_{1-\alpha, \nu}^2$, então H_0 é rejeitada e o erro grosseiro é detectado. Segundo Farias (2009) o GT está presente em praticamente todas as estratégias de medição devido a sua simplicidade, porém é necessário aplicá-lo em conjunto com outros testes, porque ele não dá nenhuma informação sobre a causa ou em qual a medição está localizado o do erro grosseiro.

2.3.2 Teste da Medição (MT)

A base deste teste são os ajustes aplicados em cada medição a partir do problema de reconciliação. Assim, é necessário realizar uma reconciliação prévia ao conjunto de dados de modo a obter os valores dos ajustes ($a_i = y_i - x_i$). A estatística utilizada para o teste é definida como:

$$Z_{a,i} = \frac{|a_i|}{\sqrt{(W_a)_{ii}}} \quad i = 1, 2, \dots, m \quad (2.10)$$

Sob H_0 , Z_a segue uma distribuição normal padrão, $N(0, 1)$. Então o teste é feito para as seguintes hipóteses:

$$H_0: Z_{a,i} = 0 \quad (2.11)$$

$$H_1: Z_{a,i} \neq 0 \quad (2.12)$$

O teste é feito comparando-se $Z_{a,i}$ a um valor crítico, Z_{ac} , definido como:

$$Z_{ac} = Z_{(1-\alpha^*/2)} \quad (2.13)$$

Sendo

$$\alpha^* = 1 - (1 - \alpha)^{1/m} \quad (2.14)$$

No caso de $Z_{a,i} \geq Z_{ac}$, a hipótese nula é rejeitada e a medição i é detectada como possuidora de EG. Neste trabalho, no entanto, este teste não é utilizado em sua forma completa, pois o objetivo é centrado na identificação dos medidores mais suspeitos de

apresentarem erros grosseiros; é, portanto, utilizado somente o valor de Z_{ai} de forma a classificar as medições como as com maior potencial de apresentarem EG, sem realmente compará-lo ao valor crítico.

3 Estudo de Caso

As seções deste capítulo abordam os seguintes temas: descrição da unidade industrial onde os dados utilizados foram coletados, apresentação do problema em estudo e descrição da metodologia de aplicação das técnicas de RD e DEG.

3.1 Descrição da Unidade Industrial

Os dados de processo utilizados neste trabalho foram obtidos numa unidade de produção de olefinas leves, localizada no Polo Petroquímico de Triunfo, Rio Grande do Sul. A unidade faz parte de um complexo de produção de petroquímicos básicos, compreendendo a primeira geração da cadeia petroquímica. A matéria-prima da unidade é a nafta petroquímica leve e os principais produtos são eteno, propeno, gasolina, corte C4, destinado à unidade de aromáticos, e gás combustível, utilizado nos próprios fornos.

A unidade pode ser dividida simplificada nos seguintes grandes blocos: área de tancagem (Bloco 0), craqueamento (Bloco 1), fracionamento primário (Bloco 2), zona de compressão (Bloco 3) e fracionamento final (Bloco 4). Um diagrama apresentando o volume de controle (VC) do problema, contemplando os diferentes blocos e as correntes envolvidas no BM de balanço (aquelas que cruzam a fronteira do VC) neles envolvidas é mostrado na Figura 3.1.

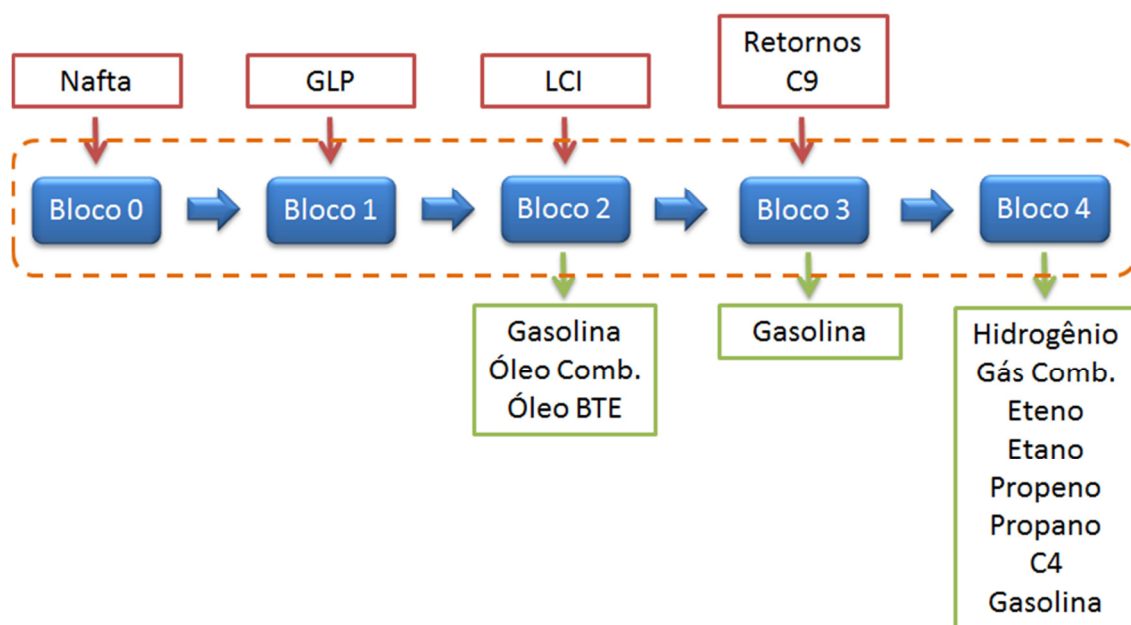


Figura 3.1 – Diferentes Blocos da unidade, com suas correntes envolvidas, e o volume de controle do problema (tracejado em laranja).

No Bloco 0, a nafta é recebida e armazenada. Após cumprir tempo de repouso, é enviada ao Bloco 1, onde sofre craqueamento pirolítico em fornos, na presença de vapor. Uma corrente de fração C3 e C4 (GLP) pode ser introduzida com a nafta nos fornos e cocraqueada. Esta corrente é gerada no próprio processo e seu processamento é realizado de acordo com os rendimentos solicitados pela área comercial da empresa. Ao sair do Bloco 1, o gás craqueado é resfriado via contato direto (*quench*) com óleo oriundo do próprio processo. No Bloco 2, ocorre o fracionamento primário do gás craqueado. Neste bloco são separadas três correntes de produto: gasolina pesada, que será misturada a outras duas correntes de gasolina, óleo BTE e óleo combustível. Um corretor de viscosidade (LCI) pode ser injetado neste bloco, dependendo das condições de operação.

O fluxo de processo segue para o Bloco 3, onde o gás é comprimido. Esta etapa é necessária para tornar possível a condensação das frações mais leves de produto. O gás passa por estágios de compressão intercalados com vasos de *flash*. Nas carcaças dos estágios do compressor é injetada uma corrente de fração C9, para “lavá-las”, eliminando possíveis massas polimerizadas. Neste ponto do processo também são introduzidas correntes de propeno, eteno e *offgas* dos reatores de polimerização das unidades coligadas de segunda geração. As frações líquidas dos vasos de *flash* compõem uma gasolina um pouco mais leve do que a obtida no fracionamento primário, chamada de gasolina intermediária. A fração gasosa segue para uma torre de neutralização dos gases ácidos formados no processo de pirólise.

No Bloco 4, os principais produtos são separados e purificados. Também são realizadas reações de metanação, para purificação de hidrogênio e hidrogenação seletiva em frações que contêm acetileno, metil-acetileno e propadieno, aumentando o rendimento em eteno e propeno. Neste bloco são originadas correntes de hidrogênio, metano (utilizado como gás combustível), eteno, etano (que pode ser recirculada aos fornos ou enviada para processamento em outra unidade), propeno, propano, corte C4 e uma terceira gasolina (gasolina leve).

3.1.1 *Determinação do Estado Estacionário da Planta*

Este trabalho está focado na aplicação de balanços materiais estacionários à unidade em estudo. Para recolher os conjuntos de dados necessários às análises aqui consideradas, foi preciso, primeiramente, determinar períodos em que a planta estivesse em regime estacionário.

Foram levantados dados da vazão de alimentação, temperatura de água de resfriamento e pressão de sucção dos dois compressores de gás de carga (a planta em estudo conta com dois trens de compressão operando em paralelo) para um grande período de tempo. Buscou-se identificar períodos em que estas variáveis estivessem aproximadamente constantes e a carga, além de constante, em patamares acima de 90% da capacidade máxima da planta.

Foram coletados conjuntos de dados das vazões relativas ao fechamento de balanço para diversos períodos. Porém, devido à própria natureza das medições, conforme já comentado, e ao grande número de correntes envolvidas, quase todos os conjuntos foram rejeitados, por apresentar medidores que estavam em manutenção ou com leitura congelada. Conseguiu-se isolar apenas um período de 48 horas de operação em estado estacionário e com todos os medidores operando. Além disto, por causa do grande número de fatores que afetam o regime da planta, é impossível determinar um regime perfeitamente estacionário, o que diminui a robustez das técnicas aqui aplicadas.

3.2 Apresentação do Problema

A indústria em questão faz parte da primeira geração da cadeia petroquímica. Os produtos principais da unidade produtiva, eteno e propeno, são a matéria prima de outras indústrias no complexo, que fazem a polimerização destes compostos em resinas plásticas. Parte das indústrias de segunda geração pertence à mesma empresa da unidade em estudo, configurando um caso de verticalização da produção. Desta forma, é de grande interesse que o rendimento obtido na primeira etapa seja maximizado, pois os principais compradores são da própria empresa e, para obter competitividade no mercado de resinas, é necessário minimizar os custos, desde a produção dos monômeros.

Minimização de perdas e consumos e aumento de rendimentos são áreas vitais na busca da redução de custos. Para tanto, é imprescindível que as informações utilizadas na operação, otimizações e transferências de custódia sejam bastante precisas.

As paradas gerais de manutenção na unidade em estudo ocorrem a cada seis anos, sendo a próxima agendada para o ano de 2014. Elas são uma grande oportunidade para troca de instrumentos. É com esta perspectiva que foi solicitada a execução de um estudo para determinar quais medidores de vazão deveriam ser trocados, de modo a minimizar os erros de balanço. Assim, serão aplicadas técnicas de DEG e RD de modo a identificar os medidores com maior suspeita de possuírem EG's. Como o interesse é minimizar os desvios no balanço global, somente uma equação de balanço foi utilizada, levando em conta os insumos e produtos da unidade como um todo, que é o volume de controle utilizado neste trabalho. A equação de balanço da unidade é apresentada abaixo:

$$\sum_{i=1}^{12} E_i - \sum_{j=1}^{14} S_j = 0 \quad (3.1)$$

Onde E_i e S_j representam, respectivamente, as diferentes correntes de entrada e saída no volume de controle e são todas variáveis medidas. Foi usada esta nomenclatura de maneira a preservar o sigilo sobre informações de capacidade e rendimentos da planta, pois os valores de vazão apresentados serão os reais. Embora sejam consideradas 26 correntes ao todo, o número de insumos e produtos é menor, pois para certos componentes as correntes estão divididas, correspondendo à vazão total destes a soma de mais de uma corrente. Estas informações ficarão mais claras mais adiante, na análise dos resultados. Além disto, cabe esclarecer que, devido à estrutura do problema, o grau de redundância das variáveis no problema é igual a 1, o que pode deteriorar as estimativas geradas pelas técnicas de RD.

3.3 Metodologia

Para a determinação do estado estacionário, foi utilizado o *software* historiador instalado na unidade, Aspen Process Explorer®. As médias horárias das variáveis citadas na seção (3.1.1) foram plotadas em função do tempo para o período de 340 dias num *trend*, no qual se avaliou a estabilidade dos valores. Para coletar os dados de vazão de fechamento de balanço, foi utilizada a ferramenta “Trend Aggregate Data” do suplemento do Aspen Process Explorer® para o Microsoft Excel 2007® e estas foram representadas pelas suas médias horárias. Da mesma forma, foram coletados dados reconciliados pelo *software* Aspen AORA®, que ficam disponíveis no servidor do Aspen Process Explorer®. Estes dados reconciliados estão disponíveis somente na forma de totais para um período de 24 horas. Conforme já comentado, apenas um *dataset* foi

utilizado, contendo dados de vazão brutos para um intervalo de 48 horas. Os valores utilizados foram as médias das vazões para o período de coleta.

Todas as técnicas, tanto de RD quanto de DEG, foram implementadas no *software* Microsoft Excel 2010®. As medidas estatísticas foram feitas utilizando as funções apropriadas do *software*. Para a comparação do parâmetro γ com a distribuição qui-quadrado do Teste Global de DEG, foi utilizada a função “INV.QUIQUA”. Para a resolução dos problemas de otimização (reconciliação), foi utilizado o suplemento “Solver”, com o método de solução “GRG Não-Linear”, alterando os valores das vazões reconciliadas de modo a minimizar as funções objetivo, sendo o problema ainda sujeito às seguintes restrições:

- Resíduo de balanço igual a zero (restrição do balanço material).
- Valores das variáveis reconciliadas maiores ou iguais a zero, para evitar a primeira restrição ser atendida simplesmente tornando negativa alguma vazão de entrada.

4 Resultados e Discussão

Neste capítulo serão apresentados os principais resultados obtidos com as técnicas implementadas com os dados coletados.

4.1 Teste Global para Detecção de Erros Grosseiros

Conforme mencionado anteriormente, os dados coletados foram as médias horárias das vazões de fechamento de balanço. Para o GT, porém, utilizaram-se dois conjuntos de dados, o primeiro referente às primeiras 24 horas e o segundo, às últimas 24 horas. Desta forma, foi possível não só ter uma amostragem razoável, com 24 medições para cada corrente avaliada, mas também executar dois testes. Os principais parâmetros e resultados dos testes estão compilados na Tabela 4.1.

Tabela 4.1 – Resumo das informações referentes aos GT's executados.

	Teste 1	Teste 2
Data dos dados	13/08/2012	14/08/2012
Resíduo	301,2	370,2
α	0,95	0,95
V	4,512	4,058
γ	20101	33765
χ^2	3,84	3,84
Existem EG's?	Sim	Sim

Verifica-se que a condição para a detecção de erros grosseiros, $\gamma \geq \chi_{1-\alpha, v}^2$, é atendida, ou seja, é rejeitada a hipótese nula do teste. Conforme discutido anteriormente, o GT não fornece nenhuma informação sobre a localização do erro, somente evidencia sua presença.

Não foi encontrada na literatura relação entre a ordem de grandeza do parâmetro γ e a quantidade ou magnitude dos EG's. Sabe-se, entretanto, que o GT é um teste para a detecção de um único EG (Farias, 2009), o que pode fazer com que sua aplicação a conjuntos de dados com mais de um EG retorne parâmetros γ com ordem de grandeza superior à da distribuição χ^2 .

Devido à sua natureza, este teste não contribui para a análise de conjuntos com muitas correntes envolvidas e onde já é esperado haver erros grosseiros. Sua aplicação é recomendável a nós com poucas correntes, para checar a existência erro grosseiro, ou

quando, pela experiência operacional, já se tem um forte suspeito, caso no qual o teste confirmaria ou não a hipótese.

4.2 Reconciliação de Dados Clássica

A aplicação da otimização da função objetivo clássica (equação 2.2) retornou estimativas para todas as variáveis. Os ajustes percentuais aplicados às correntes de entrada e saída são apresentados nas Figuras 4.1a e b, respectivamente. Os resultados completos encontram-se no Apêndice A.

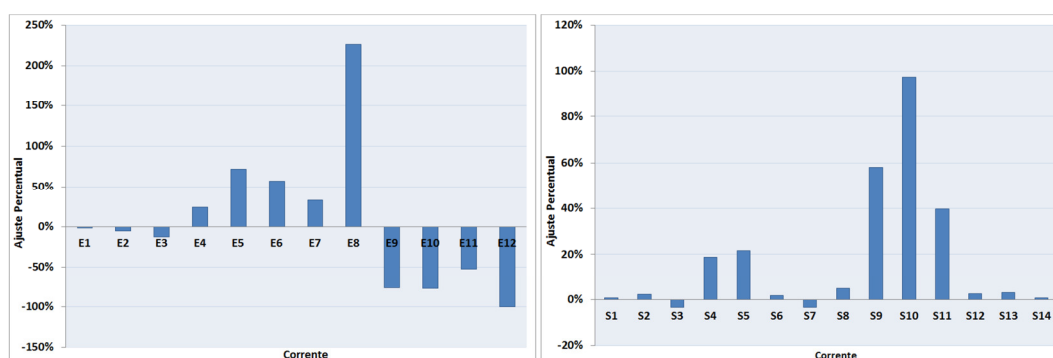


Figura 4.1 – Ajustes percentuais aplicados às correntes de entrada (a) e saída (b).

Como se pode observar, não houve padrão específico nas correções aplicadas, havendo ajustes tanto positivos quanto negativos para ambos os conjuntos (entradas e saídas). A Tabela 4.2 confronta os valores totalizados para 48 horas da medição original, aqueles obtidos pela RD clássica e aqueles registrados no *software* reconciliador comercial instalado na empresa, Aspen AORA®.

Tabela 4.2 – Diferentes resultados totalizados para um período de 48h.

Corrente	Medição Original	AORA	RD Clássica
E1	14031,2	14438,0	13725,1
E2	482,0	0,0	454,0
E4	37,9	38,2	47,0
S4+S5	146,9	120,6	177,8
S7	9,6	43,8	9,3

Verifica-se que a RD clássica acaba fornecendo valores mais distantes daqueles fornecidos pelo Aspen AORA, que é a ferramenta de fechamento de balanço empregada na empresa, presumidamente robusta. Isto se deve ao fato de o problema em questão apresentar somente uma equação de balanço. Como não existem redundâncias para as medições e a função objetivo clássica ser bastante simples, ela atua somente com o objetivo de fazer o somatório de todas as entradas se igualar ao somatório de todas as

saídas, não tendo nenhuma atuação sobre variáveis específicas. Fica claro que para este problema específico, a RD Clássica é de pouca utilidade, fornecendo valores possivelmente menos realistas do que as medições originais.

4.3 Reconciliação de Dados Robusta

Os métodos de RD robusta forneceram estimativas mais realistas, coerentes com informações de operação, dependendo do valor dos parâmetros aplicados. As duas seções a seguir mostram os principais resultados obtidos para os dois métodos. Os Resultados completos para as duas funções encontram-se nos Apêndices B e C.

4.3.1 Função Fair

Foram feitas diversas reconciliações com o mesmo conjunto de dados utilizando a função objetivo *Fair* (equação 2.4). Variou-se o parâmetro c da função objetivo de modo a verificar sua influência nas estimativas. Foram atribuídos os seguintes valores: 10; 4; 3; 2; 1; 0,1; 0,01; 0,001. Compararam-se, também os resultados das reconciliações feitas com os valores fornecidos pelo *software* Aspen AORA.

Para exibição dos resultados, foram selecionadas as quatro correntes de entrada e saída que sofreram os ajustes de maior magnitude. As Figuras 4.2 e 4.3 mostram os ajustes percentuais aplicados às correntes de entrada, e saída, respectivamente, relativos às medições iniciais em função do parâmetro c da função *Fair*. Apesar de serem mostrados os valores absolutos (para facilitar a visualização) todos os ajustes nas entradas foram negativos e, nas saídas, positivos.

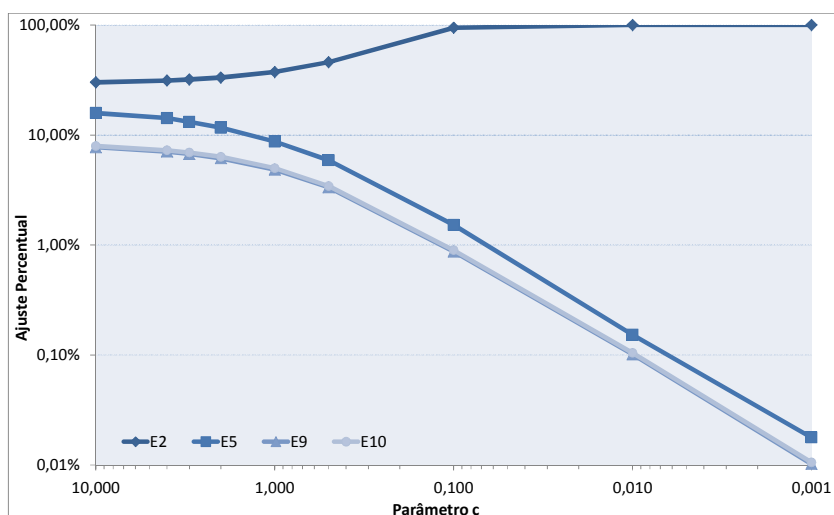


Figura 4.2 – Variação dos ajustes aplicados às correntes de entrada em função do parâmetro c .

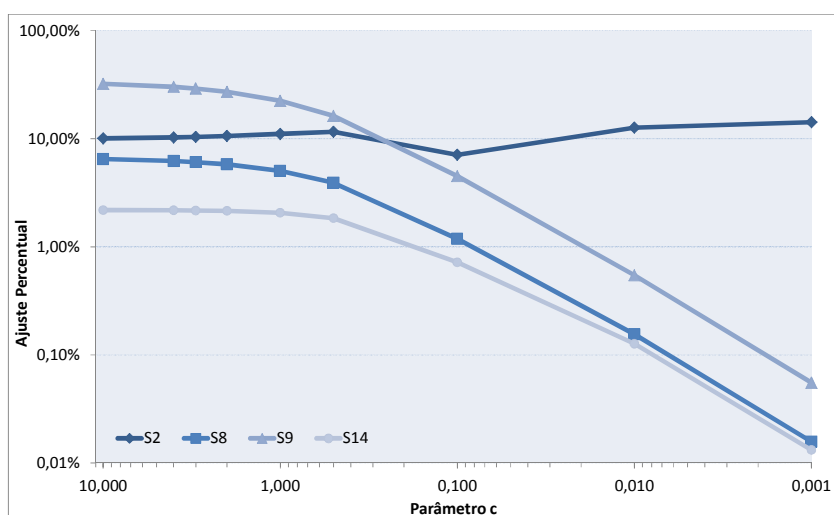


Figura 4.3 - Variação dos ajustes aplicados às correntes de saída em função do parâmetro c .

Verifica-se que a diminuição no parâmetro c tende a tornar os ajustes mais pontuais, isto é, o ajuste a variáveis específicas aumenta (caso da corrente E2) ou se mantém praticamente constante (caso da corrente S2), enquanto os ajustes às demais variáveis diminuem. De fato, segundo Albuquerque e Biegler (1996), quanto menor o valor de c , mais robusta fica a função *Fair*. Isto leva a crer que as variáveis que tiveram ajustes decrescentes com a diminuição de c não apresentam erros grosseiros. Em contrapartida, variáveis que tiveram ajustes consideráveis, mesmo com valores baixos de c , são fortes candidatas a apresentarem erros grosseiros, o que é o caso das correntes E2 e S2.

A Figura 4.4 apresenta os valores totalizados de vazão da corrente E1 em função do parâmetro c . Esta corrente é de especial interesse na análise, porque seu medidor é do tipo Coriolis, ou seja, supostamente o valor medido tem elevada precisão, característica deste tipo de instrumento. As estimativas apresentam um comportamento interessante: variam rapidamente em direção ao valor medido para valores de c imediatamente menores do que um, porém a aproximação se torna assintótica para valores ainda menores. Como já foi dito, as medições desta corrente são presumidamente bastante precisas e, portanto, os seus ajustes deveriam ser pequenos. Estes fatos estão de acordo com a ideia de que valores baixos de c deixam a função *Fair* mais robusta, fazendo com que ela “localize” e atue apenas nos erros grosseiros. Existe, entretanto, um claro *offset* em relação ao valor do *software* Aspen Aora, que supostamente é um valor mais confiável. Esta diferença provavelmente se deve ao fato de o *software* utilizar um balanço material mais detalhado, com redundâncias espaciais, chegando a um valor possivelmente mais realista. Tendo como base o *software* comercial, a corrente E1 apresenta um erro grosseiro, não identificado pela reconciliação empregando a função *Fair* a qual, entretanto, identificou a elevada precisão da corrente em questão, não identificando o *bias* devido ao formato do problema proposto.

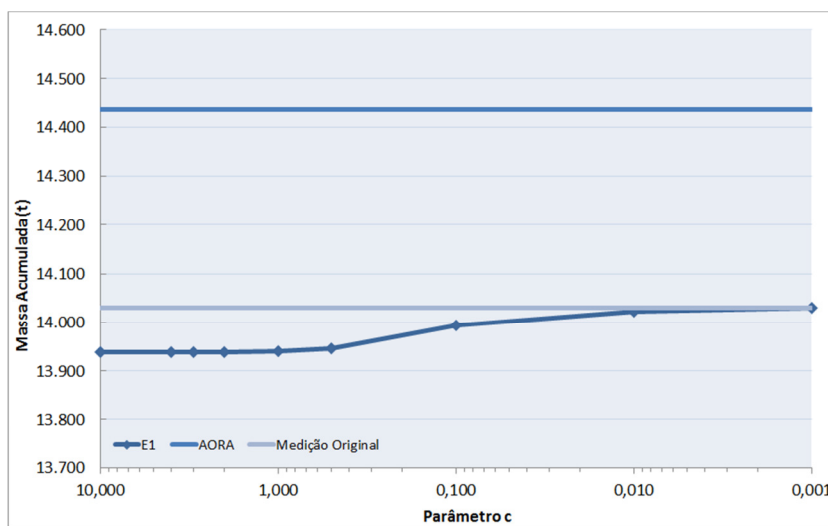


Figura 4.4 – Ajustes aplicados à corrente E1 em função do parâmetro c .

Outra análise relevante a ser feita é o caso da corrente S7. A medição original apresentou uma média de 0,200 t/h, com um desvio padrão de $3,77 \times 10^{-3}$. Totalizando este valor para o período de 48 horas, a massa que supostamente foi medida foi de 9,6 t, em contraste às 43,8 t registradas pelo Aspen AORA, que geraria uma vazão média de 0,9

t/h. A Figura 4.5 apresenta as estimativas totalizadas para 48 horas para os diferentes valores de c , juntamente com a medição original e o valor apresentado pelo Aspen AORA.

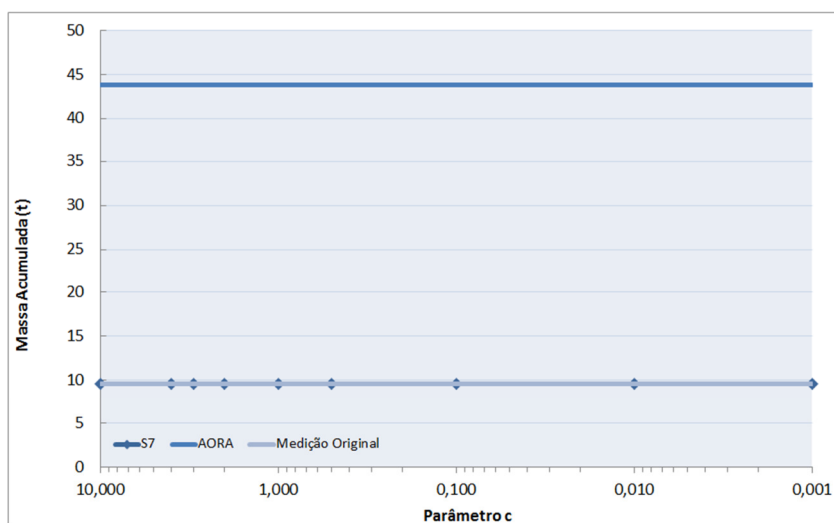


Figura 4.5 – Massas totalizadas para a corrente S7, registro do Aspen AORA, medição original e valores reconciliados para diferentes c 's.

O medidor desta corrente provavelmente estava paralisado em um certo valor. O *software* comercial, conforme já foi dito, utiliza redundâncias e por isto sua estimativa difere bastante daquela encontrada na presente investigação. Este fato demonstra que mesmo uma função robusta, quando aplicada a um balanço global com apenas uma restrição, tem suas limitações, não sendo capaz de aplicar estimativas mais realistas a um medidor “congelado”.

A Tabela 4.3 compara as massas totalizadas para o período de 48 horas para as corrente E2 e a soma S2+S3. As correntes E2 e S2 têm suspeita de apresentarem EG, conforme sugerido pela análise do comportamento dos ajustes aplicados a elas quando da variação do parâmetro c .

Tabela 4.3 – Comparação das massas totalizadas das correntes suspeitas.

Corrente	AORA (t)	Fair (c=0,001) (t)	Fair (c=0,01) (t)	Fair (c=1) (t)	Fair (c=10) (t)
E2	0	0	0	302	336
S2+S3	2330	2190	2173	2176	2172

Infelizmente, não se pode comparar isoladamente o valor da corrente S2, porém fica evidente que a soma S2+S3 tende a aproximar-se do valor do *software* ao aumentar-se a

robustez da função *Fair*. Já para a corrente E2, a função *Fair* atingiu exatamente o valor do *software*.

Examinando-se os dados, fica claro que o parâmetro c tem grande influência nas estimativas geradas pela função *Fair*. Com uma sintonia adequada, a função se torna capaz de identificar certos erros grosseiros, eliminando-os, porém ela é pouco eficaz com medições “congeladas”. Possivelmente, ao adicionarem-se redundâncias (novas equações de balanço envolvendo as mesmas variáveis), ela seja capaz de fornecer estatísticas mais realistas mesmo nestes casos.

A desvantagem de se utilizar valores pequenos para o parâmetro da equação é a eficiência diminuída do estimador. Durante as otimizações, verificou-se que o número de iterações até a convergência da função aumentava conforme c diminuía (tomou-se o cuidado de utilizar sempre a mesma aproximação inicial), porém, devido ao tamanho deste problema, este fato não chegou a requerer um grande tempo de cálculo. Caso a função venha a ser aplicada a um problema com mais vazões e restrições, pode ser necessário ter de abrir mão da robustez de forma a não tornar a resolução muito longa.

4.3.2 Função Normal Contaminada

A exemplo do que foi feito para a função *Fair*, os dados foram reconciliados para diferentes valores de parâmetros p e b da função Normal Contaminada (equação 2.6). Estes valores são apresentados na Tabela 4.4.

Tabela 4.4 – Valores arbitrados para os parâmetros p e b utilizados.

Parâmetro	Valores
p	0,1; 0,3; 0,4; 0,5; 0,6
b	1; 5; 10

Aqui também foram selecionadas as quatro variáveis que tiveram o maior ajuste absoluto (novamente, são mostrados valores absolutos, porém os ajustes aplicados às correntes de entrada foram negativos). A Figura 4.6 apresenta os ajustes aplicados às correntes de entrada com o parâmetro $b = 5$, variando-se p , a Figura 4.7 mostra os ajustes aplicados às mesmas correntes com $p = 0,5$ e b variando. As Figuras 4.8 e 4.9 são análogas às 4.6 e 4.7, respectivamente, porém mostram os efeitos sobre as correntes de saída.

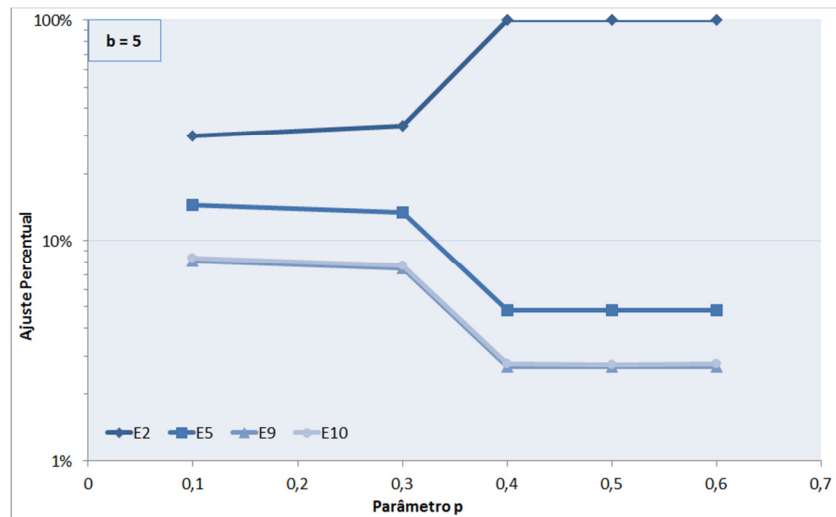


Figura 4.6 – Efeito do parâmetro p sobre os ajustes aplicados às correntes de entrada.

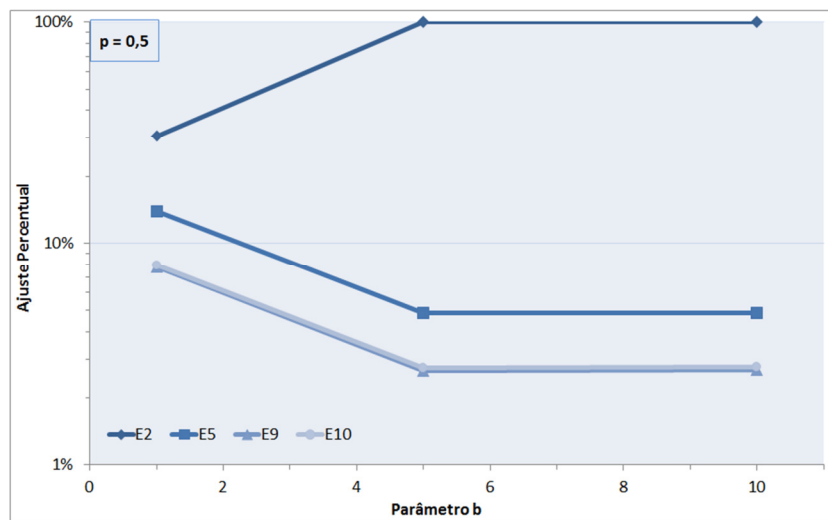


Figura 4.7 – Efeito do parâmetro b sobre os ajustes aplicados às correntes de entrada.

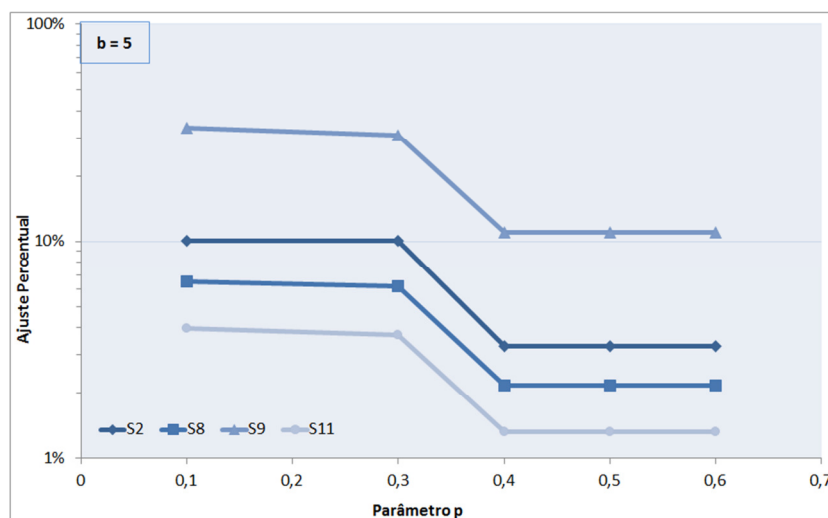


Figura 4.8 – Efeito do parâmetro p sobre os ajustes aplicados às correntes de saída.

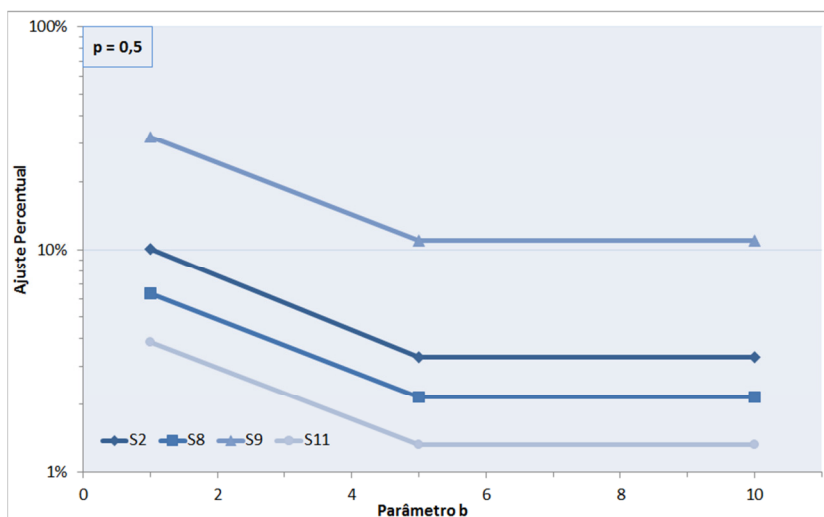


Figura 4.9 – Efeito do parâmetro b sobre os ajustes aplicados às correntes de saída.

Nota-se claramente que o aumento em qualquer um dos parâmetros p e b da função Normal Contaminada tem efeito similar ao da diminuição do parâmetro c da função *Fair*, fazendo com que os ajustes se tornem mais pontuais. Para esta função, entretanto, parece existir uma espécie de “saturação”: existe uma combinação de parâmetros crítica, após a qual os resultados fornecidos pela função não se modificam e os ajustes passam a manter a mesma distribuição sobre as variáveis.

Assim como na RD com a função *Fair*, a corrente E2 atingiu 100% de ajuste, fazendo com que fosse anulada e igualada ao valor registrado no Aspen AORA. A corrente S2, assim como todas as outras, teve o ajuste estabilizado, ao contrário do que ocorreu com a função *Fair*, onde seu ajuste tinha uma leve tendência de crescimento com a diminuição do parâmetro c . Devido ao fato de os ajustes a todas as correntes estabilizarem-se após certos valores de parâmetros p e b , não é possível identificar possíveis EG’s, exceto pela corrente E2.

A exemplo do que foi feito para a função *Fair*, será analisado o efeito dos parâmetros da função normal contaminada sobre os ajustes aplicados à corrente E1. Estas informações encontram-se na Figura 4.10.

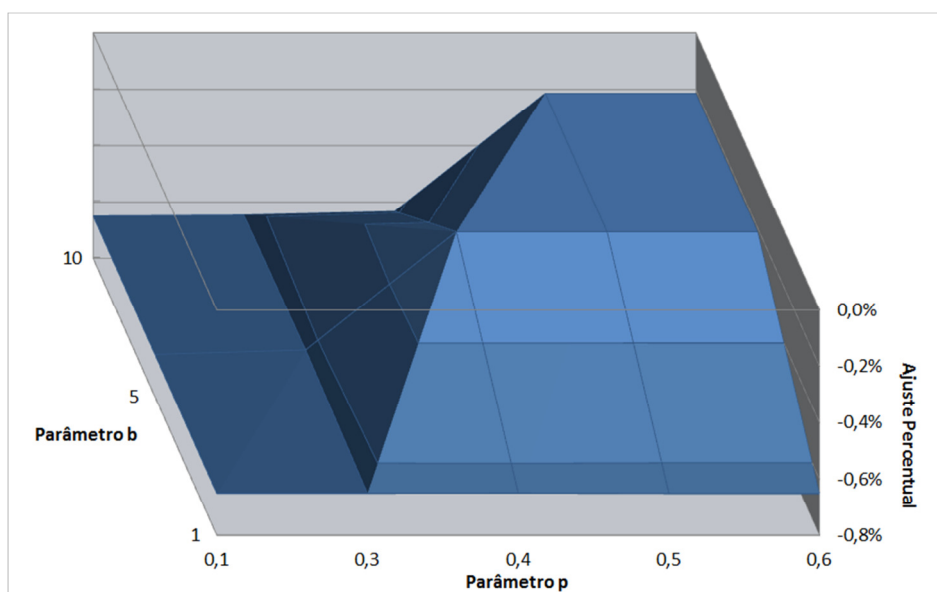


Figura 4.10 – Efeito dos parâmetros da função Normal contaminada sobre os ajustes à corrente E1.

Verifica-se um efeito inesperado para o ajuste com os parâmetros $p = 0,4$ e $b = 10$, que foi menor do que o ajuste para $p = 0,4$ e $b = 5$, por razões desconhecidas (possivelmente um problema numérico). Entretanto, fica claro que para esta corrente também é atingido um patamar de ajuste de, aproximadamente, $-0,2\%$, magnitude maior do que o ajuste nulo, que é o esperado para esta corrente. Salienta-se, contudo, que o Aspen AORA também aplicou ajuste de magnitude elevada ($2,9\%$).

A Tabela 4.5 mostra os ajustes aplicados à corrente S7 para o parâmetro $b=5$.

Tabela 4.5 – Diferentes ajustes aplicados à corrente S7.

	$p=0,1$ (%)	$p=0,3$ (%)	$p=0,4$ (%)	$p=0,5$ (%)	$p=0,6$ (%)
Ajuste	3×10^{-5}	2×10^{-5}	1×10^{-5}	1×10^{-5}	1×10^{-5}

Como se pode ver, os ajustes aplicados a esta corrente são ínfimos, demonstrando que também a função Normal Contaminada é incapaz de identificar e atuar sobre medições “congeladas”.

Diferentes evidências indicam o desempenho insatisfatório da função Normal Contaminada. É possível que problemas com maior número de restrições possam obter um desempenho melhor deste tipo de RD, porém para este caso específico, a função *Fair* gerou estimativas mais coerentes, em geral.

4.4 Teste de Medição para Detecção de Erros Grosseiros

O Teste de Medição (MT) foi executado com duas bases diferentes: uma utilizando como aproximação inicial as estimativas da RD clássica e, a segunda, as estimativas da RD *Fair* com o parâmetro $c=0,001$. O valor do parâmetro $Z_{ac} = Z_{(1-\alpha^*/2)}$ é de 1,92. Contudo, nas análises os valores dos parâmetros Z_{ai} encontrados não foram comparados com este valor crítico para determinar a presença de EG's, já que o trabalho é focado na listagem dos medidores com mais chances de apresentarem EG's. Assim, os valores absolutos de Z_{ai} somente foram ordenados de forma a identificar as correntes mais prováveis de apresentar EG's. As quatro correntes com os maiores valores de Z_{ai} para cada uma das bases são apresentadas nas Tabelas 4.6 e 4.7.

Tabela 4.6 – Maiores valores de Z_{ai} para o MT com base na RD Clássica.

Corrente	Medição original	AORA	RD Clássica	Z_a
S7	9,6	43,8	9,3	18,1
S5	118,2	-	143,6	7,7
S4	28,7	-	34,1	3,5
E1	14031,2	14438,0	13725,1	2,9
E4	37,9	38,2	47,0	2,6

Tabela 4.7 – Maiores valores de Z_{ai} para o MT com base na RD *Fair* ($c=0,001$).

Corrente	Medição original	AORA	<i>Fair</i> ($c=0,001$)	Z_a
E2	482,0	0,0	0,0	3,6
S2	1151,8	-	1316,0	1,4
E1	14031,2	14438,0	14030,3	8,3E-03
S1	4144,7	3978,6	4145,2	4,9E-03
S14	3583,6	3565,5	3584,0	4,8E-03

Conforme já discutido, a RD Clássica não apresenta bons resultados no problema em questão, porém o MT realizado com base nela identificou a corrente S7 como forte candidata a apresentar erro grosseiro. O medidor para esta corrente apresenta leitura “congelada”, o que mostra que esta detecção foi efetiva. Como não há registro no Aspen AORA para as correntes S5 e S4, não se pode concluir nada sobre elas. A detecção para a corrente E1 está correta, porém é apenas uma coincidência que ela apareça nesta posição da classificação; a corrente apresentava, sim, erro grosseiro, mas, a função clássica, que só buscou fechar o balanço, gerou uma estimativa ainda pior, fazendo com que o valor de Z_a fosse elevado, o que não significa que a reconciliação tenha gerado resultados corretos. O mesmo acontece com a corrente E4, porém neste caso, a detecção é um falso positivo.

Já o MT realizado a partir da RD *Fair* contribui para as suspeitas de que as correntes E2 e S2 contenham erros grosseiros. A corrente seguinte na classificação é a E1, que apresenta erro grosseiro. Entretanto, devido à ordem de grandeza do parâmetro Z_a , não se pode afirmar que o teste identificou o erro. Comparando-se os parâmetros Z_{ai} com o valor crítico, as correntes E1, S1 e E14 não tiveram EG's detectados.

5 Conclusões e Trabalhos Futuros

Foi selecionado um conjunto de dados reais de fechamento de balanço de uma unidade industrial, no qual foram aplicadas três técnicas de reconciliação de dados e duas para detecção de erros grosseiros. A reconciliação de dados baseada na função *Fair* se sobressaiu, aplicando ajustes coerentes, em geral, com informações sobre o processo. Já o estimador baseado na função Normal Contaminada também aplicou ajustes coerentes, porém não atingiu a magnitude de correção necessária para gerar estimativas próximas às fornecidas pelo *software* de reconciliação comercial instalado na empresa. O estimador clássico foi o que apresentou o pior desempenho dentre as técnicas aplicadas, chegando a gerar estimativas ainda mais distantes dos valores do reconciliador comercial do que as medições originais. O Teste Global para detecção de erros grosseiros apresentou resultado coerente, confirmando a presença de EG's no conjunto de dados, porém, devido à sua simplicidade, ele é pouco informativo, devido ao fato de já se presumir a presença de EG's no conjunto de dados, sendo inespecífico a respeito da sua localização. Já o Teste de Medição apresentou resultados relevantes quando aplicado baseado na RD *Fair*.

Levando em consideração os resultados da estimação com a função *Fair*, sugere-se que os medidores das correntes E2 e S2 sejam verificados, pois há fortes evidências de que estejam medindo incorretamente. Pelo resultado do *software* comercial, é sugerido verificar os medidores das correntes E1 e S7.

Salienta-se que o desempenho dos estimadores não foi superior devido à simplicidade do problema proposto. Ele foi estruturado desta maneira de forma a manter-se com pequena dimensão, pois desdobrar as vazões significaria aumentar muito a complexidade do problema. Além disto, como o interesse do trabalho é nas medições de balanço global da unidade, empregar novas equações para criar redundâncias significaria envolver muitas outras medições de pouco interesse, simplesmente para conseguir-se fechar outros volumes de controle. No caso de um problema mais elaborado, com mais restrições de processo e redundâncias de medições, possivelmente as funções robustas – e até mesmo a clássica – forneçam estimativas mais realistas, mesmo para as correntes E1 e S7.

Como sugestões para trabalhos futuros podem-se citar: emprego de uma metodologia para detecção de estado estacionário para seleccionar os períodos de dados a serem usados na análise, maior detalhamento do balanço material da unidade, com reaplicação das técnicas implementadas neste estudo, aplicação de novas técnicas de RD e DEG para a verificação dos resultados obtidos.

6 Referências Bibliográficas

ALBUQUERQUE, J. S., BIEGLER, L. T. Data Reconciliation and Gross-Error Detection for Dynamic Systems. *AIChE Journal*, v. 42, p. 2841-2856, 1996.

BUREAU INTERNATIONAL DES POIDS ET MESURES. *Guide to the Expression of Uncertainty in Measurements*. Disponível em <www.bipm.org/en/publications/guides/gum.html>. Acesso em: 08 de set. 2012.

FARIAS, A. C. *Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros*. 180 f. Dissertação (Mestrado em Engenharia Química) – DEQUI, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.

GAUSS, C. F., 1809. In: FARIAS, A. C. *Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros*, 2009, p. 15.

NARASIMHAN, S., JORDACHE, C. *Data reconciliation and gross error detection: an intelligent use of process data*. Houston: Gulf Professional Publishing, 1999. 350 p.

PERAMANU, S., WAH, J.C. Improve material balance by using proper flowmeter corrections. *Hydrocarbon Processing*, Houston, v. 90, n. 10, p. 77-84, out. 2011. Disponível em <www.hydrocarbonprocessing.com>. Acesso em: 19 de ago. 2012.

REILLY, P., CARPANI, R., 1963. In: FARIAS, A. C. *Avaliação de Estratégias para Reconciliação de Dados e Detecção de Erros Grosseiros*, 2009, p. 15.

Apêndices

A. Resultados da RD Clássica.

Corrente	Medição	σ	AORA	RD Clássica
E1	14031,2	2,2	14438,0	13725,1
E2	482,0	2,8	0,0	454,0
E3	136,7	6,2E-01	142,0	118,7
E4	37,9	7,3E-02	38,2	47,0
E5	13,8	3,3E-01	-	23,6
E6	13,9	8,2E-02	-	21,9
E7	26,6	1,2E-01	-	35,4
E8	3,1	5,9E-02	-	10,2
E9	36,9	4,0E-01	-	8,5
E10	36,8	4,1E-01	-	8,4
E11	54,3	4,2E-01	-	25,9
E12	0,2	2,6E-03	-	0,0
Soma E5 - E12	185,6	-	110,9	133,9
Total Entradas	14873,3	-	14729,1	14478,7
S1	4144,7	2,1	3978,6	4171,8
S2	1151,8	2,5	-	1178,3
S3	874,0	1,3	-	843,2
S2+S3	2025,8	-	2329,8	2021,5
S4	28,7	3,2E-02	-	34,1
S5	118,2	6,9E-02	-	143,6
S4+S5	146,9	-	120,6	177,8
S6	1646,4	5,8E-01	1733,2	1672,9
S7	9,6	3,8E-04	43,8	9,3
S8	543,5	1,4	558,0	569,6
S9	45,8	9,1E-01	44,4	72,4
S10	12,7	1,5E-01	-	25,0
S11	64,8	3,8E-01	-	90,5
S10+S11	77,4	-	75,8	115,5
S12	1043,6	1,2	-	1070,9
S13	957,8	6,0E-01	-	985,8
S12+S13	2001,3	-	2037,2	2056,7
S14	3583,6	2,1	3565,5	3611,2
Total Saídas	14225,0	-	14487,0	14478,7
Resíduo	648,3	-	242,1	6,05E-06

B. Resultados da RD baseada na função Fair.

Corrente	Medição	σ	AORA	c=0,001	c=0,01	c=0,1	c=0,5	c=1	c=2	c=3	c=4	c=10
E1	14031,181	2,2	14438,02197	14030,2981	14022,92137	13993,53348	13945,24474	13939,637	13938,751	13938,855	13939,010	13939,432
E2	481,992	2,8	0	0	0	26,50576351	260,421392	301,689	321,339	327,714	330,871	336,497
E3	136,653	6,2E-01	141,9878845	136,6430522	136,5533478	135,8084324	133,4909788	132,155	131,051	130,574	130,309	129,780
E4	37,919	7,3E-02	38,20845032	37,91908365	37,91814501	37,91002485	37,88308683	37,865	37,849	37,842	37,837	37,828
E5	13,775	3,3E-01	-	13,77230575	13,75044669	13,5646174	12,9645889	12,591	12,263	12,114	12,031	11,859
E6	13,940	8,2E-02	-	13,94031863	13,93913132	13,92887359	13,89484273	13,873	13,852	13,843	13,837	13,825
E7	26,625	1,2E-01	-	26,6245131	26,6217398	26,5978321	26,51893213	26,468	26,421	26,399	26,386	26,360
E8	3,131	5,9E-02	-	3,131086229	3,130461466	3,125054866	3,107069452	3,095	3,084	3,079	3,076	3,070
E9	36,913	4,0E-01	-	36,90881831	36,87514001	36,59031227	35,67817573	35,121	34,638	34,424	34,300	34,053
E10	36,774	4,1E-01	-	36,77024096	36,73570496	36,44372293	35,50914666	34,939	34,445	34,225	34,100	33,848
E11	54,276	4,2E-01	-	54,27216566	54,23484123	53,91964955	52,91221324	52,300	51,771	51,536	51,403	51,134
E12	0,151	2,6E-03	-	0,151023389	0,151022227	0,151012195	0,150978621	0,151	0,151	0,151	0,151	0,151
Soma E5 - E12	185,585	-	110,9171087	185,570472	185,4384877	184,3210749	180,7359475	178,536	176,625	175,770	175,285	174,301
Total Entradas	14873,331	-	14729,13542	14390,43071	14382,83135	14378,07877	14557,77615	14590	14606	14611	14613	14618
S1	4144,664	2,1	3978,648682	4145,150249	4149,323374	4170,883999	4211,673536	4219,560	4222,448	4223,121	4223,380	4223,721
S2	1151,766	2,5	-	1315,963041	1297,935131	1233,601881	1285,27354	1279,652	1273,943	1271,583	1270,342	1267,939
S3	874,004	1,3	-	874,0693549	874,6515345	879,0686909	891,1418044	896,528	900,199	901,585	902,325	903,691
S2+S3	2025,769	-	2329,838867	2190,032396	2172,586665	2112,670572	2176,415344	2176,180	2174,142	2173,168	2172,667	2171,630
S4	28,726	3,2E-02	-	28,72582551	28,72600348	28,72754883	28,73269091	28,736	28,739	28,741	28,742	28,743
S5	118,172	6,9E-02	-	118,1717644	118,1726101	118,1799398	118,2042962	118,220	118,235	118,242	118,246	118,254
S4+S5	146,897	-	120,6440926	146,8975899	146,8986136	146,9074886	146,9369871	146,956	146,974	146,983	146,988	146,998
S6	1646,436	5,8E-01	1733,216614	1646,443934	1646,518489	1647,14045	1649,088754	1650,227	1651,178	1651,592	1651,822	1652,285
S7	9,598	3,8E-04	43,77789021	9,59824518	9,598245202	9,598245416	9,598246127	9,598	9,598	9,598	9,598	9,598
S8	543,517	1,4	557,9585266	543,6029505	544,3603903	549,966007	564,8059807	571,014	575,062	576,555	577,327	578,751
S9	45,815	9,1E-01	44,38443089	45,84047141	46,06588508	47,88074764	53,26896777	56,100	58,277	59,174	59,657	60,602
S10	12,675	1,5E-01	-	12,6754103	12,67932543	12,71301734	12,82417996	12,896	12,961	12,992	13,009	13,045
S11	64,756	3,8E-01	-	64,75929552	64,78805726	65,03180146	65,81513586	66,297	66,717	66,906	67,012	67,228
S10+S11	77,431	-	75,79312897	77,43470583	77,46738269	77,7448188	78,63931582	79,193	79,678	79,897	80,021	80,273
S12	1043,560	1,2	-	1043,609097	1044,04853	1047,456068	1057,048948	1061,583	1064,807	1066,067	1066,733	1067,998
S13	957,783	6,0E-01	-	957,7919687	957,8723923	958,542068	960,6344338	961,850	962,861	963,302	963,544	964,033
S12+S13	2001,342	-	2037,204102	2001,401065	2001,920923	2005,998136	2017,683382	2023,433	2027,669	2029,369	2030,278	2032,031
S14	3583,556	2,1	3565,525314	3584,0291	3588,091383	3609,288309	3649,665635	3657,620	3660,589	3661,297	3661,573	3661,950
Total Saídas	14225,026	-	14486,99165	14390,43071	14382,83135	14378,07877	14557,77615	14590	14606	14611	14613	14618
Resíduo	648,31	-	242,14	-1,0004E-10	2,18279E-11	4,46562E-09	-2,9104E-11	0	-2,1828E-11	0	0	0,000

