



XXI Congresso Brasileiro
de Engenharia Química

Fortaleza/CE
25 a 29 de setembro



XVI Encontro Brasileiro sobre o
Ensino de Engenharia Química
Fortaleza/CE
25 a 29 de setembro

RECONCILIAÇÃO ROBUSTA DE DADOS EM COLUNA DE DESTILAÇÃO UTILIZANDO OS SOFTWARES IISE E SCILAB

D. Q. F. de MENEZES¹, R. SOARES², D. M. PRATA³, F. C. PEIXOTO³, A. R. SECCHI¹, J. C. C. S. PINTO¹.

¹ Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química

² Universidade Federal do Rio Grande do Sul, Departamento de Engenharia Química

³ Universidade Federal Fluminense, Departamento de Engenharia Química e de Petróleo

E-mail para contato: dmenezes@peq.coppe.ufrj.br

RESUMO – *Este trabalho apresenta um estudo sobre reconciliação robusta de dados em estado estacionário em colunas de destilação. É realizada uma breve revisão bibliográfica em destilação, em problemas de reconciliação de dados e nos estimadores-M robustos oriundos da estatística robusta. O caso é representado por uma coluna de destilação de hidrocarbonetos leves, cujas variáveis medidas estão corrompidas por erros aleatórios e grosseiros. O problema foi resolvido através de uma interface entre os softwares iiSE e Scilab. O Scilab foi utilizado para otimizar a função objetivo com o método de poliedros flexíveis, e as restrições do problema são representadas pelo modelo de equações MESH inserido no software iiSE. Para este caso os estimadores de Mínimos Quadrados Ponderados, Fair e Welsch foram avaliados. O estimador Welsch obteve o melhor desempenho entre os analisados, conseguindo detectar e eliminar de forma mais eficiente o efeito negativo dos erros grosseiros sobre as estimativas das variáveis.*

1. INTRODUÇÃO

Medidas de processo precisas são de extrema importância para controle, otimização, qualidade, segurança e eficiência do processo. Entretanto, essas medidas contêm erros, aleatórios e grosseiros, causados, por exemplo, pela imprecisão intrínseca dos instrumentos de medição. Desta forma, não se espera que os dados medidos obedeçam às leis de conservação. Portanto, um procedimento de retificação de dados (RTD) é essencial para recuperar satisfatoriamente a informação contida nos dados. Este é basicamente dividido em três etapas: classificação de variáveis; detecção de erro grosseiro (DEG) ou detecção de múltiplos erros grosseiros (DMEG) e reconciliação de dados (RD).

O procedimento mais utilizado na RTD é a RD, onde dados medidos são ajustados de maneira estatisticamente coerente pelo estimador (função objetivo) resultante da formulação de máxima verossimilhança sobre a distribuição estatística dos erros de medição assumida, de forma a satisfazer às leis de conservação e demais restrições impostas ao sistema (modelo matemático), obtendo estimativas confiáveis para as variáveis e parâmetros do processo (Prata *et al.*, 2010).

PROMOÇÃO

REALIZAÇÃO

ORGANIZAÇÃO





XXI Congresso Brasileiro
de Engenharia Química

Fortaleza/CE
25 a 29 de setembro



XVI Encontro Brasileiro sobre o
Ensino de Engenharia Química
Fortaleza/CE
25 a 29 de setembro

Tradicionalmente é assumida distribuição Normal, que resulta no estimador de Mínimos Quadrados Ponderados (MQP).

Uma ferramenta importante para o tratamento dos dados é a DEG, pois os erros grosseiros não segue a distribuição estatística de erros assumida. Erros grosseiros podem ser divididos em valores espúrios (*outliers*) e desvios sistemáticos (*bias*) (Narasimhan e Jordache, 2000). A literatura científica tem mostrado que estimadores-M robustos derivados da estatística robusta são ótimos candidatos para realizar procedimento de DMEG de maneira simultânea. Os estimadores-M são resultantes da formulação de máxima verossimilhança sobre uma distribuição da estatística robusta (função objetivo do problema de otimização associado à estrutura do procedimento de RD). Utilizando-se os estimadores-M robustos é possível minimizar ou eliminar os efeitos negativos dos erros grosseiros sobre as variáveis, sem a necessidade de eliminar as variáveis identificadas com erros grosseiros, simultaneamente com a RD, evitando estratégias iterativas e computacionalmente intensivas. Isto é conhecido como Reconciliação Robusta de Dados (RRD) (Prata *et al.*, 2010).

A RD tem sido aplicada nos mais diversos processos (Prata *et al.*, 2010). Um dos processos de interesse da indústria química e petroquímica é a destilação. Esta é uma antiga operação unitária, tendo como principais aplicações o fracionamento do petróleo, a obtenção de álcoois, a extração de essências e purificação de substâncias (Kister, 1992). Atualmente, com a evolução dos computadores, modelos matemáticos e projetos de colunas de destilação através de simuladores, como: Aspen/Hysys, ProII, ChemCAD, iiSE (nacional) e EMSO (nacional e acadêmico), entre outros, tem se tornado uma prática bastante eficaz e precisa, porém, bastante complexa. Isto se deve, em parte, a existência de correlações entre as variáveis de processo, e a utilização conjunta dos procedimentos de RD/DEG, controle e otimização, para fins de aplicação real nestes equipamentos em processos industriais.

Seguindo esta direção, este trabalho apresenta um estudo sobre RRD em estado estacionário em colunas de destilação, utilizando os *softwares* iiSE e Scilab.

2. REVISÃO DA LITERATURA

A aplicação de RD e DEG em problemas de engenharia química inicia-se nos anos 60. A partir daí muitos artigos e livros aplicados a processos químicos vêm sendo escritos. Entretanto, há um número pequeno de trabalhos de RD em aplicação reais (Prata *et al.*, 2010).

A complexidade de tratar com robustez a RD em um processo não trivial como a destilação juntamente com modelos fenomenológicos rigorosos de colunas de destilação, como por exemplo, as equações MESH - *Mass, Equilibrium, Summation and Enthalpy* - (Kister, 1992), faz com que se tenha uma escassez de artigos neste campo. Geralmente, os modelos são considerados em estado estacionário e representados apenas por balanço de massa, equações de normalização e balanço de massa por componente. Os dois primeiros resultam em sistemas lineares, e o último resulta em um problema não linear mais simples, chamado de bilinear. Rao e Narasimhan (1996) apresentaram um estudo de uma coluna debutanizadora com 8 componentes. Wang *et al.* (2004) avaliaram uma estratégia baseada na utilização conjunta dos testes de medida e nodal (MT-NT) para RD e DEG em

PROMOÇÃO

REALIZAÇÃO

ORGANIZAÇÃO



um conjunto de 7 colunas em série. Diferentemente, dos exemplos anteriores, que consideraram problemas em estado estacionário, Farzi *et al.* (2008) desenvolveram um modelo dinâmico e mais complexo para realizar a RD dinâmica, em uma coluna de destilação, estimando as temperaturas por meio de duas estratégias: redes neurais artificiais e filtro de Kalman estendido.

Fica claro a escassez de trabalhos envolvendo problemas de RRD em colunas de destilação utilizando como restrição o modelo de equações MESH. Isto motivou a realização deste trabalho e a pesquisa em estimadores-M robustos.

3. O PROBLEMA DE RECONCILIAÇÃO ROBUSTA DE DADOS ESTACIONÁRIA

Existem muitas classes de estimadores robustos, sendo as mais populares aquelas utilizadas nos estimadores-M, que são generalizações de um estimador de máxima verossimilhança (Prata *et al.*, 2010). Assumindo que os erros de medição não são correlacionados, o problema de RRD estacionário, de forma generalizada, adota a formulação,

$$\min \sum_i \rho \left(\frac{x_i - z_i}{\sigma_i} \right) = \min \sum_i \rho(\xi_i) \quad (1)$$

sujeito a

$$\underline{h}(x, u) = \underline{0} \quad (2)$$

$$\underline{g}(x, u) \geq \underline{0} \quad (3)$$

onde ρ é uma função razoavelmente monotônica, ξ_i e σ_i são, respectivamente, o resíduo padronizado e o desvio padrão da variável discreta medida z_i , \underline{x} e \underline{u} são os vetores das variáveis medidas reconciliadas e não medidas (observáveis) estimadas, respectivamente. Finalmente, \underline{h} e \underline{g} são as restrições algébricas de igualdade e desigualdade, respectivamente.

Apresentam-se os estimadores-M: MQP (não robusto), Fair e Welsch como possíveis escolhas para ρ descritos nas Equações (4), (5) e (6), respectivamente.

$$\rho_{MQP}(\xi_i) = \frac{\xi_i^2}{2} \quad (4)$$

$$\rho_{Fair}(\xi_i, c_F) = c_F^2 \left[\frac{|\xi_i|}{c_F} - \ln \left(1 + \frac{|\xi_i|}{c_F} \right) \right] \quad (5)$$

$$\rho_{Welsch}(\xi_i, c_W) = \frac{c_W^2}{2} \left[1 - \exp \left(-\frac{\xi_i^2}{c_W^2} \right) \right] \quad (6)$$

Nestas funções c_F e c_W são parâmetros de sintonia relacionados a eficiência relativa. Quanto mais robusto é um estimador, menos eficiente ele é (Albuquerque e Biegler, 1996). Isto é importante



XXI Congresso Brasileiro
de Engenharia Química

Fortaleza/CE
25 a 29 de setembro



XVI Encontro Brasileiro sobre o
Ensino de Engenharia Química
Fortaleza/CE
25 a 29 de setembro

para a comparação entre dois ou mais estimadores, sendo calculados em relação a uma distribuição de referência, quase sempre adotada a distribuição Normal (Prata *et al.*, 2010). Para uma eficiência relativa de 95% os valores correspondentes das constantes de sintonia são: $c_F=1,3998$ e $c_w=2,9846$.

4. RECONCILIAÇÃO ROBUSTA DE DADOS EM COLUNA DE DESTILAÇÃO

O estudo de caso deste trabalho baseou-se em uma coluna real descrita por Krishnamurthy e Taylor (1985). Os autores compararam dados medidos em uma coluna real de destilação com seus dados previstos por modelo de não equilíbrio. A coluna apresenta uma alimentação com uma mistura de 11 hidrocarbonetos, dentre eles os alcanos e seus isômeros, uma corrente de destilado e uma de fundo. Dados como razão de refluxo, temperatura de alimentação, temperatura de sub-resfriamento do condensador, concentrações de topo e fundo, vazão da alimentação, vazão de fundo, perfil de temperatura (12 estágios) e pressão da coluna foram dados pelos autores.

Assim, com base nos dados relatados por Krishnamurthy e Taylor (1985) as restrições do problema são descritas com modelo MESH estacionário inserido no simulador iiSE. O *software* iiSE então é utilizado, com os dados de entradas da coluna descrito pelos autores, para simular os valores “exatos” do problema de RD, e a partir destes dados gerou-se erros randômicos (entre $-1,96\cdot\sigma$ até $+1,96\cdot\sigma$) e erros grosseiros (magnitude acima de 10) sobre as variáveis simuladas (consideradas exatas).

O caso foi resolvido nos pacotes computacional iiSE e Scilab. O modelo MESH é resolvido no *software* iiSE a cada iteração e a função objetivo é minimizada pelo otimizador *fminsearch* (método de Neade-Mead) no *software* Scilab a cada iteração. Como a interface entre os dois softwares é possível resolver o problema utilizando uma técnica conhecida como “feasible path” (caminho viável), onde o modelo dentro do simulador iiSE recebe as variáveis de entrada (independentes) e os parâmetros de processo, resolve-se o modelo MESH e gera as variáveis de saída (variáveis dependentes). As variáveis de saída são lidas pelo software Scilab através da interface e, por fim, todas as variáveis medidas são computadas pelo otimizador *fminsearch*. O ciclo se repete até atingir um critério de parada (tolerância). A tolerância utilizada é definida pela diferença entre uma variável da iteração i e a mesma variável na iteração $i+1$, com o seu valor de parada sendo $1,0\times 10^{-4} \geq x^{(i+1)} - x^{(i)}$.

Assim, o problema de RD passa a obter um total de 48 variáveis medidas (simuladas e consideradas como exatas):

- 3 vazões (Alimentação (F), Destilado (D) e Fundo (B));
- 33 componentes (11 - F , 11 - D e 11 - B);
- 12 temperaturas (12 estágios de equilíbrio = condensador + 10 pratos + refeedor).

Com isso pode-se avaliar o problema de RRD juntamente com o modelo MESH, e, também, o método de otimização Neader-Mead.

PROMOÇÃO

REALIZAÇÃO

ORGANIZAÇÃO



4.1. Critério de avaliação

O critério TER (*Total Error Reduction*) é muito utilizado na avaliação de desempenho do procedimento de RD e DEG, quando todos os valores exatos (base da simulação) são fornecidos (Özyurt & Pike, 2004; Prata *et al.*, 2008), conforme apresentado na Equação (11).

$$TER = \frac{\sqrt{\sum \frac{(z_{medido} - z_{exato})^2}{\sigma^2}} - \sqrt{\sum \frac{(x_{reconciliado} - z_{exato})^2}{\sigma^2}}}{\sqrt{\sum \frac{(z_{medido} - z_{exato})^2}{\sigma^2}}} \quad (7)$$

O critério TER mede a redução total dos erros, tomado por base os valores medidos e os reconciliados em relação aos valores exatos. Quanto maior for o TER obtido, ou mais próximo de 1, melhor terá sido o resultado. O critério TER pode ser utilizado independentemente da função objetivo (estimador) utilizada, permitindo assim ser utilizado para comparação de estimadores-M robustos e outras estratégias nos procedimentos de RD e DEG.

5. RESULTADOS E DISCUSSÃO

Nas Tabelas de 1, 2, 3 e 4 são apresentados os resultados obtidos para o caso onde se avalia o modelo MESH (restrição do problema de RRD) na presença de variáveis com erros grosseiros.

Tabela 1 – Resultados (alimentação-F).

	Exato	Medido	Dados Reconciliados - iiSE/Scilab		
			MQP	Fair	Welsch
Vazão (F) [kmol/h]	184,0680	220,8816	196,2266	194,2708	182,0867
etano	1,0737	1,0878	1,0594	1,0637	1,0190
propano	98,9366	100,8490	105,5082	104,6521	97,9163
isobutano	39,2678	39,7448	43,1992	41,5649	38,8493
n-butano	30,8825	30,4771	32,2897	32,4061	30,4508
isopentano	10,2771	10,1372	10,3387	10,6255	10,0937
n-pentano	5,9311	5,8544	5,9271	6,0559	5,8139
hexano	8,7944	8,9321	8,9432	9,0792	8,6168
heptano	41,3130	40,7005	43,1018	43,8496	40,9154
octano	114,1733	112,6523	119,4813	120,7743	113,1773
nonano	86,6142	87,9944	93,8717	91,6915	85,7553
decano	74,0362	75,1479	81,3564	77,8784	73,1888

Valores em negrito: variável com erro grosseiro.

Tabela 2 – Resultados (alimentação-D).

	Exato	Medido	Dados Reconciliados - iiSE/Scilab		
			MQP	Fair	Welsch
Vazão (D) [kmol/h]	66,8000	67,8934	71,0359	70,3594	65,9674
etano	1,0506	1,0639	1,0592	1,0636	1,0189
propano	98,6157	100,2256	105,1350	104,2797	97,5716
isobutano	37,9457	45,5349	<u>41,7046</u>	<u>40,1201</u>	<u>37,5113</u>
<i>n</i> -butano	28,9963	29,5526	30,2678	30,3694	28,5525
isopentano	7,6627	7,5827	7,6758	7,8890	7,5025
<i>n</i> -pentano	4,0622	4,0020	4,0412	4,1300	3,9692
hexano	3,1330	3,0934	3,1764	3,2289	3,0633
heptano	3,2695	3,2232	3,4075	3,4869	3,2392
octano	0,7776	0,7876	0,8101	0,8300	0,7723
nonano	0,0399	0,0394	0,0428	0,0426	0,0396
decano	0,0024	0,0025	0,0026	0,0026	0,0024

Valores em negrito: variável com erro grosseiro.

Tabela 3 – Resultados (alimentação-B).

	Exato	Medido	Dados Reconciliados - iiSE/Scilab		
			MQP	Fair	Welsch
Vazão (B) [kmol/h]	117,2680	115,8468	125,1907	123,9114	116,1193
etano	1,40E-03	1,39E-03	1,46E-03	1,47E-03	1,39E-03
propano	3,3917	3,4489	3,7319	3,7247	3,4476
isobutano	13,2106	13,0582	14,9460	14,4473	13,3799
<i>n</i> -butano	18,8538	19,1737	20,2192	20,3666	18,9835
isopentano	26,1415	26,6229	26,6295	27,3649	25,9115
<i>n</i> -pentano	18,6869	18,4241	18,8589	19,2592	18,4467
hexano	56,6115	55,8136	57,6680	58,5029	55,5349
heptano	380,4242	387,5702	396,9428	403,6273	376,7615
octano	1133,9259	1121,2958	1186,7125	1199,4427	1124,0499
nonano	865,7203	855,3331	938,2894	916,4889	857,1577
decano	740,4767	752,0420	813,5380	778,7579	731,8637

Pode-se observar nas Tabelas 1, 2 e 3 que todos os estimadores detectaram os erros grosseiros, porem apenas o estimador de Welsch aproximou-se melhor do valor simulado (verdadeiro). O estimador não robusto MQP mostrou-se pouco eficaz na correção dos erros grosseiros.

Tabela 4 – Resultados (perfil de temperatura).

Estágio	Exato	Medido	Dados Reconciliados - iiSE/Scilab		
			MQP	Fair	Welsch
0	305,00	304,03	305,00	305,00	305,00
1	358,97	360,15	358,30	358,92	358,94
2	404,44	359,08	<u>404,07</u>	<u>404,65</u>	<u>404,51</u>
3	430,49	429,50	430,38	430,64	430,55
4	444,27	443,15	444,23	444,32	444,29
5	453,87	454,99	453,86	453,85	453,85
6	457,39	456,31	457,25	457,26	457,30
7	464,05	465,01	463,78	463,77	463,84
8	475,41	474,34	475,18	475,01	475,09
9	490,45	491,58	490,53	490,09	490,14
10	505,68	507,12	506,14	505,47	505,48
11	519,66	520,97	520,32	519,56	519,57

A Tabela 5 apresenta o critério de avaliação de desempenho utilizado neste caso.

Tabela 5 – Avaliação de desempenho (TER).

TER	MQP	Fair	Welsch
Vazões e Concentrações	0.620	0.780	0.889
Temperaturas	0.974	0.984	0.988
TOTAL	0.799	0.883	0.939

A Tabela 5 mostra claramente que o estimador Welsch apresentou um desempenho superior em relação aos estimadores MQP e Fair, o que contribui, mais uma vez, à utilização do procedimento de RRD em problemas de Engenharia Química, especificamente em colunas de destilação. Cabe ressaltar que este resultado corrobora os trabalhos apresentados de Prata *et al.* (2010), nos quais o estimador de Welsh mostrou destaque. Deve ficar claro que a utilização do estimador MQP, mesmo associado à estratégias de DMEG pode levar a conclusões errôneas (efeito “*smearing*”) sobre erros grosseiros, devido a sua natureza não robusta. Além disso, esta estratégia é iterativa e pode inviabilizar sua utilização em aplicações reais. Conforme apresentado este trabalho está alinhado a pesquisa atual de RD e DMEG com base em estimadores-M robustos.



XXI Congresso Brasileiro
de Engenharia Química

Fortaleza/CE
25 a 29 de setembro



XVI Encontro Brasileiro sobre o
Ensino de Engenharia Química
Fortaleza/CE
25 a 29 de setembro

5. CONCLUSÃO

Para o caso estudado pode-se concluir que a detecção dos erros grosseiros nas variáveis dependentes (principalmente na temperatura do segundo prato) é corroborada pelo modelo MESH, contribuindo para a regressão desses erros. Entretanto, apenas o estimador-M robusto do tipo “*redescending*” Welsch apresentou melhores regressões de erro grosseiro na variável independente (alimentação F) e na variável dependente (concentração de isobutano no destilado (x_{D3})). Assim, conclui-se, novamente, a eficiência e robustez do procedimento de reconciliação de dados e detecção de erros grosseiros simultaneamente, utilizando estimadores-M robustos do tipo “*redescending*”, juntamente com modelo rigoroso de destilação. A maior desvantagem do modelo MESH é o aumento do tempo computacional para o problema de RD, pois o grande número de equações e variáveis, juntamente ao método de otimização torna-se o problema mais dispendioso computacionalmente, mas por outro lado é possível resolver problemas de RD que seja não-observável com certa quantidade de variáveis medidas em conjunto com um modelo simples bilinear.

6. REFERÊNCIAS

- ALBUQUERQUE, J. S., BIEGLER, L. T., “Data Reconciliation and Gross-Error Detection for Dynamic Systems”, *AIChE Journal*, v. 42, pp. 2841-2856, 1996.
- FARZI, A.; MEHRABANI-ZEINABAD, A.; BOOZARJOMEHRY, R. B. Data reconciliation: Development of an object-oriented software tool. *Korean J. Chem. Eng.*, v. 25, p. 955-965, 2008.
- KISTER, H. Z. *Distillation-Design*. EUA: Editora McGraw-Hill, 1992.
- KRISHNAMURTHY, R., TAYLOR, R., “Simulation of Packed Distillation and Absorption Columns”. *Ind. Eng. Chem. Process Des. Dev.*, v. 24, pp. 513-524, 1985.
- NARASHIMHAN, S.; JORDACHE, C. *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*. Texas: Editora Gulf Professional Publishing, 2000.
- ÖZYURT, D.B.; PIKE, R.W. Theory and practice of simultaneous data Reconciliation and gross error detection for chemical process. *Comput. Chem. Engng.*, v.28, p. 381–402, 2004.
- PRATA, D. M., PINTO, J. C., LIMA, E. L., “Comparative analysis of robust estimators on nonlinear dynamic data reconciliation”, *Comput. Aided Chem. Eng.*, v. 25, pp. 501–506, 2008.
- PRATA, D. M.; SCHWAAB, M.; LIMA, E. L.; PINTO, J. C. Simultaneous Robust Data Reconciliation and Gross Error Detection through Particle Swarm Optimization for an Industrial Polypropylene Reactor. *Chem. Eng. Sci.*, v. 65, p. 4943-4954, 2010.
- RAO, R. R.; NARASIMHAN, S. Comparison of Techniques for Data Reconciliation of Multicomponent Processes. *Ind. Eng. Chem. Res.*, v. 35, p. 1362-1368, 1996.
- WANG, F.; JIA, X.; ZHENG, D.; YUE, J. An improved MT-NT method for gross error detection and data reconciliation. *Comput. Chem. Eng.*, v. 28, pp. 2189-2192, 2004.