

AValiação DA QUALIDADE DAS MEDIDAS EXPERIMENTAIS NAS REGIÕES DE CONFIANÇA DOS PARÂMETROS CINÉTICOS ESTIMADOS DE UM SISTEMA COM REAÇÕES EM SÉRIE

H. ENZWEILER¹, E. B. COUTINHO² e M. SCHWAAB³

¹ Universidade Estadual de Maringá, Departamento de Engenharia Química

² Universidade Federal do Rio Grande do Sul, Departamento de Físico-Química

³ Universidade Federal Santa Maria, Departamento de Engenharia Química

E-mail para contato: marcioschwaab@gmail.com

RESUMO – Em sistemas compostos por diversas reações em série e/ou paralelo nem sempre há a possibilidade de medir-se a concentração de todas as espécies envolvidas ou, ainda, as medidas destas variáveis não são totalmente independentes. Nestes casos, a qualidade dos valores estimados para os parâmetros cinéticos do sistema depende do número de componentes medidos e do grau de correlação entre as medições. Para um sistema reacional em série do tipo $A \rightarrow B \rightarrow C$ foi avaliado o efeito da quantidade de medidas experimentais e da presença de correlação entre as medidas experimentais sobre a região de confiança dos parâmetros estimados. Observou-se que apenas com as medidas de concentração do produto final obtém-se uma região de confiança extensa e com forma distante da elíptica. Foi verificado que o valor da correlação modifica significativamente o tamanho e a forma da região de confiança dos parâmetros cinéticos, podendo ainda causar o seu deslocamento para intervalos de valores diferentes.

1. INTRODUÇÃO

No caso de sistemas que envolvem reações em série ou em paralelo, a determinação de parâmetros cinéticos geralmente envolve a utilização de medidas da concentração dos compostos envolvidos na reação. Contudo, nem sempre há a possibilidade de obter as medidas para todos os compostos envolvidos (Box e Draper, 1965), ou mesmo estas medidas não podem ser obtidas de maneira independente, sem correlação (Santos e Pinto, 1998). Em ambos os casos, a qualidade dos parâmetros estimados sofre influência, tanto do número de medições quanto da variância e covariância dos dados.

A confiabilidade dos parâmetros pode ser verificada através da análise da região de confiança. A região de confiança é a região do espaço de variáveis que concentra uma probabilidade p das possíveis flutuações observáveis do problema (Schwaab e Pinto, 2007). A aproximação elíptica da região de confiança, rigorosamente válida somente quando os erros experimentais possuem distribuição normal e o modelo é linear, é muito usada devido a sua simplicidade, pois requer apenas informações sobre o ponto ótimo estimado e a matriz de

covariância (Schwaab *et al.*, 2008). Porém, quando o caso avaliado difere dessas condições ideais, o que ocorre em muitos casos (Larentis *et al.*, 2003), a forma elíptica fornece apenas uma aproximação da região de confiança dos parâmetros. Assim é necessário o emprego de métodos que não restrinjam a forma da região de confiança para uma melhor avaliação da confiabilidade dos valores estimados.

O objetivo principal deste trabalho é a avaliação do efeito tanto da quantidade de variáveis medidas quanto da correlação entre as variáveis medidas sobre a qualidade das estimativas dos parâmetros cinéticos de uma reação em série do tipo $A \rightarrow B \rightarrow C$.

2. METODOLOGIA

Foram utilizados dados experimentais simulados empregando-se modelos cinéticos para um sistema de reações em série. Assumindo que as taxas das duas etapas da reação R são irreversíveis e de primeira ordem e que a reação é conduzida em um reator do tipo batelada, a variação da concentração dos compostos A, B e C no tempo (t) é descrita pelo sistema de equações diferenciais apresentados na Equação 1. Considerando no tempo zero o reator contém apenas o reagente A com concentração C_{A0} . Este sistema possuiu solução analítica, sendo que as concentrações de A, B e C são funções do tempo de reação conforme Equação 2.



$$\begin{cases} \frac{dC_A}{dt} = -k_1 C_A \\ \frac{dC_B}{dt} = k_1 C_A - k_2 C_B \\ \frac{dC_C}{dt} = k_2 C_B \end{cases} \quad (1)$$

$$\begin{cases} C_A = C_{A0} \exp(-k_1 t) \\ C_B = C_{A0} \frac{k_1 [\exp(-k_1 t) - \exp(-k_2 t)]}{k_2 - k_1} \\ C_C = C_{A0} \left\{ 1 - \frac{[k_2 \exp(-k_1 t) - k_1 \exp(-k_2 t)]}{k_2 - k_1} \right\} \end{cases} \quad (2)$$

Os dados experimentais foram simulados assumindo k_1 e k_2 iguais a 0,2 e 0,1, respectivamente, e C_{A0} igual a 1 mol.L⁻¹. Foi adicionado um erro aleatório, com distribuição normal e desvio padrão de 0,05 ao valor calculado para todas as concentrações de compostos, de forma a simular o efeito de erros experimentais. Para a estimação dos parâmetros cinéticos foram consideradas seis casos quanto à quantidade de variáveis observadas experimentalmente, como apresentado na Tabela 1.

Tabela 1 – Casos considerados para a estimação de parâmetros cinéticos

Caso 1	Têm-se a medida apenas da concentração do composto B
Caso 2	Têm-se a medida apenas da concentração do composto C
Caso 3	Têm-se a medida da concentração dos compostos A e B
Caso 4	Têm-se a medida da concentração dos compostos A e C
Caso 5	Têm-se a medida da concentração dos compostos B e C
Caso 6	Têm-se a medida da concentração de todos os compostos (A, B e C)

O método da máxima verossimilhança está baseado em hipóteses relativamente simples e permite a análise de virtualmente qualquer problema experimental de forma rigorosa, desde que se conheça de forma apropriada como se comportam os erros de medição na região experimental (Schwaab e Pinto, 2007). A primeira hipótese é que os erros experimentais seguem uma distribuição conhecida na região experimental. A segunda considera o modelo perfeito, ou seja, o modelo é perfeito e capaz de descrever exatamente as relações existentes entre as variáveis do problema e que os desvios são devidos exclusivamente a erros experimentais. Pressupõe-se, ainda, que o experimento é bem feito, sem a presença de erros grosseiros ou *outliers*. Assumindo que as hipóteses são verdadeiras e que as flutuações experimentais seguem a distribuição normal, a função a ser minimizada na estimação dos parâmetros do modelo (função objetivo) é conhecida como função de mínimos quadrados generalizados, Equação 3, pois leva em consideração tanto a variância das medidas como a covariância (ou correlação) entre elas.

$$F_{obj} = \sum_{i=1}^{NE} [\mathbf{y}_i^e - \mathbf{y}_i^m(\mathbf{x}_i^e, \boldsymbol{\theta})]^T \mathbf{V}_{y,i}^{-1} [\mathbf{y}_i^e - \mathbf{y}_i^m(\mathbf{x}_i^e, \boldsymbol{\theta})] \quad (3)$$

A região de confiança elíptica dos parâmetros pode ser representada pela Equação 4 (Schwaab e Pinto, 2007). Uma maneira alternativa de encontrar a região de confiança sem restringir sua forma, conhecida como Região de Confiança de Verossimilhança (Schwaab *et al.*, 2008), é representada pela Equação 5.

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{V}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq F_{obj}(\hat{\boldsymbol{\theta}}) \frac{NP}{N - NP} F_{NP, N-NP}^{\alpha} \quad (4)$$

$$F_{obj}(\boldsymbol{\theta}) \leq F_{obj}(\hat{\boldsymbol{\theta}}) \left(1 + \frac{NP}{N - NP} F_{NP, N-NP}^{\alpha} \right) \quad (5)$$

A estimação foi realizada através da minimização da função objetivo com o uso de um algoritmo híbrido de otimização composto por um método estocástico acoplado a um método determinístico, em que a melhor solução encontrada pelo método do Enxame de Partículas (Kennedy e Eberhart, 1995) era utilizada como estimativa inicial para um método do tipo Gauss-Newton (Noronha *et al.*, 1993). A utilização do método do Enxame de Partículas, apesar de ser computacionalmente intensivo para este problema com apenas 2 parâmetros para serem estimados, permite uma fácil obtenção da Região de Confiança de Verossimilhança (Schwaab *et al.*, 2008).

3. RESULTADOS E DISCUSSÃO

3.1. Influência da quantidade de variáveis medidas

Primeiramente foram estimados os parâmetros cinéticos k_1 e k_2 , utilizando os dados experimentais simulados, considerando a variância experimental igual a 0,0025, idêntica para todas as variáveis e experimentos. Os ajustes obtidos nos casos com diferentes quantidades de variáveis medidas para as concentrações dos compostos A, B e C, podem ser observados na Figura 1. Verifica-se que medindo apenas o produto final, caso 2, o modelo não consegue prever com qualidade as concentrações dos outros dois compostos, A e B. Para os demais casos a capacidade preditiva do modelo foi idêntica.

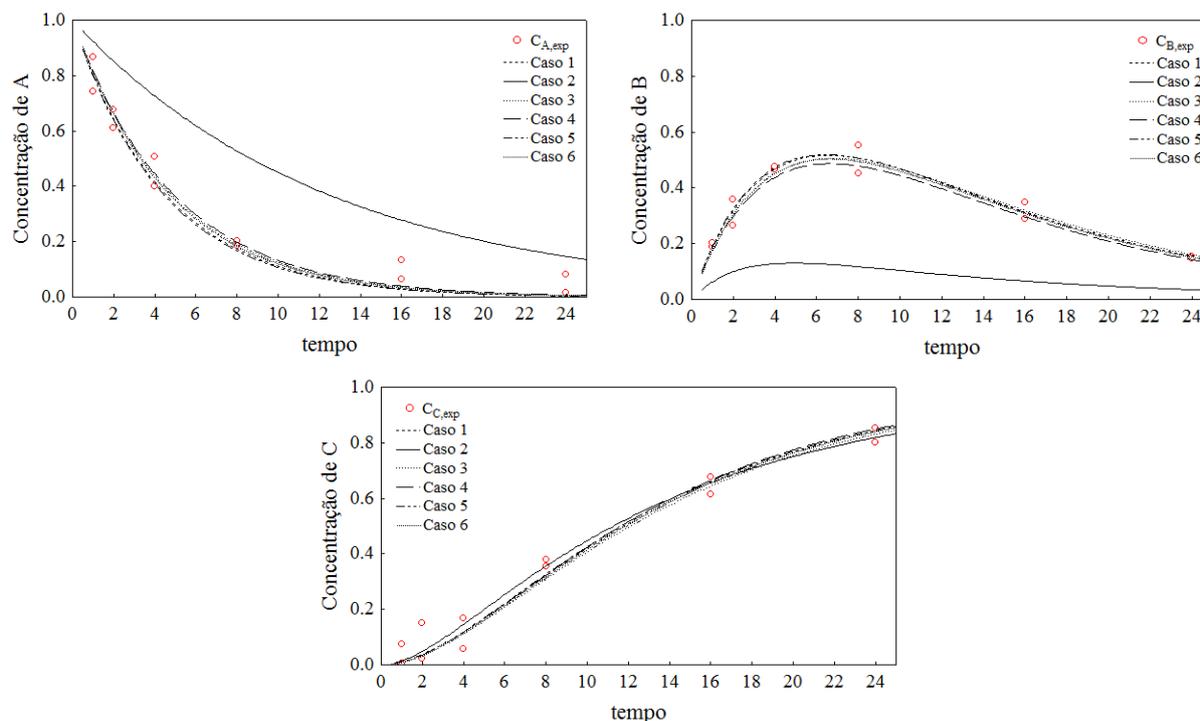


Figura 1 – Dados experimentais simulados e ajustes do modelo considerando os casos com diferentes números de variáveis medidas (concentrações dos compostos A, B e C).

Quanto aos valores dos parâmetros estimados, Tabela 2, verifica-se que há diferenças na precisão dos parâmetros cinéticos obtidos para os diferentes casos. Observa-se que os valores dos desvios padrão dos parâmetros mudam consideravelmente quando diferentes variáveis são medidas. Constatou-se que para a reação em série $A \rightarrow B \rightarrow C$, medindo-se apenas as concentrações do reagente e do produto intermediário, é possível obter os valores das constantes cinéticas k_1 e k_2 com confiabilidade equivalente àquela que seria obtida caso as concentrações de todos os compostos fossem medidas. É importante ressaltar que esse resultado só se aplica quando não há correlação entre as variáveis medidas.

Tabela 2 – Função objetivo, parâmetros cinéticos e desvio padrão para os seis casos

Caso	F_{obj}	k_1 [min^{-1}]	σ_{k_1} [min^{-1}]	k_2 [min^{-1}]	σ_{k_2} [min^{-1}]
1 - Medindo C_B	4,67	0,2195	0,0120	0,1008	0,0049
2 - Medindo C_C	10,92	0,0801	0,0112	0,4142	0,1872
3 - Medindo C_A e C_B	18,61	0,2073	0,0092	0,1003	0,0067
4 - Medindo C_A e C_C	26,36	0,2023	0,0132	0,1084	0,0098
5 - Medindo C_B e C_C	17,44	0,2239	0,0140	0,1015	0,0048
6 - Medindo C_A , C_B e C_C	31,96	0,2103	0,0093	0,1030	0,0052

A qualidade dos parâmetros obtidos também pode ser verificada através das regiões de confiança, Figura 2. À exceção do caso 2, todas as hipóteses proporcionaram a obtenção de regiões de confiança de verossimilhança semelhantes às regiões elípticas, mostrando que, para estes casos, a aproximação elíptica pode ser aplicada com segurança na estimação dos parâmetros. Porém, para o caso em que apenas a concentração do produto final é medida (Caso 2) a região de confiança apresenta forma muito distante da elíptica. Este resultado está associado a pior qualidade do ajuste, observada na Figura 1. Indicando que a confiabilidade da aproximação elíptica da região de confiança dos parâmetros está associada à qualidade do ajuste dos dados experimentais, quanto melhor o ajuste, mais a região de confiança se aproxima da forma elíptica. Assim, pode-se afirmar que medindo apenas a concentração do produto final não é possível a obtenção de parâmetros cinéticos confiáveis para uma reação em série do tipo $A \rightarrow B \rightarrow C$.

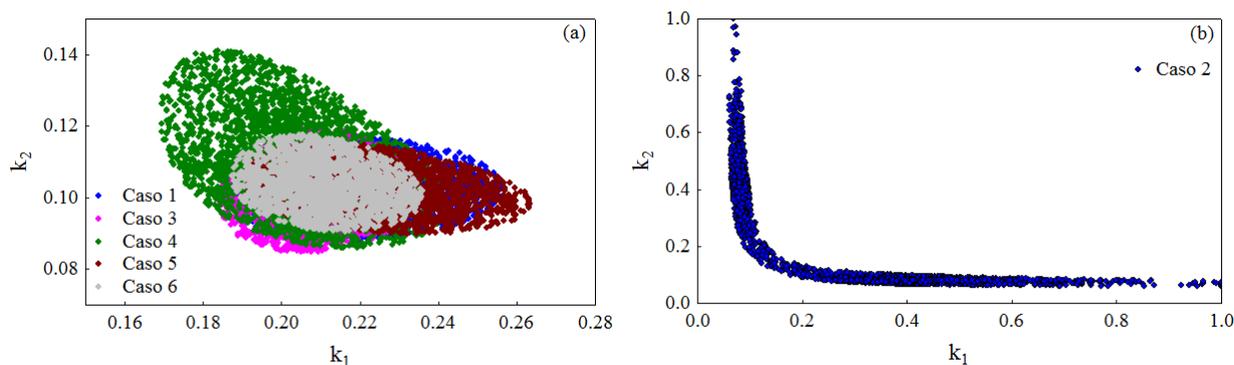


Figura 2 – Regiões de confiança dos parâmetros cinéticos para os Casos (a) 1, 3-6 (b) 2.

3.2. Influência da correlação entre as medidas experimentais

Para a avaliação da influência da correlação entre as variáveis sobre a qualidade dos parâmetros cinéticos estimados foram considerados apenas os casos em que duas concentrações eram medidas (Casos 3, 4 e 5) e o coeficiente de correlação foi variado de $-0,99$ a $+0,99$. Considerando que apenas as concentrações de A e B são medidas, Figura 3, observa-se que o aumento do coeficiente de correlação resulta no aumento do tamanho da

região de confiança de verossimilhança e na piora da aproximação elíptica para a região de confiança. Estes resultados impactam na redução da confiabilidade dos parâmetros cinéticos obtidos.

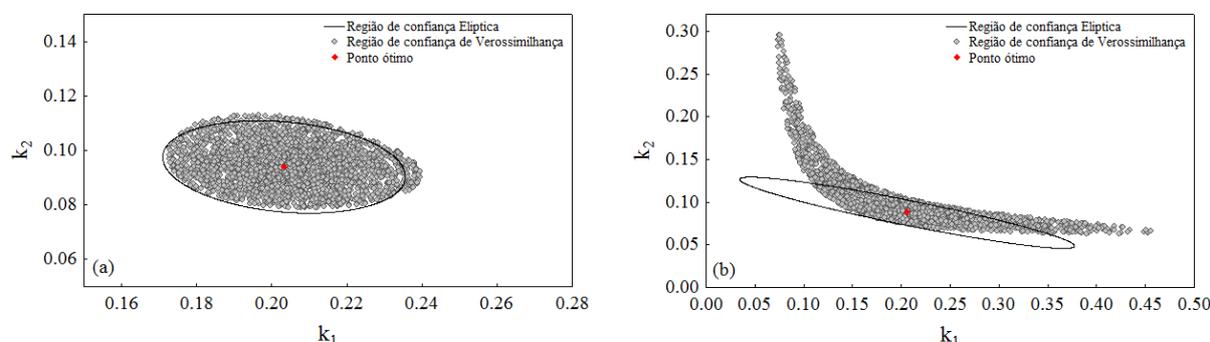


Figura 3 – Regiões de confiança elíptica e de verossimilhança para a hipótese de medirem-se as concentrações de A e B (Caso 3) com coeficiente de correlação de (a) -0,50 e (b) -0,99.

Comparando a Figura 3b com a região de confiança do caso 1 na Figura 2a, fica claro que mesmo medindo dois compostos, se a correlação entre as medidas é alta, a precisão dos parâmetros é pior do que somente medindo a concentração do composto intermediário.

Para o caso em que as concentrações de A e C (Caso 4) são medidas, os resultados obtidos, variando-se o coeficiente de correlação entre as variáveis, são apresentados na Tabela 3. Pode-se verificar que o aumento do coeficiente de correlação entre $-0,99$ e $+0,99$ ocasionou um aumento do valor estimado para k_1 e uma redução do valor de k_2 , assim como um aumento do desvio padrão do parâmetro k_1 , sendo que o desvio padrão do parâmetro k_2 passa por um máximo.

Tabela 3 – Função objetivo, parâmetros cinéticos e desvio padrão para o Caso 4, medindo-se as concentrações de A e C, com diferentes coeficientes de correlação.

Correlação	F_{obj}	k_1 [min^{-1}]	σ_{k1} [min^{-1}]	k_2 [min^{-1}]	σ_{k2} [min^{-1}]
$\rho = -0,99$	277,24	0,1929	0,0083	0,1224	0,0048
$\rho = -0,90$	39,86	0,1930	0,0089	0,1215	0,0055
$\rho = -0,50$	21,90	0,1954	0,0109	0,1168	0,0079
$\rho = 0$	26,36	0,2023	0,0132	0,1084	0,0098
$\rho = +0,50$	44,53	0,2161	0,0158	0,0953	0,0103
$\rho = +0,90$	180,13	0,2349	0,0177	0,0817	0,0090
$\rho = +0,99$	1681,13	0,2402	0,0180	0,0785	0,0086

Quando foi considerado que as concentrações dos compostos B e C (Caso 5) eram medidas, observou-se tanto o deslocamento quanto a mudança da forma da região de confiança dos parâmetros com a variação do coeficiente de correlação, Figura 4. Nesse caso, fica evidente a redução da confiabilidade dos parâmetros, pois a região de confiança de

verossimilhança aumenta consideravelmente, devido à correlação entre as variáveis medidas.

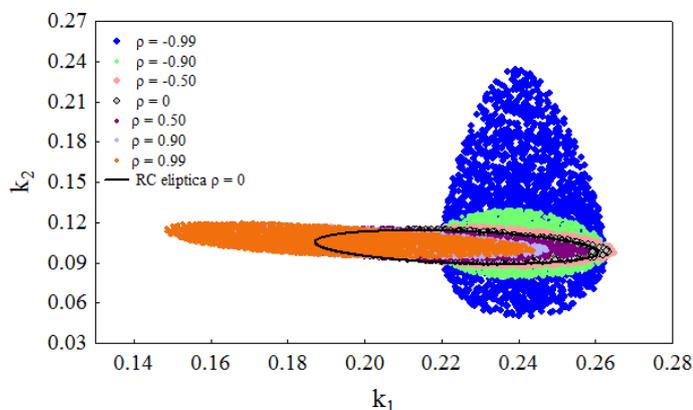


Figura 4 – Regiões de confiança para a hipótese de medirem-se as concentrações de B e C (Caso 5) com diferentes coeficientes de correlação.

4. CONCLUSÃO

Através da estimação de parâmetros cinéticos utilizando dados experimentais simulados para uma reação em série do tipo $A \rightarrow B \rightarrow C$, verificou-se a influência da quantidade de variáveis medidas experimentalmente sobre a qualidade dos resultados obtidos. Constatou-se que com as medidas das concentrações do reagente e do composto intermediário é possível a estimação dos parâmetros cinéticos com qualidade, desde que estas medidas experimentais não estejam correlacionadas. A correlação entre as variáveis impacta diretamente na precisão dos valores estimados, tanto deslocando quanto alterando a forma das regiões de confiança dos parâmetros.

5. AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) pela concessão de bolsa de estudo.

6. NOMENCLATURA

C_j	Concentração do composto j	$\mathbf{V}_{y,i}$	Matriz de covariância experimental
$F_{NP,N-NP}^\alpha$	Função de Fischer	\mathbf{V}_0	Matriz de covariância dos parâmetros
F_{obj}	Função objetivo	\mathbf{x}_i	Variável independente do experimento
GL	Graus de liberdade	\mathbf{y}_j^e	Variável dependente j do experimento i experimental
k_R	Constante cinética da reação R	\mathbf{y}_i^m	Variável dependente j do experimento i do modelo
N	Número total de medições	σ	Desvio padrão
NE	Número de medições no tempo	ρ	Coefficiente de correlação
NP	Número de parâmetros	$\hat{\theta}$	Vetor de parâmetros ótimos

θ Vetor de parâmetros

7. REFERÊNCIAS

BOX, G. E. P.; DRAPER, N. R. The Bayesian estimation of common parameters from several responses. *Biometrika*, v. 52, p. 355-365, 1965.

KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, v. 4, p. 1942-1948, 1995.

LARENTIS, A. L.; BENTES, Jr., A. M. P.; RESENDE, N. S.; SALIM, V. M. M.; PINTO, J. C. Analysis of experimental errors in catalytic tests for production of synthesis gas. *Appl Catal A-Gen*, v. 242, p. 365-379, 2003.

NORONHA, F. B.; PINTO, J. C.; MONTEIRO, J. L.; LOBÃO, M. W.; SANTOS, T. J.; 1993, *Um Pacote Computacional para Estimção de Parâmetros e Projeto de Experimentos*, Relatório Técnico PEQ/COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

SANTOS, T. J. PINTO, J. C. Taking variable correlation into consideration during parameter estimation. *Braz J Chem Eng*, v. 15, p. 1-20, 1998.

SCHWAAB, M.; PINTO, J. C. *Análise de Dados Experimentais I – Fundamentos de Estatística e Estimção de Parâmetros*. E-papers, Rio de Janeiro, 2007.

SCHWAAB, M.; BISCAIA, Jr, E. C.; MONTEIRO, J. L.; PINTO, J. C. Nonlinear parameter estimation through particle swarm optimization. *Chem Eng Sci*, v. 63, p. 1542-1552, 2008.