

## CONSEQUÊNCIAS DA ANÁLISE INCORRETA DE EXPERIMENTOS BLOCADOS

João Alexandre Bortoloti e Roy Edward Bruns\*

Departamento de Físico-Química, Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13084-971 Campinas – SP, Brasil

Recebido em 26/9/05; aceito em 16/5/06; publicado na web em 24/10/06

THE CONSEQUENCES OF INCORRECT ANALYSIS OF BLOCKED EXPERIMENTS. Errors are always present in experimental measurements so, it is important to identify them and understand how they affect the results of experiments. Statistics suggest that the execution of experiments should follow random order, but unfortunately the complete randomization of experiments is not always viable for practical reasons. One possible simplification is blocked experiments within which the levels of certain factors are maintained fixed while the levels of others are randomized. However this has a cost. Although the experimental part is simplified, the statistical analysis becomes more complex.

Keywords: optimization; split-plot design; factorial design.

## INTRODUÇÃO

Em muitos trabalhos que envolvem a realização de experimentos, é comum os químicos se depararem com a questão: como avaliar se os resultados obtidos são confiáveis? Com certeza o simples surgimento desta dúvida já é um bom sinal e indica que o experimentalista se preocupa com a qualidade de seu trabalho. Sabe-se que qualquer medida realizada é sempre afetada por erros. Erros muito pequenos não trarão grandes implicações. Contudo, se forem significativos, poderão prejudicar seriamente os resultados levando a falsas conclusões. Portanto, conhecer a natureza dos erros e preparar planejamentos que possam minimizá-los é uma estratégia que deve estar presente no dia-a-dia de todo pesquisador.

Existem dois tipos de erros, o erro sistemático e o erro aleatório. O primeiro, tem como característica afetar os resultados dos experimentos sempre na mesma direção, seja para mais ou para menos. Um exemplo simples deste tipo de situação é o caso de uma balança descalibrada que pode indicar sempre massas maiores que as reais. Mas vale notar que os erros sistemáticos podem ser identificados e, portanto, evitados. Por outro lado, há um outro tipo de erro que afeta as medidas sem nenhuma tendência clara. As medidas podem oscilar, ora para mais, ora para menos. Este tipo de erro é chamado de erro aleatório e, infelizmente, sempre estará presente em maior ou menor grau.

Ao fazer um estudo é sempre interessante realizar replicatas, pois isto permite que o erro presente nas medidas seja investigado. Além disso, com a realização de várias replicatas aumentam as chances de se aproximar mais do valor exato. Isto é evidenciado por um importante princípio da estatística: o teorema do limite central, que comprova que o erro no valor médio é menor que o erro de uma observação individual<sup>1</sup>. Todavia, para garantir a confiabilidade da análise estatística as replicatas devem ser autênticas, ou seja, a ordem em que serão realizados os experimentos e as replicatas deve ser sorteada aleatoriamente. Não é correto simplesmente realizar duas medidas do mesmo experimento de forma seqüencial, pois um erro que afetar a primeira medida certamente irá, de forma sistemática, afetar a seguinte<sup>2,3</sup>. Todas as condições para executar o experimento devem ser refeitas, para garantir a autenticidade da replicata e a qualidade da análise estatística<sup>1-4</sup>. Contudo, muitas vezes as características do procedimento expe-

rimental dificultam muito a execução de replicatas autênticas<sup>5,6</sup>. Não são raros os casos em que o ajuste das condições experimentais é extremamente trabalhoso ou lento, como, por ex., nas situações em que é necessário aguardar por equilíbrios térmicos ou efeito de memória em análise química. Portanto, parece haver um dilema, realizar medidas com duplicatas autênticas e arcar com o custo do trabalho envolvido, mas garantir a qualidade das medidas, ou evitar um grande esforço no laboratório, mas correr o risco de ter todo seu trabalho prejudicado. É neste contexto que surge uma interessante possibilidade, a blocagem dos experimentos<sup>5,7</sup>.

## BLOCAGEM DE EXPERIMENTOS

Uma possibilidade de blocagem dos experimentos é a manutenção de uma ou mais condições experimentais fixas, enquanto outras são modificadas aleatoriamente. Não é incomum que o químico algumas vezes faça a blocagem de forma até “inconsciente”, contudo, bloquear os experimentos tem um custo. A análise estatística envolvida é diferente daquela que trata dados provenientes de experimentos com replicatas autênticas, tornando-se mais complexa<sup>5,8,9</sup>. Assim, usar a análise estatística convencional em situações de blocagem poderá levar a resultados incorretos.

Uma forma de bloquear os experimentos é através de um método conhecido como “split-plot” que é empregado em diferentes situações. Um caso típico para aplicação do método “split-plot” é a otimização conjunta que envolve variáveis de processo, como por ex., temperatura, e variáveis de mistura. Uma possibilidade seria escolher um valor para a temperatura e mantê-la fixa, depois diferentes composições de misturas seriam sorteadas aleatoriamente e preparadas. E, assim, na temperatura ajustada, alguma propriedade da mistura seria medida; em seguida, outra composição de mistura deveria ser sorteada e preparada e, ainda na mesma temperatura, medida a propriedade de interesse. Note que neste caso os experimentos são executados sempre em uma mesma temperatura, portanto, esta variável foi bloqueada, já as diferentes composições de mistura foram aleatoriamente sorteadas. A variável bloqueada no método “split-plot” é chamada de “main-plot”, e em determinada condição do “main-plot” (bloco), neste caso certa temperatura, são sorteados os níveis das outras variáveis, constituindo o chamado “sub-plot”, neste caso, as diferentes composições de mistura.

\*e-mail: bruns@iqm.unicamp.br

O fato dos experimentos não serem completamente aleatórios, mas haver a blocagem de uma variável e dentro de um bloco aleatoriamente serem sorteados os níveis das outras variáveis, levará a duas fontes de erro diferentes, uma para o “main-plot” e outra para o “sub-plot”. Isto implica em uma análise estatística mais complexa, específica para o método “split-plot”. Assim, claramente há vantagem no procedimento experimental mas, em contrapartida, a análise dos resultados será mais complexa<sup>5</sup>.

**GRÁFICOS DE PROBABILIDADE ACUMULADA**

Em virtude do erro experimental, ocorre uma dispersão natural nas medidas. Ao se modelar um sistema, provavelmente, alguns parâmetros do modelo estarão descrevendo somente as variações ocasionadas por este erro. Mas, obviamente, estes parâmetros não estão descrevendo efeitos significativos das variáveis de interesse e, portanto, devem ser descartados. Sabe-se que estatisticamente os parâmetros que descrevem erros aleatórios são exemplos de hipótese nula, uma vez que o valor verdadeiro de cada um deles seria zero. Portanto, estes parâmetros, se colocados em um gráfico em papel de probabilidade normal, devem seguir uma reta centrada em zero. Os efeitos significativos não se incluem na reta, pois não fazem parte da mesma distribuição<sup>1</sup>. A Figura 1 é um exemplo de gráfico de probabilidade acumulada, onde os parâmetros significativos são representados pelos círculos preenchidos e os não significativos por círculos vazios.

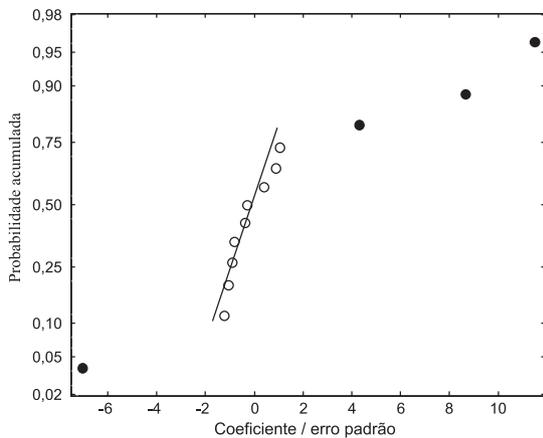


Figura 1. Gráfico de probabilidade acumulada para efeitos significativos (●) e não significativos (○)

Para se utilizar o gráfico de probabilidade acumulada deve-se tomar alguns cuidados. Primeiro, tentar plotar simplesmente os valores dos parâmetros no gráfico trará provavelmente resultados incorretos, pois muitos parâmetros que apresentam valores altos podem também possuir grandes erros associados, assim, sua significância talvez seja nula. Segundo, diferentes fontes de erro são agregadas ao cálculo dos erros dos parâmetros. Desta forma, a dimensão do erro pode variar muito entre os diferentes tipos de parâmetros (como os do “sub-plot” e “main-plot”). Devido às diferentes ordens de grandeza dos parâmetros, estes devem ser normalizados para serem comparados. Isto é feito dividindo-se os valores dos parâmetros pelos valores dos erros. Com isso, obtém-se a razão que indica o quanto um parâmetro supera seu erro<sup>9</sup>.

**Exemplo de um procedimento “split-plot”**

Um interessante trabalho foi publicado por Kowalski<sup>10</sup>, que utilizou o método “split-plot” em um planejamento fatorial 2<sup>4</sup> para com-

prender como quatro variáveis afetam a resistência de um plástico. As variáveis estudadas foram temperatura (T), porcentagem de aditivo (a), velocidade de agitação (v) e tempo de processamento (t). Para facilitar o procedimento experimental, a temperatura foi escolhida como “main-plot” e as outras três variáveis formaram o “sub-plot”. A temperatura foi fixada em dois níveis e aleatoriamente sorteu-se os níveis do “sub-plot”. A Figura 2 mostra a estratégia adotada. A Tabela 1 indica os resultados obtidos com os experimentos em duplicata.



Figura 2. Ilustração do planejamento fatorial pelo método “split-plot” para quatro variáveis, (T) temperatura, (a) porcentagem de aditivo, (v) velocidade de agitação e (t) tempo de processamento

Tabela 1. Planejamento fatorial pelo método “split-plot” para otimizar a resistência de um plástico

Temperatura	Aditivo	Velocidade	Tempo	Resposta
1	1	1	1	70,8
1	1	1	1	73,3
1	1	1	-1	66,2
1	1	1	-1	64,0
1	1	-1	1	66,8
1	1	-1	1	61,5
1	1	-1	-1	51,9
1	1	-1	-1	65,6
1	-1	1	1	68,5
1	-1	1	1	68,0
1	-1	1	-1	61,3
1	-1	1	-1	58,6
1	-1	-1	1	59,5
1	-1	-1	1	64,2
1	-1	-1	-1	58,5
1	-1	-1	-1	59,5
-1	1	1	1	63,9
-1	1	1	1	63,2
-1	1	1	-1	58,1
-1	1	1	-1	62,6
-1	1	-1	1	57,5
-1	1	-1	1	63,3
-1	1	-1	-1	57,4
-1	1	-1	-1	65,0
-1	-1	1	1	56,4
-1	-1	1	1	62,7
-1	-1	1	-1	56,5
-1	-1	1	-1	56,1
-1	-1	-1	1	53,2
-1	-1	-1	1	63,9
-1	-1	-1	-1	59,5
-1	-1	-1	-1	66,6

O procedimento adotado exige que a análise de variância seja específica, ou seja, a análise “split-plot”. Uma análise convencional, considerando experimentos executados de forma completamente aleatória, pode levar a conclusões erradas sobre os efeitos significativos nesta otimização. Para ilustrar estas situações, as duas análises de variância foram realizadas.

Inicialmente é necessário ajustar um modelo e, neste caso, o escolhido foi o bilinear,

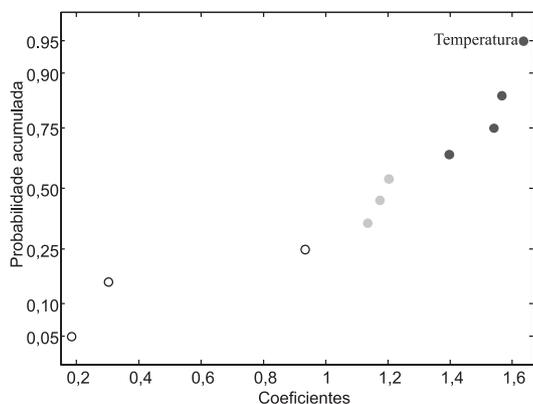
$$\hat{y} = \alpha + T + a + v + t + Ta + Tv + av + at + vt \quad (1)$$

para o qual  $\alpha$  representa a média global e  $T, a, v, t$  efeitos dos fatores identificados por seus símbolos apropriados. A primeira análise trata os dados como provenientes de experimentos completamente aleatórios e, portanto, com uma única fonte de erro. A Tabela 2 traz os coeficientes, o erro padrão e, em negrito, os parâmetros considerados significativos pelo teste  $t$ ; as equações empregadas para estes cálculos estão indicadas na referência 1 e são apresentadas no Apêndice A. Temperatura, tempo e interação temperatura-velocidade e temperatura-tempo são significativos a 95% de confiança, e com 90% inclui-se as variáveis aditivo, velocidade e interação velocidade-tempo. Com os coeficientes dos parâmetros foi construído um gráfico de probabilidade acumulada mostrado na Figura 3. Os parâmetros significativos a 95% estão preenchidos em preto, os preenchidos em cinza são significativos a 90% de confiança.

**Tabela 2.** Coeficientes, erro padrão e a razão coeficiente/erro dos parâmetros ajustados no modelo bilinear. Os parâmetros significativos a 95% estão em negrito

Efeito	GL <sup>a</sup>	Coeficientes	Erro padrão	Razão coef./erro
Intercepto	1	62,00	0,67	93,07
<b>T</b>	<b>1</b>	<b>1,63</b>	<b>0,67</b>	<b>2,45</b>
a	1	1,19	0,67	1,79
v	1	1,13	0,67	1,70
<b>t</b>	<b>1</b>	<b>1,54</b>	<b>0,67</b>	<b>2,31</b>
Ta	1	0,18	0,67	0,28
<b>Tv</b>	<b>1</b>	<b>1,57</b>	<b>0,67</b>	<b>2,35</b>
<b>Tt</b>	<b>1</b>	<b>1,40</b>	<b>0,67</b>	<b>2,10</b>
av	1	0,93	0,67	1,40
at	1	0,30	0,67	0,46
vt	1	1,17	0,67	1,76

<sup>a</sup> Graus de liberdade



**Figura 3.** Gráfico de probabilidade acumulada para os coeficientes da Tabela 2

Pela Figura 3 e Tabela 2 nota-se que a variável temperatura apresenta um efeito pronunciado, destacando-se entre os termos significativos.

Todavia, não se deve esquecer que os experimentos não foram realizados de forma completamente aleatória, mas com um procedimento “split-plot”. Portanto, não há apenas uma fonte de erro afetando os resultados. Na verdade, duas fontes de erro estão pre-

sentes: o erro “main-plot” e o erro “sub-plot”. Assim, a variável temperatura, que constitui o “main-plot”, é afetada por uma fonte de erro diferente daquela que afeta as variáveis do “sub-plot”. Então, considerando-se o tratamento adequado para o planejamento foi refeita a análise estatística.

Inicialmente foi calculada a ANOVA para os resultados dos experimentos, apresentada na Tabela 3. As equações empregadas são apresentadas no Apêndice B.

**Tabela 3.** ANOVA “split-plot” para os dados da Tabela 1

Fonte	SQ	GL	MQ
Replicatas	84,83	1	84,83
Main-plot	85,48	1	85,48
Erro main-plot	27,56	1	27,56
Sub-plot	244,63	7	34,95
Interação main-sub-plot	145,70	7	20,81
Erro sub-plot	174,81	14	12,70

A ANOVA da Tabela 3 permite separar as fontes de erro do “main-plot”, “sub-plot” e a soma quadrática devida às replicatas. Kowalski<sup>10</sup> sugere que a soma quadrática devida à replicata e ao erro “main-plot” podem ser somadas. Desta forma, o valor da média quadrática referente ao erro “main-plot” seria de 56,20 com 2 graus de liberdade. Pode-se, então, fazer um teste F entre o efeito do “main-plot” e o erro presente em suas medidas pela razão  $85,48/56,20 = 1,52$ , que não é significativo, comparado ao valor de  $F_{1;2;95\%} = 18,51$  ou  $F_{1;2;90\%} = 8,53$ . O que nos permite concluir que o efeito “main-plot” não é significativo. Por outro lado, autores como Cornell<sup>5</sup> adicionam ao erro “sub-plot” as possíveis faltas de ajuste do modelo escolhido como adequado obtendo mais graus de liberdade para tratar o erro, além de adicionar ao erro “sub-plot” o erro proveniente da interação “main-sub-plot” com o mesmo objetivo. Para exemplificar a maneira como as somas quadráticas são tratadas, um modelo com 31 termos foi gerado, o que esgota toda a falta de ajuste e toda variância é explicada pelos parâmetros ajustados. A Tabela 4 traz os 31 parâmetros do modelo e a soma quadrática explicada por cada um na regressão. O valor da soma quadrática devida à replicata ( $r$ ) é de 84,83 e a interação da replicata com o “main-plot”, erro “main-plot” ( $rT$ ), é de 27,56, plenamente concordantes com os valores da Tabela 3. O valor do efeito “main-plot” ( $T$ ) é igual a 85,48 e o efeito do “sub-plot” pode ser obtido pela somatória das somas quadráticas dos termos  $a, v, av, t, at, vt$  e  $avt$ , com 7 graus de liberdade e equivalendo a 244,64. Para determinar o erro “sub-plot”, basta somar os termos de interação entre ( $a,v,t$ ) e replicata ( $r$ ),  $ra + rv + rav + rt + rat + rvt + ravt = 117,30$ ; já o erro da interação “main-sub-plot” pode ser determinado pelas interações de  $r$  e  $t$  com  $a, v$  e  $t$ , cuja soma vale 57,51. Os valores do erro “sub-plot” e “main-sub-plot” são somados no método “split-plot”, o que equivale a 174,81 com 14 graus de liberdade.

Para comparar a ANOVA “split-plot” com a que considera todos os experimentos completamente aleatórios ajustou-se o mesmo modelo da Equação (1), que leva em conta apenas interações binárias entre os efeitos. Assim, os termos de interação superior não participam do modelo, sendo equivalentes à falta de ajuste e, portanto, adicionados ao erro “sub-plot”: erro “sub-plot” + soma quadrática de termos não ajustados =  $174,81 + 11,06 = 185,87$ , contendo 19 graus de liberdade. A média quadrática do erro será, portanto, igual a 9,78. Com as somas quadráticas dos efeitos pode-se realizar um teste F para determinar os parâmetros significativos a 95%,  $F_{1;19;95\%} = 4,38$ , e 90%,  $F_{1;19;90\%} = 2,99$ . A Tabela 5 indica as razões entre as médias quadráticas obtidas por regressão dos parâmetros e a média quadrática do erro. Comparando os valores da Tabela 2 e Tabela 5 percebem-se algumas

alterações. Primeiro, a temperatura que inicialmente era um fator muito significativo, na segunda análise não o é. Segundo, o aditivo que foi considerado significativo na primeira análise com apenas 90% de confiança passou a ser significativo com 95%. Terceiro, a interação velocidade-tempo teve um efeito significativo apenas com 90% de confiança na análise considerando os experimentos completamente aleatórios, mas na análise “split-plot” o efeito desta interação é significativo já em 95% de confiança. A Figura 4 mostra o gráfico de probabilidade acumulada para as razões dos coeficientes/erro padrão para o modelo da Equação 1 pelo método “split-plot”. Os valores dos coeficientes e seus erros são indicados na Tabela 6, as equações para obtenção dos valores presentes nesta tabela estão presentes na referência 5 e foram incluídas no Apêndice C.

**Tabela 4.** Modelo com 31 parâmetros ajustado aos dados da Tabela 1

Fonte	GL	SQ <sup>a</sup>	Fonte	GL	SQ
R	1	84,83	rt	1	0,81
T	1	85,48	Tt	1	62,44
RT	1	27,57	rTt	1	4,28
A	1	45,36	at	1	2,94
Ra	1	0,002	rat	1	43,95
Ta	1	1,09	Tat	1	0,75
RTa	1	5,04	rTat	1	0,26
V	1	41,18	vt	1	43,95
Rv	1	46,32	rvt	1	14,99
Tv	1	78,44	Tvt	1	2,82
RTv	1	0,63	rTvt	1	15,82
Av	1	27,94	avt	1	7,32
Rav	1	0,30	ravt	1	10,93
Tav	1	0,17	Tavt	1	0,003
RTav	1	0,07	rTavt	1	31,40
T	1	75,95			

<sup>a</sup> a soma quadrática, SQ, para um efeito pode ser obtida por: (efeito)<sup>2</sup>.(32), onde 32 é o número de experimentos.

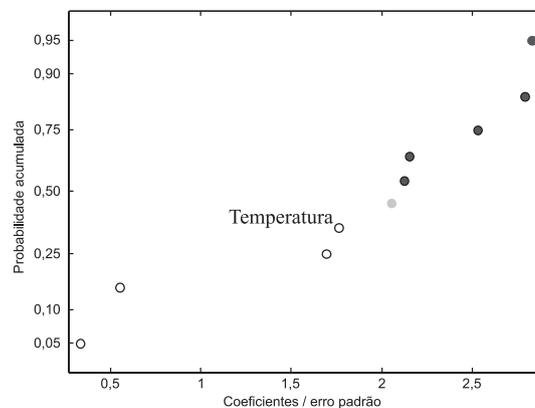
**Tabela 5.** Somas quadráticas referentes ao modelo da Equação 1 para o método “split-plot”

Efeito	SQ	GL	MQ	MQ <sub>erro</sub>	Razão MQ/MQ <sub>erro</sub>
T	85,48	1	85,48	56,20	1,52
A	45,36	1	45,36	9,78	4,63 *
V	41,18	1	41,18	9,78	4,21 <sup>∇</sup>
T	75,95	1	75,95	9,78	7,77 *
Ta	1,08	1	1,08	9,78	0,11
Tv	78,44	1	78,44	9,78	8,02 *
Tt	62,44	1	62,44	9,78	6,38 *
Av	27,94	1	27,94	9,78	2,86
At	2,94	1	2,94	9,78	0,30
Vt	43,95	1	43,95	9,78	4,49 *

\* parâmetro significativo a 95%; <sup>∇</sup> parâmetro significativo a 90%

**Tabela 6.** Coeficientes e erros para o modelo “split-plot”

Efeito	Coeficiente	Erro	Razão coef./erro
T	1,63	0,93	1,76
a	1,19	0,55	2,15
v	1,13	0,55	2,05
t	1,54	0,55	2,79
Ta	0,18	0,55	0,33
Tv	1,57	0,55	2,83
Tt	1,40	0,55	2,53
av	0,93	0,55	1,69
at	0,30	0,55	0,55
vt	1,17	0,55	2,12



**Figura 4.** Gráfico de probabilidade acumulada para os valores das razões dos coeficientes/erro padrão da Tabela 6; os círculos preenchidos em preto são significativos a 95% e os em cinza, a 90%

Comparando a Figura 3 com a Figura 4 nota-se claramente que a variável temperatura, na análise “split-plot”, após ser corrigida por seu erro, deixa de ser significativa como antes, uma vez que na análise correta o verdadeiro valor de seu erro é muito maior que na análise convencional.

## CONCLUSÕES

Realizar os experimentos em ordem aleatória é de extrema importância para que a análise estatística convencional seja válida; todavia, muitas vezes limitações de ordem experimental inviabilizam a realização de replicatas autênticas. Desta forma, uma alternativa é realizar os experimentos com métodos alternativos, como o “split-plot”, que facilitam o trabalho no laboratório, mas exigem uma análise estatística mais complexa. É importante salientar que ao planejar um conjunto de experimentos o químico também deverá se preocupar em como tratar seus resultados, pois realizar uma análise incorreta implicará, muitas vezes, na determinação de erros não verdadeiros e pode levar a conclusões injustificadas, prejudicando todo o trabalho.

## AGRADECIMENTOS

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela bolsa de estudo concedida a J. A. Bortoloti (Processo 00/14261-0) e pelos recursos fornecidos para a atividade de pesquisa (Processo 01/14371-3).

## REFERÊNCIAS

- Barros Neto, B.; Scarminio, I. S.; Bruns, R. E.; *Como fazer experimentos: pesquisa e desenvolvimento na ciência e na indústria*, Ed. da Unicamp: Campinas, 2001.
- Box, G. E. P.; Hunter W.G.; Hunter, J. S.; *Statistics for experimenters. An introduction to design, data analysis and model building*, Wiley: New York, 1978.
- Anderson, R. L.; Bancroft, T. A.; *Statistical Theory in Research*, McGraw-Hill: New York, 1952.
- Wooding, W. M.; *J. Qual. Tech.* **1973**, 5, 16.
- Cornell, J. A.; *J. Qual. Tech.* **1988**, 20, 2.
- Cornell, J. A.; *J. Am. Stat. Assoc.* **1971**, 66, 42.
- Hicks, C. R.; *Fundamental Concepts in the Design of Experiments*, 2<sup>nd</sup> ed., Holt, Rinehart and Winston: New York, 1973.
- Bortoloti, J. A.; Andrade, J. C.; Bruns, R. E.; *J. Braz. Chem. Soc.* **2004**, 15, 241.
- Bortoloti, J. A.; Bruns, R. E.; Andrade, J. C.; Vieira, R. K.; *Chem. Intel. Lab. System* **2004**, 70, 113.
- Kowalski, S. M.; Potcner, K. J.; *Qual. Progress* **2003**, 36, 60.

### APÊNDICE A - Equações empregadas para a obtenção dos valores presentes na Tabela 2

O vetor  $\mathbf{b}$  que contém os parâmetros do modelo obtido por regressão é dado por

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

onde  $\mathbf{X}$  é a matriz planejamento e  $\mathbf{y}$  o vetor resposta. Os erros dos parâmetros são obtidos extraindo-se a raiz quadrada dos elementos da diagonal principal da matriz obtida pela expressão

$$\mathbf{V}(\mathbf{b}) = (\mathbf{X}'\mathbf{X})^{-1} \sigma^2$$

em que  $\sigma^2$  é a variância do erro aleatório.

### APÊNDICE B - Equações empregadas para a obtenção dos valores presentes na Tabela 3

As somas quadráticas (SQ) são obtidas pelas equações

$$SQ_{\text{replicata}} = [(\bar{y}_{1..} - \bar{y}_{...})^2 + (\bar{y}_{2..} - \bar{y}_{...})^2 + \dots + (\bar{y}_{r..} - \bar{y}_{...})^2](m.p),$$

$$SQ_{\text{main-plot}} = [(\bar{y}_{.1.} - \bar{y}_{...})^2 + (\bar{y}_{.2.} - \bar{y}_{...})^2 + \dots + (\bar{y}_{.p.} - \bar{y}_{...})^2](r.m),$$

$$SQ_{\text{sub-plot}} = [(\bar{y}_{..1} - \bar{y}_{...})^2 + (\bar{y}_{..2} - \bar{y}_{...})^2 + \dots + (\bar{y}_{..m} - \bar{y}_{...})^2](r.p),$$

$$SQ_{\text{erro main-plot}} = [ \sum_i^r \sum_j^p (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 ] \cdot m,$$

$$SQ_{\text{interação main-sub-plot}} = \sum_j^p \sum_h^m (\bar{y}_{.jh} - \bar{y}_{.j.} - \bar{y}_{..h} + \bar{y}_{...})^2 \cdot r,$$

$$SQ_{\text{total}} = \sum_i^r \sum_j^p \sum_k^m (v_{ijk} - \bar{y}_{...})^2,$$

$$SQ_{\text{sub-plot error}} = SQ_{\text{total}} - (SQ_{\text{replicata}} + SQ_{\text{main-plot}} + SQ_{\text{sub-plot}} + SQ_{\text{erro main-plot}} + SQ_{\text{interação main-sub-plot}})$$

onde:  $i = 1, 2, 3, \dots, r$  (replicatas);  $j = 1, 2, 3, \dots, p$  (condições de processo);  $k = 1, 2, 3, \dots, m$  (misturas);  $\bar{y}_{...}$  = média global das respostas;

$\bar{y}_{ij.}$  = média da  $i$ -ésima replicata de todas as composições em uma  $j$ -ésima condição de processo;  $\bar{y}_{i..}$  = média da  $i$ -ésima replicata de todas as composições em todas as condições de processo;  $\bar{y}_{.j.}$  = média de todas as composições e replicatas na  $j$ -ésima condição de processo;  $\bar{y}_{.jk.}$  = média das replicatas na  $j$ -ésima condição de processo e  $k$ -ésima composição de mistura;  $\bar{y}_{..j.}$  = média de todas as composições e replicatas na  $j$ -ésima condição de processo;  $\bar{y}_{..k.}$  = média de todos os tratamentos e replicatas na  $k$ -ésima composição de mistura;  $n$  = número total de respostas.

### APÊNDICE C - Equações empregadas para a obtenção dos valores presentes na Tabela 6

O vetor  $\mathbf{b}$  contendo os coeficientes dos parâmetros do modelo ajustado é obtido por

$$\mathbf{b} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

onde  $\mathbf{V}$  é dada por

$$\mathbf{V} = \{ \mathbf{J}_n \otimes \mathbf{I}_r \} \sigma_R^2 + \{ \mathbf{J}_m \otimes \mathbf{I}_p \} \sigma_{RZ}^2 + \mathbf{I}_{mpr} \sigma_e^2$$

sendo que  $\mathbf{J}$  é uma matriz com blocos diagonalizados com valores unitários para todos os elementos das diagonais dos blocos e valores nulos para os elementos restantes.  $\mathbf{I}$  são matrizes identidades e  $n$ ,  $r$ ,  $m$  e  $p$  são os números de experimentos, replicatas, misturas e condições de processo, respectivamente. Os termos  $\sigma_R^2$ ,  $\sigma_{RZ}^2$  e  $\sigma_e^2$  são estimativas das variâncias dos erros das replicatas, do "main-plot" e "sub-plot", respectivamente, sendo calculados a partir dos resultados da ANOVA para planejamentos contendo replicatas.

A matriz de covariância de  $\mathbf{b}$  para determinação dos erros associados aos parâmetros do modelo, para casos com mais de uma fonte de erro, é dada por

$$\text{Cov}(\hat{\mathbf{b}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$