

AJUSTE DE UMA SUPERFÍCIE DE RESPOSTA NO DELINEAMENTO EM BLOCOS INCOMPLETOS: ANÁLISE DE VEROSSIMILHANÇA RESTRITA E BAYESIANA EM UMA SIMULAÇÃO DE ADUBAÇÃO EM CITROS

Claudiney Nunes de LIMA¹
Júlio Sílvio de Sousa BUENO FILHO¹

- RESUMO: A determinação da combinação de doses de nutrientes que resulta em produção máxima é um problema de interesse em ensaios de adubação. No presente trabalho, procurou-se ilustrar as análises de verossimilhança restrita e bayesiana de um delineamento em blocos incompletos parcialmente balanceados. Para essa ilustração, foram simulados dados de produção de uma cultura de citros, no primeiro ano de idade, com a combinação de doses de dois nutrientes. Para o ajuste do modelo e o cálculo de suas propriedades estatísticas, foi feita a inferência de verossimilhança restrita (REML) e a inferência bayesiana, usando a amostragem Gibbs, aplicadas a um modelo linear misto. Os resultados demonstram que as duas metodologias são boas ferramentas de ajuste de modelos mistos e devem ser recomendadas para ajustes de superfícies de resposta.
- PALAVRAS-CHAVE: Adubação; amostrador de Gibbs; delineamento em blocos incompletos; REML.

1 Introdução

Muitos projetos de pesquisa baseiam-se em experimentos que pretendem avaliar dois ou mais fatores de níveis quantitativos e estudar suas interações. A questão central neste caso é a de detectar a presença de níveis ótimos de “ação”, que são os níveis dos fatores que conferem pontos de máximo ou mínimo à superfície de resposta das variáveis de interesse. Embora haja exemplos reais desse tipo de experimento, pretende-se apresentar uma proposta de plano experimental e duas propostas de análise para um exemplo-modelo relativamente complexo, com dados fictícios que permitam a comparação do ajuste aos valores paramétricos simulados.

¹Departamento de Ciências Exatas

A metodologia da superfície de resposta (MSR) é um conjunto de técnicas matemáticas e estatísticas úteis na modelagem e na análise em aplicações em que se procura detectar relação funcional entre a resposta de interesse (variável dependente) e diversos fatores de estudo (variáveis independentes). O objetivo é otimizar a resposta (Box e Drapper, 1987). A abordagem geral MSR foi desenvolvida no início da década de 50, tendo sido aplicada inicialmente na indústria química com considerável sucesso, sendo que, nos últimos 20 anos, encontra-se extensa aplicação em uma ampla variedade de contextos (Montgomery, 2004).

O problema escolhido, inspirado em Ribeiro *et al.* (1999), foi o da adubação para plantas de citros no primeiro ano de idade. Supõe-se que, no plantio, tenha sido realizada a correção de fertilidade com calcário calcítico e foi detectada a necessidade de suprimento de magnésio. Para esse caso, foi proposta a adubação foliar com sulfato de magnésio. Além disso, a cultura recebe cobertura com adubo nitrogenado. A demanda do pesquisador seria estimar os níveis ótimos de adubação nitrogenada e adubação foliar que maximizam a produtividade (medida em caixas de frutos por hectare). Um problema adicional desse plano experimental é o de que o pesquisador queria comparar o efeito dos níveis médios de tratamentos com duas referências, quais sejam: ausência de qualquer adubação e efeito da água na adubação foliar. Tais testemunhas refletem o nível zero para ambos os fatores quantitativos em estudo.

O experimento fatorial para o ajuste da superfície de resposta para a produtividade envolveu muitos tratamentos e foi instalado em blocos incompletos para o adequado controle local. Isto, sabidamente, aumenta a eficiência do experimento, mas, por vezes, complica a análise. Situações desse tipo ocorrem devido à dificuldade de se encontrar áreas experimentais que sejam uniformes e que permitam blocos homogêneos contendo todos os tratamentos. Os delineamentos em blocos casualizados completos com heterogeneidade intrablocos serão pouco eficientes, pois o erro experimental tende a ser inflado, dificultando a discriminação dos tratamentos e diminuindo a precisão das estimativas de interesse (Mead, 1988).

Em alguns casos, é conveniente adicionar, em cada bloco do delineamento, um ou mais dos tratamentos, gerando delineamentos aumentados. Por exemplo, quando se deseja comparar algumas variedades novas com outras de comportamento já conhecido e que atuam como testemunhas. Isto permite que a testemunha seja mais repetida e tenha um erro padrão da média menor (Federer, 1975).

O problema estatístico enfocado é o de ajustar um delineamento em blocos incompletos, com recuperação da informação interblocos. Isto implica em modelos com efeitos de blocos tomados como aleatórios (modelo misto) e diversas (três) classes de associados.

O objetivo deste trabalho é ilustrar a comparação entre duas abordagens inferenciais para o mesmo problema: a Inferência Clássica, usando o método da Máxima Verossimilhança Restrita (REML) (Patterson e Thompson, 1971) e a Inferência Bayesiana usando o Amostrador de Gibbs (GS, do inglês *Gibbs Sampler*, Gelfand e Smith, 1990). As médias ajustadas de tratamentos foram a seguir modeladas por superfícies de resposta quadráticas, sendo calculadas as respectivas

regiões de confiança para seus pontos estacionários.

2 Metodologia

2.1 Material

Neste trabalho, foram simulados dados de um experimento em blocos incompletos, parcialmente balanceados, sendo constituído por 16 tratamentos regulares (nitrogênio e adubação foliar) e 2 comuns (testemunhas). A variável resposta é a produção dos citros (expressa em caixas/ha). O conjunto de dados simulado está apresentado na Tabela 1.

O modelo para a superfície de resposta foi:

$$S_{ijl} = 100 + (1 - \theta)(0,8N_i - 0,005N_i^2 + 0,5A_j - 0,001A_j^2 + 0,0005N_iA_j) + \theta\tau_l, \quad (1)$$

em que:

S_{ijl} são os valores simulados da variável resposta produção observada no nível i de N e j de A ; N_i é o valor da dose i de nitrogênio, sendo atribuídos os seguintes

Tabela 1 - Dados simulados para o experimento de adubação em *Citrus*

Rep.	Bloco	Trat.	Test.	Dose NO_3	Dose $MgSO_4$	Produção
1	1	1	0	30	90	158,04
1	1	6	0	60	180	190,59
1	1	11	0	90	270	204,02
1	1	16	0	120	360	203,16
1	1	17	1	0	0	93,40
1	2	2	0	30	180	180,09
1	2	7	0	60	270	201,75
1	2	12	0	90	360	201,26
1	2	13	0	120	90	163,71
1	2	18	1	0	0	101,21
1	3	3	0	30	270	181,39
1	3	8	0	60	360	185,82
1	3	9	0	90	90	173,18
1	3	14	0	120	180	190,48
1	3	18	1	0	0	102,72
1	4	4	0	30	360	173,02
1	4	5	0	60	90	170,32
1	4	10	0	90	180	194,93
1	4	15	0	120	270	198,21
1	4	17	1	0	0	99,49
2	5	4	0	30	360	161,66
2	5	7	0	60	270	195,27
2	5	10	0	90	180	196,18
2	5	13	0	120	90	165,29
2	5	18	1	0	0	97,28
2	6	3	0	30	270	188,93
2	6	6	0	60	180	201,47
2	6	9	0	90	90	175,00
2	6	16	0	120	360	196,88
2	6	17	1	0	0	98,51
2	7	2	0	30	180	181,95
2	7	5	0	60	90	177,94
2	7	12	0	90	360	188,57
2	7	15	0	120	270	208,48
2	7	17	1	0	0	94,53
2	8	1	0	30	90	161,50
2	8	8	0	60	360	192,10
2	8	11	0	90	270	197,15
2	8	14	0	120	180	198,00
2	8	18	1	0	0	105,87

...continua...

Tabela 1: Continuação.

Rep.	Bloco	Trat.	Test.	Dose NO_3	Dose $MgSO_4$	Produção
3	9	1	0	30	90	153,01
3	9	7	0	60	270	204,67
3	9	10	0	90	180	208,34
3	9	16	0	120	360	197,33
3	9	17	1	0	0	97,17
3	10	3	0	30	270	183,38
3	10	6	0	60	180	183,89
3	10	12	0	90	360	199,81
3	10	13	0	120	90	157,38
3	10	17	1	0	0	99,39
3	11	2	0	30	180	180,15
3	11	8	0	60	360	192,16
3	11	9	0	90	90	174,54
3	11	15	0	120	270	198,89
3	11	18	1	0	0	102,36
3	12	4	0	30	360	182,23
3	12	5	0	60	90	172,67
3	12	11	0	90	270	211,61
3	12	14	0	120	180	191,21
3	12	18	1	0	0	100,97

O delineamento resultante ($v^* = 18$ tratamentos) está apresentado na Tabela 2 e é a composição de um PBIB resolvido, otimizado para um fatorial 4×4 , triplo, aumentado de 1 testemunha por bloco, com $v = 16$ tratamentos em $r = 3$ superblocos (repetições), 4 blocos por superblocos, num total de $b = 12$ blocos e $k = 4 + 1 = 5$ unidades experimentais por bloco (pois a cada bloco do delineamento otimizado foi adicionada uma das testemunhas)¹. O fatorial foi otimizado para blocos incompletos conforme Bittencourt et. al. (2003), usando o aplicativo "DEW" (Edmondson, 2005).

2.3 Modelo da análise e suas pressuposições

A teoria de modelos mistos, desenvolvida em 1949 por Henderson, baseia-se em modelos lineares que contêm efeitos fixos e aleatórios. Assim, o modelo descrito pela equação 2 pode ser apresentado matricialmente por:

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{b} + \boldsymbol{\varepsilon}, \quad (3)$$

¹O processo de simulação descrito foi repetido 3 vezes, originando os superblocos considerados na análise. Desta forma, não há "efeitos" paramétricos de superblocos ou repetições, mas supõe-se que o analista de dados monte o modelo com superblocos apenas levando em conta a aleatorização

em que:

- \mathbf{y} é um vetor de dimensões (nx1) das observações, n é o número de observações;
- $\boldsymbol{\beta}$ é o vetor de dimensões (px1) dos parâmetros de efeitos fixos, p é o número de parâmetros associados aos efeitos fixos. Nesse caso específico, $\boldsymbol{\beta} = (\beta_0, r_2, r_3, t_2, \dots, t_{18})$, em que β_0 é uma constante experimental associada ao intercepto ($\mu + r_1 + t_1$) para o modelo superparametrizado, em que μ é a média geral, r_1 é o efeito da repetição 1 e t_1 é o efeito do tratamento 1;
- \mathbf{b} é o vetor de dimensões (qx1) dos parâmetros de efeitos aleatórios, sendo q é número de parâmetros associados aos efeitos aleatórios; $\mathbf{b} \sim N(\phi, I\sigma_b^2)$; em que ϕ representa um vetor nulo de dimensão igual à do vetor considerado.
- $\boldsymbol{\varepsilon}$ é o vetor de resíduos (nx1); $\boldsymbol{\varepsilon} \sim N(\phi, I\sigma_\varepsilon^2)$; em que ϕ representa um vetor nulo de dimensão igual à do vetor considerado.
- X é a matriz (nxp) de incidência dos efeitos fixos;
- Z é a matriz (nxq) de incidência dos efeitos aleatórios.

Tabela 2 - Plano experimental com o fatorial otimizado para o delineamento em blocos e a testemunha adicionada a cada bloco

Blocos	Trat. Regulares	Test	Rep
1	1,6,11,16	17	
2	2,7,12,13	18	I
3	3,8,9,14	18	
4	4,5,10,15	17	
5	4,7,10,13	18	
6	3,6,9,16	17	II
7	2,5,12,15	17	
8	1,8,11,14	18	
9	1,7,10,16	17	
10	3,6,12,13	17	III
11	2,8,9,15	18	
12	4,5,11,14	18	

2.4 Estimação REML

A equação matricial para a análise do modelo misto é definida pela equação (4) e o recurso computacional adotado para se estimarem as componentes de variância foi a função *lme* da biblioteca *nlme* do pacote estatístico $R^{\text{®}}$. Como padrão esta função executa 25 passos do algoritmo EM e ao menos 1 passo de Newton Raphson (Pinheiro e Bates, 2000). Alguns passos do algoritmo EM (Dempster e Laird, 1977)

úteis na determinação das matrizes de variância e covariância das estimativas serão apresentados, apenas por questões didáticas e de continuidade com outros elementos do texto.

Tomando-se valores iniciais para a estimativa da componente de variância σ_b^2 , relativa a \mathbf{b} e um valor inicial para a variância do erro σ_e^2 , tem-se que as soluçõ

1

Td (n)T 4.94115

Assim, a função de verossimilhança, pressupondo independência, é dada por:

$$L(\mathbf{y}|\boldsymbol{\beta}, \mathbf{b}, \sigma_e^2, \sigma_b^2) = \left(\frac{1}{2\pi\sigma_e^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma_e^2}(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})'(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})\right\}. \quad (6)$$

O passo seguinte corresponde à definição das distribuições *a priori* para os diferentes parâmetros do modelo.

2.5.1 Especificação das distribuições *a priori*

É comum encontrar-se, na literatura, a suposição de que não haja informação prévia a respeito dos efeitos fixos $\boldsymbol{\beta}$. Assim, pode-se definir uma *priori* não informativa por meio de uma distribuição constante ou uniforme, o que é uma maneira de caracterizar total desconhecimento da distribuição de $\boldsymbol{\beta}$, segundo Sorensen (1996). Ou seja:

$$p(\boldsymbol{\beta}) \propto K,$$

em que K é uma constante.

Em geral, para manter a consistência, ao se supor que os efeitos fixos tenham distribuição tão vaga, é razoável supor, também, que as componentes da variância terão distribuição correspondentemente pouco informativas. Assim, neste estudo, foram assumidas distribuições *a priori* Qui-Quadrado Inversa Escaladas para as componentes da variância. Os valores dos hiperparâmetros para σ_e^2 e σ_b^2 foram escolhidos conforme se apresenta a seguir:

$$\sigma_e^2 \sim Inv - \chi^2(2, 5);$$

$$\sigma_b^2 \sim Inv - \chi^2(2, 5).$$

Para essas distribuições *a priori* de componentes da variância, a correlação intraclasse $\frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}$ fica sendo relativamente pouco informativa, conforme se apresenta na Figura 1.

Adotou-se esta abordagem pouco informativa para fins de comparação com o método REML, pois de acordo com Bernardo e Smith (1994), esta é uma maneira conveniente de comparar a análise bayesiana com a de verossimilhança correspondente.

Finalmente, para o vetor dos efeitos aleatório de blocos, empregou-se a distribuição normal, como se segue:

$$\mathbf{b} \sim N(\boldsymbol{\phi}, A\sigma_b^2),$$

em que:

$\boldsymbol{\phi}$ é o vetor nulo de mesma dimensão do aleatório;

A é a matriz que representa os coeficientes das correlações entre os efeitos aleatórios. Nesse caso, será suposta como sendo a matriz identidade (I).

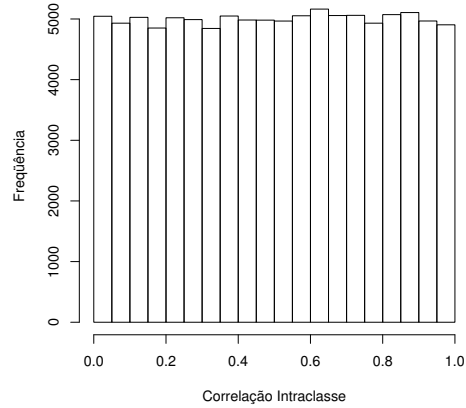


Figura 1 - Histograma de uma amostra de 100000 valores da distribuição *a priori* da correlação intraclass. d 1. Tf . da 2 correlação intraclass. d 1 BT 1 . 2 Tf 1 0 0 1 1 . 1 2 (t)T 0 T d d 1. 1 0012227 1 hm.0 d 0 0 227 1 hl BT 1 . 2 Tf 1 0 0 1 1 . 1 2 (t)T 0 T

2.5.2 Distribuição *a posteriori* conjunta. do ao c c t j b u i a o a o r i
 da

s)Tj 3.6936 0 T10.2b

) Td (t)Tj . 21 0 Td (r)Tj . 2 0 Td (i)Tj 2. 01 0 Td (b)Tj 7. 0 72070 Td (c)Tj .1 0 Td (i)Tj . 0 Td (,)Tj . 272070 Td
 0 Td (r)Tj . 0 7d (j .7 2 0 Td(t)Tj . 21 0 T Td (,)Tj 0.00 770 Td (u)Tj . 27 0 Td (d)Tj .20 7 0 . Tj . 27 0 Td (

2.5.3 Distribuições condicionais completas

A distribuição condicional completa *a posteriori* para os efeitos fixos (média e tratamentos), segundo a equação (7), é dada por:

$$\begin{aligned}
 p(\boldsymbol{\beta} \mid \mathbf{b}, \sigma_b^2, \sigma_e^2, \mathbf{y}) &\propto \exp\left\{-\frac{1}{2\sigma_e^2}(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})'(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})\right\} \\
 &\propto \exp\left\{-\frac{1}{2\sigma_e^2}\left[(\mathbf{y} - Z\mathbf{b})' - (X\boldsymbol{\beta})'\right]\left[(\mathbf{y} - Z\mathbf{b}) - (X\boldsymbol{\beta})\right]\right\} \\
 &\propto \exp\left\{-\frac{1}{2\sigma_e^2}\left[(\mathbf{y} - Z\mathbf{b})'(\mathbf{y} - Z\mathbf{b}) - (\mathbf{y} - Z\mathbf{b})'(X\boldsymbol{\beta}) - (X\boldsymbol{\beta})'(\mathbf{y} - Z\mathbf{b}) + (X\boldsymbol{\beta})'(X\boldsymbol{\beta})\right]\right\} \\
 &\propto \exp\left\{-\frac{1}{2\sigma_e^2}\left[-(\mathbf{y} - Z\mathbf{b})'X\boldsymbol{\beta} - \boldsymbol{\beta}'X'(\mathbf{y} - Z\mathbf{b}) + \boldsymbol{\beta}'X'(X\boldsymbol{\beta})\right]\right\} \\
 &\propto \exp\left\{-\frac{1}{2\sigma_e^2}\left[\boldsymbol{\beta} - (X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b})\right]'(X'X)\left[\boldsymbol{\beta} - (X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b})\right]\right\} \\
 &\quad \exp\left\{-\frac{1}{2\sigma_e^2}\left[-(\mathbf{y} - Z\mathbf{b})'X(X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b})\right]\right\} \\
 &\propto \exp\left\{-\frac{1}{2\sigma_e^2}\left[\boldsymbol{\beta} - (X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b})\right]'(X'X)\left[\boldsymbol{\beta} - (X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b})\right]\right\}. \quad (8)
 \end{aligned}$$

Portanto:

$$\boldsymbol{\beta} \mid (\mathbf{b}, \sigma_e^2, \sigma_b^2, \mathbf{y}) \sim N\left((X'X)^{-1}X'(\mathbf{y} - Z\mathbf{b}), \sigma_e^2(X'X)^{-1}\right). \quad (9)$$

A distribuição *a posteriori* condicional completa para os efeitos aleatórios (blocos) pode ser encontrada de maneira análoga à distribuição condicional completa *a posteriori* dos efeitos fixos, ou seja, dada a distribuição *a posteriori* conjunta (7), tem-se:

$$p(\mathbf{b} \mid \boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \mathbf{y}) \propto \exp\left\{-\frac{1}{2\sigma_e^2}\left(\frac{\mathbf{b}'A^{-1}\sigma_e^2\mathbf{b}}{\sigma_b^2} + (\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})'(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})\right)\right\}. \quad (10)$$

Sendo $\gamma = \sigma_e^2/\sigma_b^2$, logo:

$$\mathbf{b} \mid (\boldsymbol{\beta}, \sigma_e^2, \sigma_b^2, \mathbf{y}) \sim N\left((Z'Z + \gamma A^{-1})^{-1}Z'(\mathbf{y} - X\boldsymbol{\beta}), (Z'Z + \gamma A^{-1})^{-1}\sigma_e^2\right). \quad (11)$$

A distribuição condicional completa para a componente da variância σ_e^2 , em (7), é dada a seguir:

$$p(\sigma_e^2 \mid \boldsymbol{\beta}, \sigma_b^2, \mathbf{b}, \mathbf{y}) \propto (\sigma_e^2)^{-(\frac{n+\nu}{2}+1)} \exp\left\{-\frac{1}{2\sigma_e^2}\left[(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})'(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b}) + \nu s^2\right]\frac{(n+\nu)}{(n+\nu)}\right\}, \quad (12)$$

ou seja, uma Qui-Quadrado Inversa Escalada.

$$\sigma_e^2 \mid (\boldsymbol{\beta}, \sigma_b^2, \mathbf{b}, \mathbf{y}) \sim Inv - \chi^2\left(n + \nu, \frac{(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})'(\mathbf{y} - X\boldsymbol{\beta} - Z\mathbf{b})}{n + \nu} + \frac{\nu s^2}{n + \nu}\right). \quad (13)$$

Da mesma forma, para a componente da variância σ_b^2 , pode-se obter sua distribuição condicional completa por meio de (7), como se segue:

$$p(\sigma_b^2 | \beta, \sigma_e^2, \mathbf{b}, \mathbf{y}) \propto (\sigma_b^2)^{-\left(\frac{J+\nu_b}{2}+1\right)} \exp \left\{ -\frac{1}{2\sigma_b^2} \left(\frac{\mathbf{b}' A^{-1} \mathbf{b} + \nu_b s_b^2}{J + \nu_b} \right) (J + \nu_b) \right\}, \quad (14)$$

que é também uma Qui-Quadrado Inversa Escalada:

$$\sigma_b^2 | (\beta, \sigma_e^2, \mathbf{b}, \mathbf{y}) \sim Inv - \chi^2 \left(J + \nu_b, \frac{\mathbf{b}' A^{-1} \mathbf{b} + \nu_b s_b^2}{J + \nu_b} \right). \quad (15)$$

2.6 Avaliação da convergência do amostrador de Gibbs

No processo de verificação de convergência, utilizou-se a biblioteca BOA (Bayesian Output Analysis) do programa R[®] v.2.1.1(R, 2005).

Inicialmente, foi gerada uma cadeia de 4000 iterações para todos os parâmetros, com o objetivo de obter uma referência de tamanhos de “queima” das primeiras observações e “pulo” entre as observações, calculados pelo critério de Raftery e Lewis (1992), o qual indicou, dentre todos os parâmetros, os maiores valores de 2 e 1 para queima e pulo, respectivamente.

Para garantir a segurança do processo amostral, a partir de duas cadeias de 150000 iterações em paralelo, foram obtidas duas amostras de tamanho igual a 4000, de cada uma dessas cadeias, levando-se em conta o descarte das 50000 primeiras iterações, para reduzir a influência dos valores iniciais. Além disso, numa frequência de 25, ou seja, a cada 25 observações, retirava-se um valor de cada parâmetro para tentar diminuir a autocorrelação existente.

A verificação da convergência das duas cadeias resultantes, para cada parâmetro, foi feita por meio do critério de convergência de Gelman e Rubin (1992a), observando-se se os valores do fator de redução de escala potencial (\hat{R}) estavam próximos ao valor unitário. Gelman (1995) sugere aceitar, como garantia de convergência, valores de \hat{R} menores que 1,1 ou 1,2.

2.7 Estimação por intervalo

2.7.1 Intervalos de confiança para predições (método REML)

Existem dois tipos de predição: de resposta média e de observações futuras, sendo que, neste trabalho, adotou-se somente a primeira. No modelo fixo, a estimativa da combinação linear $\lambda' \beta$ é dada por $\lambda' \hat{\beta}$ e sua variância é dada por $\lambda' (X' X)^{-1} \lambda \sigma^2$. No modelo misto, $(X' X)^{-1} \sigma^2$ é substituída por V^{-1} da equação 5. Note-se que o vetor λ não precisa necessariamente ser observado, λ é um vetor de pesos para os parâmetros, tal que $\lambda' \beta$ é uma combinação linear de interesse, não necessariamente observada em algum tratamento do experimento. Então, a variância do estimador do vetor de parâmetros é dada por:

$$V(\hat{\beta}) = (V)^{-1}. \quad (16)$$

O estimador dessa variância é obtido substituindo-se a variância paramétrica pelo estimador da variância ($S^2 = \text{QMResíduo}$). Assim, tem-se o estimador da variância do estimador dos parâmetros dado por:

$$\hat{V}(\hat{\beta}) = (V)^{-1}. \quad (17)$$

Dessa forma, o estimador do erro padrão da predição para o intervalo de confiança para a combinação linear $\lambda' \theta$ é dado por:

$$S(\hat{Y}_i) = \sqrt{\lambda'(V)^{-1}\lambda}. \quad (18)$$

Por fim, o intervalo de confiança para predição da média é:

$$\hat{Y}_i \pm t_{n-p}^{(\frac{\alpha}{2})} S(\hat{Y}_i), \quad (19)$$

em que n é o tamanho da amostra e p é o número de parâmetros do modelo.

2.7.2 Intervalos de credibilidade

O conceito de intervalo de credibilidade (o equivalente bayesiano do intervalo de confiança) é baseado em resumos da distribuição *a posteriori*.

Definição: C é um intervalo de credibilidade de $100(1-\alpha)\%$ para θ se $P(\theta \in C) \geq 1 - \alpha$.

Note que a definição expressa de forma probabilística a pertinência ou não de θ ao intervalo. Assim, quanto mais concentrada for a distribuição do parâmetro, menor é o tamanho do intervalo. A exigência de que a probabilidade acima possa ser maior do que o nível de confiança é essencialmente técnica pois queremos que o intervalo seja o menor possível, o que em geral implica em usar uma igualdade. No entanto, a desigualdade será útil se o parâmetro tiver um distribuição discreta onde nem sempre é possível satisfazer a igualdade.

O intervalo de credibilidade de menor comprimento é usualmente conhecido como o intervalo de "máxima densidade *a posteriori*" ou "HPD", do inglês "*highest posterior density interval*", Gelman et al. (2003). A obtenção destes intervalos geralmente é feita via implementação de métodos computacionais.

2.8 Erro quadrático médio

O erro quadrático médio (EQM) de um estimador W de um parâmetro θ é a função de θ definida por $E_{\theta}(W - \theta)^2$. Segundo Casella e Berger (1990), essa medida tem pelo menos duas vantagens em relação a outras medidas de distância: primeiro, ela é bastante "tratável" analiticamente; segundo, possui a interpretação $E_{\theta}(W - \theta)^2 = \text{Var}_{\theta}W + (\text{Viés}_{\theta}W)^2$, em que o viés de um estimador pontual W , de um parâmetro θ , é a diferença entre o valor esperado de W e θ . Portanto, o EQM incorpora dois componentes, um medindo a variabilidade do estimador (precisão) e o outro medindo seu viés (acurácia).

3 Resultados e discussão

3.1 Superfície de resposta – análise REML

A parte superior da Tabela 3 da análise de variância (obtida pelo modelo 1.m, como descrito no Apêndice) revelou efeito não significativo da interação entre os fatores **dose de nitrogênio** e **adubação foliar** para a variável produção².

Como o interesse deste estudo foi analisar os fatores conjuntamente, considerando-se que seus níveis são quantitativos, ajustou-se um modelo de regressão de acordo com a equação (20), citada a seguir, para os níveis de dose de nitrogênio e de adubação foliar por meio da superfície de resposta.

Ilustrou-se a aplicação da superfície de resposta usando-se a REML para a estimação das médias de tratamentos ajustadas apresentada na Tabela 4, juntamente com a análise de variância. Tal análise (obtida pelo modelo 1.s, como descrito no Apêndice), com a decomposição dos efeitos de tratamentos sob a forma de superfície de resposta, encontra-se apresentada na parte inferior da Tabela 3.

Tabela 3 - Quadro-resumo da análise de variância para a variável resposta produção (simulada)

FV	GL	SQ	QM	F _c	Pr>F
R	2	15	8		
R/B	9	286	32		
Trat	17	82514	4854	191,4	2x10 ⁻¹⁶
Test	1	230	230	9,2	0,00486
FAT	15	9623	641,5	25,7	2,4x10 ⁻¹³
N	3	2725	908	34,4	3,899x10 ⁻¹⁰
A	3	6555	2185	82,8	3,582x10 ⁻¹⁵
N*A	9	343	38	1,4	0,21038
T vs FAT	1	72661	72661	2865,4	2,2x10 ⁻¹⁶
n	1	47658	47658	1946,3	2,2x10 ⁻¹⁶
a	1	19880	19880	811,9	2,2x10 ⁻¹⁶
n*a	1	273	273	11,7	0,001
n ²	1	10005	10005	408,6	2,2x10 ⁻¹⁶
a ²	1	4312	4312	176,1	2,2x10 ⁻¹⁶
Falta de ajuste	12	386	32,2	1,3	0,27
Resíduo	31	786	25		
Total	59	83601			

²Todas as análises estatísticas desenvolvidas nesta seção foram realizadas utilizando-se o software R[®] (R 2005).

Observa-se que, pela Tabela 3, a variável analisada (produção) mostrou-se significativamente influenciada pelas fontes de variação: efeito linear da dose de nitrogênio, efeito linear da dose de adubação foliar e efeito linear da interação entre as duas doses. Tal fato leva à conclusão da existência de efeitos das doses de nitrogênio e das doses de adubação foliar, como também de um comportamento diferenciado da aplicação dos níveis de nitrogênio nas diversas combinações com os níveis da adubação foliar.

Tabela 4 - Valores paramétricos, estimativas e respectivos limites de confiança para a análise REML; estimativas e limites de credibilidade para a análise bayesiana

Trat.	VP	REML			bayesiana		
		LI	Estimativa	LS	LI	Estimativa	LS
t ₁	157,75	151,07	157,17	163,27	145,62	157,61	169,28
t ₂	179,8	174,49	180,59	186,69	169,89	182,00	193,59
t ₃	185,65	178,56	184,06	190,76	174,72	186,19	198,42
t ₄	175,3	166,59	172,69	178,79	162,98	175,05	187,11
t ₅	169,6	167,41	173,51	179,61	163,42	174,74	186,70
t ₆	193	185,75	191,85	197,95	180,76	192,83	204,34
t ₇	200,2	194,69	200,79	206,89	191,71	203,01	215,76
t ₈	191,2	183,96	190,06	196,17	179,71	191,29	203,36
t ₉	172,45	168,06	174,16	180,26	163,18	174,15	186,87
t ₁₀	197,2	194,27	200,37	206,47	191,60	203,11	215,42
t ₁₁	205,75	197,66	203,76	209,86	191,99	203,80	215,62
t ₁₂	198,1	190,48	196,58	202,68	185,75	198,33	209,48
t ₁₃	166,3	156,58	162,68	168,78	153,68	165,62	177,28
t ₁₄	192,4	186,85	192,95	199,06	181,49	193,57	205,13
t ₁₅	202,3	195,94	202,04	208,14	192,44	203,81	216,42
t ₁₆	196	192,56	198,66	204,77	186,90	198,97	210,59
t ₁₇	100	92,71	97,03	101,36	87,84	98,27	109,61
t ₁₈	105	97,46	101,79	106,11	91,82	103,38	114,07

A equação de regressão ajustada, utilizando-se a função *lme* do software $R^{\text{®}}$, seguindo um modelo misto, com blocos aleatórios, e os demais componentes como fixos é dada por:

$$\hat{y}_{ij} = 98,897 + 0,86364N_i - 0,00575N_i^2 + 0,49425A_j - 0,00104A_j^2 + 0,00075N_iA_j, \quad (20)$$

que é análoga a uma função f de duas variáveis (N e A), por meio da qual podem-se estimar as produções (imagem) para quaisquer combinações dos níveis dos fatores doses de nitrogênio $0 \leq N_i \leq 120$ e doses de adubação foliar $0 \leq A_j \leq 360$, que são o domínio da função f .

Os pontos críticos dessa função, que são os níveis ótimos, foram $N = 92,78$ Kg/ha e $A = 271,07$ mg/l, os quais proporcionam uma produção de 205,95 caixas/ha, que, de acordo com o teste da segunda derivada, conclui-se tratar-se de um ponto de máximo relativo, o qual, mediante a Figura 2, pode-se observar, trata-se de um máximo local.

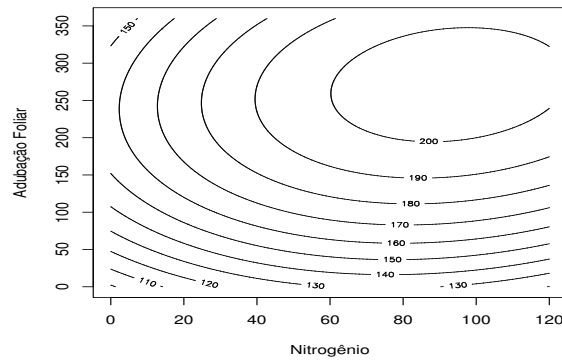


Figura 2 - Curvas de nível para a produção de citros (caixas/ha), em função das doses de nitrogênio (kg/ha) e adubação foliar (mg/l). Análise REML.

Observa-se que, de maneira geral, houve influência de doses de nitrogênio e de adubação foliar sobre a produção, indicando que maiores produtividades foram obtidas com as combinações das doses nas faixas de $60 \leq N_i \leq 100$ e $250 \leq A_j \leq 300$ para nitrogênio e adubação foliar, respectivamente.

Pelo esboço da Figura 3, tem-se, para a superfície ajustada (a curva situada na região central), o intervalo de confiança dos preditores da resposta média, sendo as curvas acima e abaixo dessa as superfícies de resposta superiores e inferiores, que correspondem aos limites superiores e inferiores de um intervalo de confiança. A obtenção das equações dessas superfícies de confiança foi feita mediante o ajuste de alguns pontos que se situavam tanto na região superior como na inferior. Assim:

Para o cálculo da elipse de confiança, que pode ser vista na Figura 6, elevou-se o plano formado pelos eixos (N, A) , de modo que ficasse tangente ao ponto de máximo (205,95), que passou a ser a origem do novo sistema. Igualando-se a superfície de resposta superior a essa origem, encontrou-se a equação da elipse de confiança projetada no plano (N, A) , para os níveis ótimos de nitrogênio e de adubação foliar. Convém destacar que tal metodologia foi desenvolvida especificamente para aplicação neste trabalho.

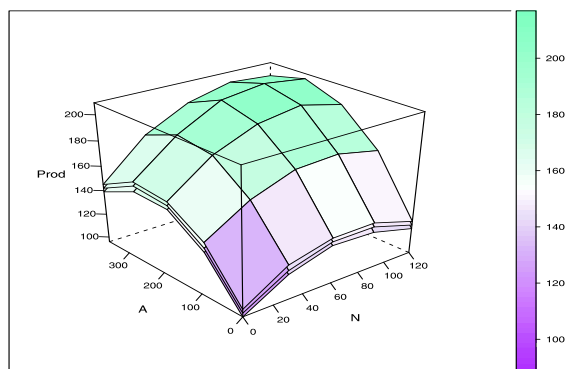


Figura 3 - Ajuste de superfícies para as estimativas das médias dos tratamentos e limites inferiores e superiores do IC para cada média. Prod: produção de citros (caixas/ha); N: doses de nitrogênio (kg/ha); A: doses de adubação foliar (mg/l)(Análise REML).

3.2 Diagnóstico de convergência do amostrador de Gibbs

Pode-se observar que os valores para o fator de redução de escala potencial, do último critério citado, para todos os parâmetros estimados (efeitos fixos e componentes da variância) estão apresentados na Tabela 5. A avaliação da convergência indicou que o algoritmo amostrador de Gibbs convergiu para todos os parâmetros, em ambas as amostras, uma vez que os valores de \hat{R} ficaram muito próximos do valor unitário, havendo diferenças apenas de algumas casas decimais, situação que exprime, segundo Gelman e Rubin (1992), a convergência do processo iterativo envolvido na análise. Dessa forma, pode-se ter plena confiança de que as estimativas dos parâmetros são verdadeiramente as médias das distribuições marginais, pois essas distribuições encontram-se em uma situação estável, ou seja, sob a condição de estacionariedade.

Também pela Tabela 5, podem-se notar, pelo critério de Raftery e Lewis (1992a), aplicado na amostra válida, indicações para pulo e queima mínimos para a retirada do efeito de dependência entre os valores amostrados, tornando o processo amostral estável e independente.

3.3 Análise bayesiana

Para o ajuste da superfície de resposta da análise bayesiana foram tomadas as médias ajustadas dos tratamentos demonstradas na Tabela 4 das distribuições *a posteriori*, cuja equação ajustada é:

$$\hat{y}_{ij} = 100,9 + 0,8607N_i - 0,0055N_i^2 + 0,4943A_j - 0,00102A_j^2 + 0,00063N_iA_j \quad (21)$$

Tabela 5 - Estimativas dos indicadores de convergência para os critérios de Gelman e Rubin; Raftery e Lewis e do erro de Monte Carlo para cada uma das cadeias dos parâmetros do modelo

Parâmetro	Raftery e Lewis		Gelman e Rubin	Erro MC
	Descarte	Pulo	\hat{R}	
σ_e^2	1	2	1,0001888	0,0884011
σ_b^2	1	2	0,9998877	0,07271616
μ	1	2	1,000242	0,06229866
r ₂	1	2	1,0001362	0,06840015
r ₃	1	2	1,0001372	0,06700255
t ₂	1	2	1,001617	0,0656974
t ₃	1	2	1,0004186	0,06719994
t ₄	1	2	1,0000252	0,07902311
t ₅	1	2	1,0000639	0,0677919
t ₆	1	2	0,9999436	0,06249842
t ₇	1	2	0,999811	0,07538871
t ₈	1	2	1,00082	0,13567047
t ₉	1	2	0,9999046	0,05300587
t ₁₀	1	2	1,0014557	0,06540558
t ₁₁	1	2	1,0008425	0,06307528
t ₁₂	1	2	1,0037092	0,05663703
t ₁₃	1	2	1,0003743	0,05771956
t ₁₄	1	2	1,0023839	0,04784242
t ₁₅	1	2	1,0039859	0,05578633
t ₁₆	1	2	1,0011513	0,06662412
t ₁₇	1	2	1,0022488	0,06608614
t ₁₈	1	2	1,0023457	0,07057838
b ₁	1	2	1,0010881	0,04997043
b ₂	1	2	1,0003638	0,05092744
b ₃	1	2	1,0006943	0,05515476
b ₄	1	2	0,9998241	0,0559402
b ₅	1	2	1,0014523	0,06276761
b ₆	1	2	1,0022157	0,07314327
b ₇	1	2	1,0036355	0,05631986
b ₈	1	2	1,0009586	0,07491498
b ₉	1	2	1,0001834	0,07824042
b ₁₀	1	2	0,999933	0,06997551
b ₁₁	1	2	1,0001401	0,07408732
b ₁₂	1	2	1,000013	0,06412052

A forma de obtenção dos pontos críticos e a obtenção das superfícies (superior e inferior) é análoga à da REML. Os pontos críticos encontrados para as duas doses são: $N = 93,78 \text{ Kg/ha}$ e $A = 271,27 \text{ mg/l}$, com um ponto de máximo igual a 208,30 caixas/ha. O esboço das curvas de nível de sua superfície de resposta está apresentado na Figura 4.

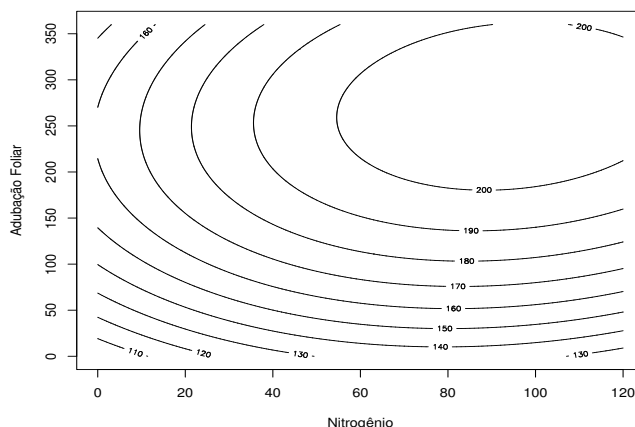


Figura 4 - Curvas de nível para a produção de citros (caixas/ha), em função das doses de nitrogênio (kg/ha) e adubação foliar (mg/l). (Análise bayesiana).

A Figura 5 mostra os gráficos das superfícies que correspondem aos ajustes para os limites inferiores e superiores do intervalo de credibilidade para cada média dos tratamentos, bem como da própria estimativa. Por essa figura fica claro que a superfície tem concavidade para baixo, sendo possível determinar ótimo para as combinações de níveis dos fatores.

A Figura 6 apresenta as elipses de confiança para a análise REML e bayesiana, na região dos níveis ótimos para a combinação das doses dos dois nutrientes em estudo. Note que o valor paramétrico para o ponto de máximo corresponde a 93,67 (kg/ha) da fonte de nitrogênio e 273,42 (mg/l) de adubação foliar, estando próximo ao centro de ambas as elipses.

3.4 Comparação das metodologias

A Tabela 6 abaixo apresenta o quadro resumo das estimativas dos parâmetros do modelo pelo análise bayesiana e ,como referência, os valores paramétricos e as estimativas REML dos parâmetros correspondentes.

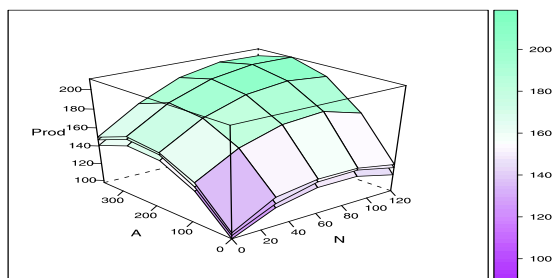


Figura 5 - Ajuste de superfícies para as estimativas das médias dos tratamentos e limites inferiores e superiores do intervalo de credibilidade para cada média. Prod: produção de citros (caixas/ha); N: doses de nitrogênio (kg/ha); A: doses de adubação foliar (mg/l) (Análise bayesiana).

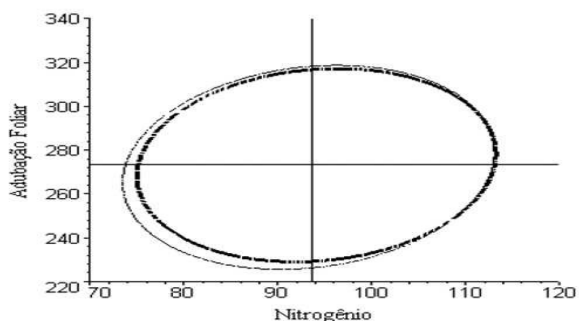


Figura 6 - Regiões de 95% de credibilidade para os níveis ótimos das doses de nitrogênio (Kg/ha) e de adubação foliar (mg/l). (A elipse interior refere-se à análise REML e a exterior refere-se à análise bayesiana.)

Quanto aos coeficientes de correlação de Pearson entre os valores paramétricos e as estimativas (médias) pelas duas formas de análise, foram calculados os seguintes valores: 0,99789 (para a análise REML) e 0,99787 (para a análise bayesiana). Isso indica que ambos os métodos geram estimativas muito correlacionadas ao modelo paramétrico. Para os erros quadráticos médios, obteve-se os valores 4,86 (para a análise REML) e 5,79 (para a análise bayesiana). Isso poderia indicar um ligeira superioridade da análise REML em relação à bayesiana, no entanto esta diferença é muito pequena e os métodos geram estimativas muito semelhantes.

Tabela 6 - Quadro-resumo da análise do modelo experimental pela análise bayesiana: média, desvio padrão, intervalos de credibilidade com 95% de probabilidade e valores das estimativas pontuais REML dos efeitos do modelo e seus valores paramétricos

Parâmetro	Média	EP	IC		REML	Val.Par.
			LI	LS		
σ_e^2	25,04	6,4	14,08	38,03	24,89	25
σ_b^2	17,39	10,84	3,32	38,82	2,33	9
intercepto	155,48	4,27	147,05	163,75	156,51	-
r ₂	0,88	4,25	-7,66	8,87	0,79	-
r ₃	1,26	4,25	-6,81	9,53	1,22	-
t ₂	24,39	4,6	15,45	33,87	23,41	179,8
t ₃	28,57	4,61	19,81	38,1	27,49	185,65
t ₄	17,44	4,57	8,29	26,1	15,52	175,3
t ₅	17,12	4,64	8,56	26,72	16,34	169,6
t ₆	35,22	4,4	26,83	44,14	34,68	193
t ₇	45,4	4,43	36,58	53,57	43,62	200,2
t ₈	33,68	4,52	25,17	43,17	32,89	191,2
t ₉	17,54	4,67	8,54	26,81	16,98	172,45
t ₁₀	45,5	4,48	37,3	54,71	43,19	197,2
t ₁₁	46,19	4,27	37,66	54,3	46,59	205,75
t ₁₂	40,72	4,55	31,87	49,54	39,41	198,1
t ₁₃	8,01	4,58	-0,86	17,11	5,5	166,3
t ₁₄	35,96	4,44	27,28	44,67	35,78	192,4
t ₁₅	46,2	4,64	37,4	55,52	44,87	202,3
t ₁₆	41,35	4,3	32,85	49,54	41,49	196
t ₁₇	-59,35	3,78	-66,73	-51,7	-60,14	100
t ₁₈	-54,24	3,94	-62,04	-46,81	-55,39	105

Conclusões

As duas metodologias empregadas neste trabalho, análises bayesiana e de verossimilhança restrita, mostraram-se boas ferramentas para a estimação de parâmetros de um modelo linear misto, visto que os resultados obtidos em ambas são semelhantes.

Levar em conta as distribuições *a priori* que resultassem em pouca informação para a correlação intraclassa não apresentou vantagem em relação à análise de verossimilhança restrita.

Ambas as formas de análises são recomendadas para ajustes de superfícies de resposta.

Agradecimentos

À CAPES pela bolsa de doutorado concedida ao primeiro autor; ao CNPq pela bolsa de Produtividade em Pesquisa concedida ao segundo autor; aos Professores Carlos Alberto Silva e João José G. S. M. Marques, do Departamento de Solos da Universidade Federal de Lavras, pelas críticas e sugestões ao exemplo simulado.

LIMA, C. N. de; BUENO FILHO, J. S. de S. Fitting a response surface in an incomplete block design: Restricted likelihood and Bayesian analysis applied to a manuring simulation in citrus. *Rev. Mat. Estat.*, São Paulo, v.24, n.4, p.133-157, 2006.

■ **ABSTRACT:** *Finding the combination of nutrient levels that yield maximum production is an interesting problem in manuring assays. The aim of this work was to illustrate the restrict likelihood analysis (REML) and Bayesian analysis (GS, or via Gibbs Sampler) of a partially balanced incomplete block design. For this illustration, production data for the first year of a citrus culture were simulated, using combinations of levels for two nutrients. Model fitting and statistical properties were worked out through REML and GS analysis of a mixed model. Results from both methodologies were equivalent, indicating that they are good tools for mixed models fitting that should be recommended for response surface estimation.*

■ **KEYWORDS:** *Gibbs sampling; incomplete block designs; manuring; REML.*

Referências

BERNARDO, J. M.; SMITH, A. F. M. *Bayesian Theory*. New York: John Wiley & Sons, 1994. 608p.

BOX, G. E.; DRAPPER, N. R. *Empirical Model-Building and Response Surfaces*. New York: John Wiley & Sons, 1987. 424p.

BITTENCOURT, F. ; BUENO FILHO, J. S. S. ; FERREIRA, D. F. Comparação entre métodos de construção de delineamentos fatoriais com dois fatores em blocos incompletos. *Revista de Matemática e Estatística*, Jaboticabal, v.21, n.1, p.49-65, 2003.

CASELLA, G.; BERGER, R. L.; *Statistical inference*. Belmont: Wadsworth, 1990. 650p.

DEMPSTER, A.; LAIRD, N.; RUBIN, D. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.*, série B, London, v.39, n.1, p.1-38, 1977.

EDMONDSON, R. N. *Design of experiments website*. 2005. Disponível em: <<http://biometrics.hri.ac.uk/experimentaldesigns/website/hri.htm>>. Acesso em: mar. 2006.

- FEDERER, W. T. On Augmented Designs. *Biometrics*, Washington, v.31, n.1, p.29-35, 1975.
- GELFAND, A. E.; SMITH, A. F. M. Sampling based approaches for calculating marginal desities. *J. Am. Stat. Assoc.*, London, v.85, n.410, p.398-409, June 1990.
- GELMAN, A.; CARLIN, J. B.; STER, I. S.; RUBIN, D. B. *Bayesian data analysis*, London: Chapman and Hall, 2003. 668p.
- GELMAN, A. Inference and monitoring convergence. In: GILKS, W.R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Ed). *Markov chain Monte Carlo in practice*. New York: Chapman and Hall, chap.8, p.131-144. 1995
- GELMAN, A.; RUBIN, D. B. Inference from interative simulation using multiple sequences. *Stat. Sci.*, Hayward, v.7, n.4, p.457-511, 1992a.
- HENDERSON, C.R. Estimation of variance and covariance components. *Biometrics*, Washington, v.9, n.2, p.226-252, 1953.
- HINKELMANN, K.; KEMPTHORNE. *O.Design and analysis of experiments: Advanced experimental design*. New Jersey: John Wiley & Sons, 2005. v.2, 780p.
- MEAD, R. *The design of experiments: statistical principles for practical applications*. Cambridge: Cambridge University Press. 1988. 634p.
- MONTGOMERY, D. C.; MYERS, R. H. *Response surface methodology: process and product optimization using designed experiments*. New York: John Wiley & Sons, 1995. 728p.
- O'HAGAN, A. *Kendall's advanced theory of statistics: Bayesian inference*. [S.1.]: Edward Arnold, 1994. v.2b. 330p.
- PATTERSON, H. D.; THOMPSON, R. Recovery of interblock information when block sizes are unequal. *Biometrika*, London, v.58, n.3, p. 545-554, 1971.
- PINHEIRO, J. C.; BATES, D. M. *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag, 2000. 528p.
- R DEVELOPMENT Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2005. Disponível em: <<http://www.R-project.org>>. Acesso em: 2006.
- RAFTERY, A. E.; LEWIS, S. Comment: One long run with diagnostics: implementation strategies for markov chain Monte Carlo. *Stat. Sci.*, Hayward, v.7, n.4, p. 493-497, 1992a.
- RIBEIRO, A. C.; GUIMARÃES, P. T.; ALVAREZ V. V. H. (Ed). *Recomendações para o uso de corretivos e fertilizantes em Minas Gerais: 5. Aproximação*. Viçosa: Comissão de Fertilidade do Solo do Estado de Minas Gerais, 1999. 359p.
- SORENSEN, D.; *Gibbs sampling quantitative genetics*. Foulun:[s.n], 1996. 186p.

Recebido em 10.05.2006.

Aprovado após revisão em 20.02.2007.

Apêndice

Regressões múltiplas com dados desbalanceados por meio de modelos de superfície de resposta
por meio de procedimentos de mínimos quadrados - L (são definidas lme()) e de regressão
Análise de regressão múltipla com dados desbalanceados por meio de modelos de superfície de resposta
são definidas as funções de regressão múltipla com dados desbalanceados por meio de modelos de superfície de resposta
propostas por eles

```
# Leitura de dados e definição do delineamento #

valpar <- read.table("arquivo de dados", header=TRUE)

R      <- factor(valpar$Repet)
B      <- factor(valpar$Bloco)
Trat   <- factor(valpar$Trat)
Test   <- factor(valpar$Test)
Nitro  <- factor(valpar$Nitro)
AdFol  <- factor(valpar$AdFol)
Prod   <- valpar$Prod
N      <- valpar$N
A      <- valpar$A
simul  <- data.frame(R,B,Trat,Test,Nitro,N,AdFol,A,Prod)
attach(simul)

# Análise do modelo misto - REML #

library(nlme) modelo.1.m <- lme(Prod ~ R + Trat, data = simul,
random = ~1| B) anova (modelo.1.m) summary (modelo.1.m)

# Para visualizar o gráficos das interações #

par(mfrow=c(1,2))
interaction.plot(Nitro,AdFol,Prod)
interaction.plot(AdFol,Nitro,Prod)

# Superfície de Resposta #

N2      <- N*N
A2      <- A*A
modelo.1.s <- lme(Prod ~ R+ N*A + N2 + A2, random = ~ 1 | B)
summary(modelo.1.s)

# Delineamento #
XZ      <- model.matrix(modelo.1.m)
```

```

X      <- XZ[,1:20]
xz     <- matrix(c(rep(1,5),rep(0,55)))
Z      <- cbind(xz,XZ[,21],XZ[,24],XZ[,27],XZ[,31],
               XZ[,34],XZ[,37],XZ[,40],XZ[,44],
               XZ[,47],XZ[,50],XZ[,53])

As     <- svd(X,nu=0)
rank.X <- sum(As$d>0.000001*As$d[1])
rank.X

# Inferência Bayesiana #
iter   <- 150000
BI     <- 50000
TH     <- 25
SS     <- (iter-BI)/TH
cadeia <- matrix(0,iter,(2+20+12))
amostra <- matrix(0,SS,(2+20+12))

# Prioris #
rb     <- 2
sb     <- 5
re     <- 2
se     <- 5

# Pontos iniciais (Chutes) #
b      <- ginv(t(Z)%*%Z)%*%t(Z)%*%y
beta   <- ginv(t(X)%*%X) %*% t(X)%*% y
e      <- y - (X)%*%beta + Z)%*%b
vb     <- as.real((t(b)%*%b)/11)
ve     <- as.real((t(e)%*%e)/31)
F      <- ginv(t(X)%*%X)
cF     <- chol(F)

# Agora vamos percorrer a cadeia e gerar todas as amostras #

# Loop MCMC #

for (i in 1:iter)
{
  gamma <- as.real(ve/vb)

  # Sendo A=I #
  Ag    <- gamma *diag(12)

```



```

# Condicional completa para efeitos fixos #

meanf <- ginv(t(X)%*%X)%*%t(X)%*%(y-Z %*% b)
kf <- rnorm(20)
beta <- sqrt(ve) * t(cF) %*% kf + meanf

# Condicional completa para efeitos aleatórios #

M <- ginv(t(Z)%*%Z + solve(Ag))
meanr <- M%*%t(Z)%*%(y-X%*%beta)
kr <- rnorm(12)
sr <- sqrt(ve) * chol(M)
b <- t(sr) %*% kr + meanr

# Condicional completa para variância residual #

e <- y - X%*%beta-Z%*%b
c1 <- (n+re)/2
c2 <- (t(e)%*%e+re*se)/2
ve <- rgamma(1,c1,c2)
ve <- as.real(1/ve)

# Condicional completa para variância de blocos #

c1 <- (12+rb)/2
c2 <- (t(b) %*% b + rb*sb)/2
vb <- rgamma(1,c1,c2)
vb <- as.real(1/vb)

# Gerando a cadeia de resultados do Amostrador de Gibbs #

cadeia[i,] <- cbind(ve,vb,t(beta),t(b))

for (j in 1:SS)
{
  if (round((i-BI-1)/TH) == j)
  {
    amostra[j,] <- cbind(ve,vb,t(beta),t(b))
  }
}
}

```