

# Análise espacial e de vizinhança no melhoramento genético de plantas

Janaina Ribeiro Costa<sup>(1)</sup>, Júlio Sílvio de Sousa Bueno Filho<sup>(2)</sup> e Magno Antônio Patto Ramalho<sup>(2)</sup>

<sup>(1)</sup>Embrapa Agrobiologia, BR 465, Km 47, CEP 23851-970 Seropédica, RJ. E-mail: janaina@cnpab.embrapa.br <sup>(2)</sup>Universidade Federal de Lavras, Caixa Postal 37, CEP 37200-000 Lavras, MG. E-mail: jssbueno@ufla.br, magnoapr@ufla.br

Resumo – Três formas de análise espacial foram comparadas à análise do modelo linear Gauss-Markov normal em experimentos de Genética, tendo-se suposto os efeitos de progênies como aleatórios: médias móveis nos dados brutos (MM), médias móveis nos dados residuais (Papadakis – PPD) e análise espacial por meio de modelagem de covariâncias residuais (AE). Inicialmente, ignorou-se a informação do controle local, para testar a efetividade da análise espacial. Posteriormente, foi verificado se haveria melhoras com as diferentes formas de análise espacial aplicadas ao modelo completo, considerando-se o controle local do delineamento em látice. Os valores médios de razões, entre estimativas de componentes de variância e de herdabilidade, foram usados como guia de discussão sobre qual a melhor forma de análise. Em geral, ignorar o delineamento experimental e usar somente a informação espacial resultou em análises ineficientes. Os modelos MM e PPD, em média, melhoraram o modelo original justificado pelo delineamento, embora a AE não o tenha melhorado. A AE, apesar de ineficiente, não mudou as estimativas dos componentes de variância e de herdabilidade. Esta propriedade garante que a combinação de efeitos aleatórios para progênies e a AE não violam as suposições (algumas delas justificadas pelo delineamento). Isto é especialmente útil com experimentos amplos, com grande número de progênies.

Termos para indexação: análise estatística, médias móveis, Papadakis.

## Neighborhood and spatial analysis in plant breeding

Abstract – Three forms of spatial analysis were compared to the analysis of the normal Gauss-Markov model in genetical experiments, supposing progeny effects as random: moving averages on raw data (MM), moving averages on residual data (Papadakis – PPD), and spatial analysis modelling with residual covariances (AE). The local control information was initially ignored to test the effectiveness of spatial analysis. Thereafter, the different forms of spatial analysis were applied to the complete model, considering the local control of lattice design. The average values of proportions between estimates of variance components and of heritability were used as a discussion guide to determine the best form of analysis. Results showed that ignoring the experimental design, using spatial information was not effective, in general. In average, MM and PPD improved the original model justified by design, in contrast to AE. AE, although ineffective, did not change variance component estimates and heritability. This property guarantees that the combination of random effects for progenies and AE does not violate the assumptions (some of these justified by the design). This is specially useful in large experiments, with a huge number of progenies.

Index terms: statistical analysis, moving average, Papadakis.

## Introdução

O sucesso dos programas de melhoramento depende da avaliação de grande número de progênies, na expectativa de se detectarem diferenças cada vez menores entre elas. Nesse contexto, mesmo com pequeno número de repetições, os experimentos tornam-se grandes e dependem de formas mais sofisticadas de

planejamento e de análises que levem a estimativas precisas dos parâmetros de interesse.

Na tentativa de se manter número adequado de repetições, sem sacrificar a precisão experimental, são utilizadas várias alternativas pelos melhoristas, e entre elas destacam-se os métodos de análise que levam em conta a posição das parcelas na área experimental, chamados métodos espaciais. Em geral, nesse tipo de aná-

lise, considera-se que os efeitos das parcelas experimentais (erro experimental) são tão mais semelhantes, quanto mais próximas sejam as parcelas.

Entre as estratégias baseadas em modelagens espaciais que têm sido sugeridas no melhoramento de plantas, para exercer o controle local e aumentar a precisão experimental, destacam-se os métodos de análise de vizinhança, denominados médias móveis e método de Papadakis, que modelam a dependência espacial na forma de análise de covariância. Outra estratégia que tem merecido atenção é a análise espacial, em que a dependência espacial é modelada na forma de matriz de variâncias e covariâncias residuais, considerando-se a distância entre parcelas no campo experimental. Cressie & Hartfield (1996), Souza (1997) e Duarte (2000) abordaram os métodos de análise espacial no melhoramento de plantas.

Papadakis (1937) foi o primeiro a introduzir o método do vizinho mais próximo (Nearest-Neighbor Method – NN), para reduzir o erro experimental. O método de Papadakis consiste em corrigir o valor observado de cada parcela, pelo efeito médio do tratamento a ela aplicado e, então, a média dos valores corrigidos das parcelas vizinhas é utilizada como covariável.

Rickey (1924), pioneiro na utilização do método das médias móveis, ajustava médias de progênies de milho a partir da razão entre as médias das fileiras vizinhas e, visto que cada fileira possuía vizinhos diferentes, o autor denominou a técnica de médias móveis. O autor detectou reduções no erro padrão das médias, após o ajuste por essa técnica (Rios, 1997). Assim, no método das médias móveis, o valor de cada parcela é corrigido pela média dos valores das parcelas vizinhas, e não pelo efeito de tratamento, como no método de Papadakis.

Na análise espacial, a dependência espacial é associada à Geoestatística, que surgiu na África do Sul quando Krige (1951) trabalhava com dados de concentração de ouro. Ele concluiu que as covariâncias de duas amostras dependiam da distância entre elas.

Esses métodos têm sido aplicados em algumas oportunidades, por meio de simulação (Vivaldi, 1990), ou utilizando-se dados de campo (Souza, 1997; Duarte, 2000) que evidenciem a sua utilidade. Contudo, ainda há poucos estudos sobre esses métodos, para que sejam recomendados rotineiramente, nos programas de melhoramento.

O objetivo deste trabalho foi buscar alternativas que possam melhorar a eficiência dos experimentos de avaliação de linhagens e de progênies de feijão e milho, por

meio de técnicas de análise que contribuam para maior precisão experimental.

## Material e Métodos

Os dados utilizados neste trabalho referem-se a experimentos distintos de progênies de feijão e de milho, com relação ao número de repetições, tratamentos e ao tamanho das parcelas. Os experimentos foram conduzidos em Lavras, MG, a 910 m de altitude, 21°14' S e 45°0' W. Foram analisados os dados de produtividade de grãos (kg ha<sup>-1</sup>) de 36 experimentos de feijão e dois de milho, implementados em látices.

Como formas de análise, consideraram-se: o modelo de análise clássica, representado pelo modelo de Gauss-Markov Normal (GMN), e os modelos espaciais, representados pelos modelos de Médias Móveis (MM), o de Papadakis (PPD) e o de Análise Espacial (AE). Em cada modelo foram, ainda, consideradas as situações de ausência e de presença de controle local. Os modelos utilizados estão descritos a seguir:

Modelo 1 (GMN – ausência de controle local):

$$y_{ij} = \mu + p_i + e_{ij};$$

Modelo 2 (MM – ausência de controle local):

$$y_{ij} = \mu + p_i + \beta I_{ij} + e_{ij};$$

Modelo 3 (PPD – ausência de controle local):

$$y_{ij} = \mu + p_i + \beta \bar{I}_{ij} + e_{ij};$$

Modelo 4 (AE – ausência de controle local):

$$y_{ij} = \mu + p_i + \varepsilon_{ij};$$

Modelo 5 (GMN – presença de controle local):

$$y_{ijk} = \mu + r_k + b_{j(k)} + p_i + e_{ijk};$$

Modelo 6 (MM – presença de controle local):

$$y_{ijk} = \mu + r_k + b_{j(k)} + p_i + \beta I_{ij(k)} + e_{ijk};$$

Modelo 7 (PPD – presença de controle local):

$$y_{ijk} = \mu + r_k + b_{j(k)} + p_i + \beta \bar{I}_{ij(k)} + e_{ijk};$$

Modelo 8 (AE – presença de controle local):

$$y_{ijk} = \mu + r_k + b_{j(k)} + p_i + \varepsilon_{ijk};$$

em que  $y_{ij}$  ou  $y_{ijk}$  é o valor observado na parcela que recebeu a  $i$ -ésima progênie, na  $j$ -ésima repetição (Modelos 1 a 4) ou no  $j$ -ésimo bloco (Modelo 5 a 8) e, ainda, na  $k$ -ésima repetição (Modelos 5 a 8);  $\mu$  é a constante associada a todas as observações;  $p_i$  é o efeito

aleatório da  $i$ -ésima progênie com  $i = 1, 2, \dots, I$  e  $p_i \sim N(0, \sigma_p^2)$ ;  $r_k$  é o efeito fixo da  $k$ -ésima repetição com  $k = 1, 2, \dots, K$ ;  $b_{j(k)}$  é o efeito aleatório do  $j$ -ésimo bloco dentro da  $k$ -ésima repetição;  $\beta$  é o coeficiente de regressão linear entre a variável produção de grãos e a covariável índice ambiental;  $I_{ij}$  e  $\bar{I}_{ij}$  são os índices ambientais da parcela que recebeu a  $i$ -ésima progênie na  $j$ -ésima repetição para os Modelos 2 e 3, respectivamente;  $I_{ij(k)}$  e  $\bar{I}_{ij(k)}$  são os índices ambientais da parcela que recebeu a  $i$ -ésima progênie, no  $j$ -ésimo bloco e na  $k$ -ésima repetição para os Modelos 6 e 7, respectivamente;  $e_{ij}$  ou  $\varepsilon_{ij}$  é o erro experimental associado à observação  $y_{ij}$ ;  $e_{ijk}$  ou  $\varepsilon_{ijk}$  é o erro experimental associado à observação  $y_{ijk}$ .

### Modelo GMN versus AE

O que distingue os modelos GMN e AE são as pressuposições acerca dos erros experimentais. Assim, nas análises dos modelos GMN (Modelos 1 e 5), considerou-se que os erros são variáveis aleatórias independentes e têm distribuição normal com média zero e variância constante  $\sigma^2$ , matricialmente simbolizados por  $e \sim N(0, I\sigma^2)$ . Nas análises dos modelos AE (Modelos 4 e 8), assumiu-se que os erros são dependentes e têm distribuição normal com média zero e matriz de variâncias e covariâncias residuais ( $R$ ) dada por  $f(d)\sigma^2$ , em que  $f(d)$  é uma função da distância ( $d$ ) entre duas parcelas, e  $\sigma^2$  é a variância residual. Assim, no modelo AE, o erro é simbolizado matricialmente por  $e \sim N(0, R)$ . A função  $f(d)$  utilizada no ajuste das covariâncias pode ser vista mais adiante no item Modelo Misto.

### Modelo MM versus PPD

Os modelos MM e PPD modelam a dependência espacial na forma de análise de covariância. O que diferencia um do outro é que, no primeiro, o índice ambiental ( $I_{ij}$  ou  $I_{ij(k)}$ ), utilizado como covariável, corresponde à média dos valores observados de parcelas vizinhas e, no segundo, esse índice ( $\bar{I}_{ij}$  ou  $\bar{I}_{ij(k)}$ ) corresponde à média dos resíduos de parcelas vizinhas, e é o vetor de resíduos, matricialmente estimado, por meio da seguinte equação:  $\hat{e} = y - \hat{y}$ , em que  $\hat{e}$  é o vetor das estimativas de erros aleatórios;  $y$  é o vetor das observações; e  $\hat{y}$  é o vetor das estimativas das observações.

No estudo dos Modelos 2, 3, 6 e 7 foram utilizadas, quando possível, no máximo quatro parcelas vizinhas, uma à direita, uma à esquerda, uma acima e uma abaixo da referida parcela. Para as parcelas das extremidades

foram utilizados os valores das observações (no modelo MM), ou dos resíduos médios (no modelo PPD) de uma, duas ou três parcelas imediatamente próximas a elas.

### Modelo misto

Os Modelos 1 a 8 podem ser matricialmente representados pelo modelo linear misto:

$$Y = X\beta + Zp + e \quad (1)$$

em que  $y$  é o vetor das observações, de dimensões  $(n) \times (1)$ ;  $X$  é a matriz de incidência dos efeitos fixos do delineamento, de dimensões  $(n) \times (u)$ ;  $\beta$  é o vetor de efeitos fixos desconhecidos, a serem estimados, de dimensões  $(u) \times (1)$ ;  $Z$  é a matriz de incidência dos efeitos aleatórios do delineamento, de dimensões  $(n) \times (v)$ ;  $p$  é o vetor de efeitos aleatórios desconhecidos, a serem preditos, de dimensões  $(v) \times (1)$ ;  $e$  é o vetor de erros aleatórios, de dimensões  $(n) \times (1)$ ;  $n$  é o número de observações;  $u$  é o número de efeitos fixos; e  $v$  é o número de efeitos aleatórios. Assumiu-se que:  $p \sim N(0, G)$ ,  $e \sim N(0, R)$ ,  $E(y) = X\beta$ ,  $\text{Var}(y) = V = ZGZ' + R$ , em que  $0$  é um vetor nulo;  $G$  é a matriz de variâncias e covariâncias genéticas, dada por:  $G = A\sigma_p^2$ , em que  $A$  é a matriz de coeficientes de parentesco ou similaridade genética entre progênies;  $\sigma_p^2$  é a variância de progênies;  $R$  é a matriz de variâncias e covariâncias residuais.

Nos experimentos em estudo não se conhecia a genealogia das progênies. Assumiu-se, então, que  $A$  é igual à matriz identidade ( $I$ ) e, portanto:  $G = I\sigma_p^2$ .

Para as análises de variância por meio dos Modelos 1 e 5, que pressupõem independência espacial entre as parcelas (modelagem GMN), tem-se:  $R = I\sigma^2$ , em que  $\sigma^2$  é a variância residual. Para a análise espacial por meio dos Modelos 4 e 8, que pressupõem correlação espacial entre as parcelas, tem-se:  $R = [f(d)]\sigma^2$ , em que  $f(d)$  é uma função da distância ( $d$ ) entre duas parcelas.

Portanto, na modelagem espacial, a covariância é assumida como uma função da distância entre as posições físicas das observações, sendo cada posição especificada por duas coordenadas: linha ( $l_{ij}$ ) e coluna ( $c_{ij}$ ). Assim, as distâncias entre as parcelas ( $i, j$ ) e ( $i', j'$ ) foram calculadas por:  $d = [(l_{ij} - l_{i'j'})^2 + (c_{ij} - c_{i'j'})^2]^{0,5}$ .

Na abordagem de modelos mistos tem-se, então, que:

$$\text{Cov}(e_{ij}, e_{i'j'}) = \begin{cases} \sigma^2, & \text{se } d = 0 \\ \sigma_{ij,i'j'} = [f(d)]\sigma^2, & \text{se } d > 0. \end{cases}$$

Por não apresentar alterações em relação a outros modelos testados, quanto às estimativas dos componentes de variância, foi escolhido o modelo esférico, para o ajuste das covariâncias (elementos da matriz R), encontrado em Littell et al. (1996), dado por:

$$f(d) = [1 - 1,5(d/\rho) + 0,5(d/\rho)^3]I \quad (2)$$

em que  $\rho$  é o alcance ("range"). Do ponto  $\rho$  em diante, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras torna-se invariante com a distância. Assim, na equação (2), se  $d < \rho$ , então  $I = 1$ , e pode-se concluir que as observações são espacialmente correlacionadas. Ao contrário, se  $d \geq \rho$ , então  $I = 0$ , e conclui-se que a correlação espacial é efetivamente zero.

O vetor de soluções do modelo misto, na equação (1), é obtido matricialmente por:

$$\begin{bmatrix} \beta^o \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}. \quad (3)$$

Nesse estudo, como os componentes de variância  $\sigma_p^2$  e  $\sigma^2$  (e, portanto, G e R) eram desconhecidos, utilizou-se o procedimento "Proc mixed" do SAS, para obtenção de suas estimativas, por meio de algoritmos iterativos, que utilizam o método da máxima verossimilhança restrita (REML), de forma a se obter, quase simultaneamente, o vetor de soluções na equação (3), composto de BLUEs e BLUPs, representadas pelos vetores  $\beta^o$  e  $\hat{p}$ , respectivamente. A sigla BLUE (Best Linear Unbiased Estimation) significa a melhor estimativa linear não tendenciosa dos efeitos fixos; e BLUP (Best Linear Unbiased Prediction) significa a melhor predição linear não tendenciosa dos efeitos aleatórios (Henderson, 1950; Robinson, 1991).

### Comparação das formas de análise

A comparação das diferentes formas de análise se deu por meio de valores médios de razões ( $\phi$ ) entre estimativas de variâncias residuais ( $\hat{\sigma}^2$ ), de variâncias de progênies ( $\hat{\sigma}_p^2$ ) e de herdabilidade no sentido amplo ( $\hat{h}^2$ ), tendo-se sempre tomado como padrão, no denominador, as estimativas oriundas do Modelo 5. Assim, se as diferentes formas de análise apresentarem eficiência semelhante àquela proveniente do Modelo 5, espera-se que as razões sejam próximas ou igual a 100%.

As expressões de razões utilizadas foram as seguintes:

(a) Razões ( $\phi_{\hat{\sigma}^2}$ ) entre estimativas de  $\sigma^2$ :

$$\phi_{\hat{\sigma}^2} (\%) = \left( \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \right) 100,$$

em que  $\hat{\sigma}_1^2$  é a estimativa da variância residual de um dos oito modelos avaliados, e  $\hat{\sigma}_2^2$  é a estimativa da variância residual do Modelo 5.

(b) Razões ( $\phi_{\hat{\sigma}_p^2}$ ) entre estimativas de  $\sigma_p^2$ :

$$\phi_{\hat{\sigma}_p^2} (\%) = \left( \frac{\hat{\sigma}_{p1}^2}{\hat{\sigma}_{p2}^2} \right) 100,$$

em que  $\hat{\sigma}_{p1}^2$  é a estimativa da variância de progênie de um dos oito modelos avaliados, e  $\hat{\sigma}_{p2}^2$  é a estimativa da variância de progênie do Modelo 5.

(c) Razões ( $\phi_{\hat{h}^2}$ ) entre estimativas de  $h^2$ :

$$\phi_{\hat{h}^2} (\%) = \left( \frac{\hat{h}_1^2}{\hat{h}_2^2} \right) 100, \text{ em que } \hat{h}_1^2 \text{ é a estimativa da}$$

herdabilidade de um dos oito modelos avaliados; e  $\hat{h}_2^2$  é a estimativa da herdabilidade do Modelo 5. A estimativa de herdabilidade no sentido amplo, que é a proporção da variância fenotípica devida à variância genética, foi obtida por meio do seguinte estimador:

$$\hat{h}^2 (\%) = \left( \frac{\hat{\sigma}_p^2}{\hat{\sigma}_p^2 + \hat{\sigma}^2} \right) 100.$$

### Correlação de Spearman

Para os 38 experimentos analisados, conforme o Modelo 5 (modelo do látice), foram também obtidas as correlações de Spearman ( $r$ ) entre os efeitos estimados fixos e aleatórios de progênies, tendo-se considerado o efeito de repetição como fixo e o efeito de bloco como aleatório, para ambas as situações de efeitos de progênies.

### Resultados e Discussão

O que se espera ao analisar os valores de  $\phi_{\hat{\sigma}^2}$  (Tabela 1) é revelar a eficiência do controle local (comparando-se o Modelo 1 com o Modelo 5); verificar se as formas de análise espacial recuperaram a eficiência da análise (caso o controle local tenha sido eficiente), o que pode ser feito analisando-se os Modelos 2, 3 e 4; e, finalmente,

verificar se as formas de análise espacial melhoraram, de alguma maneira, o delineamento com controle local (comparando-se os Modelos 6, 7 e 8 com o Modelo 5).

Valores de  $\phi_{\sigma^2}$  maiores que 100% (Tabela 1) indicam que houve redução em  $\hat{\sigma}^2$  de  $(\phi_{\sigma^2} - 100)\%$ , quando se utilizou o Modelo 5, comparativamente a algum outro modelo.

Desse modo, o valor de  $\phi_{\sigma^2}$ , proporcionado pelo Modelo 1, de 111,8%, indica que houve redução média na variância residual de 11,8%, quando foi utilizado o controle local do delineamento na análise.

Após os resultados anteriores terem indicado que, em média, o controle local do látice foi eficiente, verificou-se que as formas de análise espacial provenientes dos Modelos 2, 3 e 4, em média, não recuperaram a eficiência da análise, pois apresentaram valores de  $\phi_{\sigma^2}$  superiores a 100% (102,7%, 101,6% e 111,8%, respectivamente). Resultados contrários a esses foram encontrados por Cargnelutti Filho et al. (2003), que ao comparar, em cinco ensaios de competição de cultivares de milho, o método de Papadakis aplicado a um delineamento inteiramente casualizado, sobre um delineamento em blocos ao acaso em que não havia sido aplicado o método, obtiveram uma redução média de 48,1% no erro experimental, ao considerar, no cálculo da covariável pelo método de Papadakis, a parcela de referência e as quatro vizinhas mais próximas a ela.

Valores de  $\phi_{\sigma^2}$  inferiores a 100%, apresentados na Tabela 1 pelos Modelos 7 (93,8%) e 6 (94,5%), indicam que tais modelos, em média, contribuíram 6,2% e 5,5%, respectivamente, para melhorar a eficiência do controle local. Contudo, a estimativa de  $\phi_{\sigma^2}$  do Modelo 8 foi a

mesma que a do Modelo 5 (100%), apenas tendo surgido outro parâmetro (o alcance da correlação espacial,  $\rho$ ). Os Modelos 6 e 7 não apresentaram menores valores de  $\hat{\sigma}^2$  (menores  $\phi_{\sigma^2}$ ), e o Modelo 8 não alterou essa estimativa; tal fato sugere que modelar a média via covariável (por meio de PPD ou MM) é, em geral, um pouco mais eficiente que modelar a própria covariável (por meio da AE).

Observa-se que, nos casos estudados, a AE não trouxe vantagens sobre os delineamentos planejados. Os modelos PPD e MM, apesar de terem contribuído para melhorias, necessitam de adaptações aos métodos espaciais, por exemplo, quanto ao número de vizinhos considerado no cálculo da covariável, pois qualquer ganho no melhoramento é de grande valia, no sentido de se detectar diferenças genéticas e reduzir os custos. Cargnelutti Filho et al. (2003) avaliaram cinco formas diferentes no cálculo da covariável pelo método de Papadakis, com variações nas tendências geográficas (quanto à posição e ao número de parcelas utilizadas), e concluíram que a forma mais eficiente na redução do erro experimental é a que considera cinco parcelas: a de referência e as quatro mais próximas a ela.

É de se esperar também que, no caso de haver maior heterogeneidade intrablocos, o método PPD passe a ser o mais eficiente, por utilizar diretamente os resíduos como covariável.

Os métodos PPD e MM afetaram, no entanto, as estimativas dos componentes das variâncias de progênies ( $\hat{\sigma}_p^2$ ) e de herdabilidade ( $\hat{h}^2$ ), em comparação ao Modelo 5, pois provocaram alterações na estimativa de  $\phi_{\sigma_p^2}$  e  $\phi_{h^2}$  (Tabela 1). Quando se observa os modelos com controle local, verifica-se que a maior alteração em  $\phi_{\sigma_p^2}$  ocorreu no Modelo 7 (6,3%), e que no Modelo 6 encontra-se a maior alteração em  $\phi_{h^2}$  (4,3%). Pode-se suspeitar, então, que os modelos de Papadakis e de médias móveis são menos robustos do que o da análise GMN para o ajuste aos dados, pois o modelo GMN é justificado pela aleatorização do experimento, enquanto MM e PPD são formas particulares de análise espacial, em modelos do tipo auto-regressivo, conforme elucidado por Atkinson (1969).

Os modelos MM e PPD são modelos não-lineares complexos e, além disso, pode estar havendo um “overfitting”, situação em que o controle local extremamente eficiente, realizado com os vizinhos, ocorre devido, parcialmente, a diferenças genotípicas (vale lembrar que foram considerados apenas quatro vizinhos por parcela).

**Tabela 1.** Valores médios de razões ( $\phi$ ) entre as estimativas de variâncias residuais ( $\hat{\sigma}^2$ ), de variâncias de progênies ( $\hat{\sigma}_p^2$ ) e de herdabilidade ( $\hat{h}^2$ ), para cada modelo estudado, considerando-se os 38 experimentos avaliados.

| Modelo           | Descrição do modelo <sup>(1)</sup> | Razão (%)         |                     |              |
|------------------|------------------------------------|-------------------|---------------------|--------------|
|                  |                                    | $\phi_{\sigma^2}$ | $\phi_{\sigma_p^2}$ | $\phi_{h^2}$ |
| 1                | GMN – ausência de controle local   | 111,8             | 96,5                | 82,3         |
| 2                | MM – ausência de controle local    | 102,7             | 99,8                | 91,0         |
| 3                | PPD – ausência de controle local   | 101,6             | 102,6               | 93,4         |
| 4                | AE – ausência de controle local    | 111,8             | 96,5                | 82,3         |
| 5 <sup>(2)</sup> | GMN – presença de controle local   | 100,0             | 100,0               | 100,0        |
| 6                | MM – presença de controle local    | 94,5              | 100,5               | 95,7         |
| 7                | PPD – presença de controle local   | 93,8              | 106,3               | 98,1         |
| 8                | AE – presença de controle local    | 100,0             | 100,0               | 100,0        |

<sup>(1)</sup>GMN: Gauss-Markov Normal; MM: Médias Móveis; PPD: Papadakis; AE: Análise Espacial. <sup>(2)</sup>Modelo considerado como padrão.

Os resultados das altas correlações de Spearman ( $r$ ), encontradas entre os efeitos estimados fixos e aleatórios de progênies, indicaram que praticamente não houve alterações no ordenamento das progênies para seleção (Tabela 2). Isto implica em seleção das mesmas progênies, em termos de produção de grãos, quando seus efeitos são considerados fixos ou aleatórios. Tais resultados sugerem que o número de progênies a serem avaliadas não influencia na escolha da natureza do efeito do tratamento, se fixo ou aleatório, pelos melhoristas.

Duarte (2000) utilizou dados de produção de grãos de 32 experimentos de avaliação de linhagens de soja e comparou, por meio de correlações de Pearson e de Spearman, diferentes modelos, tendo considerado ora

**Tabela 2.** Número de progênies ( $t$ ) e correlações de Spearman ( $r$ ) entre as estimativas de efeitos fixos e aleatórios de progênies, para os experimentos de feijão e milho avaliados em látice.

| Experimento | Cultura | t   | r                 |
|-------------|---------|-----|-------------------|
| 1           | Feijão  | 256 | 1,00**            |
| 2           | Feijão  | 400 | 1,00**            |
| 3           | Feijão  | 400 | 1,00**            |
| 4           | Feijão  | 169 | 1,00**            |
| 5           | Feijão  | 100 | 0,98**            |
| 6           | Feijão  | 36  | 0,99**            |
| 7           | Feijão  | 36  | 0,99**            |
| 8           | Feijão  | 25  | 0,99**            |
| 9           | Feijão  | 25  | 0,99**            |
| 10          | Feijão  | 20  | 0,99**            |
| 11          | Feijão  | 226 | 1,00**            |
| 12          | Feijão  | 81  | 1,00**            |
| 13          | Feijão  | 25  | 0,99**            |
| 14          | Feijão  | 25  | 1,00**            |
| 15          | Feijão  | 16  | 0,97**            |
| 16          | Feijão  | 16  | 1,00**            |
| 17          | Feijão  | 25  | 1,00**            |
| 18          | Feijão  | 25  | 1,00**            |
| 19          | Feijão  | 25  | 1,00**            |
| 20          | Feijão  | 25  | 1,00**            |
| 21          | Feijão  | 25  | 1,00**            |
| 22          | Feijão  | 25  | 0,97**            |
| 23          | Feijão  | 25  | 0,99**            |
| 24          | Feijão  | 25  | -( <sup>1</sup> ) |
| 25          | Feijão  | 25  | 0,98**            |
| 26          | Feijão  | 25  | 1,00**            |
| 27          | Feijão  | 25  | 0,99**            |
| 28          | Feijão  | 25  | 1,00**            |
| 29          | Feijão  | 25  | -( <sup>1</sup> ) |
| 30          | Feijão  | 25  | 1,00**            |
| 31          | Feijão  | 25  | 1,00**            |
| 32          | Feijão  | 25  | 0,98**            |
| 33          | Feijão  | 25  | 0,98**            |
| 34          | Feijão  | 25  | 0,99**            |
| 35          | Feijão  | 25  | 1,00**            |
| 36          | Feijão  | 25  | 1,00**            |
| 37          | Milho   | 144 | 1,00**            |
| 38          | Milho   | 81  | 1,00**            |

(<sup>1</sup>)Valores não calculados, pois as estimativas dos efeitos aleatórios de progênies, para os diferentes tratamentos, foram nulas. \*\*Significativo a 1% de probabilidade.

efeito de linhagens como fixo, ora como aleatório. Na maioria dos experimentos, os modelos selecionaram praticamente os mesmos genótipos, por terem apresentado altas correlações. Contudo, o autor acrescentou que apesar da concordância entre os modelos, quanto ao ordenamento dos genótipos, os valores das respostas genótípicas médias poderão diferir substancialmente entre esses modelos, se a herdabilidade for baixa (inferior a 0,20). O autor ainda alertou que uma escolha equivocada do modelo de análise poderia comprometer a eficiência do programa de melhoramento, no sentido de que materiais genéticos pouco promissores poderiam ser avaliados e mantidos em próximos ciclos seletivos, o que acarretaria em desperdício de tempo e de recursos.

Bueno Filho (1997) também obteve praticamente o mesmo ordenamento dos genótipos, pelas quatro formas de análise avaliadas em seu estudo, tendo considerado as combinações de efeitos para progênie e bloco, respectivamente: 1) fixo, aleatório; 2) fixo, fixo; 3) aleatório, aleatório; e 4) aleatório, fixo.

## Conclusões

1. O modelo de Gauss-Markov Normal (GMN) pode continuar sendo utilizado em delineamentos com controle local, sem perda de precisão, em comparação ao modelo de Análise Espacial (AE); os modelos de Papadakis (PPD) e de média móveis (MM) melhoram a eficiência do controle local.

2. Apesar de ineficiente, a análise espacial não altera as estimativas de componentes da variância de progênies ( $\hat{\sigma}_p^2$ ), nem a herdabilidade ( $\hat{h}^2$ ), em relação ao delineamento planejado, o que leva a uma vantagem conceitual desse modelo. O mesmo não ocorre com os modelos de médias móveis e Papadakis, que alteram essas estimativas e podem levar a efeitos indesejáveis no ordenamento dos genótipos.

3. Representar progênies por efeitos fixos ou aleatórios, nos processos de estimação, não provoca alterações no ordenamento das progênies para seleção.

## Referências

- ATKINSON, A.C. The use of residuals as a concomitant variable. *Biometrika*, v.56, p.33-41, 1969.
- BUENO FILHO, J.S. de S. **Modelos mistos na predição de valores genéticos aditivos em testes de progênies florestais**. 1997. 118p. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

- CARGNELUTTI FILHO, A.; STORCK, L.; LÚCIO, A.D. Ajustes de quadrado médio do erro em ensaios de competição de cultivares de milho pelo método de Papadakis. **Pesquisa Agropecuária Brasileira**, v.38, p.467-473, 2003.
- CRESSIE, N.; HARTFIELD, M.N. Conditionally specified gaussian models for spatial statistical analysis of field trials. **Journal of Agricultural, Biological, and Environmental Statistics**, v.1, p.60-77, 1996.
- DUARTE, J.B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal**. 2000. 293p. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.
- HENDERSON, C.R. Estimation of genetic parameters. **Annals of Mathematical Statistics**, v.21, p.309-310, 1950.
- KRIGE, D.G. A statistical approach to some basic mine evaluation problems on the Witwatersrand. **Journal of the Chemical, Metallurgical and Mining Society of South Africa**, v.52, p.119-139, 1951.
- LITTELL, R.C.; MILLIKEN, G.A.; STROUP, W.W.; WOLFINGER, R.D. **SAS system for mixed models**. Cary, North Caroline: SAS Institute, 1996. 633p.
- PAPADAKIS, J.S. **Méthode statistique pour des expériences sur champ**. Thessalonike: Institut d'Amélioration des Plantes à Salonique, 1937. 30p. (Bulletin, 23).
- RICKEY, F.D. Adjusting yields to their regression on a moving average, as a means of correcting for soil heterogeneity. **Journal of Agricultural Research**, v.27, p.79-90, 1924.
- RIOS, M.C.D. **Alternativas de procedimentos estatísticos para avaliação de genótipos em programas de melhoramento genético vegetal**. 1997. 59p. Dissertação (Mestrado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.
- ROBINSON, G.K. That BLUP is a good thing: the estimation of random effects. **Statistical Science**, v.6, p.15-51, 1991.
- SOUZA, E.A. de. **Alternativas experimentais na avaliação de progênies em programas de melhoramento genético vegetal**. 1997. 122p. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.
- VIVALDI, L.J. Comparação entre métodos de análise espacial de experimentos de campo. **Pesquisa Agropecuária Brasileira**, v.25, p.77-84, 1990.

---

Recebido em 8 de dezembro de 2004 e aprovado em 4 de maio de 2005