Crop Breeding and Applied Biotechnology 10: 147-153, 2010
Brazilian Society of Plant Breeding. Printed in Brazil

Comparison of estimation and prediction methods of genotypic means in maize variety trials

Cristiane Rachel de Paiva Felipe[1*], and João Batista Duarte[2]

Received 14 October 2009

Accepted 28 January 2010

ABSTRACT - *The objective of this study was to assess different statistical approaches to estimating or predicting genotypic means of maize varieties in lattice experiments. The following models were evaluated: fixed model (FF), mixed model with random block effect (AF), mixed model with random treatment effect (FA), random model (AA), and shrinkage James-Stein estimator (JS). Forty-seven experiments were analyzed, each one with three replications and 15 to 28 treatments. The mean of two check cultivars (controls) per growing season was used as reference for the selection. In most experiments, the rate of genotypes selected by the shrinkage approaches (FA, AA and JS) was lower than by the FF and AF models, which also tended to select low-yielding genotypes, even when the genotypic determination coefficient (h^2') was low. At high h^2' levels the genotypes selected by the different approaches were quite coincident, although the ranking differed.*

Key words: Linear mixed model, breeding value, lattice, James-Stein estimator, shrinkage.

INTRODUCTION

In plant breeding programs, the estimation of genotypic means from statistical analysis based on a fixed model is a common practice, even in situations where the model is naturally mixed, eg., when the treatments are obtained by sampling a population (Duarte and Vencovsky 2001). According to Robinson (1991), the BLUP (Best Linear Unbiased Predictor), used in mixed model approaches, is an unbiased predictor ($E[\hat{u}]=E[u]$) with shrinkage effect, since its components are less dispersed than the corresponding least square estimators used to assesses fixed effects. This shrinking results from the fact that the predictor considers a relation of co-variances ($C'V^{-1}$), ranging from 0 to 1, which corresponds to a determination coefficient such as heritability (h^2).

According to Efron and Morris (1977), the James-Stein estimator also has a shrinkage effect on the sample (phenotypic) mean, but no assumptions are required regarding the fixed or random effects, or the distributions of the means to be estimated. Only the unbiasedness condition has to be weakened.

This study aimed to compare different statistical approaches of estimation or prediction of genotypic means and assess their effect on the ranking of test genotypes and the proportion of the selected genotypes to the control cultivars. Five statistical approaches were evaluated: fixed model, mixed model with random block effect, mixed model with random treatment effect, random model and James-Stein estimator.

[1] Emater-GO, Estação Experimental de Senador Canedo, 75250-000, Senador Canedo, GO, Brazil. *E-mail: cristianerachel@yahoo.com.br.
[2] Escola de Agronomia da Universidade Federal de Goiás, CP 131, 74001-970, Goiânia, GO, Brazil.

MATERIAL AND METHODS

Grain yield data from 47 experiments of a network of regional field trials of maize varieties (summer growing seasons of 2002/03, 2003/04, 2004/05, and 2005/06) were used. This trial network is coordinated by the Secretaria da Agricultura, Pecuária e Abastecimento do Estado de Goiás (Seagro-GO), in partnership with Fundação de Desenvolvimento Assistência Técnica e Extensão Rural de Goiás (Fundater).

Three replications were used per treatment in each experiment of the different growing seasons, with the following designs: 5 x 3 lattice, with 15 treatments in 2002/03; 6 x 4 lattice with 24 treatments in 2003/04; 7 x 4 lattice, with 28 treatments in 2004/05; and 6 x 3 lattice with 18 treatments in 2005/06. Each plot consisted of two rows with the following variations: 4.0 m to 5.0 m long, spaced 0.4 m to 1.0 m apart. Individual data of grain yield per plot were adjusted by covariance analysis, as suggested by Vencovsky and Barriga (1992), for an initial population of 60,000 plants per hectare.

To facilitate computer analysis of the different statistical approaches, the lattice design model was rewritten by grouping the block and replication effects in a single variation source of local control, namely:

$$Y_{ij} = m + b_j + t_i + e_{ij} \quad (1)$$

where: Y_{ij} is the observed value of treatment i in block j; m is the overall mean; t_i is the treatment (genotype) effect *i*; b_j is the block effect; and e_{ij} is the experimental error associated with observation Y_{ij}. Thus, index j covers all blocks of the experiment, regardless of the replication.

The statistical approaches were labeled as follows: fixed model (FF), i.e., effects of block and treatment assumed as fixed, corresponding to the intra-block analysis; the mixed model with random effect of blocks (AF), i.e., analysis with recovery of interblock information; the mixed model with random treatment effects (FA), where intergenotypic information is recovered; and the random model (AA), which recovers both information types. The James-Stein estimator was labeled JS.

In model (1), independence among m, b_j, t_i and e_{ij}, and the following distributions were assumed for the respective approaches: $FF - e_{ij} \sim N(0,s^2_e)$; $AF - b_j \sim N(0,s^2_b)$ *and* $e_{ij} \sim N(0,s^2_e)$; $FA - t_i \sim N(0,s^2_g)$ *and* $e_{ij} \sim N(0,s^2_e)$; *and* $AA - b_j \sim N(0,s^2_b)$, $t_i \sim N(0,s^2_g)$ *and* $e_{ij} \sim N(0,s^2_e)$.

In matrix language, the FF model, with fixed block and treatment effects, can be written as: $y = X\beta + \varepsilon$, where *y* is the observation vector; *X* is the incidence design matrix; *b* is the vector of unknown fixed effects (m, b_j and t_i); and *e* is the vector of random error effect, with var(*e*) = $I\sigma^2$. Consequently, var(*y*) = $V = I\sigma^2$. In this case, the vector of genotypic means is a parametric function of the effects on *b*, i.e., L'*b*, of which the Best Linear Unbiased Estimator (BLUE) can be presented as: BLUE(L'β) = L', where L'$\hat{\beta}$ is the matrix of coefficients of fixed effects in that function.

The general description of the mixed and random models (AF, AF and AA) can be given as: $y = X\beta + Zu + \varepsilon$; where *y* is the observation vector; *X* is the incidence matrix of the effects in β, the vector of fixed effects (only μ in model AA) to be estimated; *Z* is the incidence matrix of the effects in *u*, the vector of random effects to be predicted, assumed $u \sim N(\phi,G)$; and ε is the error vector, assumed to be $e \sim N(\phi,R)$. Therefore, E(*y*) = $X\beta$, and var(*y*) = $V = ZGZ' + R$, where *G* and *R* are the variance-covariance matrices of random vectors *u* and *e*, respectively. Thus, the covariances between different vectors are assumed as zero (Henderson 1984). This modeling generalizes any correlation structure between observations, be it by the inclusion of new random factors into the model (*G* matrix), or by the presence of spatial or temporal correlation between the units of observation (*R* matrix). In this case, the following simplified structures were used for these matrices: $G = I\sigma^2_b$, in AF; $G = I\sigma^2_g$, in FA; $G = [I\sigma^2_b \oplus I\sigma^2_g]$ (where $\oplus$ is the direct sum operator) in AA; and also $R = I\sigma^2_e$, in all these cases.

The BLUE of β (vector of fixed effects), obtained as solution of generalized least squares, given by $\hat{\beta} = (X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}y$, results from the solution of the mixed model equations of Henderson (MME). The prediction of random effects (BLUP of *u*) is also derived from this solution expressed as: $\hat{\mu} = \hat{C}'V\text{-}1\ (y - X\hat{b})$, where: $\hat{C}$ is the matrix of covariance between genotypic values (not observable) and observable values (data), also given by: $\hat{C} = \hat{G}Z'$, and $\hat{V}^{-1}$ is the inverse of an estimate of matrix $\hat{V}$. Thus, the vector of genotypic means is a linear function that combines fixed and random effects, i.e., $w = L'\beta + u$, where *L'* is the matrix of coefficients of the fixed effects on the respective function. Searle et al. (1992) showed that, for estimable L'*b*, the predictor of *w* has BLUE properties (minimum mean-square error, linearity for *y* and unbiasedness) and can be calculated as: BLUE $(w) = L'\hat{\beta} + \hat{\mu}$.

The means of the treatments with the James-Stein estimator (JS) were obtained by the expression: $\bar{Y}_i^{JS} = k(\bar{Y}_{i.} - \bar{Y}..) + \bar{Y}$, where *k* is the effect of the shrinkage factor of the sample genotypic mean on the overall mean, given by: $k = 1 - x/F$, where *F* is the value of the Snedecor *F* test for

the variation source of treatments in the analysis of variance applied to model (2) assumed as fixed. The different approaches, with exception of the James-Stein estimator (of easy computation), were performed with proc mixed (procedure for mixed models) of SAS software, using restricted maximum likelihood (REML).

For comparison of the statistical approaches the following aspects were evaluated: i) percentage of selected genotypes, i.e., genotypes with higher estimated or predicted means than the check cultivars (controls); ii) coincidence in genotypic selection with different pairs of statistical approaches; and iii) ranking of genotypes according to the means estimated or predicted by the different procedures. The mean yield of two check cultivars, calculated from their estimates obtained by the fixed model (FF), was used as selection criterion in each growing season. These cultivars were Al Bandeirante, tested in all growing seasons, plus another with high mean yield in each growing season (Saracura in 2002/03, IPR114 in 2003/04, Emcapa202 in 2004/05, and BR106 in 2005/06). The rates of genotypes selected by each approach and the coincidence of these selections between the pairs of procedures were also calculated.

With the estimates of the AA model, the genotypic determination coefficient ($h^{2'}$), a genetic parameter analogous to the intrapopulation heritability in the broad sense, was computed as well:

$$h^{2'} = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_b^2 + \hat{\sigma}_e^2}$$

where $\hat{\sigma}_g^2$, $\hat{\sigma}_b^2$ and $\hat{\sigma}_e^2$ were the estimates of the variance among cultivars, variance among blocks, and residual variance, respectively.

RESULTS AND DISCUSSION

In the 47 experiments, the coefficients of variation ranged from 5.6% to 18.2%, and the mean grain yield from 3,564 kg ha^{-1} to 9,207 kg ha^{-1}. These results are compatible and acceptable in relation to the requirements for maize variety trials in Brazil (Brasil 1998).

In the 2002/03 growing season, in the trials in which the mean yield of the control cultivars was lower than the mean of the experiment, as observed in Morrinhos, Senador Canêdo 2, Inhumas, and Goiatuba, the approaches with shrinkage of the genotypic means (FA, AA and JS) generally tended to result in a higher rate of selected genotypes than in cases where the control means exceeded the experimental means, regardless of the $h^{2'}$value (Figure 1). On the other hand, it should be noted that the use of the models FF and AF always resulted in some selection rate different from zero in the different experiments, even if the $h^{2'}$values were extremely low and the controls and experimental means were reduced.

In the 2003/04 growing season, in the experiments conducted in Porangatu, Senador Canêdo and Ipameri, where the experimental and/or control means were high and $h^{2'}$ values exceeded 43%, the selection rates were the same in the different approaches (Figure 2). Only in Itaberaí, where the experimental mean exceeded the control mean, the selection rate of approaches with shrinkage effect was greater than of the others, although the $h^{2'}$ value was the lowest (17%) of that growing season (17%). In the other cases, selection rates tended to be higher for the models FF and AF, even in the experiments of Campo Alegre and Morrinhos, for example, where the mean yields were close to 4,000 kg ha^{-1}.

In the 2004/05 growing season, the selection rates of the shrinkage approaches were close to the models FF and AF in the trials of Morrinhos and Orizona, where the means of the control cultivars were below the experimental means (Figure 3). Apart from these two experiments, in general, the highest percentages of selected genotypes resulted from the application of the models FF and AF. However, in Porangatu, where the experimental mean exceeded the mean of the control cultivars, the selection rates of the different approaches did not differ from each other.

As in the other growing seasons, in most trials in 2005/06, the highest selection rates resulted from the use of models FF and AF (Figure 4). In the experiment of Porangatu, the only trial where the experimental mean exceeded the control mean, the shrinkage approaches resulted in higher selection rates. The trial of Senador Canêdo, with $h^{2'}$ high values, was also the only experiment where the selection rates of the procedures did not differ.

A comparison of the rates of genotypes selected by the different procedures in each growing season showed that, except in 2002/03, the James-Stein estimator usually produced the lowest rates, with similar values to the FA and AA models. Thus, using these three approaches reduced the selection rate of genotypes in experiments with a mean below the respective control mean. On the other hand, as confirmed by Duarte and Vencovsky (2001), when the overall mean of genotypes under selection exceeded the control mean, the shrinkage estimators

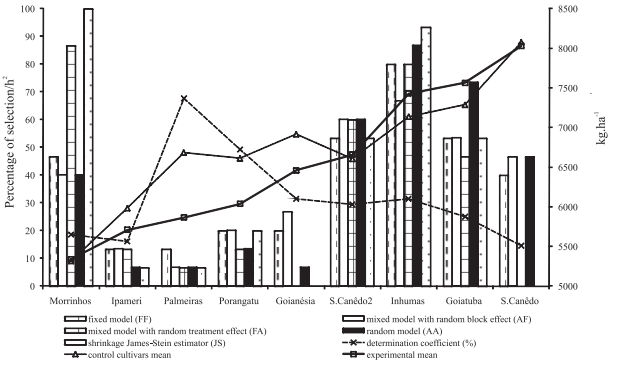

Figure 1. Rate of genotypes selected in nine maize variety trials conducted in the state of Goiás, Brazil (2002/03), by five statistical approaches to estimate or predict genotypic means, in comparison with the experimental mean, control cultivar mean and determination coefficient (h^2').

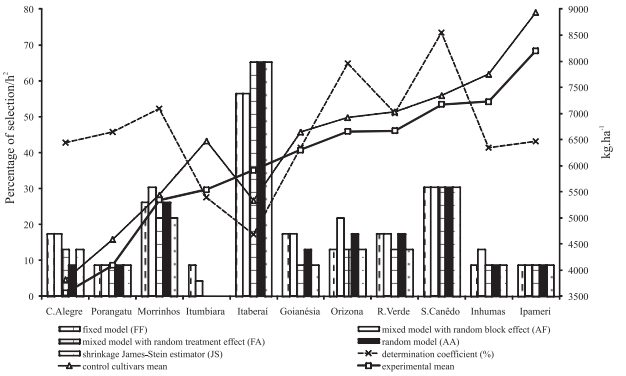

Figure 2. Rate of genotypes selected in 11 maize variety trials conducted in the state of Goiás, Brazil (2003/04), by five statistical approaches to estimate or predict genotypic means, in comparison with the experimental mean, control cultivar mean and determination coefficient (h^2').

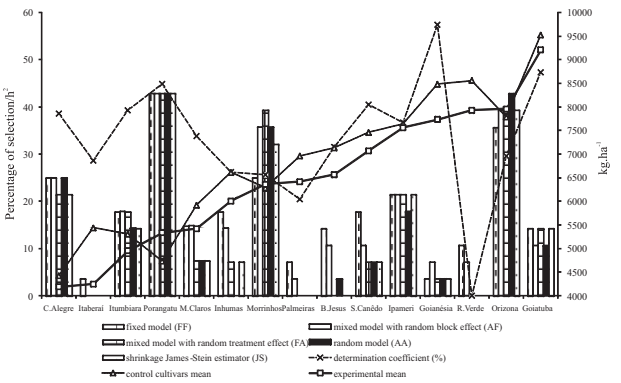

Figure 3. Rate of genotypes selected in 15 maize variety trials conducted in the state of Goiás, Brazil (2004/05), by five statistical approaches to estimate or predict genotypic means, in comparison with the experimental mean, control cultivar mean and determination coefficient (h^2').

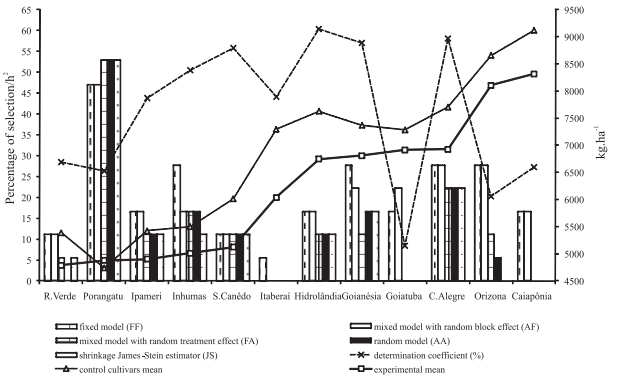

Figure 4. Rate of genotypes selected in 12 maize variety trials conducted in the state of Goiás, Brazil (2005/06), by five statistical approaches to estimate or predict genotypic means, in comparison with the experimental mean control cultivar mean and determination coefficient (h^2').

resulted in a higher percentage of selected genotypes. According to the authors, the approach of random treatments tends to protect promising populations, selecting more genotypes, in contrast to those of low potential (lower mean of genotypes under selection than the mean of control cultivars). In these cases, these models tended to exclude a higher proportion of test genotypes from the selection. This behavior was observed in all growing seasons, indicating that such approaches result in a greater preservation of genotypes, when their means are higher than of the check cultivars. For this reason, these approaches should be preferred, since they indicate a higher selection rate of promising genotypes. Several studies emphasize the preference for shrinkage estimators or predictors, even when the effects under study are assumed as fixed by traditional approaches (Duarte and Vencovsky 2001, Smith et al. 2001, Costa 2002, Resende and Duarte 2007).

Comparing the results of the different trials, it was concluded that in situations of high h^2' values (above 70%), no differences are observed in the selection rates of the different statistical approaches. This highlights the importance of knowing the heritability estimates of the traits to be assessed, since the use of models in which the genotype effects are accepted as fixed increases the risk of selecting low-yielding genotypes in trials with low heritability. This may lead to a waste of resources in the breeding program, reducing the efficiency of genotypic evaluation and increasing the time required for the release of a cultivar.

Regarding the similarities between the genotypes selected by the different approaches, when the experiments were grouped in terms of h^2', it was found that the FF and AF models were most coincident. This was not surprising, since the treatment effects in both models had been estimated under the same assumption. For h^2'>70%, the coincidence balance of selected genotypes between the different pairs of approaches is greater, with 100% coincidence of FF, AF, AA, and JS.

In all experiments, the different statistical approaches led to some change in the ranking of genotypic means, confirming findings of other authors (Federer and Wolfinger 1998, Duarte and Vencovsky 2001, Resende and Duarte 2007). Summing up, the breeder's choice of one approach instead of another can determine the exclusion of different genotypes from the breeding program. If a less efficient approach is chosen, there is a higher risk of carrying genotypes with low probability of success to the following selection cycle, due to the low accuracy in the inference of genotypic means, associated to the chosen method.

In situations as in the experiment of Senador Canêdo (Figure 5), where h^2'=15%, the shrinkage effect of the JS estimator and FA model would lead to the exclusion of all test genotypes. Thus, aside from the differences in the number of selected genotypes by the approaches (the models AF and FF achieved highest rates), there was also change in the ranking of genotypic means, as reported in other studies as well (Federer and Wolfinger 1998, Duarte and Vencovsky 2001, Resende and Duarte 2007).

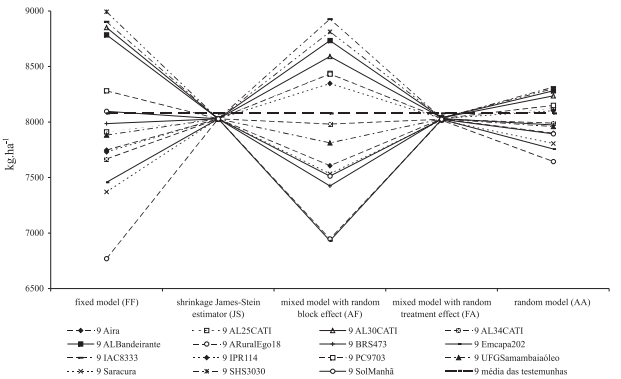

Figure 5. Mean yields of genotypes (kg ha^{-1}) in a variety experiment conducted in Senador Canêdo, state of Goiás, in the 2002/03 growing season, obtained by the following statistical approaches: fixed model (FF), mixed model with random block effect (AF), mixed model with random treatment effect (FA), random model (AA) and estimator James-Stein (JS), compared with the mean of two control cultivars.

In the case of zero value for h^2', as in the trial of Rio Verde (Figure 6), where the absence of genotypic variability indicated by JS, FA and AA would lead to the exclusion of all genotypes, the FF and AF models still selected six and four cultivars, respectively. This fact would be inadmissible if the parametric heritability were really zero. However, as the genotypes were derived from different breeding programs and due to the differences in the h^2'values observed in other experiments, in the same growing season, it is likely that the adjustments of means by JS, FA and AA approaches are not valid for this situation.

Tomé et al. (2002) compared different statistical approaches to data analysis in lattice through simulation, and observed that the FA and AA models have the highest percentages of non-valid adjustments. However, increases in the percentage of valid adjustments based on the FA model were observed with increasing experimental precision; in other words, the non-zero estimates of variance components increased, unlike in the case of AA and AF. In a simulation study associated with this research (Felipe et al. 2008), the FA, AA and JS procedures (with shrinkage effect) resulted in 14.5%, 11.9% and 13.5% of non-valid adjustments, respectively, when the experiments were small (with 15 treatments) and h^2' values between 6% and 48%. On the other hand, when the h^2' values were between 63% and 82% (experiments with higher genotypic variance), the occurrence of non-valid adjustments in these approaches was no longer observed.

These results suggest that, in the Rio Verde experiment, despite the coefficient of variation of 13.98%, the low variances of genotypes and/or blocks led to problems in estimating the means by FA, AA and JS approaches. Thus, in this context such approaches should perhaps be revoked for the options FF and FA. However, based on the findings of Tomé et al. (2002) and Felipe (2010), the breeder must also remember that the FF model never leads to invalid adjustments, but due to its nature rather than because of its ability to estimate parameters. Therefore, if the parametric genotypic variance is actually zero under these environmental conditions, the choice of the approaches FF and AF would lead to the maintenance of low-yielding genotypes in the breeding program, that is, of genotypes with means below the mean of the control cultivars. As already mentioned, this would result in losses in human and financial resources for the experimentation, and prevent the release of promising cultivars at the end of the selection cycle.

In general, there was a strong shrinkage effect on the genotypic means obtained by the approaches AA, AF and JS, in the different experiments and growing seasons. Only in the experiments in which h^2' reached values close to or above 70% this phenomenon was no longer significant. An example is the experiment of Senador Canêdo, in 2003/04, with h^2'= 73%. Such situations, where the coefficient h^2' is high and the means estimated or predicted by different statistical approaches are identical, or nearly identical, are also reported in the literature (Duarte and Vencovsky 2001, Tomé et al. 2002).

The following conclusions of this study are noteworthy:

i) Different statistical approaches to estimate or predict genotypic means commonly result in different rates of selected genotypes in relation to check cultivars, as well as in differences in the genotype ranking based on these means.

ii) The number of genotypes selected relation to check cultivars by shrinkage estimators or predictors (with random genotypic effect) is smaller than by models with fixed genotypic effect, especially when the mean this cultivars exceeds the experimental mean.

iii) In experiments with high h^2' values, the different approaches result in quite similar selections, although differences in the ranking are possible.

Comparação de métodos de estimação e predição de médias genotípicas em ensaios varietais de milho

RESUMO - *O objetivo deste trabalho foi avaliar diferentes abordagens estatísticas para estimação/predição de médias genotípicas de variedades de milho de experimentos delineados em látice. Foram avaliados: o modelo fixo (FF); o modelo misto com efeito aleatório de blocos (AF); o modelo misto com efeito aleatório de tratamentos (FA); modelo aleatório (AA); e o estimador shrinkage de James-Stein (JS). Foram analisados 47 experimentos, com três repetições e número de tratamentos variando entre 15 e 28. A média de duas cultivares (testemunhas) por safra foi utilizada como referência para a seleção. Na maioria dos experimentos, a adoção de abordagens shrinkage (FA, AA ou JS) levou a menor taxa de genótipos selecionados do que os modelos FF e AF, os quais ainda tiveram tendência de selecionar genótipos pouco promissores, mesmo sob reduzidos coeficientes de determinação genotípica (h^2'). Sob elevados valores de h^2', a seleção de genótipos pelas diferentes abordagens foi bastante coincidente, porém, com diferenças no ordenamento dos genótipos.*

Palavras-chave: Modelo linear misto, valor genético, látice, estimador James-Stein, *shrinkage*.

REFERENCES

Brasil (1998) Portaria nº 294, de 14 de outubro de 1998. Estabelece os critérios mínimos a serem observados nos ensaios para determinação do Valor de Cultivo e Uso - VCU de cultivares de algodão, arroz, batata, feijão, milho, soja, sorgo e trigo e os respectivos formulários de solicitação de inscrição de cultivares no Registro Nacional de Cultivares - RNC. **Diário Oficial [da] República Federativa do Brasil**. Available at <http://extranet.agricultura.gov.br/>. Access on Feb 19, 2010.

Costa RB, Resende MDV, Gonçalves PS and Silva MA (2002) Individual multivariate REML/BLUP in the presence of genotype x environment interaction in rubber tree (*Hevea*) breeding. **Crop Breeding and Applied Biotechnology 2**: 131-140.

Duarte JB and Vencovsky R (2001) Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. **Scientia Agricola 58**: 109-117.

Efron B and Morris C (1977) Stein's paradox in statistics. **Scientific American 236**: 119-127.

Federer WT and Wolfinger RD (1998) SAS code for recovering intereffect information in experiments with incomplete block and lattice rectangle designs. **Agronomy Journal 90**: 545-551.

Felipe CRP, Duarte JB and Camarano LF (2010) Estratificação ambiental para avaliação e recomendação de variedades de milho no Estado de Goiás. **Pesquisa Agropecuária Tropical 40**: 14p.

Henderson CR (1984) **Applications of linear models in animal breeding**. University of Guelph, Guelph, 462p.

Resende MDV and Duarte JB (2007) Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical 37**: 182-194.

Robinson GK (1991) That BLUP is a good thing: the estimation of random effects. **Statistical Science 6**: 15-51.

Searle SR, Casella G and Mcculloch CE (1992) **Variance components**. John Wiley, New York, 528p.

Smith A, Cullis B and Gilmour A (2001) The analysis of crop variety evaluation data in Australia. **Australian & New Zealand Journal of Statistics 43**: 129-145.

Tomé MA, Bueno Filho JSS and Ferreira DF (2002) Blocos fixos ou aleatórios? O caso dos ensaios em látice no melhoramento vegetal. **Ciência e Agrotecnologia 26**: 576-584.

Vencovsky R and Barriga P (1992) **Genética biométrica no fitomelhoramento**. Revista Brasileira de Genética, Ribeirão Preto, 496p.