



## Use of REML/BLUP for the selection of sugarcane families specialized in biomass production

Márcio Henrique Pereira Barbosa<sup>1\*</sup>, Marcos Deon Vilela de Resende<sup>2</sup>, Luiz Alexandre Peternelli<sup>3</sup>, José Antônio Bressiani<sup>4</sup>, Luis Cláudio Inácio da Silveira<sup>1</sup>, Felipe Lopes da Silva<sup>1</sup>, and Izabel Cristina Rodrigues de Figueiredo<sup>1</sup>

Received 16 December 2003

Accepted 25 July 2004

**ABSTRACT** - One hundred and fourteen full-sib families were evaluated at two sites in Minas Gerais in experiments in augmented blocks and the data analyzed by the REML/BLUP method. Results showed a low genetic variability among families, low dominance variance, and absence of families  $\times$  sites interaction. The mean components estimated via BLUP permitted the selection of superior families and parents. The use of the augmented block design without replication in family selection experiments did not prove adequate due to the low estimates of selective accuracy and family mean heritability. The adoption of other designs that use a higher number of replications is desirable to increase the efficiency of family selection for this population. Owing to the large number of families and the use of the REML/BLUP procedure, using or not using the checks in the augmented block design did not alter the estimated genetic parameters.

Key-words: *Saccharum* spp., mixed model, family selection, parent selection, components of variance.

### INTRODUCTION

The sugar and alcohol sector holds an expressive potential for the production of renewable biomass. An efficient thermoelectric power plant can set an approximate 40% overspill of bagasse aside for other uses, which may include the sale of the surplus to other thermoelectric plants. Apart from the bagasse excess, the total or partial removal of mechanical trash has been a further aim. There are research reports in Brazil on the viability of collecting the residues (Copersucar 1998) as well as on the quantification of available energy in harvest residues (Furlani Neto et al. 1997) for thermal recycling. The residues are also suitable for alcohol production.

On these premises, sugarcane cultivars with a high biomass production potential should be developed by genetic improvement programs. It is therefore necessary to define populations for reciprocal recurrent selection programs, in view of the importance of the non-additive effects for stalk yield per hectare (Bastos et al. 2003). The identification of new superior families is also important, which will furnish superior clones for commercial purposes.

Clustering clones in different specialized populations and the definition of the crosses could be realized by a continuous process of obtaining estimates of the general (GCA) and the specific (SCA) combining ability by means

[Departamento de Fitotecnia, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brasil. \*E-mail: barbosa@ufv.br

<sup>2</sup>Embrapa Florestas, Estrada da Ribeira, km 111, 83411-000, Colombo, PR, Brasil

<sup>3</sup>Departamento de Informática, UFV, 36570-000, Viçosa, MG, Brasil

<sup>4</sup>Centro de Tecnologia Copersucar, C.P. 162, 13400-970, Piracicaba, SP, Brasil

of the evaluation experiments of full-sib families along a sequence of years. Such experiments are normally unbalanced. Consequently, the best alternative would be to use the Best Linear Unbiased Prediction-BLUP method to obtain the GCA and SCA estimates. The BLUP or mixed model methodology was developed by animal breeders and has been applied in forest improvement and also to some fruit species (Resende 2001). In sugarcane, there are first reports on the use of BLUP in the BSES program for sugarcane improvement in Australia (Stringer 2000).

To estimate variance components or genetic parameters with unbalanced data, the standard procedure is the restricted maximum likelihood (REML) method. Certain experimental designs like the augmented blocks are, by construction, unbalanced. REML is particularly indicated for such a design.

The present study aimed at an application of REML/BLUP to estimate the variance components and parent selection of sugarcane families specialized for biomass production and evaluated in the augmented block design.

## MATERIAL AND METHODS

### Plant material

The progenies used in this study were samples from crosses realized by COPERSUCAR, in Camamu, State of Bahia, and at the Universidade Federal de Alagoas, in Murici, State of Alagoas, in 1999. The obtained progenies were planted at the Centro de Pesquisa e Melhoramento da Cana-de-Açúcar - CECA of the Universidade Federal de Viçosa, State of Minas Gerais, in the county of Oratórios (lat 20° 25' S, long 42° 48' W, altitude 494m asL).

Seedlings were obtained and transplanted to the field according to the methodology described by Barbosa (2000); the latter process occurred in December 1999. The plants developed up to a height of approximately one meter and were cut without selection in July 2000. In February 2001, one stalk per clump with more than six buds each was sampled for the experiments. The stalks were collected from 72 clumps from each one of the 112 families used in the experiments. This procedure was repeated twice to obtain seedlings for both experiments. Besides these 112 families, seedlings from the families SP80-180 x SP847017 and SP84-1192 x RB855035 were obtained, and used as checks.

After extracting the seedlings for the experiments, the plants on the mentioned fields were left to grow. In May 2001, mass selection was realized for ratoons. The selected genotypes could have been planted in experiments for the evaluation and selection of clones, though such experiments were not used in the present study.

### Experimental details

Two experiments were installed, one in the distillery Atenas, in São Pedro dos Ferros, MG, and the other on the experimental campus for agronomy of the Universidade Federal de Viçosa.

The experimental design was of augmented blocks (Federer 1956) with equal features at both sites, described in the following. The plots consisted of three four-meter rows. For all treatments, buds were planted in a density of 18 a meter, using the previously prepared canes with six buds each. Hence, each family was represented in the plot by a mixture of stalks from 72 different genotypes.

Each block consisted of 14 regular treatments represented by families of full-sibs plus four checks. Two kinds of checks were used: two clones represented by the cultivars RB72454 and SP80-1816, and two families, SP80-180 x SP847017 and SP84-1192 x RB855035.

The rows were spaced 1.40 m apart, while the plots were interrupted by a two-meter alley at every four meters. Only outer border plots were used. The plantation was fertilized with 600kg per hectare of the formula containing 4% N, 30% P<sub>2</sub>O<sub>5</sub>, and 16% K<sub>2</sub>O in the Distillery Atenas, and 500kg per hectare of the formula containing 0% N, 25% P<sub>2</sub>O<sub>5</sub>, and 25% K<sub>2</sub>O in Viçosa. For ratoon fertilization, 600 kg per hectare of the formula containing 20% N, 0% P<sub>2</sub>O<sub>5</sub>, and 20% K<sub>2</sub>O were used at both sites. Weeds were controlled by herbicides and manual weeding.

In July 2003, all plants of the plot were cut after previous clearing by burning. The stalks were collected manually and their tops removed. The total weight of the plot was measured by a dynamometer and then transformed to tons of stalks per hectare - TSH. The sugarcane plant was not evaluated. This means that the evaluations at both sites were limited to ratoon.

### Data analysis

The statistical analyses were carried out with the genetics and statistics computer software Selegen-Reml/Blup (Resende 2002a).

The mixed model equations (Resende 2002b) were used to calculate the BLUPs of the genetic values and the specific combining ability - SCA of each family for TSH, considering the genetic relationship matrix described below.

### Mixed linear model

$y = Xl + Za + Wc + Qga + Ub + e$ , where:

$y$ ,  $l$ ,  $a$ ,  $c$ ,  $ga$ ,  $b$  and  $e$ : data vectors of the fixed effects of the sites, of the random additive genetic effects, of the random SCA effects, of the random effects of the interaction genotype x

environment, of the random block effects, and random errors, respectively.

$X, Z, W, Q$  and  $U$ : matrices of incidence of  $l, a, c, ga, b$ , and  $e$ , respectively.

**Distributions and structures of means and variances**

- $y \mid l, V \sim N(X, V)$
- $a \mid A, \sigma_a^2 \sim N(0, A\sigma_a^2)$
- $c \mid \sigma_c^2 \sim N(0, I\sigma_c^2)$
- $b \mid \sigma_b^2 \sim N(0, I\sigma_b^2)$
- $ga \mid \sigma_{ga}^2 \sim N(0, I\sigma_{ga}^2)$
- $e \mid \sigma_e^2 \sim N(0, I\sigma_e^2)$
- $Cov(a, e) = 0; Cov(a, ga) = 0; Cov(a, b) = 0; Cov(a, e) = 0$
- $Cov(c, ga) = 0; Cov(c, b) = 0; Cov(c, e) = 0; Cov(ga, b) = 0$
- $Cov(ga, e) = 0; Cov(b, e) = 0, \text{ ou seja:}$

$$E \begin{bmatrix} y \\ a \\ c \\ ga \\ b \\ e \end{bmatrix} = \begin{bmatrix} Xl \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad e \quad Var \begin{bmatrix} a \\ c \\ r \\ b \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 & 0 & 0 & 0 \\ 0 & I\sigma_c^2 & 0 & 0 & 0 \\ 0 & 0 & I\sigma_{ga}^2 & 0 & 0 \\ 0 & 0 & 0 & I\sigma_b^2 & 0 \\ 0 & 0 & 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

$Var(y) = ZA\sigma_a^2Z' + WI\sigma_c^2W' + QI\sigma_{ga}^2Q' + UI\sigma_b^2U' + I\sigma_e^2$ , where  $A$  is the matrix of the additive genetic relationship between the parents used in the cross.

**Mixed model equations**

$$\begin{bmatrix} X'X & X'Z & X'W & X'Q & X'U \\ Z'X & Z'Z + A^{-1}\lambda_1 & Z'W & Z'Q & Z'U \\ W'X & W'Z & W'W + I\lambda_2 & W'Q & W'U \\ Q'X & Q'Z & Q'W & Q'Q + I\lambda_3 & Q'U \\ U'X & U'Z & U'W & U'Q & U'U + I\lambda_4 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{c} \\ \hat{ga} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ W'y \\ Q'y \\ U'y \end{bmatrix}, \text{ where}$$

$$\lambda_1 = \frac{\sigma_a^2}{\sigma_e^2} = \frac{1 - h^2 - c^2 - ga^2 - b^2}{h^2};$$

$$\lambda_2 = \frac{\sigma_c^2}{\sigma_e^2} = \frac{1 - h^2 - c^2 - ga^2 - b^2}{c^2};$$

$$\lambda_3 = \frac{\sigma_{ga}^2}{\sigma_e^2} = \frac{1 - h^2 - c^2 - ga^2 - b^2}{ga^2};$$

$$\lambda_4 = \frac{\sigma_b^2}{\sigma_e^2} = \frac{1 - h^2 - c^2 - ga^2 - b^2}{b^2};$$

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_c^2 + \sigma_{ga}^2 + \sigma_b^2 + \sigma_e^2} : \text{ narrow-sense individual heritability}$$

$$c^2 = \frac{\sigma_c^2}{\sigma_a^2 + \sigma_c^2 + \sigma_{ga}^2 + \sigma_b^2 + \sigma_e^2} : \text{ determination coefficient of the effects of the specific combining ability}$$

$ga^2 = \frac{\sigma_{ga}^2}{\sigma_a^2 + \sigma_c^2 + \sigma_{ga}^2 + \sigma_b^2 + \sigma_e^2}$  : determination coefficient of the effects of the genotype x environment interaction

$b^2 = \frac{\sigma_b^2}{\sigma_a^2 + \sigma_c^2 + \sigma_{ga}^2 + \sigma_b^2 + \sigma_e^2}$  : correlation due to the common environment of the block

The variance components were obtained by the REML method and used to calculate the heritability estimates at the individual level and at the level of full-sib family means.

**Iterative estimators of the variance components by REML via EM algorithm**

$$\hat{\sigma}_e^2 = [y'y - \hat{l}'X'y - \hat{a}'Z'y - \hat{c}'W'y - \hat{ga}'Q'y - \hat{b}'U'y] / [N - r(x)]$$

$$\hat{\sigma}_a^2 = [\hat{a}'A^{-1}\hat{a} + \hat{\sigma}_e^2 \text{tr}(A^{-1}C^{22})] / q$$

$$\hat{\sigma}_c^2 = [\hat{c}'\hat{c} + \hat{\sigma}_e^2 \text{tr} C^{33}] / s_1;$$

$$\hat{\sigma}_{ga}^2 = [\hat{ga}'\hat{ga} + \hat{\sigma}_e^2 \text{tr} C^{44}] / t;$$

$$\hat{\sigma}_b^2 = [\hat{b}'\hat{b} + \hat{\sigma}_e^2 \text{tr} C^{55}] / s_2, \text{ where:}$$

$C^{22}, C^{33}, C^{44}$  and  $C^{55}$  come from the inverse of  $C$

$C$  : matrix of the coefficients of the mixed model equations

$\text{tr}$  : trace of a matrix operator

$r(x)$  : rank of the  $X$  matrix

$N, q, s_1, t$  and  $s_2$ : total number of data, of parents, of crosses, of elements of  $g \times e$  interaction, and of blocks, respectively

The estimator of the component of dominance variance between families is given by  $\hat{\sigma}_d^2 = \hat{\sigma}_c^2$ , in other words, it is equal to the variance component associated to the specific combining ability. In this case,  $\hat{\sigma}_d^2$  is equivalent to 1/4 of the genetic variance of total dominance in the population.

**Selection procedures**

Two selection procedures were considered: of parents and of families. The parents were organized and selected by their underlying predicted additive genetic values ( $a$ ). The families were selected by means of their genotypic values, predicted by  $(1/2)(\hat{a}_i + \hat{a}_j) + \hat{c}_{ij}$ , where  $\hat{a}_i$  is the predicted additive genetic value of parent  $i$ , and  $\hat{c}_{ij}$  is the SCA of the cross between parents  $i$  and  $j$ .

**Normality data test**

To verify the normality, the Quantile-Quantile plot test proposed by Johnson and Wichern (1988) was applied to the data set.

**Strategies of statistical analysis**

Three strategies were considered for the REML/BLUP joint analysis in relation to the use of the check treatments. In

The first analysis strategy, the checks were not applied. In the second, cultivars RB72454 and SP80-1816 were used as checks. For the third analysis strategy, two full-sib families, SP80-180 x SP847017 and SP84-1192 x RB855035 were used as checks. Analyses with and without the use of the genetic relationship matrix were carried out.

## RESULTS AND DISCUSSION

### Family selection

The genotypic values for TSH of each family were calculated considering the genetic relationship matrix based on the REML/BLUP joint analysis of the two sites. In this case, 114 full-sib families were evaluated in the augmented block design, where the families SP80-180 x SP84-7017 and SP84-1192 x RB855035 were used as checks in all blocks. For this analysis, all effects were considered random, with exception of the sites.

Figure 1 illustrates the data dispersion of the evaluated population. The histogram displays reasonable symmetry anticipating that the genotypic values for TSH adjusted to the normal distribution. The correlation between the normal standardized values and the genotypic values for TSH was 0.9868, which indicates that the data adjusted to a normal distribution at  $P > 0.05$ .

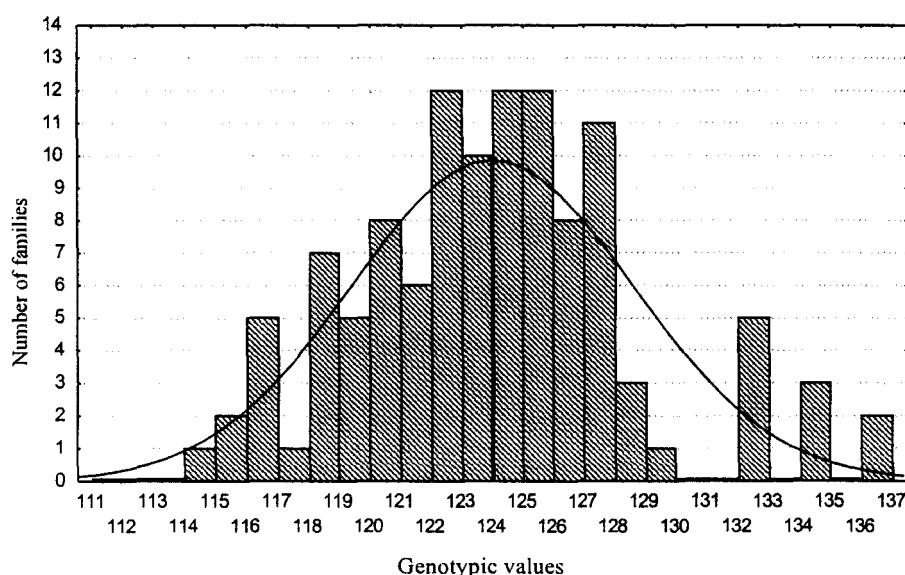
Based on the frequency distribution (Figure 1), 33 families with TSH genotypic values above 126 were detected. The probability to obtain families in this population with a

productivity above 126 TSH was 35.57%, using the properties of the normal distribution.

It is worth underlining that in the present study no data were collected at the individual plant level, in view of the fact that the plots consisted of a clone mixture sampled from a control population. It was therefore not possible to draw any conclusion on the genetic variability among individuals within families for the selection effect. Sampling at the plant level is too labor-intensive to be implanted as routine procedure in improvement programs. On this account, an individual selection using individual and family information (individual BLUP) is restricted at this stage of the sugarcane improvement, as pointed out.

An alternative method to the use of the combined is the selection used by McRae et al. (1998) in a genetic improvement program for sugarcane in Australia. The productivity data are obtained at the sugarcane plant stage by mechanical harvest. A visual selection is realized and clones within the best families obtained for the ratoon. As proposed by Cox et al. (2000), the selection percentage among families is 40%. Four groups were formed from this total, classifying the families based on the sugar productivity per hectare. For each group of superior families, a selection of 40%, 30%, 20%, and 10% within families, respectively, was used. This proceeding was designated sequential selection.

In these premises, and furthermore stressing the difficulties of implementing selection by the individual BLUP in sugarcane improvement, Bressiani (2001) suggested a



modification of the method of sequential selection. The author estimated different percentages of within-family selection for the main components of stalk yield and also for Brix, associated to each family. In his doctoral thesis, he therefore presented the straight lines of linear equations based on the family classification for each evaluated trait, by means of which it is possible to obtain the percentages of within selection for each family. The expected gains with the original sequential, the modified and the individual BLUP selection (via selection by an index) were largely similar to the conditions of that experiment, when considering individual selection in direct form.

#### Parent selection

An additional advantage of the evaluation experiments of families is that, besides providing estimates of the genotypic value for selection of the best families, they also allow the estimation of the additive genetic effects of the crosses' parents.

This last alternative is very important for the planning of crosses. The additive genetic effects can be obtained in evaluation experiments of half-sib as well as in experiments with full sib families. In the present study, not all possible cross combinations between the parents displayed in Table 1 were established, and the parents participated in different numbers of crosses due to the difficulty of coincidence in flowering. This fact made it troublesome to follow the traditional designs of crosses and partial and circulant

diallels. Therefore, the procedure used to obtain the genetic additive values in the present study was BLUP.

The additive genetic effects allow conclusions on the general combining ability - OCA of the parent in focus. The parents with the largest OCA for TSH were observed to be those with the highest additive genetic values (Table I). Clone SP82-6108 with the highest additive genetic value (14.64) in this study was also the genotype with the greatest OCA identified by Bastos et al. (2003). In this case, the authors used, besides SP82-6108, the clones SP84-2029, SP86-96, SP88-754 and SP88-819 and the cultivars IAC86-2210, SP81-1763, SP87-365 in a complete diallel scheme.

Cultivar IAC86-2210 presented the second-highest additive genetic effect (Table I) and was one of the four most used clones in the cross campaign of the genetic improvement program of Copersucar in 1997. In the article of Bastos et al. (2003), this parent did not contribute to the TSH increase of its progenies. Cultivar SP87-365, which was the third classified in the present study regarding the additive genetic value, was one of the six most used genotypes in the cross programs of Copersucar in 1999 and 2000. On the other hand, Bastos et al. (2003) verified that this cultivar contributed significantly to reduce the TSH mean of its progenies.

As reported earlier, the unbalance is evident in this study. As an example, it would be enough to cite that the parent SP81-5450 participated in 13 crosses, five times as female and eight times as male parent. On the other hand,

**Table 1.** Additive genetic effects (a) predicted by the BLUP procedure for the parents of 114 full-sib families of sugarcane evaluated at two sites for tons of stalks per hectare

Genotype	a	Genotype	a	Genotype	a	Genotype	a	Genotype	a
SP82-6108	14.64	SP84-7017	1.70	RB855598	-0.39	RB855036	-2.32	RB855002	-5.08
IAC86-2210	10.89	SP88-721	1.56	SP80-1836	-0.60	SP83-1483	-2.56	NCO310	-5.34
SP87-365	10.21	SP86-155	1.35	SP88-607	-0.86	SP83-2847	-2.81	SP84-5124	-5.79
SP88-754	9.93	SP80-3480	1.29	RB835089	-0.88	RB75126	-2.82	SP88-819	-7.18
RB835205	7.66	SP79-2233	1.24	NA56-79	-1.03	RB855584	-2.92	SP82-3530	-7.24
SP80-185	6.46	SP80-3280	1.04	RB855046	-1.13	SP86-45	-3.37	SP81-3250	-7.32
SP81-1763	6.38	SP83-1226	0.81	SP80-1816	-1.53	RB815521	-3.45	TUC71-7	-7.48
SP80-1230	5.98	RB855035	0.76	SP80-1842	-1.55	RB855536	-4.14	RB855113	-7.55
RB825548	5.55	SP88-797	0.73	SP70-1143	-1.73	RB72454	-4.20	RB845239	-7.58
SP85-162	4.89	RB835486	0.70	SP80-4439	-1.73	RB855181	-4.22	SP84-2025	-7.96
SP86-96	4.87	SP81-5450	0.65	RB825336	-1.75	RB815627	-4.31	L60-14	-8.05
SP80-180	4.59	RB855156	0.61	SP84-5019	-1.78	SP84-2029	-4.60	SP85-7227	-8.88
RB865526	4.43	SP79-1011	0.60	SP81-306	-1.84	SP87-425	-4.64	SP86-91	-9.85
SP77-5181	3.35	SP80-144	0.43	SP84-2268	-2.24	SP81-3251	-4.97	RB845210	-12.85
TUC77-42	2.12	SP84-1192	0.22						

clones NA56-79, NC0310, RB72454, RB75126, RB815521, TUC77-42, RB835089, RB855002, RB855113, RB855156, RB85518L, RB855598, SP70-1143, SP80-1836, SP83-1226, SP83-1483, SP86-155, SP86-45, SP88-797, and TUC71-7 participated in only one cross each. Despite the results of the BLUP procedure are unbiased, and its predicted genetic values adjusted, the aforementioned unbalance contributed to different accuracies in the prediction of the genetic values and increased the predicted genetic values of the most tested compared to the least tested parents (Resende 2002b).

In this context, the practical importance of obtaining the additive genetic values becomes evident. These should, in turn be used to choose future crosses concomitantly with other criteria such as endogamy and the association of traits of agronomical interest. However, the best estimates of additive genetic values are those obtained in evaluation experiments of half-sib families or even of controlled crosses, though with the same number of crosses per parent. In this case, the associated error would be the same for all estimates.

This goal could be attained by *Topcross* experiments, similarly to achievements in genetic improvement of maize. For the interpopulational improvement of sugarcane, clones could be used in crosses with 2 or 4 testers of other populations, instead of the lines. The clones with the greatest combining ability in the *Topcross* experiments would be used in the next cross campaign to obtain hybrid families. At this point, the sequential selection system would be applied as suggested by McRae et al. (1998), or the modified sequential system according to Bressiani (2001).

**Strategies of statistical analysis of evaluation family experiments in augmented block design**

The three strategies for REMLIBLUP joint analysis in relation to the use of checks are presented in Table 2. It is worth highlighting that, since the original proposition of the augmented block design by Federer in 1956 until the 90s, the intrablock analysis prevailed strongly. Bearzoti (1994) showed that the genetic variances are underestimated in the

**Table 2.** Estimates of the variance components and genetic parameters in the joint analysis of full-sib families of sugarcane design at two sites

Parameters	REML/BLUP analysis strategy for an augmented block design using or not using checks with an equal or different genetic structure to the regular treatments		
	(Strategy 1) Absence of checks	(Strategy 2) Clones RB72454 and SP80-1816 as checks	(Strategy 3) Two full-sib families as checks
$\hat{\sigma}_a^2$	188.6377	198.5425	143.9033
$\hat{\sigma}_d^2$	0.0130	0.0112	0.1956
$\hat{\sigma}_{ga}^2$	0.0145	0.0118	0.2670
$\hat{\sigma}_e^2$	1082.9894	1034.0240	1164.0347
$\hat{h}_r^2$	0.1249±0.09	0.1373±0.09	0.0967±0.08
$\hat{h}_a^2$	0.1250	0.1373	0.0973
$\hat{h}_m^2$	0.1485	0.1610	0.1100
$r_g$	0.9998	0.9998	0.9963
General mean	124.5766	130.8237	124.2614
VC <sub>e</sub> %	26.41	24.57	27.45
VC <sub>g</sub> /VC <sub>e</sub>	0.4176	0.4382	0.3518
N° evaluated families	112	112	114
Type of the model effects:			
Checks		Fixed	Random
Treatment regular	Random	Random	Random
Blocks	Random	Random	Random
Sites	Fixed	Fixed	Fixed

$\hat{\sigma}_a^2$ : additive genetic variance;  $\hat{\sigma}_d^2$ : dominance genetic variance among families;  $\hat{\sigma}_{ga}^2$ : variance of the families x site interaction;  $\hat{\sigma}_e^2$ : residual variance;  $\hat{h}_r^2$ : individual narrow-sense heritability;  $\hat{h}_a^2$ : individual broad-sense heritability;  $\hat{h}_m^2$ : broad-sense heritability at the family mean level;  $r_g$ : genotypic correlation across the sites; VC<sub>e</sub>: environmental variation coefficient; VC<sub>g</sub>: genotypic variation coefficient

intrablock analysis and suggests their recovery by the interblock information. Some approached proceedings were proposed and used by Barbosa (1996) in an estimate of the effective error derived by Ferreira in 1995. To by-pass these problems, some alternatives have recently been presented which increase the efficiency in the analyses of the augmented block design via mixed models (Federer 1998, Federer et al. 2001). Thereby, it is possible to adjust to the environmental effects, recover the genetic interblock information, as well as estimate the variance components by the REML method.

In the context of the mixed model analysis, the purely environmental effects are considered fixed and the genetic effects random (Smith et al. 2001). In augmented block designs, the blocks are not purely environmental effects, since they are incomplete. There is, therefore, genetic information that should be recovered in the blocks and the block effects must be considered random.

For the first and third analysis strategies, all effects were considered random, except for the sites. For the second analysis strategy, the effects of the checks and sites were considered fixed and the others random (Table 2).

The use of the augmented block design for the evaluation of sugarcane families did not prove suitable due to the low heritability values at the level of family means, regardless of the analysis strategy in use. For analysis strategy 1, 2, and 3 (Table 2), the heritabilities at the level of family means were 0.1485, 0.161, and 0.1100 with accuracies of 38%, 40%, and 33%, respectively. The use of replication is particularly essential to raise the heritability at the mean level, in order to increase the response with selection. Assuming the use of four replications, the heritabilities would be 0.4109, 0.4343, and 0.3308 with accuracies of 64%, 65%; and 57%, respectively, for the analysis strategy 1, 2, and 3.

The variance of the interaction families x site was low, and the genotypic correlation across sites high (Table 2). This points to the greater importance of increasing the number of replications than the number of sites for the experiments of family selection. Similarly, the families x sites interaction for TSH in the sugarcane improvement program in Australia observed by McRae and Jackson (1995) was not significant. On the other hand, in Brazil, Bressiani (2001) verified a strong families x sites interaction, predominantly of a complex character, for TSH.

The dominance variance was extremely low (Table 2), which demonstrates the low genetic variability in the population. Likewise, the relation  $VC/VC_e$  is an indicator that little success can be obtained with family selection in this population. This somehow indicates the need to reevaluate the strategies of population formation for selection,

as well as the need to increase the number of replications for a successful selection.

There is a strong predominance of the additive effects in detriment of the dominance effects for this population. These results agree with those presented by Mariotti et al. (1999), where the authors verified predominance of the additive effects when determining the genetic variability among full-sib families. Likewise, Bressiani et al. (2002) showed that the GCA was superior to the SCA for TSH. On the other hand, Bressiani (2001) verified that 92.4% of the total genetic variation was within family for TSH, giving evidence of a strong component due to the dominance variance in that population of 33 families, where 18 were full and 15 half-sibs.

In the following, results of the REML/BLUP analysis without the genetic relationship matrix are discussed. For the analysis strategies 1, 2, and 3 (Table 2), the additive variance rose from 188.63 to 209.10, from 198.54 to 218.07, and from 143.90 to 169.33, respectively. This somehow reflects the parentage present in this population, and highlights the importance and power of the REML/BLUP analysis for such situations. The method proved adequate for an estimation of the genetic parameters and prediction of the genotypic and additive values for sugarcane improvement. It allowed an analysis in a complex situation of a non-orthogonal experimental design (augmented blocks), disequilibrium regarding the number of parent crosses, disequilibrium regarding the representation of the progenies by blocks and sites, and genetic relationship between the crossed parents. It would be difficult to obtain efficient analyses in this situation by other methods.

## CONCLUSIONS

1. Low genetic variability among the evaluated families and absence of the interaction families x sites were observed.
2. The additive effects were predominant to explain the genetic variation among families.
3. The augmented block design (without replication) did not prove appropriate for the experiments of family selection for the evaluated population.
4. The adoption of other designs that use a larger number of replications is desirable to increase the efficiency of the family selection for this population.
5. When the number of families was high and the REML/BLUP method was used, the inclusion or exclusion of the checks in the augmented block design did not alter the estimated genetic parameters.
6. The use of clones or families as checks did not influence

the magnitude of the genetic parameters estimated in the evaluation of full-sib families in the augmented block design.

7. The REMLIBLUP method proved to be suitable for the estimation of genetic parameters and the prediction of important genotypic and additive values for family and parent selection in sugarcane improvement.

#### ACKNOWLEDGEMENTS

To the CNPq, FAPEMIG, RIDESA, and sugar and alcohol factories in Minas Gerais for supporting the genetic improvement program of sugarcane of the Universidade Federal de Viçosa. Seeds courtesy of COPERSUCAR.

## Emprego do REML/BLUP na seleção de famílias de cana-de-açúcar especializadas na produção de biomassa

**RESUMO** - Cento e quatorze famílias de irmãos germanos foram avaliadas em dois locais em Minas Gerais por meio de experimentos em blocos aumentados. A análise dos dados foi feita pelo método REMUBLUP. Observou-se baixa variabilidade genética entre famílias, baixa variância de dominância e ausência da interação famílias por locais. Os componentes de médias estimados via BLUP possibilitaram a seleção de famílias e de genitores superiores. O emprego do delineamento em blocos aumentados sem repetição em experimentos de seleção de famílias não se mostrou adequado em virtude das baixas estimativas de acurácia seletiva e herdabilidade ao nível de médias de família. Adotar outros delineamentos que empreguem maior número de repetições é desejável para incrementar a eficiência da seleção de famílias para esta população. Com o grande número de famílias e uso do procedimento REMLIBLUP, o emprego ou não de tratamentos comuns no delineamento de blocos aumentados não alterou os parâmetros genéticos estimados.

Palavras-chave: *Saccharum* spp., modelos mistos, seleção de famílias, seleção de pais, componentes de variância.

#### REFERENCES

- Barbosa MHP (1996) Capacidade combinatória e comparação entre critérios de seleção de clones de batata (*Solanum tuberosum* L.). PhD Thesis, Universidade Federal de Lavras, Lavras, 141p.
- Barbosa MHP (2000) Perspectivas para o melhoramento da cana-de-açúcar. In: IV Simpósio de atualização em genética e melhoramento de plantas: genética e melhoramento de espécies de propagação vegetativa. Universidade Federal de Lavras, Lavras, p. 1-17.
- Bastos IT, Barbosa MHP, Cruz CO, Burnquist WL, Bressiani JA and Silva FL (2003) Análise dialélica em clones de cana-de-açúcar. *Bragantia* 62: 199-206.
- Bearzoti E (1994) Comparação entre métodos estatísticos na avaliação de clones de batata em um programa de melhoramento. MSc. Thesis, Universidade Federal de Lavras, Lavras, 128p.
- Bressiani JA (2001) Seleção sequencial em cana-de-açúcar. PhD Thesis, Universidade de São Paulo, Piracicaba, 134p.
- Bressiani JA, Burnquist WL, Fuzatto SR, Bonato AL and Geraldi 10 (2002) Combining ability in eight selected clones of sugarcane (*Saccharum* spp.). *Crop Breeding and Applied Biotechnology* 2: 411-416.
- Copersucar (1998) Projeto BRA/96/G3 J: Geração de energia por biomassa, bagaço da cana-de-açúcar e resíduos. STAB 16: 44-46.
- Cox M, Hogarth M and Smith G (2000) Cane breeding and improvement. In: Hogarth M and Allsopp P (eds.) *Manual of cane growing*. PK Editorial Service, Brisbane, p. 91-108.
- Federer WT (1956) Augmented (or Hoonuiaku) designs. *The Hawaiian Planters' Record* 55: 191-208.
- Federer WT (1998) Recovery of interblock, intergradient, and intervariety information in incomplete block and lattice rectangle designed experiments. *Biometrics* 54: 471-481.
- Federer WT, Reynolds M and Crassa J (2001) Combining results from augmented designs over sites. *Agronomy Journal* 93: 389-395.
- Furlani Neto VL, Ripoli TC and Villa Nova NA (1997) Biomassa de cana-de-açúcar: Energia contida no palhico remanescente de colheita mecânica. STAB 15: 24-26.
- Johnson EM and Wichern OW (1988) *Applied multivariate statistical analysis*. 2<sup>th</sup>, Prentice Hall, Wisconsin, 607p.
- Mariotti JA, Cuenya MI and Salas MBG (1999) Análisis de componentes familiares e intra-familiares em progenies de combinaciones biparentales de canã de azucar (*Saccharum* spp.). *Revista Industrial y Agrícola de Tucumán* 76: 52-57.



- McRae TA, Erquiaga DL, Jensen LF, Rattey AR and Stringer JK (1998) BSES sugarcane breeding program in the Burdekin. In: Proceedings of the Australian society of sugar cane technologists. Watson Ferguson, Brisbane, p. 196-203.
- McRae TA and Jackson PA (1995) Selection of sugarcane families for the Burdekin River irrigation area. In: Proceedings of the Australian society of sugar cane technologists. Watson Ferguson, Brisbane, p. 134-141.
- Resende MDV (2002a) Software Selegen-REMLJBLUP. Embrapa Floresta, Curitiba, 67p. (Documentos 77).
- Resende MDV (2002b) Genética biométrica e estatística no melhoramento de plantas perenes. Embrapa Informação Tecnológica, Brasília, 975p.
- Resende MDV (2001) Melhoramento de espécies perenes. In: Nass LL, Valois ACC, Melo 15 and Valadares-Ingles MC (eds.) Recursos genéticos e melhoramento de plantas. Fundação MT, Rondonópolis, p. 357-421.
- 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.
- Stringer JK (2000) Best linear unbiased prediction (BLUP) is a faster and more cost effective method of estimating breeding value in sugarcane. In: Hybrid breeding and genetics of forest trees. QFRIICRC/SPF, Brisbane, p. 505-509.