



Effects of plot size on the estimation of genetic parameters in sugarcane families

Mauro Sérgio de Oliveira Leite¹, Luiz Alexandre Peternelli^{1*}, and Márcio Henrique Pereira Barbosa²

Received 24 June 2005

Accepted 12 April 2006

ABSTRACT - An optimization of the use of resources in sugarcane breeding programs that would not interfere with the genetic evaluations is desirable and of great importance. The costs could be reduced and the chances of success would be higher through the possibility of investing rationally in a larger number of available genetic resources. Our study evaluated the smallest number of rows per plots of sugarcane families for yield traits; the variations in the parameter estimates for different plot sizes were discussed; and the ideal plot size was estimated by the maximum curvature method. Plot reduction leads to gradual CV increases and interferes with the ratio CV_g/CV_e , which in some cases dropped to values below one unit, indicating experimental inaccuracy. Heritability estimates decrease with plot size. The maximum curvature method showed that 2-row plots are enough to evaluate all characteristics.

Key words: genetic variation, heritability, *Saccharum* spp, selection.

INTRODUCTION

Family selection is routinely used in a number of sugarcane breeding programs around the world (Cox et al. 1996), primarily for traits whose heritability based on family means is greater than the heritability of individual plants.

When a character is controlled by several genes, which is expected for most traits of economical value, the probability of finding all favorable alleles together in one individual is minimal. Numerous studies have demonstrated that the combination of family selection with mass selection within selected families, in the early stages of breeding programs, resulted in greater genetic gains and frequency of superior clones (McRae et al. 1993, Cox et al. 1996).

However, for an effective family selection the experimental precision must be high. Several factors such as soil heterogeneity, plant material and managing systems affect experimental accuracy (Gómez and Gómez 1984).

Some strategies can be used to minimize the effects of soil heterogeneity, for instance the experimental design, number of replications, plot size and trait sampling.

Each program establishes the best option regarding selection methods based on the available resources and regional or national peculiarities of the cane production system (Hogarth et al. 1997, Matsuoka et al. 1999, Barbosa and Silveira 2000).

To determine the appropriate plot size, it is essential to know the number of genotypes representing the test families of a trial. According to Durner (1989), the plot size plays a decisive role in field experiments. Considering the same experimental area, small plots allow a higher number of replications, whereas larger plots guarantee smaller variance and are statistically more advantageous.

Recommendations found in literature on the minimum representative sample size for sugarcane are however

¹ Área de Estatística, Departamento de Informática, Universidade Federal de Viçosa (UFV), 36.570-000, Viçosa, MG, Brasil. * E-mail: peternelli@dpi.ufv.br
² Departamento de Fitotecnia, UFV

somewhat variable. Suggestions range from 20 to 150 plants per family (Wu et al. 1977, Wu et al. 1978, Mariotti et al. 1981, Peixoto and Ricci 1984, Barbosa et al. 2001). The suggestions for sample size vary according to the parameter to be estimated as well as the type of trait under evaluation. Aside from the differences in recommendations, different estimation methodologies of the plot or sample size are used. It is a fact that, if the sample size (number of individuals) or even rows representing the family is inadequate, the experimental precision will sink and the establishment of genetic and phenotypic parameter estimates is, consequently, affected.

In view of the large number of families tested in the early stages of a sugar-cane breeding program, the optimization of resources without affecting the genetic evaluations is desirable and of great importance. The costs of programs could be reduced and the success more assured through a rational investment in a larger number of available genetic resources.

In this context, it is important to define how far the plot size in the family trials can be reduced in order to reduce the costs, to test a larger number of families or to increase the number of replications with the same available resource. The optimization of these resources associated to family selection can make the selection of superior clones more effective, increasing the relative gains obtained in a sugar-cane selection program, which is a decisive point in the initial phase of a breeding program.

One objective of this study was the evaluation of the minimum number of basic units (rows) in plots of sugarcane families which would allow an effective estimation of the genetic parameters for yield traits required in support of the selection of families with potential for the development of superior clones; and secondly the discussion about the variations in the parameter estimates for the different plot sizes.

MATERIAL AND METHODS

The varieties IAC86-2210, SP81-1763 and SP87-365, and clones SP82-6108, SP84-2029, SP86-96, SP88-754, and SP88-819 were used as parents in a balanced diallel crossing at the Experimental station of the company Copersucar in Camamu, state of Bahia. Seeds produced in these crosses were planted at the Sugarcane Breeding Center (CECA), Universidade Federal de Viçosa-MG, in the municipal district of Oratórios – MG (lat 20° 24' S, long 42° 40' W, lowland area, alt 400 m asl, red-yellow latosol soil).

Some crosses did not produce seeds due to self-

incompatibility and/or male-sterility, while others failed to produce enough seeds to establish the plots. These problems became significant in reciprocal crosses and selfings. Therefore, out of 64 possible treatments, only 47 produced enough seeds for field evaluations. Two treatments that were not part of the diallel crosses were included, to make the 7 x 7 lattice analysis possible. The last treatment had the peculiarity that each row in its plots bore a different genotype (clone). Since the area was not sufficiently homogeneous but presented some soil stains and the number of treatments was high the experiment was set up in a 7 x 7 partially lattice design with three replications (Gomes 1985). Seedlings were obtained and transplanted as described by Barbosa and Silveira (2000).

The experiment comprised 147 plots, covering a total area of 0.5084 ha. Each plot consisted of six 3.5 m rows with seven plants, with 1.40 m inter-row and 0.5 m intra-row spacing, and 29.4 m² usable area. Two rows with variety RB72454 were used as lateral border. The experimental area was fertilized with 500 kg ha⁻¹ of a formula containing 5% of N, 25% P₂O₅ and 25% of K₂O at planting.

In July 1999, all plants were manually harvested with machete and the possible clones subjected to natural selection for ratooning ability in unfavorable environmental conditions (dry and cold season). In July 2000, the ratoon data were collected by the following procedures:

a) Brix measurements of every plant, one by one, were read at the 5th internode, from base to apex with a Brix hand refractometer.

b) Counting of the total stalk number per row.

After the first part of data collection, the plants were burned and cut, the leafy tops of the cane stalks removed, and a new stage of collection initiated:

a) Random sampling and weight of 10 stalks per row;

b) Total stalk weight per row.

Once data collection was completed, the variables to compose the analyses were obtained. The variable mean Brix per plot or treatment was obtained from individual Brix measurements. Weight of 10 stalks per row was also transformed into mean stalk weight per plot. The number and total stalk weight per plot were calculated by the total stalk number and total stalk weight per row. From total stalk weight per plot, the trait tons of cane per hectare (TCH) (cane yield) was estimated by the expression: (total plot weight in kg x 10)/29.4. Tons of Brix per hectare (TBH) was found by the equation: (TCH x mean Brix)/100.

On the whole, five variables or traits were obtained:

total stalk number per plot (NS); mean stalk weight (kg) per plot (MW); Mean Brix per plot (BRIX); tons of cane per hectare (TCH); and tons of Brix per hectare (TBH).

For the statistical analysis, where families were considered random, the commonly used procedure in breeding known as modified lattice was adopted (Silva et al. 1999). This approach considers mainly treatment and residue as sources of variation. The mean sum of squares for treatment was obtained from the adjusted means of the analysis recovering interblock information, and the residue was given by the effective error. To estimate the components of variance, it was assumed that EMS is similar to those obtained for experiments in random blocks.

Based on the analysis of variance the estimates for the experimental coefficient of variation (CV_e) and the following genetic parameters were calculated: genotypic ($\hat{\sigma}_g^2$) and phenotypic ($\hat{\sigma}_f^2$) variance, genotypic coefficient of variation (CV_g) and broad-sense heritability (h^2) for each appraised trait and each different plot size. The estimators used for these parameters are described as follows:

$$\hat{\sigma}_g^2 = \frac{QMT - QMR}{r}; \quad \hat{\sigma}_f^2 = \frac{QMT}{r}; \quad CV_e(\%) = \frac{\sqrt{QMR}}{\bar{X}} \times 100;$$

$$CV_g(\%) = \frac{\sqrt{\hat{\sigma}_g^2}}{\bar{X}} \times 100; \quad h^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_f^2}.$$

To determine the new plot sizes, all possible combinations among the basic units (the rows), previously numbered from 1 to 6 within each 6-row plot, were considered first. Importantly, it was assumed in this study that a 6-row plot would be enough to estimate the family mean, since the expected genetic effect of the i^{th} family is the same in the three replications.

Due to the great data volume and number of analyses, some functions were developed using a software package called R for automation of the procedures. The objective of the first function, once all data in the area had been collected, was to simulate values for the new plot sizes, according to each possibility of combination among the rows. This resulted in 63 new data groups for each trait, which were analyzed and organized rationally for posterior comparison and interpretation.

It was also necessary to develop another function for running the analysis of variance for each new data group. This function runs analyses in the 7 x 7 partially balanced lattice design (modified lattice). It was created to generate the complete table of analysis of variance and further to estimate the parameters needed for the comparisons in this study. It has the advantage of working

with any data group since it is a general algorithm that follows the order of tabulation for insertion into the R program and agrees with the aforementioned design particularities. The developed functions were tested and validated using data obtained from classic papers of statistics (for instance, Cochran and Cox 1957).

Several methods of estimation of the experimental plot size are presented in literature (see, for example, Viana et al. 2002). The minimum plot size to identify the best families for each appraised trait was determined in this study by the maximum curvature method (Federer 1955). The plot size was further evaluated according to the variation and effects in the genetic parameter estimations of each trait.

RESULTS AND DISCUSSION

All results displayed here were obtained in an experiment with 3 replications. The best plot size found for stalk number (SN) would be two rows, i.e., with 14 plants (Figure 1A). In a more recent study, developed in the sugarcane breeding program at the UFV, Barbosa et al. (2001) concluded that 50-plant plots would be enough to estimate the stalk number in trials with families. It was observed that the CV_e increases with decreasing plot size (Table 1 and Figure 1A). This is the case because a reduction in the plot size raises the relative variation among the data within a given treatment. The CV_e increments indicate an increased experimental error, in other words, a lower precision, which is not advantageous.

This fact does not rule out the use of plots with only two rows for this trait, because the CV_e increments are relatively small and the estimates for the components of genotypic and phenotypic variances kept a certain proportion (Table 1), which can be demonstrated by the heritability values for this trait. Despite the fact that the heritability of 0.61, relative to 6 rows is relatively greater than 0.47 for two rows (Table 1), this difference may not necessarily be due to genetic factors, but mainly to environmental factors at family level, since according to Zhang et al. (1994), smaller plots are more susceptible to errors, caused mainly by the soil heterogeneity. These errors can increase the phenotypic variance estimate, reducing the heritability in small plots, as it was the case for stalk number in this study.

Table 1 shows that the CV_g was almost constant, and that by reducing the plot from 6 to two rows for instance, for stalk number, the existing genetic variation among families remains practically the same. Another

Table 1. Mean estimates of experimental coefficient of variation (CV_e), genotypic variance ($\hat{\sigma}_g^2$), phenotypic variance ($\hat{\sigma}_f^2$), coefficient of genetic variation (CV_g) and heritability (h^2) as a function of plot size per row number, for stalk number (SN), mean stalk weight (MW), mean Brix (BRIX), tons of cane per hectare (TCH), and tons of Brix per hectare (TBH)

Row #	Trait	CV_e	$\hat{\sigma}_g^2$	$\hat{\sigma}_f^2$	CV_g	CV_g/CV_e	h^2
SN							
1		7.47	20.10	50.88	3.52	0.47	0.38
2		5.73	69.27	142.88	3.28	0.57	0.47
3		5.02	147.95	275.97	3.22	0.64	0.53
4		4.63	254.92	448.99	3.19	0.69	0.56
5		4.37	389.78	661.57	3.17	0.73	0.59
6		4.19	552.45	913.59	3.15	0.75	0.61
MW							
1		7.52	0.03	0.04	5.25	0.70	0.57
2		5.55	0.10	0.14	5.12	0.92	0.69
3		4.70	0.21	0.28	5.07	1.08	0.75
4		4.22	0.38	0.49	5.05	1.20	0.79
5		3.90	0.59	0.72	5.04	1.29	0.81
6		3.67	0.84	1.01	5.03	1.37	0.83
BRIX							
1		1.52	0.21	0.59	0.65	0.43	0.34
2		1.19	0.71	1.66	0.62	0.52	0.41
3		1.05	1.51	3.19	0.60	0.57	0.46
4		0.98	2.59	5.20	0.60	0.61	0.49
5		0.93	3.97	7.68	0.60	0.65	0.51
6		0.90	5.64	10.62	0.59	0.66	0.53
TCH							
1		9.34	47.34	91.96	5.81	0.62	0.50
2		7.30	178.94	288.94	5.70	0.78	0.62
3		6.48	394.02	589.95	5.65	0.87	0.67
4		6.02	691.98	994.39	5.62	0.93	0.70
5		5.74	1072.73	1502.12	5.60	0.98	0.71
6		5.54	1536.24	2113.12	5.59	1.01	0.73
TBH							
1		9.42	0.29	0.55	6.08	0.65	0.52
2		7.29	1.11	1.73	5.94	0.81	0.64
3		6.42	2.43	3.52	5.88	0.92	0.69
4		5.94	4.26	5.93	5.85	0.98	0.72
5		5.64	6.60	8.95	5.82	1.03	0.74
6		5.42	9.45	12.58	5.81	1.07	0.75

noteworthy point is the ratio CV_g/CV_e , which for this trait was proportional to plot size (0.75 for 6 rows and 0.57 for two rows), indicating that the environmental variation at experiment level surpasses the genetic variation, which is a sign of experimental inaccuracy.

Figure 1B shows that two rows would also be ideal for mean stalk weight according to the maximum curvature method. The CV_e for mean stalk weight performed similarly to stalk number, with approximated values. However, higher CV_g values are important for mean stalk weight, contributing to the increase of the CV_g/CV_e ratio (between 1.37 and 0.92 for 6 and two rows, respectively) (Table 1). This ratio indicated that the genetic variation for this trait exceeded the environmental variation in 6-row plots.

Nevertheless, with the gradual reduction of the plot, the environmental variation, even when low, surpasses the genetic variation at the experiment level.

For mean stalk weight, the heritability estimates were the highest (from 0.83 with 6 rows to 0.69 with two rows). Comparing the estimates of genotypic and phenotypic variances, it was observed that great part of the phenotypic variation assessed at family level may be due to the variation of genotype effects, which is clearly expressed in the heritability found for this trait.

The CV_e estimates for mean Brix (Table 1) were the smallest compared with other traits, indicating that a good precision for this trait can be attained in family trials, with lower experimental errors. In a study on broad-sense

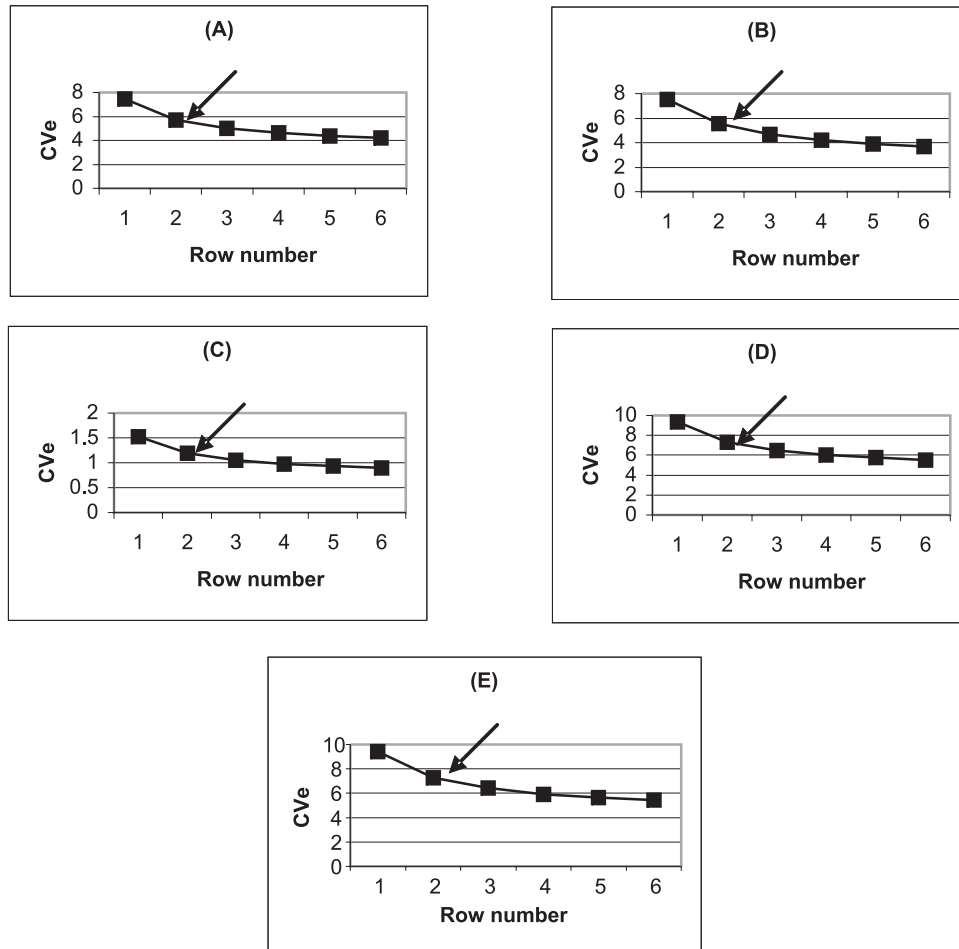


Figure 1. Relationship between the experimental coefficient of variation and plot size (in rows), for stalk number (A), mean stalk weight (B), mean Brix (C), tons of cane per hectare (D) and tons of Brix per hectare (E). The arrows indicate the maximum curvature point

heritability of the main traits used in sugar-cane breeding, Skinner et al. (1987) found that Brix estimates, based on individual plants, were 0.27 (Hawaii) and 0.65 (Australia), whereas for families these values rose, respectively, to 0.53 and 0.90. These results suggest that in the early program stages the individual selection for Brix can also be quite efficient. This can be affirmed due to the high heritabilities in families found for Brix by this author. Since family selection is mainly preferred when the selected trait shows low heritability, greater advantages are gained in this selection strategy when the environmental deviations constitute a great part of the phenotypic variation, that is, when the heritability is low.

Figure 1C shows that the plot size for mean Brix can also be reduced to two rows, in other words, 14 plants. Similarly to this, in a study mentioned above, Barbosa et

al. (2001) found that family Brix could be accurately estimated with a 10-plant plot sample.

The parameter estimates, except for genotypic and phenotypic variances, found for tons of stalk per hectare (TCH) and tons of Brix per hectare (TBH) were very close (Table 1). Two rows would be enough for the experimental plot size for these two parameters as well (Figures 1D and 1E). The values found for CV_e were inversely proportional to the plot size, in agreement with the other traits. It should be mentioned that small increases in CV_e would be justified by the optimization of available resources and the more practical evaluation in smaller plots. This makes a difference particularly for TCH, where the great difficulty is to cut and weigh the whole plot, and for TBH, which depends on the mean Brix obtained in individual plant readings. Hatheway and Williams (1958) affirmed that the

plot size also depends on the costs, in other words, the available resources should be optimized for each case.

For the two traits, the ratio CV_g/CV_e was greater than one unit with 6 rows, but as the plot decreased to two rows, this ratio dropped to 0.78 in TCH and 0.81 in TBH, showing that the environmental surpasses the genetic variation as the plot size decreases.

Skinner et al. (1987), who studied broad-sense heritability for cane yield, found values of 0.75 in Australia and 0.48 in Fiji, with family data. Similar results were found in the present study, where the heritability estimate for TCH in 6 rows was 0.73, but decreased when the plot size was reduced (Table 1), for the same reason as in the case of stalk number.

With regard to the trait used to evaluate families, Mariotti and Cuenya (1992) demonstrated that stalk number is the most stable component among them and important to predict expected gain with selection. Mariotti et al. (1981) also concluded that a sample of 20 genotypes per family would be enough to estimate family mean accurately for most traits. The same can be affirmed in this study according

to the results presented, since for both traits 2-row plots were indicated through the maximum curvature method.

CONCLUSIONS

1. When evaluating sugarcane families in an experiment with three replications, 2-row plots (3.5m; 7 plants per row) are ideal for estimating stalk number, mean stalk weight, mean Brix, tons of cane and tons of Brix per hectare;

2. Plot reduction affects the relationship between CV_g and CV_e directly, and may lower the heritability, affecting the identification of the best families. This can be compensated by optimizing available resources and effective cost-cutting in this kind of experiments.

ACKNOWLEDGEMENTS

The authors wish to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG), respectively, for the scholarship and for the financial aid that supported this study.

Efeitos do tamanho da parcela na estimação de parâmetros genéticos em famílias de cana-de-açúcar

RESUMO - Estudos visando otimizar os recursos em programas de melhoramento de cana-de-açúcar, sem comprometer as avaliações genéticas são desejáveis e importantes, permitindo redução de custos e garantia de êxito pela possibilidade de se investir, racionalmente, em maior número de recursos genéticos disponíveis. Este trabalho objetivou avaliar o número mínimo de sulcos em parcelas de famílias de cana-de-açúcar, para caracteres de produção; discutir as variações ocorridas nas estimativas dos parâmetros genéticos em vários tamanhos de parcelas e usar o método de Máxima Curvatura para estimar o tamanho ideal de parcela. A diminuição da parcela causa aumentos gradativos no CV_e e interfere na relação CV_g/CV_e , que em alguns casos atingiu valores menores que uma unidade, evidenciando imprecisão experimental. As estimativas de herdabilidade decrescem com a diminuição da parcela. Pelo método de Máxima Curvatura, dois sulcos são suficientes para avaliar todas as características.

Palavras-chave: variação genética, herdabilidade, *Saccharum* spp, seleção.

REFERENCES

Barbosa MHP and Silveira LCI (2000) Metodologias de seleção, progressos e mudanças no programa de melhoramento genético da cana-de-açúcar da Universidade Federal de Viçosa. **STAB 18:** 30-32.

Barbosa MPH, Peternelli LA and Silveira LCI (2001) Plot size in sugarcane family selection experiments. **Crop Breeding and Applied Biotechnology 1:** 271-276.

Cochran WG and Cox GM (1957) **Experimental designs**. 2nd ed., John Wiley & Sons, New York, 611p.

Cox MC, McRae TA, Bull JK and Hogarth DM (1996) Family selection improves the efficiency and effectiveness of sugar cane improvement program. In: Wilson JR, Hogarth DM, Campbell JA and Garside AL (eds.) **Sugarcane: Research towards efficient and sustainable production**. CSIRO Division of Tropical Crops and Pasture, Brisbane, p. 42-43.

Durner EF (1989) OPS: a computer program for estimating optimum plot size for field research. **Hortscience 24:** 1040.

Federer WT (1955) **Experimental design**. MacMillan, New York, 544p.

MSO Leite et al.

- Gómez KA and Gómez AA (1984) **Statistical procedures for agricultural research**. John Wiley, New York, 680p.
- Gomes FP (1985) **Curso de estatística experimental**. ESALQ, Piracicaba, 466p.
- Hatheway WH and Williams EJ (1958) Efficient estimation of the relationship between plot size and the variability of crop yields. **Biometrics** **14**: 207-222.
- Hogarth DM, Cox MC and Bull JK (1997) Sugarcane improvement: past achievements and future prospects. **Crop Improvement for the 21st Century**, p. 29-55.
- McRae TA, Hogarth DM, Foreman JW and Braihwaite MJ (1993) Selection of sugarcane seedling families in the Burdekin district. In: Imre BC and Hacker JB (eds.) Focused plant improvement: towards responsible and sustainable agriculture. **Proceedings of the 10th Australian Plant Breeding Conference**. Gold Coast, Austrália, p. 77-82.
- Mariotti JA and Cuenya MI (1992) Efectos familiares y su implicancia en la eficiencia de la selección en caña de azúcar. **Revista Industrial y Agrícola de Tucumán** **69**: 35-51.
- Mariotti JA, Scandaliaris J, Ahmed MA and Cuenya MI (1981) Análisis del comportamiento familiar en progenies híbridas de caña de azúcar. **Revista Industrial y Agrícola de Tucumán** **58**: 1-14.
- Matsuoka S, Garcia AAF and Arizono H (1999) Melhoramento da cana-de-açúcar. In: Borém A (ed.) **Melhoramento de espécies cultivadas**. Editora UFV, Viçosa, 817p.
- Peixoto TC and Ricci JRA (1984) Amostra mínima representativa para caracterização da variabilidade expressa na descendência de progenitores de cana-de-açúcar. **Boletim Técnico Copersucar** **28**: 27-30.
- Silva HD, Regazzi AJ, Cruz CD and Viana JMS (1999) Análise de experimentos em látice quadrado com ênfase em componentes de variância. **Pesquisa Agropecuária Brasileira** **34**: 1811-1821.
- Skinner JC, Hogarth DM and Wu KK (1987) Selection methods, criteria and indices. In: Heinz DJ (ed.) **Sugar cane improvement through breeding**. Elsevier, Amsterdam, p. 409-453.
- Viana AES, Sedyama T, Cecon PR, Lopes SC and Sedyama MAN (2002) Estimativas de tamanho de parcela em experimentos com mandioca. **Horticultura Brasileira** **20**: 58-63.
- Wu KK, Heinz DJ, Meyer HK and Ladd SL (1977) Selection studies in sugar cane (*Saccharum* spp. hybrids). III. A method to determine sample size for the estimation of population variance. **Theoretical and Applied Genetics** **51**: 29-33.
- Wu KK, Heinz DJ, Meyer HK and Ladd SL (1978) Minimum sample size for estimating progeny mean and variance. **Crop Science** **18**: 57-62.
- Zhang R, Warrick AW, Myers DE (1994) Heterogeneity, plot shape effect and optimum plot size. **Geoderma** **62**: 183-197.