

MÉTODOS DE ESTIMAÇÃO DO COEFICIENTE DE REPETIBILIDADE NO MELHORAMENTO DA SERINGUEIRA¹

MARIA ELIZABETH DA COSTA VASCONCELLOS, PAULO DE SOUZA GONÇALVES,
JOÃO RODRIGUES DE PAIVA² e AFONSO CELSO CANDEIRA VALOIS³

RESUMO - O presente estudo foi conduzido com o objetivo de testar vários métodos de estimação dos coeficientes de repetibilidade, bem como determinar o número adequado de minitestes de produção na seleção de plântulas em viveiro de seringueira (*Hevea* sp). Foram utilizados três minitestes em plântulas oriundas de polinização aberta do clone RRIM 600 e estimados os coeficientes de repetibilidade, considerando os métodos da análise de variância e componentes principais, segundo metodologia apresentada por Mansour et al. (1981). A equação que relaciona os três minitestes de produção explicou cerca de 87% da variação existente entre os testes. Não foram detectadas grandes diferenças nas estimativas do coeficiente de repetibilidade entre os processos testados, optando-se, assim, pelo ANOVA, por ser de mais fácil utilização.

Termos para indexação: *Hevea*, componentes principais, modelos lineares, análise multivariada.

METHODS OF ESTIMATION OF COEFFICIENT OF REPEATABILITY IN RUBBER TREE BREEDING

ABSTRACT - This study was undertaken to test various repeatability coefficient estimation procedures, as well as to determine the adequate number of minitest of production in the selection of seedlings in rubber tree nursery. Three minitests were run on open-pollinated progeny seedlings of the RRIM 600 clones and the coefficient of repeatability estimated by using both the analysis of variance and principal components techniques, described by Mansour and co-workers. The equation that relates all three minitests of production explained about 87% of the existing variations among test. Since no marked differences among tests procedures were found in the estimation of the coefficient of repeatability, the ANOVA technique is recommended, due to its simplicity.

Index terms: *Hevea*, principal components, linear models, multivariate analysis.

INTRODUÇÃO

A seleção de plântulas em viveiro de seringueira é feita através de vários testes precoces de produção, visando reduzir o tempo necessário para determinar clones para uso comercial. Dentre os vários tipos de testes existentes, Mendes (1971) desenvolveu o miniteste de produção (MTP), cujas características permitem quantificar o conteúdo de borracha seca de plantas jovens. Entretanto, existe pouca informação acerca do número de testes necessários na mesma planta, para que se obtenham resultados mais concretos em relação à sua eficiência (Gonçalves et al. 1982b). Para tanto, necessário se torna, do ponto de vista experimental,

tomar mais de uma medida de cada genótipo como critério para seleção de plântulas ou clones.

Um fato muito importante é que, observando-se seringueiras jovens em viveiro, durante o ano todo, verifica-se a emissão de novos fluxos de folhas, caracterizando uma periodicidade no crescimento, e que uma mesma planta pode apresentar diferentes estádios foliares na época de aplicação dos testes, havendo alteração nos mecanismos fisiológicos da planta, e, conseqüentemente, exercendo influência direta sobre a produção de látex.

Segundo Paranjothy (1980), a planta em estágio de novo fluxo foliar apresenta baixa produção causada pela demanda de nutrientes e carbono para o desenvolvimento das folhas do novo fluxo. Em seringueiras adultas é evidenciada a redução de borracha seca durante o reenfolhamento (Adjuwana & Soerianegara 1970).

O coeficiente de repetibilidade (ρ), segundo Falconer (1975), por ser o limite superior do coeficiente de herdabilidade (h^2), é bem mais simples de ser estimado, pois este último exige cruzamentos controlados e estudo de progênies. Estatística-

¹ Aceito para publicação em 12 de dezembro de 1984. Trabalho realizado com a participação de recursos financeiros do Convênio SUDHEVEA/EMBRAPA.

² Eng. - Agr., M.Sc., EMBRAPA/Centro Nacional de Pesquisa de Seringueira e Dendê (CNPDS), Caixa Postal 319, CEP 69000 Manaus, AM.

³ Eng. - Agr., Dr., Chefe do CNPDS.

mente o coeficiente de repetibilidade usado pelos melhoristas é a correlação intraclasse (Abeywardena 1972).

Este coeficiente permite, também, determinar quantas observações fenotípicas devem ser feitas em cada indivíduo para que a seleção fenotípica seja feita com eficiência e um mínimo de trabalho.

Um problema prático, envolvendo medidas tomadas em que o fator periodicidade ocorre, é que estas são afetadas por algumas mudanças regulares, irregulares ou sistemáticas de ordem fisiológica que ocorrem nas plantas ou animais.

Exemplo muito comum é o caso de coco que possui flutuações bienais de produção (Shrikhande 1957 e Abeywardena 1972). Na área animal, temos o caso do estudo de lactação para a seleção de matrizes (Lush 1945). Nestes casos, se o coeficiente de repetibilidade for calculado pela análise da variância, este necessariamente será subestimado. Portanto, deve-se procurar identificar os casos em que esse fator ocorre e com isso não se trabalhar com medidas tendenciosas.

O tradicional método para se estimar o coeficiente de repetibilidade tem sido o método da análise da variância (ANOVA). Entretanto, recentes métodos multivariados usando componentes principais tem sido propostos por Abeywardena (1972) e Rutledge (1974).

Numa situação em que o fator periodicidade ocorre, este não pode ser isolado na análise da variância. Com isto, este componente é incluído em σ_e^2 , elevando o seu valor e, muitas vezes, tornando esta estimativa falsa. Nestas condições, a repetibilidade é subestimada.

O índice de repetibilidade deve considerar estes períodos e isto somente é possível através da análise de componentes principais (Kendall 1975).

O presente estudo teve como objetivo comparar diferentes métodos multivariados de estimar o coeficiente de repetibilidade com o da análise da variância, bem como determinar o número de minutos necessários para a seleção de plantas de seringueira em viveiro de um ano de idade, através da análise de componentes principais.

MATERIAL E MÉTODOS

Os coeficientes de correlação e de repetibilidade, cal-

culados segundo método da análise de variância, foram retirados do trabalho de Gonçalves et al. (1982b). Trata-se de um estudo em viveiro de um ano de idade, instalado no campo experimental do Centro Nacional de Pesquisa de Seringueira e Dendê, no km 28 da Rodovia AM-010, em Manaus, AM, pertencente à Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA.

O modelo matemático de uma planta foi $Y_{ij} = \mu + p_i + t_j + e_{ij}$, onde Y_{ij} é a produção de borracha, μ é a medida geral, p_i é o efeito aleatório de plantas, t_j é o efeito fixo dos testes e $e_{ij} \sim N(0, \sigma^2)$ é o erro experimental, onde $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, k$.

As estimativas dos componentes de variância foram obtidas pela esperança matemática dos quadrados médios, segundo o esquema da Tabela 1. O estimador do coeficiente de repetibilidade pelo método de análise da variância foi:

$$\hat{\rho}_{ao} = \hat{\sigma}_p^2 / (\hat{\sigma}_p^2 + \hat{\sigma}_e^2)$$

Os estimadores do coeficiente de repetibilidade pelos métodos multivariados utilizando a análise de componentes principais foram:

$$\hat{\rho}_{pv} = (\hat{\lambda}_v / \hat{\sigma}_F^2 - 1) / (K - 1) \quad (1)$$

onde $\hat{\rho}_{pv}$ é um estimador de repetibilidade baseado nos componentes principais da matriz de covariância da amostra dos três testes (matriz S). A maior raiz característica ($\hat{\lambda}_v$), segundo Morrison (1978) é:

$$\hat{\lambda}_v = \hat{\sigma}_F^2 [1 + (K - 1)\rho] \quad (2)$$

onde $\hat{\sigma}_F^2$ é a média dos termos da diagonal da matriz

$$S \text{ e } K = 1, \dots, 3$$

Se os componentes principais forem extraídos da matriz de correlação, a maior raiz característica será a mesma de (2) com $\hat{\sigma}_F^2$ igual a 1. Neste caso, o estimador de repetibilidade proposto por Rutledge (1974) é:

$$\hat{\rho}_{pr} = (\hat{\lambda}_r - 1) / (K - 1) \quad (3)$$

onde $\hat{\lambda}_r$ é a maior raiz característica da matriz de correlação da amostra (matriz R).

Um outro possível estimador de repetibilidade tem como componente principal o vetor b, onde:

$$b' = [1/\sqrt{K}, \dots, 1/\sqrt{K}] \quad (4)$$

resultando no estimador

$$\hat{\rho}_{sv} = (b' S b / \hat{\sigma}_F^2 - 1) / (K - 1) \quad (5)$$

TABELA 1. Esquema de análise da variância utilizado na avaliação da produção de borracha pelo MTP em plantas jovens de seringueira. Manaus, AM, 1984.

| | G.L. | QM | E (QM) |
|-----------------|-----------------|--|---|
| Teste | (k - 1) | $\frac{k}{n} \sum_{j=1}^k (Y_{.j} - Y_{..})^2 / (k - 1)$ | $\sigma_e^2 + \frac{n}{k-1} \sum_{j=1}^k t_j^2$ |
| Plantas | (n - 1) | $\sum_{i=1}^n (Y_i - Y_{..})^2 / (n - 1)$ | $\sigma_e^2 + k \sigma_p^2$ |
| Teste x plantas | (k - 1) (n - 1) | Subtração | σ_e^2 |

Este mesmo procedimento, utilizando-se a matriz de correlação nos fornece o seguinte estimador:

$$\hat{\rho}_{sr} = [\underline{b}' R \underline{b} - 1] / [K - 1] \quad (6)$$

Um outro estimador usado é o de máxima verossimilhança, com $\hat{\sigma}_F^2$ obrigatoriamente positivo. Segundo Searle (1971), este estimador será:

$$\hat{\rho}_{mv} = \text{Max} \{ (QmP - QmE) / [QmP + (K - 1) QmE], 0 \}$$

$$\hat{\rho}_{mv} = \text{Max} [\hat{\rho}_{ao}, 0] \quad (7)$$

A técnica, basicamente, tenta explicar a estrutura de variâncias e covariâncias das K variáveis, construindo um conjunto menor de combinações lineares ortogonais, ou seja, os componentes principais das variáveis ortogonais (Kendall 1975).

O primeiro componente principal é aquela combinação ponderada das variáveis originais que carrega a máxima quantidade da variação total das variáveis originais. O segundo componente principal é a combinação que, entre todas as combinações ortogonais ao primeiro componente principal, é responsável pela máxima quantidade de variação remanescente, e assim sucessivamente.

Matematicamente, o problema é o de explicar a estrutura de covariância das variáveis Y_1, Y_2, \dots, Y_k por meio de m ($m < k$) combinações lineares $Z_n = a_{1n} Y_1 + a_{2n} Y_2 + \dots + a_{kn} Y_k$, onde $h = 1, \dots, m$ e a_{jh} ($j = 1, \dots, k$) são os elementos componentes do h-ésimo maior vetor característica extraído da matriz amostral S de covariâncias ou da matriz R de correlações.

RESULTADOS E DISCUSSÃO

Na Tabela 2 estão os estimadores de repetibilidade estudados, com os respectivos valores obtidos

e seus processos de obtenção. O estimador $\hat{\rho}_{pv}$ (0,7967), obtido através da análise de componentes principais da matriz de covariância (matriz S), teve praticamente o mesmo valor de $\hat{\rho}_{ao}$ (0,7946), o mesmo acontecendo com $\hat{\rho}_{mv}$ (0,7946), obtido pelo método da máxima verossimilhança. Como $\hat{\rho}_{mv}$ depende do valor de $\hat{\rho}_{ao}$, o primeiro assume grande importância nos casos em que se obtenham componentes de variância negativos, pois $\hat{\rho}_{mv}$ considera $\sigma_p^2 = 0$ se este for negativo. O estimador calculado pelo método de análise da variância ($\hat{\rho}_{ao}$), obtido quando retirado o efeito fixo dos testes no modelo matemático, é a correlação intraclasse (Mansour et al. 1981).

Abeywardena (1972) demonstra que a correlação intraclasse é uma estimativa grosseira do coeficiente de repetibilidade. Portanto, a relação entre $\hat{\rho}_{ao}$ e $\hat{\rho}_{mv}$ não é mantida exatamente, se o termo t_j é retirado do modelo.

Os estimadores $\hat{\rho}_{pr}$ (0,8106) e $\hat{\rho}_{sr}$ (0,8086), obtidos através dos componentes principais da matriz de correlação, praticamente alcançaram o mesmo valor, sendo que a diferença em relação aos demais estimadores foi pequena. Comparando-se esses estimadores com o proposto por Falconer (1975), verifica-se que também não houve diferença, pois este coeficiente foi de 0,7968. Entretanto, esses estimadores adquirem real importância no caso de existência de heterogeneidade de variância. Nestes casos, segundo Mansour et al. (1981), esses dois estimadores têm melhor performance que os demais, principalmente quando existe correlação alta entre as variáveis, pois todos os estima-

TABELA 2. Estimativas do coeficiente de repetibilidade por seis diferentes processos para o caráter produção de borra-cha avaliado pelo MTP em plantas jovens de seringueira. Manaus, AM, 1984.

| Estimador | Processo | Repetibilidade |
|-------------------|--|----------------|
| $\hat{\rho}_{ao}$ | ANOVA | 0,7946 |
| $\hat{\rho}_{pv}$ | Componentes principais da matriz de covariância | 0,7967 |
| $\hat{\rho}_{pr}$ | Componentes principais da matriz de correlação | 0,8106 |
| $\hat{\rho}_{sv}$ | Estimador estrutural da matriz covariância usando o vetor característica | 0,7981 |
| $\hat{\rho}_{sr}$ | Estimador estrutural da matriz de correlação usando o vetor característica | 0,8086 |
| $\hat{\rho}_{mv}$ | Máxima verossimilhança | 0,7946 |

dores que não forem obtidos através da matriz de correlação apresentam uma tendenciosidade negativa. Outro caso em que existe uma esperada diferença entre os estimadores, é quando o coeficiente de repetibilidade é pequeno.

Gonçalves et al. (1982a), estudando quatorze clones, verificaram que somente três mostraram significância no primeiro MTP com o teste Hamaker Morris-Mann modificado (HMM-m), concluindo que, nesta amostra, o MTP não foi eficaz na seleção de clones. Essas conclusões fazem com que se procure utilizar outras metodologias mais refinadas, onde fontes de variação não controladas sejam consideradas e, com isso, se chegue na seleção de material de grande potencial.

Abeywardena (1972) considera que o método de obtenção de estimadores através da análise de componentes principais é o critério mais eficiente de seleção de plantas superiores, pois a média aritmética de observações pode ser tendenciosa nas condições em que haja uma periodicidade em seu crescimento. Conclui que, através de componentes principais, se obtém os verdadeiros estimadores, pois esta considera todas as situações possíveis.

Na Tabela 3, encontram-se os vetores da análise de componentes principais baseados na matriz de correlação. Nesta situação, somente o primeiro autovetor é significativo, explicando 87,4% da variação total. Baseado nestes vetores e calculando os pesos correspondentes a cada miniteste de produção de acordo com Kendall (1975), observou-se

TABELA 3. Componentes principais da matriz de correlação entre testes (MTP) aplicados em plantas jovens de seringueira. Manaus, AM, 1984.

| Raiz característica | Variação (%) | Vetores | | |
|---------------------|--------------|----------|----------|----------|
| | | l_1 | l_2 | l_3 |
| 2,62 | 87,37 | + 0,5783 | + 1,0318 | - 1,4668 |
| 0,26 | 8,60 | + 0,5924 | + 0,2860 | + 1,9785 |
| 0,12 | 4,03 | + 0,5638 | - 1,1614 | - 0,5729 |

que os três minitestes são necessários na seleção precoce de plântulas produtivas.

Gonçalves et al. (1982b), calculando o coeficiente de repetibilidade através da análise da variância, concluíram que somente dois testes seriam necessários para se obter um bom resultado na seleção precoce de plantas jovens de seringueira. Segundo estes autores, há uma tendência de o coeficiente de repetibilidade ser inversamente proporcional ao número de testes aplicados em função do aumento de σ_e^2 causado pela periodicidade no crescimento da seringueira. Entretanto, ao considerarem os diferentes estádios foliares das plantas na época de aplicação dos testes, verificaram que as melhores plantas num teste não foram necessariamente as melhores em outro teste, caracterizando um efeito de interação teste x planta. No entanto, este componente não foi isolado e testado na análise da variância, pois esta interação teve função de resíduo. Para se testar o efeito teste x plan-

ta seria necessário separar as plantas em cada miniteste, de acordo com o seu estádio foliar, e este fator entra como fonte de variação. Como o coeficiente de repetibilidade calculado através de componentes principais foi baseado nestes dados, não se detectaram grandes diferenças nas estimativas do coeficientes de repetibilidade entre os processos utilizados.

Este fato caracteriza, mais uma vez, que, em virtude de periodicidade de crescimento das plantas, pode-se deixar de selecionar determinada planta por esta coincidentemente se encontrar emitindo novos fluxos foliares em dois testes aplicados. Portanto, seria de grande importância considerar em que fase de desenvolvimento foliar encontra-se a planta.

Convém ressaltar que Gonçalves et al. (1982a) evidencia uma possível ineficácia do MTP na seleção de plantas, ou então, a necessidade de se utilizar maior número de plantas de determinado clone a serem submetidas ao teste. Com isso, haveria uma diminuição nas flutuações da produção teste após teste, causadas pela irregularidade dos estádios foliares das plantas. Entretanto, com a utilização da análise de componentes principais, onde as plantas são separadas em cada teste de acordo com o seu estádio foliar, pode-se estudar a eficiência do MTP em futuros trabalhos.

CONCLUSÕES

1. A equação considerando os três minitestes na seleção de plântulas em viveiro foi mais eficiente pelo fato de a mesma explicar 87% da variação total ocorrida.

2. Não foram detectadas grandes diferenças nas estimativas do coeficiente de repetibilidade entre os diferentes processos utilizados, indicando que o processo ANOVA pode ser utilizado por ser menos trabalhoso.

3. A periodicidade no crescimento de seringueira não invalida os métodos tradicionais de estimação do coeficiente de repetibilidade do caráter

produção de borracha, avaliado pelo miniteste de produção.

REFERÊNCIAS

- ABEYWARDENA, V. An application of principal component analysis in genetics. *J. Genet.*, 61:27-51, 1972.
- ADIJUWANA, H. & SOERIANEGARA, I. Photosynthesis and latex production of tree *Hevea brasiliensis* clones. *Menara Perkebunan*, 39(5/6):77-86, 1970. Texto em indonésio. Resumo em inglês.
- FALCONER, D.S. *Introduction to quantitative genetics*. London, Longman, 1975. 365p.
- GONÇALVES, P. de S.; PAIVA, J.R. de & ROSSETTI, A.G. Eficiência do miniteste de produção na seleção precoce de plantas de seringueira em relação ao teste Hamaker Morris-Mann. *Pesq. agropec. bras.*, Brasília, 17(8):1145-53, ago. 1982a.
- GONÇALVES, P. de S.; ROSSETTI, A.G. & PAIVA, J.R. de. Coeficiente de repetibilidade e eficiência do miniteste de produção na seleção de plantas de seringueira. *Pesq. agropec. bras.*, Brasília, 17(2):233-7, fev. 1982b.
- KENDALL, M.G. *Multivariate analysis*. New York, MacMillan, 1975. 210p.
- LUSH, J.L. *Animal breeding plants*. s.l., Iowa State College Press, 1945.
- MANSOUR, H.; NORDHEIM, E.V. & RUTLEDGE, J.J. Estimators of repeatability. *Theor. Appl. Genet.*, 60:151-6, 1981.
- MENDES, L.O.T. Poliploidização da seringueira; um novo teste para determinação da capacidade de produção de seringueira jovem. *Polímeros*, São Paulo, 1(1):22-30, 1971.
- MORRISON, D.F. *Multivariate statistical methods*. 2.ed. New York, McGraw-Hill, 1978. 415p.
- PARANJOTHY, K. Physiological aspects of wintering, flower induction and fruit-set in *Hevea*. In: RUBBER RESEARCH INSTITUTE OF MALAYA, Kuala Lumpur, Malásia. *Hevea breeding course; lecture notes*. Kuala Lumpur, 1980. 12p.
- RUTLEDGE, J.J. A scaling which removes bias of Abeywardena's estimator or repeatability. *J. Genet.*, 61:247-50, 1974.
- SEARLE, S.R. *Linear models*. New York, J. Wiley, 1971. 532p.
- SHRIKHANDE, V.J. Some considerations in designing experiments on coconut trees. *J. Indian Soc. Agric. Stat.*, 9:82-99, 1957.