UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA "LUIZ DE QUEIROZ"

MODELOS MISTOS NA PREDIÇÃO DE VALORES GENÉTICOS ADITIVOS EM TESTES DE PROGÊNIES FLORESTAIS

Júlio Sílvio de Sousa Bueno Filho

(Eng. Agrônomo)

Orientador: **Prof. Dr. Roland Vencovsky**

Tese apresentada à Escola Superior de
Agricultura "Luiz de Queiroz", da Univer-
sidade de São Paulo, para a obtenção do título
de Doutor em Agronomia, Área de Concen-
tração: Genética e Melhoramento de Plantas.

PIRACICABA
Estado de São Paulo - Brasil
março de 1997

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - Campus "Luiz de Queiroz"/USP

Bueno Filho, Júlio Silvio de Sousa

Modelos mistos na predição de valores genéticos aditivos em testes de progênies florestais / Júlio Silvio de Sousa Bueno Filho. - - Piracicaba, 1997.
118 p. : il.

Tese (doutorado) -- Escola Superior de Agricultura Luiz de Queiroz, 1997.
Bibliografia.

1. Genética florestal 2. Melhoramento genético vegetal 3. Modelo matemático
4. Seleção genética 5. Teste de progênie I. Título

CDD 634.9567

MODELOS MISTOS NA PREDIÇÃO DE VALORES GENÉTICOS ADITIVOS
EM TESTES DE PROGÊNIES FLORESTAIS

Júlio Sílvio de Sousa Bueno Filho

Aprovado em: 29.04.1997

Comissão Julgadora:

Prof. Dr. Roland Vencovsky	ESALQ/USP
Prof. Dr. José Branco de Miranda Filho	ESALQ/USP
Prof. Dr. Irineu Umberto Packer	ESALQ/USP
Dr. Rui da Silva Verneque	CNPGL/EMBRAPA
Prof. Dr. Mário Luiz Teixeira de Moraes	FEIS/UNESP

Prof. Dr. Roland Vencovsky
(Orientador)

"(...) Sabe-se que era proibido aos judeus investigar o futuro. Ao contrário, a Torá e a prece se ensinam na rememoração. Para os discípulos a rememoração desencantava o futuro, ao qual sucumbiam os que interrogavam os adivinhos. (...)"

WALTER BENJAMIN

Sobre o Conceito da História

"...

Que tempos são esses, em que

Falar de árvores é quase um crime

Pois implica silenciar sobre tantas barbaridades?

..."

BERTHOLD BRECHT

excerto de "Aos que vão nascer"

Poemas de Svendborg

À memória de

Durvalina, João, Yolanda e Júlio.

Aos presentes

Jacqueline, João Pedro e Jonas.

AGRADECIMENTOS

Ao Dr. Roland Vencovsky, pelo estímulo e percepção como orientador e sobretudo pela paciência e compreensão (muito além do dever) como pessoa.

Aos Professores dos Departamentos de Genética e de Matemática e Estatística da ESALQ pela experiência no trato com questões de melhoramento e análise de experimentos que suscitaram esta investigação .

Ao Dr. Mário L. T. de Moraes, pela gentileza da cessão de cópias das planilhas de campo originais com os dados analisados e ao Dr. Paulo Yoshio Kageyama pelo apoio.

Aos Doutores Carlos Alberto Arrabal Arias, Cláudio Takeda, Daniel Furtado Ferreira e Nelson Fonseca Junior, pelo coleguismo e a ajuda cotidiana com as disciplinas, pelo estímulo e interesse pelo tema da investigação e pelo apoio computacional e bibliográfico sem o qual este trabalho não poderia ser levado a termo.

Aos demais amigos da graduação e da pós-graduação na ESALQ, na figura do Dr. Rodrigo Rocha Latado, colega de longo curso. Aos funcionários da ESALQ e do CIAGRI, direta ou indiretamente envolvidos na realização deste trabalho, nas pessoas das bibliotecárias Aparecida Elisabeth dos Santos da Silva e Silvana Machizelli Gregório, da Biblioteca Setorial do Departamento de Genética.

Aos colegas professores da Universidade de Taubaté, Dra. Ana Aparecida da Silva, Dr. Edson Furtado, Dr. Marcos R. Furlan e Dr. Sigismundo Bialoscorski e aos seus alunos de Agronomia. Aos Professores Augusto Ramalho de Moraes e Nelson Willibaldo Werlang, pela leitura do resumo e do “summary”. Ao Dr. Marcos Deon Vilela de Resende, pela leitura e sugestões valiosas.

Aos moradores de Piracicaba que tornaram nossa estada muito agradável e dos quais sentiremos muitas saudades, nas figuras de D. Maria e Sr. José Furlanete, de D. Valdelice Ribeiro e Sr. Alípio Ribeiro, de D. Neusa Corrêa Ferreira e Sr. Ademir Crisóstomo Ferreira, Sra. Benedita Elias e o Eng. André Elias Netto, da Sra. Aldenir Dias dos Santos, do Eng. Jorge Luiz Mangolini Neves e da Dra. Raimunda Ferreira de Almeida.

Aos meus pais e familiares.

A Jacqueline M. A. Bueno, pela ajuda inestimável mas previsível.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

Ao CNPq.

SUMÁRIO	página
LISTA DE FIGURAS	viii
LISTA DE TABELAS	ix
RESUMO	xii
SUMMARY	xiv
1. INTRODUÇÃO	01
2. REVISÃO DA LITERATURA	05
2.1. REFERENCIAL DO MELHORAMENTO FLORESTAL	05
2.2. TESTES DE PROGÊNIES FLORESTAIS	08
2.3. MODELOS MISTOS	11
2.3.1. Introdução e definições	11
2.3.2. Métodos de Predição	13
2.3.3. Tipos de Preditores	15
2.3.3.1. Forma geral dos Índices de seleção	17
2.3.3.2. Forma geral dos BLP	17
2.3.4. Estimação de BLUEs e Predição de BLUPs	19
2.3.5. Componentes da Variância	21
2.3.5.1. Métodos dos Momentos (ou da ANAVA)	23
2.3.5.2. Métodos de Máxima Verossimilhança (ML)	25
2.3.5.3. Máxima Verossimilhança Restrita (REML)	26
2.3.5.4. Outros métodos de estimação das compo- nentes da variância	27
2.4. PREDIÇÃO DE VALORES GENÉTICOS	29
3. MATERIAL E MÉTODOS	35
3.1. DESCRIÇÃO DO MATERIAL VEGETAL	35
3.2. MODELO DA ANÁLISE CONVENCIONAL	36
3.3. MODELO MISTO DE PREDIÇÃO	39
3.4. BLP	40
3.4.1. Comparações Entre Métodos através dos BLP	41
3.4.2. Caso A: Seleção com base em médias marginais (gerais) de progênies	41
3.4.3. Caso B: Seleção com base em médias de progênies por local	42

3.4.4. Caso C: Seleção com base em médias de parcelas	42
3.5. ÍNDICE DE SELEÇÃO	44
3.5.1. Comparações Entre Métodos Pelo Índice	44
3.6. ESTIMAÇÃO de BLUEs e PREDIÇÃO de BLUPs	45
3.7. COMPONENTES DA VARIÂNCIA E BLUP ITERATIVO	46
4. RESULTADOS E DISCUSSÃO	48
4.1. COMPARAÇÕES TEÓRICAS	48
4.1.1. Comparações Entre os Pesos	48
4.1.2. Comparações Entre os Efeitos do Modelo	55
4.2. MÉDIAS FENOTÍPICAS GERAIS (BV) versus BLP (a partir de médias de parcelas)	57
4.2.1. Vetores C, Matrizes V e vetores de pesos $C'V^{-1}$	57
4.2.2. Progênies Seleccionadas Através do BLP	61
4.3. BLUP ITERATIVO	65
4.3.1. Estimativas REML das Componentes da Variância	65
4.3.2. Progênies Seleccionadas Através do BLUP	68
4.4. SELEÇÃO SEQUENCIAL versus ÍNDICE DE SELEÇÃO	75
4.4.1. Índices de Seleção Calculados	75
4.4.2. Árvores Seleccionadas Através do Índice e da Seleção Sequencial	79
4.5. PREDIÇÃO NO MELHORAMENTO FLORESTAL	85
4.5.1. Caracteres em Questão	
4.5.2. Seleção de Matrizes (pomar de 1 ^a geração)	89
4.5.3. Seleção de Árvores (pomar de 2 ^a geração)	94
4.5.4. Perspectivas de Utilização	98
5. CONCLUSÕES	102
6. REFERÊNCIAS BIBLIOGRÁFICAS	104
APÊNDICES	113

LISTA DE FIGURAS	página
Fig. 1: Dispersão das predições para o DAP pelos três métodos (BV, BLP e BLUP). Teste de Progênie de <i>E. grandis</i> , sete anos.	86
Fig. 2: Dispersão das predições para a altura pelos três métodos (BV, BLP e BLUP). Teste de Progênie de <i>E. grandis</i> , sete anos.	86
Fig. 3: Dispersão das predições para o volume pelos três métodos (BV, BLP e BLUP). Teste de Progênie de <i>E. grandis</i> , sete anos.	87
Fig. 4: Dispersão das predições para a densidade pelos três métodos (BV, BLP e BLUP). Teste de Progênie de <i>E. grandis</i> , sete anos.	87
Fig. 5: Predições de valores aditivos pelos três métodos, para o DAP. Teste de Progênie de <i>E. grandis</i> , sete anos.	90
Fig. 6: Predições de valores aditivos pelos três métodos, para a altura. Teste de Progênie de <i>E. grandis</i> , sete anos.	91
Fig. 7: Predições de valores aditivos pelos três métodos, para o volume. Teste de Progênie de <i>E. grandis</i> , sete anos.	92
Fig. 8: Predições de valores aditivos pelos três métodos, para a densidade. Teste de Progênie de <i>E. grandis</i> , sete anos.	93
Fig. 9: Progressos genéticos e respostas correlacionadas pelos cinco métodos considerados, para o DAP. Teste de Progênie de <i>E. grandis</i> , sete anos.	95
Fig.10: Progressos genéticos e respostas correlacionadas pelos cinco métodos considerados, para a altura. Teste de Progênie de <i>E. grandis</i> , sete anos.	95
Fig.11: Progressos genéticos e respostas correlacionadas pelos cinco métodos considerados, para o volume. Teste de Progênie de <i>E. grandis</i> , sete anos.	96
Fig.12: Progressos genéticos e respostas correlacionadas pelos cinco métodos considerados, para a densidade. Teste de Progênie de <i>E. grandis</i> , sete anos.	96

LISTA DE TABELAS

página

Tabela 1: Valores das variâncias das predições, acurácia seletiva e quocientes teóricos entre as variâncias das predições para dois casos hipotéticos de seleção em três ambientes pela média marginal (BV) e pelo BLP a partir de médias por ambiente. 52

Tabela 2: Pesos para cada ambiente-alvo de seleção e cada caráter por local, aplicáveis a cada porção aleatória (no nível de repetições dentro de locais) e estimativas de herdabilidades na seleção entre médias de progênie por ambiente e entre médias gerais de progênie, para os quatro caracteres em estudo. Teste de Progênie de *E. grandis*, aos sete anos. 58

Tabela 3: Variâncias das predições, variâncias dos erros das predições e acurácia seletiva para cada caractere, conforme o método considerado (BLP ou BV) e o ambiente-alvo da seleção. Teste de Progênie de *E. grandis*, sete anos. 59

Tabela 4: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para o DAP. Teste de Progênie de *E. grandis*, sete anos. 62

Tabela 5: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para a altura. Teste de Progênie de *E. grandis*, sete anos. 62

Tabela 6: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para o volume. Teste de Progênie de *E. grandis*, sete anos. 63

Tabela 7: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para a densidade. Teste de Progênie de *E. grandis*, sete anos. 63

Tabela 8: Estimativas de componentes da variância e respectivos erros padrão obtidas através do método REML para cada um dos quatro caracteres em estudo, por locais e conjuntas, no modelo misto. Estimativas de médias gerais por locais e conjuntas. Teste de Progênie de *E. grandis*, aos sete anos. 66

Tabela 9: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para o DAP. Teste de Progênie de *E. grandis*, sete anos. 69

Tabela 10: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para a altura. Teste de Progênes de *E. grandis*, sete anos. 69

Tabela 11: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para o volume. Teste de Progênes de *E. grandis*, sete anos. 70

Tabela 12: Valores aditivos preditos em ordem decrescente dos BV e número de ordem (BLP), para a densidade. Teste de Progênes de *E. grandis*, sete anos. 70

Tabela 13: Diferenças de ordenamento entre os látices individuais e os BV (análise conjunta dos látices), em relação aos BLUP. Seleção de dez progênes dentre 52. Teste de Progênes de *E. grandis*, sete anos. 71

Tabela 14: Variâncias das predições, variâncias (média) do erro das predições para o BLUP, como porcentagens da variância aditiva estimada. Teste de Progênes de *E. grandis*, sete anos. 71

Tabela 15: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de 10 progênes das 52, baseada em suas médias marginais (BV). Teste de Progênes de *E. grandis*, aos sete anos. 73

Tabela 16: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de 10 progênes das 52, baseada em suas médias marginais (BLP). Teste de Progênes de *E. grandis*, aos sete anos. 73

Tabela 17: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de 10 progênes das 52, baseada em suas médias marginais (BLUP). Teste de Progênes de *E. grandis*, aos sete anos. 73

Tabela 18: Predições de valores genéticos aditivos para a seleção para um ambiente intermediário pelo método BLP, para os quatro caracteres, para as progênes não comuns e testemunhas. Teste de Progênes de *E. grandis*, sete anos. 74

Tabela 19: Predições de valores genéticos aditivos para a seleção para um ambiente intermediário pelo método BLUP, para os quatro caracteres, para as progênes não comuns e testemunhas. Teste de Progênes de *E. grandis*, sete anos. 74

Tabela 20: Pesos calculados para o índice de seleção adotado, segundo os caracteres e níveis de observação; variância das predições e acurácia: a partir das estimativas das componentes da variância apresentadas por MORAES (1987).
Teste de Progênes de *E. grandis*, sete anos. 76

Tabela 21: Pesos calculados para o índice de seleção adotado, segundo os caracteres e níveis de observação; variância das predições e acurácia: a partir das estimativas das componentes da variância estimadas sob o modelo misto. Teste de Progênes de *E. grandis*, sete anos. 76

Tabela 22: Progênes e número de árvores por progênie selecionados pela seleção sequencial e pelo índice e tamanho efetivo populacional resultante dos dois métodos para o DAP. Teste de Progênes de *E. grandis*, sete anos. 80

Tabela 23: Progênes e número de árvores por progênie selecionados pela seleção sequencial e pelo índice e tamanho efetivo populacional resultante dos dois métodos para a altura. Teste de Progênes de *E. grandis*, sete anos. 80

Tabela 24: Progênes e número de árvores por progênie selecionados pela seleção sequencial e pelo índice e tamanho efetivo populacional resultante dos dois métodos para o volume. Teste de Progênes de *E. grandis*, sete anos. 81

Tabela 25: Progênes e número de árvores por progênie selecionados pela seleção sequencial e pelo índice e tamanho efetivo populacional resultante dos dois métodos para a densidade. Teste de Progênes de *E. grandis*, sete anos. 81

Tabela 26: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção sequencial de árvores dentro das melhores progênes. Teste de Progênes de *E. grandis*, aos sete anos. 83

Tabela 27: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de árvores baseada no índice de seleção. Teste de Progênes de *E. grandis*, aos sete anos. 84

MODELOS MISTOS NA PREDIÇÃO DE VALORES GENÉTICOS ADITIVOS EM TESTES DE PROGÊNIES FLORESTAIS

Autor: JÚLIO SÍLVIO DE SOUSA BUENO FILHO

Orientador: Prof. Dr. ROLAND VENCovsky

RESUMO

Diferentes formas de predição de valores genéticos aditivos foram comparadas na seleção de matrizes e de árvores dentro de progênies, a partir de testes de progênies florestais de polinização livre, em diferentes ambientes. Foram analisados um caso teórico e outro com dados reais envolvendo quatro caracteres: diâmetro medido a 1,30m de altura (DAP), altura das árvores, volume cilíndrico e densidade básica da madeira, e considerada apenas a análise univariada, de forma a se ressaltar a comparação entre os métodos de melhoramento.

A análise conjunta pelo modelo de látice com recuperação da informação interblocos foi comparada à análise por modelos com efeitos de progênies supostos aleatórios e de ambientes e blocos, supostos fixos. Abordou-se a seleção de matrizes para um pomar de sementes por clones de primeira geração e na seleção de árvores dentro de progênies para um pomar de sementes por clones de segunda geração.

Comparou-se teoricamente a seleção pelo método usual de ordenamento com base em médias marginais de progênies, que representam o valor genético aditivo da matriz, "BV", (do inglês "breeding value") com a seleção pelo melhor preditor linear (BLP), tomado a partir de médias de progênies por ambiente, que leva em conta apenas as diferenças de variabilidade fenotípica entre os ambientes. Estabeleceram-se expressões do BLP, das variâncias das predições e da acurácia seletiva em termos das componentes da variância usuais do

melhoramento vegetal. Verificou-se que a abordagem BLP foi sempre igual ou superior à BV: igual para ambientes com variância fenotípica uniforme e superior nos demais casos. Tais resultados permitem generalizações para casos mais complexos (maior número de efeitos fixos, aleatórios e interações) e mostram que deverão ocorrer erros maiores na seleção pelo método convencional para três ou mais ambientes.

Comparou-se o BV com o BLP e também com o BLUP (melhor preditor linear não viciado) em um ensaio real de *Eucalyptus grandis* (Hill) Maiden. Ambos, BLP e BLUP, apresentaram propriedades estatístico-genéticas superiores ao BV, como maior dispersão das predições e maior acurácia seletiva. Ambos apontaram diferenças de ordenamento quanto ao BV. As estimativas do progresso genético pelo BLP foram pouco superiores às do BV e os maiores ganhos foram os do BLUP.

Comparou-se a seleção de árvores individuais pelo método sequencial (entre progênies e dentro de parcelas) com um índice de seleção multiefeitos, no mesmo teste de progênies de *E. grandis*. As árvores selecionadas pelos dois métodos foram muito diferentes, com forte concentração de árvores selecionadas nas melhores progênies na seleção para os caracteres de crescimento pelo índice, o inverso ocorrendo para a densidade. As melhoras nos ganhos genéticos e nas respostas correlacionadas com o índice foram pouco expressivas se comparadas à forte redução no tamanho efetivo populacional que pode acarretar. A predição de valores genéticos individuais pode orientar, todavia, uma estratégia de seleção de árvores dentro de um número de progênies mínimo pré-estabelecido que preserve o tamanho efetivo final.

Foram discutidas situações em que o BLP, o BLUP e o índice de seleção, proporcionam maior vantagem com respeito aos métodos tradicionais e enfatizadas as possibilidades do emprego dos modelos mistos na predição de valores genéticos aditivos no melhoramento florestal.

MIXED MODELS FOR PREDICTING BREEDING VALUES IN FOREST TREE PROGENY TRIALS

Author: JÚLIO SÍLVIO DE SOUSA BUENO FILHO

Adviser: Prof. Dr. ROLAND VENCovsky

SUMMARY

Different methods of predicting breeding values from open pollinated progeny trials in multiple environments were compared considering: a theoretical evaluation and an analysis of real data. Four traits were analyzed: diameter at 1.30m height (DBH), tree height, tree volume and wood density. A single trait analysis was used to restrict comparisons to breeding methods only.

Combined analysis of lattice designs with recovery of interblock information was compared with an alternative model in which progeny effects were taken as random effects and environmental and block as fixed effects, in a backward selection approach to set up a first generation clonal seed orchard and in a forward selection (second generation clonal seed orchard).

Several explicit expressions were worked out, which involve only variance components usually estimated in plant breeding for: the best linear prediction of breeding values of progenies (BLP method); the breeding value for selection based on progeny means (BV method); the variance of predictions and the accuracy of selection by both methods. The BLP approach was found to be equal to or better than BV when phenotypic variances are uniform. These results hold generalizations for more complex cases (models concerning higher number of fixed parameters, random variables, and mixed interactions). Higher selection errors should be

expected with the BV method applied to three or more environments in forest breeding programs.

The BLP and an iterative best linear unbiased prediction method (BLUP) were compared with the BV method using data of an *Eucalyptus grandis* (Hill) Maiden progeny trial. Both BLP and BLUP showed superior statistic-genetical properties, as higher dispersion of predictions and higher precision. Both also showed that the BV method leads to higher selection errors. For comparison of genetic gain predictions, BLP was slightly superior to BV. BLUP, on the other hand, resulted in greater expected gains.

Sequential selection (among progenies and within plots) was compared with an index that estimates the individual breeding values, in the same *E. grandis* progeny trial. Results showed that different trees were selected by each of the methods. For growth traits the index led to a high concentration of selected trees belonging to a small number of superior progenies. The opposite was found for wood density. Additional gains with index selection were inexpressive due to the strong limitation in effective population sizes. Predicting individual breeding values can however, orient in selecting superior trees belonging to a pre established minimum number of progenies to maintain a sufficiently high final effective size.

Situations of forest improvement where discussed in which prediction of breeding values by BLP, BLUP and the related index are better tools than conventional ones. This may help in clarifying the advantages and potential of applying mixed models in progeny test analysis.

1. INTRODUÇÃO

A genética quantitativa tem contribuído com a maior parte dos avanços nos programas de melhoramento ao longo do século. Os testes de progênie são uma forma de se predizer o valor genético dos pais com base no fenótipo de sua descendência (ALLARD, 1971). Esta abordagem, mais que centenária na prática dos melhoristas, foi ajustada a modelos estatísticos como a análise da variância (ANAVA), talvez a maior responsável pelo surto de desenvolvimento da área desde sua proposição por Fisher nos anos vinte (Van BUIJTENEN, 1988).

EISENHART (1947)¹ mostrou que as pressuposições da análise da variância são adequadas à resolução de duas classes de problemas, ambos de estimação, quais sejam: estimar os valores constantes de relações entre efeitos de tratamentos (modelo fixo, ou modelo I) ou das chamadas componentes da variância das variáveis aleatórias (modelo aleatório, ou modelo II). A decisão sobre o uso de um ou outro modelo depende do propósito da estimação.

Tais modelos e formas de estimação mostraram-se bastante eficientes na orientação de estratégias de melhoramento, tanto para espécies anuais como para o melhoramento florestal e animal. No entanto, a necessidade de se elaborar modelos

¹ EISENHART, C. "The assumptions underlying the analysis of variance" *Biometrics*, v.1, n.3, pp1-21. 1947.: "I- a detecção e estimação de relações fixas (constantes) entre as médias de subclasses de um universo de objetos considerados; e II- a detecção e estimação de componentes da variação (aleatória) associados a uma população composta." (trad.)

com objetivos múltiplos para a seleção em delineamentos fortemente desbalanceados, com diferentes efeitos perturbadores para a seleção (como os do melhoramento animal) originou a busca de generalizações da ANAVA em modelos lineares mistos, cada vez mais livres de restrições.

A aplicação de delineamentos estatísticos mais eficientes que contornem os problemas oriundos da grande heterogeneidade ambiental em ensaios florestais, além da possibilidade da realização da predição de valores genéticos (aditivos) individuais faz com que muitos dos desenvolvimentos do melhoramento animal ganhem a atenção dos melhoristas florestais.

Dentre estes desenvolvimentos, destaca-se o conceito de que testes de progênies podem ser melhor descritos como modelos estatísticos mistos, e como tal analisados (COCHRAN, 1951; HENDERSON, 1975, entre outros). Esta abordagem implica em uma mudança brusca nos métodos de estimação que em geral tratam os modelos como fixos para fins de seleção (com base nas suas estimativas de médias de progênies) e aleatórios ou mistos para fins da estimação das componentes da variância (igualando-se esperanças de somas de quadrados ou quadrados médios da ANOVA, realizada num contexto fixo, em que o único efeito aleatório é o erro experimental, de variância homogênea) e realizações de testes exatos para as comparações entre os efeitos fixos levando em conta as esperanças matemáticas das formas quadráticas.

No contexto dos modelos mistos, procura-se **estimar** os efeitos fixos, ou seja, estabelecer as esperanças das constantes como as relações entre fatores fixos e as variâncias das variáveis aleatórias, bem como **predizer**, ou seja, achar a esperança condicional (futuras ou em outros ensaios) das variáveis aleatórias, dadas suas realizações (observações medidas nos ensaios) e os efeitos fixos já estimados.

Os termos estatísticos associados a cada uma destas técnicas são, respectivamente, determinar os BLUE ("Best Linear Unbiased Estimators") dos efeitos fixos e os BLUP ("Best Linear Unbiased Predictors") dos efeitos aleatórios, através das equações dos modelos mistos MME ("Mixed Model Equations").

Há aqui implícita a pressuposição de que as estimativas das variâncias das variáveis aleatórias influem nas estimativas dos efeitos fixos e ambas as estimativas influem na predição dos efeitos aleatórios. Há portanto o risco de a decisão sobre a natureza dos efeitos (se fixos ou aleatórios) levar a diferenças de seleção.

A aplicação prática dos modelos mistos envolve, no entanto, desenvolvimentos de estatística que não são ainda familiares ao melhoramento vegetal, bem como uma escolha criteriosa de que modelos genéticos podem ser assim representados. Boa parte destes desenvolvimentos já se encontram prontos para aplicações no melhoramento florestal e compilados em livros-texto, como o de WHITE e HODGE (1989), trabalho fundamental para a compreensão do estado da arte no setor.

Neste sentido, evitou-se enfocar a adequação dos modelos aos fenômenos, a sua robustez estatística, ou ainda o desenvolvimento de novos e complexos modelos incluindo muitos efeitos genéticos e ambientais. Tais tarefas têm ocupado a pesquisa em genética quantitativa animal, com a aceitação geral dos modelos mistos de predição como a mais completa contribuição dos modelos lineares para o melhoramento.

Procurou-se enfocar os métodos mais diretos de predição, quase todos já abordados anteriormente por WHITE e HODGE (1989), aplicando-os a situações teóricas e práticas comuns ao melhoramento florestal. O objetivo geral deste trabalho foi, portanto, a identificação de situações em que o contexto de modelos mistos seja superior às abordagens tradicionais no

melhoramento florestal brasileiro, em especial ao melhoramento do eucalipto; situações nas quais o esforço adicional do emprego da nova metodologia seja claramente recompensado.

Abordou-se a seleção em um só caráter por vez, em testes de progênies florestais realizados em mais de um local, envolvendo progênies comuns. Para estes ensaios, procurou-se identificar os "erros de seleção" (e os possíveis decréscimos no ganho genético) pelo emprego das formas usuais de seleção quando comparadas ao modelo misto de predição equivalente.

Para tanto, procurou-se atingir alguns objetivos específicos, quais sejam:

- Comparar teoricamente a seleção de matrizes através da predição de valores aditivos de progênies obtidos a partir de médias de progênies por locais e por parcelas com o método tradicional da seleção pela média marginal das progênies;
- Comparar os efeitos da seleção de matrizes através da predição de valores aditivos de progênies, obtidos a partir das médias de progênies por parcelas, com o método tradicional da seleção pela média geral das progênies em um teste de progênies real de *Eucalyptus grandis*;
- Na seleção de árvores para um pomar de segunda geração a partir do mesmo teste de progênies de *E. grandis*, comparar um índice de seleção (a predição de valores genéticos de árvores individuais) com o método tradicional da seleção sequencial entre e dentro de progênies (seleção da melhor árvore de cada parcela de progênies selecionadas por suas médias gerais).

2. REVISÃO DA LITERATURA

2.1. REFERENCIAL DO MELHORAMENTO FLORESTAL

A genética florestal tem se caracterizado pela melhor compreensão da biologia reprodutiva e da genética de populações florestais, pelo avanço das aplicações da biotecnologia e pelo uso mais rigoroso da teoria da genética quantitativa na definição de estratégias de melhoramento, centrada em aumentos nos ganhos genéticos e diversificação dos objetivos (e espécies alvo) dos programas de seleção.

No Brasil, tem sido dada maior ênfase a espécies de importância industrial, tais como as de *Eucalyptus* e *Pinus*. As linhas gerais que orientam estudos sobre estratégias de melhoramento resumem-se, em flexibilidade na definição dos objetivos (melhorar caracteres de crescimento e outros complementares) e simplicidade no emprego de métodos de melhoramento, sendo que o grande desenvolvimento dos índices de seleção abre possibilidades para uma maior clareza na execução dos programas (KANOWSKI, 1993).

Apesar da grande exposição aos índices de seleção na literatura florestal não se tem encontrado a necessária correspondência numa definição precisa dos objetivos do melhoramento. Assim, é comum que se disponha de teoria para índices complexos em diferentes caracteres e com o uso de informações sobre relações de parentesco, sem que se adote estes índices na prática das empresas de melhoramento. No entanto, para as espécies que se dispõe de suficiente

informação de natureza econômica e genética, as vantagens dos índices tornam-se mais claras (WHITE e HODGE, 1989).

No Brasil, o emprego de modelos mistos no melhoramento florestal é ainda incipiente (ver a respeito, RESENDE, 1997), destacando-se os trabalhos de índices de seleção de RESENDE et alii (1993) e RESENDE e HIGA (1994); o de BUENO FILHO e VENCovsky (1993) que estuda a eficiência de um índice clássico (SMITH-HAZEL) aplicado à seleção combinada entre e dentro de famílias, em um só caráter e o de RESENDE e BERTOLUCCI (1993), também com um índice de seleção (FALCONER, 1981), além de RESENDE et alii (1995), que apresentam aplicações do tipo BLP e BLUP.

O uso da genética quantitativa permite a definição de critérios adequados para se julgar os métodos de seleção. Têm sido propostos neste sentido comparações com ênfase no ganho genético por ano (e.g. COTERRIL, 1986a, 1986b); ênfase nos custos associados e na eficiência econômica (Friedman, 1992 e Shelbourne 1992, citados por KANOWSKI, 1993); ou ainda avaliações das possibilidades de ganhos contínuos no melhoramento pela manutenção de ampla base genética para ideótipos não estritamente definidos, mais voltadas para espécies não industriais.

A seleção baseada exclusivamente nos índices pode ser bastante restritiva em termos do tamanho efetivo populacional resultante, mas o uso dos próprios índices é a forma mais flexível e eficiente de se maximizar os ganhos seja para uma dada base genética, para um dado ciclo de melhoramento, um montante de investimento, ou outro critério de comparação dos métodos (HODGE e WHITE, 1992).

Entre as estruturas populacionais nas quais o melhoramento tem se concentrado, destacam-se: núcleos elite de cruzamento dentro de uma população principal, constituídos pelos indivíduos com maior valor genético; populações múltiplas com objetivos distintos e complementares; "siblines" mantidas separadas para controlar o parentesco e

posteriormente cruzados na população de produção (LOWE e Van BUIJTENEN, 1986); ou ainda formas populacionais mais complexas como as derivadas dos híbridos interespecíficos, responsáveis pela maior parte do reflorestamento por clones em espécies de *Eucalyptus* e *Pinus* em que a propagação vegetativa é eficiente (NIKLES 1992, citado por KANOWSKI, 1993; LOODINKS, 1992; FERREIRA, 1992¹; MELLO, 1996²).

Sobre a eficiência do processo seletivo, a principal contribuição é ainda a maior compreensão das correlações adulto-juvenis (ver por exemplo KANG, 1991) sendo que o desenvolvimento da metodologia de modelos mistos para a análise BLUP dos ganhos por seleção é visto como uma grande aquisição mas ainda pouco empregado pelos melhoristas.

A predição de valores genéticos de árvores individuais é vista como uma ferramenta que permite comparar mais facilmente as consequênciass de procedimentos seletivos alternativos (pode-se estabelecer primeiro as predições e depois selecionar conforme qualquer método, não necessariamente a seleção truncada com base nos valores preditos, obtendo-se assim a melhor predição de ganho genético para o método em questão).

Como exemplo das propriedades que contribuem para a superioridade dos métodos de predição, temos a seguinte propriedade do BLUP ("Best Linnear Unbiased Predictor", ou melhor preditor linear não viesado): o BLUP de combinações lineares é a combinação linear dos BLUPs individuais.

Desta forma, a seleção em caracteres complexos pode se dar, inicialmente, pela predição BLUP de seus componentes mais simples, com o ajuste posterior de pesos econômicos (WHITE e HODGE, 1989; KENNEDY e SORENSEN, 1988; MAGNUSSEN, 1991).

¹ Mário Ferreira. (ESALQ/USP, Piracicaba, SP) Comunicação pessoal. 1992.

² Eduardo José de Mello e Shinitiro Oda. (Comercial e Agrícola Paineiras, Itapetininga, SP) Comunicação pessoal, 1996.

2.2. TESTES DE PROGÊNIES FLORESTAIS

O estudo de delineamentos complexos como dialelos tem sido de pouco reflexo na prática do melhoramento florestal, sua eficiência é questionada, por exemplo por COTERRIL (1986a e 1986b); porém podem ser de maior utilidade no melhoramento para objetivos complementares ou para o aproveitamento do crescente desenvolvimento das técnicas associadas aos marcadores moleculares como a seleção assistida por marcadores (LANDE, 1992) ou mesmo o estabelecimento de padrões mais precisos de identidade gênica que os coeficientes de parentesco.

Como exemplo, BERNARDO (1994) usou marcadores RFLP ("Restriction Fragment Length Polimorphism", ou polimorfismo de tamanho de fragmentos de restrição) para a obtenção de dados de identidade alélica e desta forma estabelecer os coeficientes de parentesco no uso dos BLUP para a predição de médias de híbridos em cruzamentos dialélicos incompletos em milho.

Segundo KANOWSKI (1993) o princípio de se limitar os objetivos dos testes de progênie florestal (estimar componentes da variância ou orientar a seleção) tem prevalecido na literatura, porém a prática de testes de progênies e procedências conjuntos, com o comprometimento das estimativas de componentes da variância genética em função dos objetivos de seleção parecem prevalecer nos ensaios das empresas, bem como na aceleração dos programas para espécies de menor interesse industrial. A maioria dos testes envolve pequeno número de progênies (raramente mais que cinquenta), que são muitas vezes um conjunto já selecionado de progênies resultando em estimativas imprecisas dos parâmetros genéticos e fenotípicos da população.

Em algumas espécies de coníferas a indução ao florescimento precoce tem tido bons resultados possibilitando a realização de cruzamentos controlados de tipos diversos, o

que torna viável pensar em delineamentos de cruzamentos envolvendo progênies de tipos diversos para o estudo e aproveitamento da ação gênica não aditiva, tanto para o melhoramento clássico como para a seleção assistida por marcadores.

De um modo geral, no entanto, os delineamentos se desenvolvem para tirar vantagens das características reprodutivas da espécie em questão e de acordo com as possibilidades que oferecem para o melhoramento (LINDGREN, 1986; NAMKOONG e KANG, 1990; BURDON, 1992).

Desta forma, prevalecem os testes de progênies de polinização livre, assumidas como sendo de meios-irmãos, combinados ou não com informações sobre procedências e locais. Tais delineamentos são relativamente pouco trabalhosos por não exigirem polinização controlada e permitem uma eficiente estimação da variância genética aditiva, parâmetro de maior importância em programas de seleção recorrente intrapopulacional.

A ausência de maior elaboração nos delineamentos e o emprego generalizado de progênies de polinização livre sofreu críticas por parte de VAN BUIJTENEN (1988), pois mesmo em testes com gerações mais avançadas de melhoramento esta tendência se mantém, podendo levar a distorções e reduções no potencial de ganho das populações causadas pelo descontrole da endogamia, hipótese que, segundo o autor, é em geral subestimada tanto por melhoristas práticos (interessados em mostrar predições de ganho bastante elevadas) quanto por geneticistas quantitativos interessados em mostrar resultados impressionantes pela adoção de técnicas sofisticadas de seleção.

Quanto aos delineamentos estatísticos, são bastante empregados os de blocos casualizados, blocos incompletos parcialmente balanceados (látices) ou arranjos de parcelas subdivididas no esquema experimental de blocos. O esforço para contornar a heterogeneidade ambiental em experimentos

envolvendo muitas árvores e anos em grandes áreas sugere o uso de delineamentos estatísticos eficientes como blocos incompletos, para o que também se destaca o advento de computadores capazes de analisar rapidamente tais experimentos (MAGNUSSEN, 1991).

A competição inter-árvores tem recebido renovado interesse. Destacam-se as sugestões de se reduzir o tamanho de parcelas, utilizar parcelas não contíguas ou modelos móveis como os de análise de vizinhança de forma a melhorar a precisão experimental (LOO-DINKINS, 1992).

O desenvolvimento de rotinas computacionais diretamente relevantes para a genética florestal como SAS VARCOMP, GENSTAT REML e mais recentemente, SAS MIXED, vem facilitar análises de maior qualidade estatística, porém introduzem dificuldades adicionais de interpretação. Assim, embora poderosas, muitas das ferramentas da genética quantitativa recente ainda apresentam consideráveis entraves à sua utilização, em especial nos casos em que as soluções ótimas são não triviais.

Um importante condicionante do melhoramento florestal é ainda o tempo decorrido entre o planejamento dos ensaios e a obtenção de resultados. Tal dificuldade pode levar a decisões de seleção pouco consistentes com ambientes de cultivo e prioridades econômicas que continuamente se alteram.

Estas circunstâncias provocam o aparecimento de interações complexas do tipo genótipos por ambientes que levam à redução na eficiência do processo seletivo. Estas interações são também relacionadas às diferenças no desenvolvimento e na fisiologia dos genótipos considerados nos ambientes de teste e de cultivo (MELLO e ODA, 1996³). Para TALBERT (1992) este fato, aliado às incertezas e à competição, típicas do reflorestamento comercial na era da informação, aumenta a importância das decisões de seleção.

³ Eduardo José de Mello e Shinitiro Oda. (ESALQ/USP, Piracicaba, SP) Comunicação pessoal, 1996.

Neste sentido, a metodologia de modelos mistos pode melhorar a qualidade da informação sobre efeitos fixos e também aleatórios em delineamentos complexos, podendo ainda resultar em informações flexíveis e na integração de resultados de várias fontes (por exemplo, vários ensaios e várias gerações da mesma população) de forma a melhorar continuamente (e até mesmo corrigir) as decisões de seleção nos programas de melhoramento.

2.3. MODELOS MISTOS

2.3.1. Introdução e Definições

Um modelo matemático é uma tentativa de se explicar as observações de uma variável (dependente) através dos "efeitos" diferenciais que se atribui a outra série de variáveis independentes (parâmetros do modelo). Tais efeitos podem ser de natureza fixa ou aleatória, conforme representem, respectivamente, constantes a serem estimadas ou realizações de uma variável aleatória com distribuição de probabilidade conhecida.

Modelos lineares estatísticos são os que apresentam ao menos um efeito aleatório (comumente chamado erro experimental). Se um modelo apresenta todos os demais componentes fixos, é chamado de "modelo fixo", se no entanto, todos os demais fatores forem aleatórios (a menos de uma constante, para modelos outros que não o de médias de caselas) o modelo é chamado de "aleatório"; quando o modelo apresenta tanto efeitos aleatórios como fixos, é denominado de "modelo misto".

Em geral, o estabelecimento de modelos está ligado à possibilidade de se **estimar** parâmetros associados às definições do modelo ou de se **predizer** o comportamento das suas variáveis aleatórias para um dado conjunto de observações; isto porque se diz que os *parâmetros populacionais* (esperança de efeitos fixos e variâncias populacionais de variáveis aleatórias) são *estimáveis*,

enquanto que as *variáveis aleatórias* "v.a." podem ser *preditas*, mas não estimadas, pois não possuem valor fixo "per se", mas numa amostra dos seus possíveis valores, pode-se obter indicadores de sua esperança, conhecida sua distribuição e na suposição de o efeito aleatório estar correlacionado ao caráter observado (WHITE e HODGE, 1989).

Assim, na análise de modelos fixos deve-se estimar os próprios efeitos fixos dos parâmetros, enquanto que os modelos aleatórios prestam-se para estimar as componentes da variância (das *variáveis aleatórias*), bem como para a predição das próprias v.a.. Ou seja, os modelos mistos podem servir tanto à estimação de médias de um modo mais preciso (modelo generalizado, levando em conta a influência das componentes da variância), estimativas das componentes da variância, ou ainda para a predição.

O conceito consagrado na prática agrônômica para o enquadramento de efeitos como fixos ou aleatórios é relacionado ao objetivo da análise: se os níveis do efeito constituem-se amostras de uma população sobre a qual se quer tirar conclusões (efeitos aleatórios) ou se são parâmetros constantes sobre os quais se quer identificar diferenças e magnitudes (efeitos fixos). Muitos autores, no entanto, entendem tal diferença como banal, sem consequências para o processo de estimação-predição.

Assim, VIEIRA e HOFFMAN (1988), por exemplo, afirmam que, com efeitos aleatórios ao invés de fixos, apenas as esperanças matemáticas dos quadrados médios e testes "F" se alteram. É implícito que não se reconhecem alterações de conjunto na análise de um experimento por se pressupor algum dos fatores como aleatórios ao invés de fixos (num contexto de estimação de mínimos quadrados ordinários "OLSE").

Em verdade, para o rigor da pressuposição, mesmo para dados completamente balanceados, as componentes da variância presentes transformariam tal modelo num modelo generalizado

para a estimação dos BLUE, e numa análise BLUP para os fatores aleatórios (SCHWARZ, 1993).

Portanto, considerar fatores como fixos ou aleatórios, longe de ser prática inofensiva, pode afetar a estimação e suas consequências pela presença das variâncias das v.a. perturbadoras, ou seja, pode alterar os testes de hipóteses sobre contrastes paramétricos, como a magnitude de testes "F", e conclusões sobre parâmetros, como estimativas de componentes da variância. Ver a este respeito (SCHEFFÉ, 1959; HOCKING, 1985; VENCOVSKY e BARRIGA, 1992).

Por outro lado, ignorar as covariâncias entre observações de um mesmo nível da v.a. pode também afetar a qualidade da predição, pela alteração no ordenamento dos valores das predições dos efeitos aleatórios ou pela simples mudança na magnitude destes valores (WOLFINGER, 1993).

Mais diretamente, ignorar o fato que as progênies de polinização livre de um teste de progênie devem ser representadas por efeitos aleatórios para fins do próprio modelo da análise, pode levar a alterações no ordenamento dos valores aditivos entre as progênies, ou seja, a erros de seleção.

2.3.2. Métodos de Predição

Pode-se restringir as diferentes metodologias de predição (de valores genéticos) sob modelos mistos aos índices de seleção, os BLPs ("Best Linear Predictors", ou melhores preditores lineares), os BLUPs ("Best Linear Unbiased Predictors", ou melhores preditores lineares não-viesados), e os BLUPs iterativos. As formas de aplicação de cada uma destas técnicas aparentemente não se relacionam, mas constituem na verdade um conjunto de generalizações sucessivas, historicamente encadeadas pelo trabalho da genética quantitativa animal desde a década de 1940 (QUAAS e POLLAK, 1980; HENDERSON, 1984 e 1988).

Tais metodologias permitiam inicialmente apenas a predição de valores genéticos aditivos supondo conhecidas suas médias e variâncias, a validade do modelo linear, a normalidade da distribuição de y e a homogeneidade de variâncias (o caso dos índices de seleção).

Pelo emprego de métodos derivados do BLUP no melhoramento animal são possíveis predições de valores genéticos aditivos, dominantes, epistáticos, efeitos de sexo e maternos, bem como estimativas das suas componentes de variância, concatenando-se diferentes dados (caracteres, gerações e delineamentos), pressupondo apenas a linearidade do modelo e normalidade na distribuição da variável aleatória (HENDERSON, 1988).

A própria apresentação dos modelos de predição passou a ser conhecida como "modelo animal". Enfatizava tipicamente a matriz "**A**", de variâncias e covariâncias genéticas entre os animais, assumindo a possibilidade (e até mesmo a frequência) de se usar um tal modelo como explicativo de observações em animais de gerações sobrepostas, diferentes graus de consanguinidade e observados em ensaios muito diferentes, concatenados por estimativas gerais de efeitos fixos (como sexo, raça, ano, etc...) e pelas progênies comuns (frequentemente poucas). Para estas situações, uma necessidade computacional imediata era o estabelecimento de processos de construção e inversão da matriz **A** (ver Van VLECK, 1993).

Para as aplicações do melhoramento florestal, no entanto, tais procedimentos, até desejáveis, são ainda pouco explorados. Uma primeira e radical diferença de conceito é a assunção geral do modelo aleatório (para estimar componentes da variância e progressos pela seleção truncada). Outra diferença, mais substancial, é a qualidade dos testes pois, em geral, é possível instalar ensaios com grande número de progênies de polinização livre ("maternas") em cada ambiente, com muitas progênies comuns a todos os ambientes, todas de

uma mesma geração e idade, oriundas de uma população assumida como não endogâmica.

Estas características deram origem a uma abordagem genética e estatisticamente mais simples para o emprego dos modelos mistos no melhoramento florestal (e também agrícola) cujo potencial foi evidenciado através de trabalhos pioneiros como WHITE e HODGE (1989); WHITE, HODGE e POWELL (1989) (1989); MAGNUSSEN e YEATMAN (1990).

Tais abordagens ainda não atingiram ampla divulgação, seja pela relativa dificuldade dos melhoristas em aceitar mesmo os preditores mais diretos (índices e BLPs), seja pela falta de segurança em utilizar uma tal metodologia em substituição aos métodos tradicionais derivados da análise da variância (ANAVA), uma vez que esta, afinal, tem permitido um bom trabalho de seleção no melhoramento prático.

2.3.3. Tipos de preditores

O melhor preditor para qualquer tipo de modelo requer que a distribuição de y e seus momentos sejam conhecidos. Desta forma o melhor preditor é a média condicional $E(K'\beta + Z'u|y)$, que é não-viezada e possui variância mínima entre todas as funções possíveis preditoras das observações y (COCHRAN, 1951; RAO, 1965). Qualquer destas funções poderá ser chamada de "best predictor", sendo que o termo "best" indica mínima variância do erro entre todos os preditores. Se dentre estas funções forem consideradas apenas as que são lineares nos parâmetros, é preciso conhecer apenas os primeiros e segundos momentos de y [médias, $X\beta$, e variâncias, $V(y)$] e não sua distribuição. Desta forma o melhor preditor linear passa a ser escrito: $BLP(u) = E(K'\beta + Z'u) + C'V^{-1}(y - X\hat{\beta})$, em que $C' = Cov[(K'X\beta + Z'u), y]$; assim, quando y tiver distribuição multivariada normal, BLP será BP.

Em geral, $X\beta$ e V são desconhecidas, porém pode-se assumir que se conhece $\hat{V}$ (uma estimativa de V). Neste caso,

o preditor é não viesado com respeito às médias, obtendo-se assim

$$BLUP(u) = K' \hat{\beta} + C' V^{-1}(y - X\hat{\beta}),$$

em que:

$$\hat{\beta} = (X' V^{-1} X)^{-1} X' V^{-1} y,$$

em que o sinal " $()^{-1}$ " simboliza a operação de inversa generalizada da matriz que está dentro do parênteses.

De todos os preditores lineares não viesados de y , BLUP é o que minimiza a variância do erro. No entanto, caso y não possua distribuição normal, pode haver outras funções $K'\hat{\beta} + Z'u$ com menor variância do erro que a do BLUP.

Deve-se notar que há uma leve diferença, aparentemente apenas teórica, entre BLUP, BLP e BP, basicamente, para diferenciar BLP de BLUP. Deve-se perceber que apenas as componentes da variância são supostas conhecidas no caso dos BLUPs, as estimativas de efeitos fixos são portanto BLUEs. No caso dos BLPs, parte-se da suposição que, tanto as componentes da variância quanto as médias dos efeitos fixos são conhecidas (tomadas como sem erro). Assim, BLP é um caso particular de BLUP.

Ao processar a predição, obtém-se novos valores para $X\hat{\beta}$ e $Z\tilde{u}$, que permitem nova estimação das componentes da variância, num processo iterativo (um exemplo destas iterações pode ser visto em SILVA, 1982). Esta é uma possível maneira de se incrementar sucessiva e assintoticamente a qualidade das estimativas das componentes da variância, das esperanças dos efeitos fixos e das predições, baseada nos resultados de HENDERSON (1975).

Esta última abordagem pode ser vista como um novo processo (BLUP iterativo) e é a maneira com que rotinas computacionais como SAS-MIXED promovem a estimação-predição no contexto de modelos mistos (ou aleatórios superparametrizados) (SAS, 1992).

2.3.3.1. Forma geral dos Índices de seleção

A realização da predição através de índices pode ser descrita pelo ajuste de um índice do tipo:

$$I = b_n [y - E(y)],$$

em que b_n é o "peso" e $E(Y)$ é a esperança matemática, correspondentes ao nível de observação considerado (y).

Além da sugestão de WHITE e HODGE (1989), podem-se encontrar exemplos de aplicação de índices na seleção de árvores individuais no modelo misto ou aleatório, a partir de estimativas de componentes da variância tradicionais do método dos momentos nos trabalhos de Coterril, para espécies de eucaliptos na Austrália [ver COTERRIL e JACKSON (1985), COTERRIL, 1986a], e também em KLEIN (1995), para *Pinus banksiana*, no Canadá, num trabalho cujo objetivo foi apresentar e discutir os BLP de árvores individuais. Tais trabalhos têm em comum a exploração da capacidade dos índices de representar formas de seleção combinada, entre e dentro de progênies.

2.3.3.2. Forma geral dos BLP

A predição de valores genéticos através dos BLP envolve a determinação de vetores de pesos que se aplicam ao nível de observação considerado de maneira análoga aos índices. Para enfatizar que se quer uma combinação linear preditora dos valores genéticos, faz-se o modelo da predição ser apresentado para cada progênie (ou valor genético a ser predito) como: $BLP(g) = \hat{g}_i = a_i + b_i' y$, de forma que o valor de a e b' sejam tais que $E(g_i - \hat{g}_i)$ seja mínimo.

Desta forma, os BLP são uma generalização dos índices clássicos Smith-Hazel, só que nos índices um único conjunto de pesos b' aplica-se a todos os dados, enquanto que nos BLP diferentes índices podem surgir para cada progênie (ou valor genético aditivo a ser predito).

A equação habitual da predição BLP como apresentada por HENDERSON (1963) é a que segue:

$$\mathbf{BLP}(\mathbf{g}) = \hat{\mathbf{g}}_i = \boldsymbol{\tau} + \mathbf{C}'\mathbf{V}^{-1}(\mathbf{y} - \boldsymbol{\alpha}), \text{ em que:}$$

$\boldsymbol{\tau} = \mathbf{E}(\mathbf{g})$ é o vetor de valores genéticos esperados dos valores genéticos a serem preditos (no melhoramento florestal pode significar, por exemplo, o valor médio da procedência, ou pode ser tomado como 0 para uma mesma população);

$\mathbf{C}'$ = matriz de covariâncias entre o valor genético a ser predito e as observações;

$\mathbf{V}^{-1}$ = inversa da matriz de variâncias e covariâncias das observações;

$\mathbf{y}$ = vetor de observações;

$\boldsymbol{\alpha} = \mathbf{E}(\mathbf{y})$ é o vetor de valores esperados das observações (associado aos efeitos fixos, que no caso dos BLP são supostos conhecidos).

A porção relativa aos pesos de cada observação é $\mathbf{b}' = \mathbf{C}'\mathbf{V}^{-1}$ o que faz a analogia entre BLP e índices Smith-Hazel: em ambos os casos, supõem-se conhecidos os efeitos fixos e as componentes da variância. Caso se admitam conhecidas apenas as componentes da variância, tem-se a abordagem BLUP. Caso os efeitos fixos e componentes da variância sejam estimados no processo de predição, tem-se a abordagem BLUP-iterativo, também chamada BLUP_e (SAS) ou ainda "predição-REML", quando as componentes da variância forem estimadas pelo método da máxima verossimilhança restrita.

WHITE e HODGE (1989) e HODGE e WHITE (1992) apresentam as propriedades dos BLP, análogas às dos BLUP, envolvendo apenas a restrição adicional de supor conhecidos (sem erro) os efeitos fixos. Os autores apresentam possíveis vantagens dos BLP sobre os BLUP na prática florestal. Tais vantagens estariam associadas ao menor esforço computacional e à produção de soluções para problemas como:

- a) a seleção para um ambiente-alvo em ensaios com diversos ambientes, utilizando informações de todos os ambientes [também enfocada por RESENDE et alii (1993)];
- b) a seleção decompondo a estimativa da variância devida a interações do tipo genótipos por ambientes [também enfocada por HELGADOTTIR et alii (1995)];

Para o emprego de preditores BLP consideram-se conhecidos os efeitos fixos (sem erro), no entanto, com a obtenção de estimativas de efeitos fixos no modelo generalizado (com heterogeneidade da matriz V), não há diferença entre BLP e BLUP para fins dos primeiros momentos, apenas a estimativa da precisão do processo de predição aparece como superestimada para os BLP em relação aos BLUP (WHITE e HODGE, 1989).

KLEIN (1995) apresenta uma alternativa interessante de análise BLP para a seleção combinada de árvores individuais (uma matriz V para cada árvore) que antes apenas tinha sido proposta para índices ou BLUP (WHITE e HODGE, 1989; MARTINS et alii, 1993). O uso de tal método pode propiciar ordenamento tão correto quanto o BLUP, com menor esforço computacional, pois trata-se de matrizes menores, da ordem do número de árvores por progênie.

2.3.4. Estimação de BLUEs e Predição de BLUPs

Para o modelo misto $y = X\beta + Zu + e$, em que u é o vetor dos valores genéticos aditivos (tal modelo é assim empregado no melhoramento animal), a predição pode ser vista como sendo a esperança condicional destes valores, dado um conjunto de realizações fenotípicas. O problema é então determinar a esperança $E(u|y)$, que é chamada "melhor preditor de u ", que sob normalidade assume a forma:

$$\tilde{u} = E(u) + \text{cov}(u, y') [\text{var}(y)]^{-1} [y - E(y)].$$

Esta equação do BLUP permite visualizar a analogia com BLP, sendo $E(u)$ análoga a τ , $\text{cov}(u, y') [\text{var}(y)]^{-1}$, igual a $C'V^{-1}$ e supondo conhecida a estimativa $E(y)$ (análoga a α).

Para o caso de modelos mistos com a variância das variáveis aleatórias dada por $\text{var}(u) = G\{\sigma_i^2 q_i\} = G$ (matriz), com $\text{cov}(u, y') = \text{cov}(u, u'Z') = GZ'$, (análoga ao vetor C apresentado para o BLP) e $\text{var}(y) = ZGZ' + \sigma_e^2 I = V$, sendo a matriz de variâncias e covariâncias das observações y , tem-se que:

$$\text{BLUP}(u) = \tilde{u} = GZ'V^{-1}(y - X\hat{\beta}), \text{ e para a parte fixa:}$$

$$\text{BLUE}(X\beta) = X\hat{\beta} = X(X'V^{-1}X)^{-1}X'V^{-1}y.$$

Com G e σ^2 conhecidas o único problema passa a ser inverter matrizes da ordem do número de observações. Esta dificuldade pode ser contornada pelo uso das chamadas "equações de modelos mistos" (equivalentes às equações normais para modelos fixos particionados, com a devida correção para os efeitos aleatórios):

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \sigma_e^2 G^{-1} \end{bmatrix} \begin{bmatrix} \beta^* \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}.$$

Basta então inverter matrizes de ordem $p+q$ (número de efeitos fixos e aleatórios, respectivamente). Segundo SEARLE (1987) esta igualdade foi primeiro demonstrada por Henderson et alii (1959).

Se G (e portanto V) não for conhecida, pode-se utilizar suas estimativas de forma análoga para obter BLUE_V e BLUP_V , ou seja, estimativas e predições tão adequadas quanto possíveis para um dado ensaio.

Note-se que, para dados balanceados, no modelo fixo, os BLUEs reduzem-se a estimadores de mínimos quadrados ordinários (OLSE). Neste caso, o processo de estimação para componentes de variância associa-se à estimação de médias de efeitos fixos e à predição de (médias) de variáveis aleatórias, numa construção teórica mais robusta que a simples reinterpretação da ANAVA (para modelos mistos ou

modelos aleatórios superparametrizados). Os modelos aleatórios superparametrizados podem ser vistos como casos especiais de modelos mistos com a média sendo sempre fixa, logo, a média geral é um efeito fixo que pode estar mal estimado em função de perturbações nas componentes da variância.

Pode ser demonstrado também que, na análise tradicional dos látices com recuperação da informação interblocos, o modelo considerado é o de tratamentos fixos e blocos aleatórios, assim, os BLUEs dos tratamentos correspondem às médias ajustadas do látice, estimadores para comparações interclasse e intraclasse (SEARLE, 1987).

2.3.5. Componentes da Variância

Um dos problemas centrais da aplicação dos modelos mistos é a estimação das componentes de variância (e covariâncias) dos efeitos aleatórios. Esta questão é muitas vezes tratada como acessória para os interesses últimos do melhoramento prático, em que o maior interesse, como se sabe é o ordenamento de progênies (ou indivíduos) para a seleção. Nas abordagens comumente empregadas, as componentes da variância não influem no ordenamento das progênies, limitando sua influência à estimativa do progresso pela seleção e à precisão do processo seletivo.

Na abordagem de modelos mistos, uma vez estimada $\mathbf{V}$, toma-se a estimativa pelo parâmetro na expressão do $\text{BLUE}_{\mathbf{V}}$ e $\text{BLUP}_{\mathbf{V}}$ para que se proceda à estimação e à predição. Kackar e Harville (1981; 1984), citados por SEARLE (1987), mostraram que $E\{\text{BLUE}_{\mathbf{V}}\} = \text{BLUE}$, para toda uma gama de estimadores de $\mathbf{V}$; os autores também mostraram como calcular aproximações de $\text{var}[\text{BLUE}_{\mathbf{V}}(\lambda' \mathbf{X} \boldsymbol{\beta})]$, ou seja, dos BLUEs das funções estimáveis.

Desta forma, uma determinação precisa das estimativas das componentes da variância melhora tanto o processo de estimação como o de predição, podendo resultar em alterações

no ordenamento e na magnitude, tanto dos efeitos fixos como das variáveis aleatórias (base do ordenamento para a seleção).

A estimação de componentes da variância em modelos balanceados é o caso mais simples e de onde deriva boa parte da metodologia para dados desbalanceados. O método dos momentos (ou da ANAVA) é o mais comumente empregado, consistindo em se igualar formas quadráticas (expressões para cálculo de somas de quadrados ou quadrados médios) a suas esperanças (combinações lineares das componentes de variância), obtendo-se assim um conjunto de equações que permite a estimação.

Se o desbalanceamento ocorre apenas ao nível das subclasses (dentro das parcelas, por exemplo) o método dos momentos produz estimativas relativamente precisas. O desbalanceamento torna-se mais problemático quando se dá nos níveis das caselas, como no delineamento em blocos incompletos (em que há blocos que não recebem todos os tratamentos) ou no caso de parcelas perdidas no delineamento em blocos completos.

Há uma multiplicidade de métodos de estimação (mesmo entre os chamados "Métodos dos Momentos"), e o desenvolvimento de pacotes computacionais tornou disponível para fins práticos uma série de metodologias antes consideradas de difícil cômputo. Para os mais importantes, SEARLE (1987) e SEARLE et alii (1992) apresentam os procedimentos de estimação e os méritos e deméritos dos estimadores. Serão abordados aspectos dos estimadores de "Máxima Verossimilhança" (ML) e os de "Máxima Verossimilhança Restrita" (REML), que têm sido os mais recomendados para fins do melhoramento.

2.3.5.1. Método dos momentos (ou da ANAVA)

A aplicação do método da ANAVA para dados balanceados em modelos mistos (MM) é direta. Casos particulares podendo ser encontrados em textos padrão de Modelos Lineares, como GRAYBILL (1976) e HOCKING (1985). Note-se que, mesmo para MM balanceados, surgem polêmicas sobre o modelo básico e o significado das componentes da interação de efeitos fixos com aleatórios (ver por exemplo, SCHEFFÉ 1959, SEARLE 1971 e KEMPTHORNE 1975).

Uma generalização da ANAVA para dados desbalanceados funciona de forma análoga aos dados balanceados. Assim a estimativa de um conjunto de formas quadráticas " q " pode ser dada por:

$$E(q) = M \hat{\sigma}^2 \Leftrightarrow \hat{\sigma}^2 = M^{-1} q$$

para alguma matriz M , não-singular, e com os elementos de q sendo quaisquer formas quadráticas das observações, tais que valham as equações acima. Dessa forma, pode haver muitos conjuntos diferentes de equações para estimar σ^2 e o problema torna-se determinar "o melhor". Para dados desbalanceados, pode não haver um conjunto melhor (SEARLE, 1987).

Os três métodos de HENDERSON (1953) são também formas de aplicação do método dos momentos, embora as formas quadráticas que empregam não se constituam sempre somas de quadrados clássicas da ANAVA.

Sob normalidade, variâncias amostrais dos estimadores são de fácil obtenção, pois os quadrados médios da ANAVA são independentes e seguem distribuições proporcionais a χ^2 . Matricialmente, a dispersão de $\hat{\sigma}^2$ para um delineamento qualquer desbalanceado é dada por (SEARLE, 1987):

$$\text{Var}(\hat{\sigma}^2) = M^{-1} \text{var}(q) M^{-1'}.$$

em que os elementos de $\text{var}(\mathbf{q})$ são variâncias e covariâncias entre as formas quadráticas usadas como elementos de $\mathbf{q}$.

Os estimadores oriundos do método dos momentos são em geral não-viesados, seu principal problema é que não previnem contra o aparecimento de estimativas negativas (variâncias são parâmetros sempre positivos). Não obstante, estimativas negativas são relativamente comuns de se obter (SEARLE, 1987).

Mesmo sob normalidade, a distribuição dos estimadores ANAVA é desconhecida, isto porque estes estimadores não são apenas somas de variáveis independentes com distribuição χ^2 , mas, em geral, incluem diferenças, não seguindo distribuição exata de χ^2 e não apresentando, portanto, intervalos de confiança exatos (ver ROBINSON, 1965).

Há uma relativa facilidade na obtenção de estimadores não-viesados e suas respectivas variâncias e covariâncias amostrais pelo método dos momentos. No entanto SEARLE (1987) relaciona alguns deméritos do(s) método(s), quais sejam:

(a) Mais de um conjunto de formas quadráticas pode ser empregado para a obtenção de $\mathbf{q}$;

(b) Muitas formas quadráticas, no caso de modelos mistos, conterão funções quadráticas de tais efeitos que nada têm a ver com as variâncias que se quer estimar, desta forma a equação $E(\mathbf{q}) = \mathbf{M} \hat{\sigma}^2$ não se aplica;

(c) Não há como distinguir entre as possíveis formas quadráticas que geram estimadores não-viesados;

(d) Podem ocorrer estimativas negativas;

(e) Não se conhecem as propriedades distribucionais dos estimadores, (em geral só as de $\hat{\sigma}^2$ são conhecidas).

2.3.5.2. Métodos de Máxima Verossimilhança (ML)

Segundo SEARLE (1987), o método da Máxima Verossimilhança, idealizado por Fisher, foi primeiro aplicado a estimativas de componentes da variância (em modelos mistos) por HARTLEY e RAO (1967). Assumindo que os erros são aleatórios e normalmente distribuídos, e utilizando a mesma matriz das variâncias e covariâncias entre os efeitos, já apresentada, tem-se:

$$y = Xb + Zu + e \sim N(Xb, V), \text{ com}$$

$$V = \sum_{i=0}^r Z_i Z_i' \sigma_i^2, \text{ para as "r" componentes da variância.}$$

A função de verossimilhança é então:

$$L = (2\pi)^{-N/2} |V|^{-1/2} \exp\left[-\frac{1}{2} (y - X\beta)' V^{-1} (y - X\beta)\right].$$

Maximizando L com respeito aos elementos de β e de V , obtêm-se equações que permitem o cálculo dos estimadores ML dos mesmos. Estas equações podem ser escritas de uma forma variada atendendo a diferenças em numerosos detalhes (ver por ex. SEARLE, 1987). Sua forma geral apresenta dois tipos de equações, quais sejam:

(a) $X' \hat{V}^{-1} X\hat{\beta} = X' \hat{V}^{-1} y$, que para V conhecido transforma-se nas equações normais do BLUE dos efeitos fixos;

(b) E equações para o traço da seguinte matriz:

$$\text{tr}(\hat{V}^{-1} Z_i Z_i') = (y - X\hat{\beta})' \hat{V}^{-1} Z_i Z_i' \hat{V}^{-1} (y - X\hat{\beta}).$$

Nota-se que estas equações são formas análogas das equações do método dos momentos em que o primeiro termo refere-se ao produto da contagem da ocorrência das variáveis aleatórias pela sua variância e o segundo termo é a forma quadrática correspondente.

Estas últimas equações são não lineares nos parâmetros dos estimadores das componentes da variância. Isto indica que, exceto para alguns casos e em dados completamente balanceados, não há soluções analíticas para as componentes da variância, devendo-se recorrer a soluções numéricas (iterativas). Assim, seja o projetor $P = V^{-1} - V^{-1} X(X'V^{-1}X)^{-1} X'V^{-1}$, com $I = V^{-1} \sum_{i=0}^I Z_i Z_i' \sigma_i^2$, a operação de traço da matriz (para variâncias), que passa a ser escrita como:

$$\left\{ \text{tr}(\hat{V}^{-1} Z_i Z_i' \hat{V}^{-1} Z_j Z_j') \right\} \{\hat{\sigma}_i^2\} = \{y' \hat{P} Z_i Z_i' \hat{P} y\}.$$

Assim, pode-se visualizar uma iteração, com valores iniciais de $\hat{\sigma}_i^2$ em V (e portanto em P) para resolver a equação do traço e utilizar a resposta para novos V .

Existem vários problemas numéricos na obtenção destes estimadores, que não serão aqui abordados. É importante frisar que "pacotes" computacionais como as rotinas "Proc VARCOMP" e "Proc MIXED" (SAS) e o "BMDP3V", procedem a estimativas ML.

Algumas das propriedades desta classe de estimadores são fornecidas por SEARLE (1987). Para amostras grandes têm distribuição normal e com variâncias amostrais conhecidas, sendo preferíveis aos estimadores ANAVA para dados desbalanceados. Estimadores ANAVA são especialmente inferiores aos ML no caso de modelos mistos: os estimadores ML fornecem o BLUE dos efeitos fixos a cada passo da iteração, enquanto que não existe um tratamento adequado para o reajuste das estimativas dos efeitos fixos no método dos momentos.

2.3.5.3. Máxima Verossimilhança Restrita (REML)

É uma variação da ML usada primeiro por PATTERSON e THOMPSON (1971) para modelos de blocos. A verossimilhança

maximizada corresponde à porção invariante para os parâmetros fixos, ou seja: $\mathbf{X}\boldsymbol{\beta}$. Diz-se que é feito um ajuste prévio dos dados para as estimativas dos efeitos fixos, seguido do ajuste à função de verossimilhança para as componentes da variância. Usa-se para a estimação portanto apenas o conjunto (b) de equações apresentado no item anterior. As equações resultantes são:

$$\left\{ \text{tr}(\hat{\mathbf{P}} \mathbf{Z}_i \mathbf{Z}_i' \hat{\mathbf{P}} \mathbf{Z}_j \mathbf{Z}_j') \right\} \{ \sigma_i^2 \} = \{ \mathbf{y}' \hat{\mathbf{P}} \mathbf{Z}_i \mathbf{Z}_i' \hat{\mathbf{P}} \mathbf{y} \},$$

ou seja, as mesmas de ML com $\mathbf{P}$ no lugar de $\mathbf{V}^{-1}$.

As propriedades dos estimadores REML são semelhantes às da ML, embora não estimem $\boldsymbol{\beta}$ (SEARLE, 1987). No entanto, uma vez estimadas as variâncias de $\mathbf{V}$, procede-se à estimação de efeitos fixos no modelo generalizado usual.

Sob normalidade e para dados balanceados, estimadores REML são idênticos a estimadores ANAVA, podendo ser obtidos de forma analítica. Tanto ML como REML produzem estimadores viesados para dados muito desbalanceados mas as propriedades destes estimadores em amostras grandes sob normalidade são as mais desejáveis.

2.3.5.4. Outros métodos de estimação das componentes da variância

O método MINQUE (Minimum Norm Quadratic Unbiased Estimation), foi desenvolvido independentemente por LA MOTTE (1970, 1971, 1973) e RAO (1970, 1971a,b, 1972) (ambos citados por SEARLE, 1987). Os autores descrevem uma classe de estimadores que são funções não-viesadas das observações e de norma euclidiana mínima. O importante é notar que a solução das equações de estimação apresenta forma semelhante às de REML, com $\mathbf{P}_w$ no lugar de $\mathbf{P}$, em que w são pesos pré conhecidos para as relações entre as componentes da variância.

O método não envolve pressuposições sobre normalidade, porém produz uma gama infinita de soluções, dependendo da matriz (numérica) P_w escolhida.

O método MIVQUE (Minimum Variance Quadratic Unbiased Estimation) consiste na obtenção de solução MINQUE de variância mínima, propriedade que advém de se invocar a pressuposição de normalidade para os MINQUE.

Considerando a expressão para os MINQUE, pode-se visualizar uma iteração envolvendo as soluções MINQUE para a construção de nova matriz P_w . Este procedimento faz com que I-MINQUE (Iterative-MINQUE) e REML se confundam num só procedimento de estimação, pois qualquer solução MINQUE pode ser um primeiro passo da iteração REML (o que ocorre como "default" das rotinas "Proc VARCOMP" e "Proc MIXED" do SAS).

BROWN (1976) mostra que, mesmo sem a pressuposição da normalidade, estimadores I-MINQUE (e REML) possuem distribuição normal em amostras grandes o que faz com que estes sejam os estimadores mais promissores, reunindo o maior conjunto de propriedades desejáveis para estimativas das componentes da variância.

VALERIO FILHO (1991) mostrou através de simulação, que estimativas REML são superiores às da ANAVA, MIVQUE0 e ML, especialmente para dados desbalanceados.

SMITH e GRASER (1986) propuseram a maximização por métodos numéricos da função de verossimilhança restrita, a partir da busca direta ("*grid search*") de valores do domínio dos parâmetros que maximizem L . Este método torna desnecessária a obtenção das expressões para os estimadores (derivadas parciais do processo de maximização analítica de L) e passou a ser conhecido como "Derivative-Free-REML" (DF-REML).

Tais estimadores são REML e apresentam todas as suas propriedades desejáveis. Em sua aplicação deve-se evitar fenômenos de convergência para máximos locais, comuns a este

tipo de estimação numérica. Um conjunto de pacotes recentes de estimação DF-REML de componentes da variância (e covariâncias) em problemas uni e multivariados foi então reunido para fins do melhoramento animal no recente pacote MTDFREML (BOLDMAN et alii, USDA, 1995).

VERNEQUE (1994) apresenta as características de três destes procedimentos de estimação de componentes da variância e covariâncias do tipo não derivativos, sendo necessária ainda uma maior experiência com este tipo de estimação nas estimativas das componentes da variância para o melhoramento florestal.

2.4. PREDIÇÃO DE VALORES GENÉTICOS

O objetivo geral dos testes de progênie é a predição do potencial de um indivíduo de transmitir bons genes à sua descendência (FALCONER, 1981; ALLARD, 1971); avalia-se progênies com vistas à capacidade geral de combinação dos progenitores, ou ao chamado "breeding value" (valor genético aditivo). Certamente esta característica depende da composição genética da população a qual o indivíduo pertence e com a qual será cruzado.

Para KEMPTHORNE (1988) um tal modelo, levando em conta apenas a ação gênica aditiva, é geneticamente muito pobre, constituindo, no entanto, uma primeira aproximação de desenvolvimentos muito importantes no melhoramento animal. Modelos mais complexos como o assim chamado "modelo animal", modelos aditivo-dominantes, ou ainda modelos que levam em conta interações epistáticas são possíveis sob a mesma abordagem BLUP que gerou as predições do chamado "modelo de touros" (para o valor aditivo do pai num teste de progênie). O problema central é acessar a conveniente matriz de parentescos (os coeficientes " ρ " de Malécot) entre os indivíduos dois a dois, de forma a obter a adequada relação de covariâncias entre todos os indivíduos. Problemas adicionais ligam-se ao grande número de parâmetros (tanto

fixos como aleatórios) que se precisa estimar (ou predizer) e à determinação das estimativas das componentes da variância (elas mesmas ou relações entre elas) a serem empregadas.

Já se apresentou solução teórica para modelos de maior complexidade genética como em HENDERSON (1988) e KENNEDY e SORENSEN, (1988). Modelos complexos passaram a ser aplicados a outras e variadas formas de estimação-predição como a predição de híbridos simples não realizados em delineamentos dialélicos parciais num modelo aditivo-dominante, com o parentesco entre as linhagens definido probabilisticamente através de esquemas de análise da genealogia como o de parentesco de Malécot (FALCONER, 1981), ou de Kempthorne (KEMPTHORNE, 1957) ou por similaridade quanto a marcas de RFLP (BERNARDO, 1994 e 1995).

Um primeiro preditor do valor aditivo de uma progênie para a seleção de matrizes seria a diferença entre o valor da progênie e a média geral dos cruzamentos, mas as diferentes formas de expressão gênica, penetrâncias e expressividades dos diversos caracteres fazem com que esta não seja em geral uma medida confiável do valor aditivo (KEMPTHORNE, 1988).

Uma aproximação inicial do problema da predição é considerar a herdabilidade como coeficiente de predição, de forma a associar um valor genético aditivo a cada progênie a partir de sua média e da média geral. Esta abordagem corresponde a predizer um ganho genético pela seleção individual de cada progênie, com base no produto do diferencial de seleção pela herdabilidade correspondente (HODGE e WHITE, 1992).

Embora a teoria do BLUP tenha sido inteiramente desenvolvida e aplicada com sucesso ao "modelo animal", sua extensão para testes de progênies florestais foi demonstrada e tem sido proposta (e.g. WHITE e HODGE, 1989 e SOH, 1994). A

predição de valores genéticos se daria em modelos mistos onde todos os parâmetros ambientais fossem estimáveis (tidos como fixos).

Seguindo a linha geral de observação destes autores, os testes de progênies florestais apresentam maiores problemas de desbalanceamento dentro de subclasses (parcelas com diferentes números de árvores), ou em delineamentos com desbalanceamento programado ou ainda para grandes comparações entre ensaios (com ou sem progênies ou testemunhas comuns). Nestes casos, BLUP (ou também BLP), têm intervalos de confiança bem mais restritos que para o melhoramento animal e podem ser úteis, tanto para a construção de índices de seleção combinada para a predição do valor genético individual (levando em conta o mérito da família e do indivíduo) quanto para a melhoria das predições de efeitos de progênies e do progresso esperado sob seleção.

Os BLUPs e BLPs têm sido também propostos para a seleção de progênies na prática agrícola, por exemplo, na seleção entre e dentro de progênies de meios-irmãos milho (PACHECO et alii, 1995), progênies autofecundadas em soja (PANTHER e ALLEN, 1995), análise de cruzamentos dialélicos em feijão (PALOMARES e SANJUAN, 1996), além da predição de cruzamentos em dialelos incompletos (BERNARDO, 1994, 1995). Este fato reforça os métodos de modelos mistos propostos para o melhoramento animal como padrão de análise em ensaios genéticos.

Quanto à adequação do modelo ao fenômeno examinado, pode-se relatar os trabalhos de BENIWAŁ et alii (1995a; 1995b), entre outros, que encontraram grande concordância entre as respostas preditas e observadas pela seleção de progênies de irmãos completos ao longo de quarenta gerações em ratos (usando BLUPs iterativos com estimativas REML).

Quanto à robustez teórica e à verificação de propriedades estatísticas, ensaios de simulação realizados desde os anos 1980 (e.g. QUAAS e POLLACK, 1980; POLLACK e QUAAS, 1981a, 1981b) comprovam a superioridade dos BLUP sobre os quadrados mínimos ordinários (no modelo fixo).

ZHU e WEIR (1996) compararam métodos de predição como o BLUP (a partir de componentes da variância conhecidas), BLUP-REML (estimação simultânea das componentes da variância) e um método referido como "Adjusted Unbiased Prediction" (AUP), através de simulação de dados e para ensaios reais com várias gerações em ratos e bicho da seda. Os autores encontraram boa adequação para os BLUP em modelos bastante complexos (por exemplo, para dialelos com efeitos maternos e ligados ao sexo), no entanto os BLUP com estimativas simultâneas das componentes da variância mostraram-se menos eficientes. Os autores propõem o método AUP como o melhor em qualquer caso.

Conclui-se que para fins da aplicação imediata, os BLUP ou os BLP são uma excelente forma de predição, desde que boas estimativas das componentes da variância estejam disponíveis. São ainda preferíveis a métodos como a seleção por médias fenotípicas gerais, derivados do modelo fixo (OLSE), pois estes fazem pressuposições (em geral falsas) sobre as componentes da variância, embora pareçam ser delas independentes.

Os BLUPs aplicados ao chamado "modelo animal" ou outros modelos genéticos possuem propriedades estatísticas desejáveis que podem ser sumariadas a partir de SEARLE (1987), HENDERSON (1988), KENEDY e SORENSEN (1988), WHITE e HODGE (1989), MARTINS (1993) e SILVA (1995), como segue:

- a) Estimação e predição não-viesadas e num procedimento linear único;
- b) Pode levar em conta efeitos de seleção e de endogamia ao longo das gerações, desde que conhecida a genealogia

ou outra estimativa do parentesco (com ou sem gerações intercaladas na composição dos dados);

c) Pode se adequar a modelos genéticos aditivos, aditivo-dominantes, ou que incluam epistases (bem como modelos de clones e de medidas repetidas);

d) Permite estimação simultânea de efeitos maternos e de ambiente, tratados como fixos;

e) Possibilitam a análise multivariada, com a predição de ganhos e demais resultados em termos de agregados genéticos de diversas variáveis ou de uma variável-chave de maior interesse;

f) Permitem a completa unificação dos dados genéticos de uma população (no tempo e no espaço), mesmo os oriundos de ensaios de diferentes tipos, uma vez estimados os efeitos fixos e as propriedades distribucionais que os diferenciam.

Além disso, permitem algumas aplicações de especial importância para o melhoramento de espécies perenes, como:

g) Melhoram a qualidade das próprias estimativas de componentes da variância, bem como permitem que se obtenha estimativas com distribuição de probabilidade conhecidas para a construção de intervalos de confiança exatos (baseados, por exemplo em estimativas REML);

Isto é possível especialmente porque na prática florestal cada ensaio é relativamente balanceado e o efeito de cada local ou bloco pode ser estimado por um número suficientemente grande de árvores (e parcelas).

h) Permitem estabelecer mais facilmente comparações entre estratégias de melhoramento (por exemplo, na seleção a uma mesma base genética por dois métodos diferentes), bem como corrigir possíveis distorções do processo seletivo em gerações pregressas.

Estes procedimentos carecem de abordagens unificadoras, pois na abordagem convencional é praticamente impossível avaliar o valor genético de um pomar de sementes, constituído de clones (ou mudas) de diferentes gerações, selecionados por critérios diferentes ao longo dos anos (seja pela mudança de ambiente de cultivo, pela mudança da importância relativa dos caracteres, pelo surgimento de novos problemas tecnológicos, novas doenças, etc.).

i) BLUPs e BLPs produzem estimativas das variâncias das predições e do erro a elas associado.

Tais estimativas podem ou não trazer embutidas informações sobre os erros nas pressuposições (BLUP ou BLP, respectivamente), mas são sempre mais livres de pressuposições que as do modelo fixo correspondente, em que apenas a variância do erro influi nas estimativas dos parâmetros.

Assim, os modelos mistos de predição permitem estabelecer a confiabilidade da seleção de maneira mais geral e precisa que nas abordagens convencionais, o que se constitui em mais um aspecto de sua superioridade teórica.

3. MATERIAL E MÉTODOS

A notação empregada para as descrições que se seguem baseia-se na notação "ponto-traço" apresentada em FEDERER (1955), na notação habitualmente empregada pela genética quantitativa florestal (e.g. NAMKOONG, 1979), na descrição de modelos mistos para o melhoramento florestal, que indica efeitos fixos com maiúsculas e efeitos aleatórios com minúsculas (WHITE e HODGE, 1989) e na notação matricial apresentada por SEARLE (1987).

3.1. Descrição do material vegetal

O material vegetal empregado constituiu-se de ensaios de progênies obtidas de matrizes selecionadas fenotipicamente de uma população de *Eucalyptus grandis* (Hill) Maiden, estabelecida com sementes da procedência de Coffs' Harbour, (New South Walles, Austrália).

Os ensaios foram avaliados em três locais, quais sejam: Anhembi, Brotas e Lençóis Paulista, no Estado de São Paulo. Para cada local, as progênies e duas testemunhas constituíram os 64 tratamentos que, plantados em delineamentos experimentais do tipo blocos incompletos parcialmente balanceados com três repetições por local (látices triplos 8X8).

Algumas das progênies e testemunhas não foram plantadas em todos os locais. O número de tratamento (do látice) para as progênies também variou conforme o local. No apêndice 1 é

apresentado um quadro com os nomes de progênies e seus números de tratamento por local.

As planilhas originais de resultados das 5676 árvores foram gentilmente cedidos pelo Prof. Mário L. T. Moraes da Universidade Estadual Paulista (UNESP), Campus de Ilha Solteira. Maiores detalhes quanto à origem do material e características dos locais de ensaio, bem como um croqui da área experimental (casualização do látice), podem ser encontrados em MORAES (1987).

As parcelas constituíram-se de vinte e cinco árvores individuais (parcela quadrada 5X5 árvores), sendo avaliadas as nove árvores centrais mais uma, num total de dez árvores, a menos de falhas, caso em que árvores suplementares eram avaliadas à medida da necessidade.

As medidas tomadas à idade de corte (sete anos) consistiam de altura de árvore (**H**), diâmetro à altura do peito (**DAP**, tomado a 1,30m) e densidade básica da madeira (**D**) (conforme metodologia descrita por MORAES, 1987). Para a obtenção do volume cilíndrico utilizou-se a expressão $V = (\pi/4) \times (DAP^2) \times (H)$.

3.2. Modelo para a análise convencional

A seleção de progênies através das médias gerais sobre os diferentes ambientes foi chamada BV (do inglês "breeding value"). Para fins de comparação procedeu-se ao cálculo do valor genético aditivo num contexto de predição (BLP) sobre estas médias. A rotina para a análise convencional deste tipo de ensaio envolve:

- a) analisar os látices por local (análise com recuperação da informação interblocos, dados tomados no nível de médias de parcelas);
- b) utilizar as médias ajustadas de progênies para o cálculo direto da variância de progênies;

c) formar uma tabela de dupla entrada (médias ajustadas de progênies por locais) com as progênies comuns a todos os locais e a partir dela calcular a interação progênies por locais;

d) calcular independentemente a variância dentro de parcelas (média das variâncias amostrais de parcelas ponderada pelos graus de liberdade de parcelas, divididos pela média harmônica do número de árvores por parcela).

Os modelos da análise da variância com as respectivas esperanças de quadrados médios utilizados neste roteiro de análise são os indicados por MORAES (1987). O método consiste em igualar as médias ajustadas do látice e o erro efetivo às médias de tratamentos e ao erro no modelo de blocos ao acaso, obtendo-se assim estimativas de componentes da variância pelo método dos momentos ao se igualar as variâncias determinadas às esperanças dos quadrados médios do modelo aleatório de blocos ao acaso.

Procede-se então ao cálculo de estimativas de herdabilidade e a seguir dos progressos esperados pela seleção conforme o método de seleção considerado e a partir das intensidades de seleção programadas (e dos diferenciais de seleção estandardizados correspondentes).

No presente estudo, os dois casos considerados foram: a) seleção de progênies para um pomar de sementes clonal; e b) seleção sequencial de árvores individuais dentro de progênies selecionadas.

Exemplos desta análise completa, bem como as fórmulas necessárias para se proceder a estes cálculos podem ser encontrados em KAGEYAMA (1980) e MORAES (1987), entre outros.

Tal abordagem corresponde à estimação dos valores genéticos aditivos de progênies a partir de observações de árvores individuais no seguinte modelo:

$$Y_{ijkl} = m + E_i + P_k + (PE)_{ik} + b_j + e_{jk} + d_{ijkl} ,$$

Em que cada observação Y pode ser explicada por m , sua média geral; E_i , o efeito do local i (suposto fixo); P_k , o efeito da progênie k (também fixo); $(PE)_{ik}$, o efeito (fixo) da interação da progênie k no local i ; b_j o efeito do bloco j (nos látices, contido dentro de repetições, suposto aleatório); e finalmente por e_{ijk} e d_{ijkl} , que são os efeitos dos erros experimentais da parcela ijk e da árvore l dentro da parcela ijk , respectivamente, (com " a " locais, " s " progênies, " b " repetições e " n " árvores por parcela).

Para fins das estimativas das componentes da variância o modelo é então considerado aleatório, (o que é, como já relatado na revisão, usual: considerar modelo fixo para a análise e comparações entre médias e em seguida aleatório para estimar as componentes da variância).

Para a realização desta análise foi empregada a rotina "Proc Mixed" do SAS (cuja sintaxe é exemplificada no apêndice 3), com o objetivo de verificar a concordância com a análise apresentada por MORAES (1988) para os mesmos dados. Uma mínima diferença ocorreu nas médias ajustadas e foi atribuída a não se ter adotado o procedimento de estimação de parcelas perdidas para a realização da análise do látice (segundo COCHRAN e COX, 1957), método empregado por MORAES (1988).

A abordagem central para as diversas comparações propostas é a predição do valor genético aditivo, que proporciona uma generalização para as estimativas de progresso convencionais, ou seja, partindo-se da expressão geral do progresso genético apresentada por FALCONER (1981) $G_s = ds.h^2$, o BV do valor genético de uma progênie constitui-se, didaticamente, na predição do progresso esperado na seleção daquela progênie, ou seja, a somatória dos BV de todas as progênies é nula e a soma dos BV das progênies selecionadas é o próprio progresso genético predito pela seleção.

O tamanho efetivo resultante da seleção de árvores através dos diferentes métodos foi calculado pela expressão desenvolvida por VENCovsky (1976):

$$N_e = \frac{4 N_1}{\frac{s_{k_s}^2}{k_s} + \bar{k}_s + 3}$$

em que k_s é o número de árvores selecionado por progênie, $s_{k_s}^2$ é a variância deste número e N_1 é o número total de árvores selecionado.

3.3. Modelo misto de predição

Para a predição de valores genéticos aditivos de matrizes seguiu-se um modelo estatístico usual (efeitos aleatórios de progênies e efeitos abientais fixos). Uma análise coerente deste tipo de modelo (na metodologia de modelos mistos) em testes de progênies florestais é apresentada por WHITE e HODGE (1989), qual seja:

$$Y_{ijkl} = m + E_i + B_j + p_k + (pE)_{ik} + e_{jk} + d_{ijkl},$$

em que m é a média geral, E_i é o efeito do ambiente i (efeito fixo), B_j é o efeito do bloco j dentro de cada local (efeito fixo), p_k é o efeito da progênie k (aleatório), $(pE)_{ik}$ é o efeito da interação da progênie k no local i (aleatório), e_{jk} e d_{ijkl} são os efeitos dos erros experimentais (aleatórios) da parcela ijk e da árvore l dentro da parcela $ijkl$, respectivamente, (com "a" ambientes, "s" progênies, "b" repetições, $b \neq J$, e "n" árvores por parcela).

Para este modelo as covariâncias entre pares de fatores foram consideradas nulas e os efeitos aleatórios considerados normalmente distribuídos. Os primeiros e segundos momentos dos efeitos fixos e aleatórios são descritas como se segue:

$$E(E_i) = E_i ; E(E_i^2) = E_i^2 ; E(B_{ij}) = B_{ij} ; E(B_{ij}^2) = B_{ij}^2 ;$$

$$E(p_k) = 0 ; E(p_k^2) = \sigma_p^2 ; E(pE_{ik}) = 0 ; E((pE_{ik})^2) = \sigma_{pE}^2 ;$$

$$E(e_{jk}) = 0 ; E(e_{jk}^2) = \sigma_e^2 ; E(d_{ijkl}) = 0 ; E(d_{ijkl}^2) = \sigma_d^2 .$$

O efeito "p" equivale à metade do valor genético aditivo da progênie (assumida como de meios-irmãos), ou seja, o valor

aditivo de progênie é dado por $g_i = 2p_i$ e a variância genética aditiva é dada por $V(g_i) = 4V(p_i) = \sigma_s^2 = \sigma_A^2 = 4\sigma_p^2$.

Tais especificações permitem a construção dos diferenciais de seleção teóricos e dos pesos que sobre eles incidem na abordagem BLP.

3.4. BLP

A predição de valores genéticos através dos BLP envolve a determinação de vetores de pesos que se aplicam ao nível de observação considerado de maneira análoga aos índices. O modelo da predição pode ser apresentado para cada progênie como: $BLP(g) = \hat{g}_i = a_i + b_i' y$, com **a** e **b'** tais que $E(g_i - \hat{g}_i)^2$ seja mínimo (derivação pelo método dos mínimos quadrados) ou ainda, que a covariância entre os valores genéticos aditivos e seus preditores $Cov(g_i, \hat{g}_i)$ seja máxima (derivação pelo método da máxima verossimilhança). Estas derivações para o BLP podem ser encontradas em HODGE e WHITE (1992) e WHITE e HODGE (1989), respectivamente.

A equação habitual da predição BLP como apresentada por HENDERSON (1963) é a que segue:

$$BLP(g) = \hat{g}_i = \tau + C'V^{-1}(y - \alpha), \text{ em que:}$$

$\tau = E(g) = 0$, é o vetor de valores genéticos esperados dos valores genéticos a serem preditos (média do valor genético populacional);

C' = matriz de covariâncias entre o valor genético a ser predito e as observações;

V^{-1} = inversa da matriz de variâncias e covariâncias das observações;

y = vetor de observações;

$\alpha = E(y)$ é o vetor de valores esperados das observações, suposto conhecido, na prática, estimado pelo BLUE ($m + E_i + B_j$), no contexto de quadrados mínimos ordinários ("OLSE", ver item 3.4.3.).

O "peso relativo" de cada observação é $b' = C'V^{-1}$, explicitado em termos do modelo como segue:

$C' = \text{Cov}(\text{valor genético a ser predito, observação considerada});$

$V = \text{Cov}(\text{matriz de variâncias e covariâncias das observações consideradas}).$

Para qualquer destes modelos considerou-se apenas as análises univariadas (embora todos permitam generalização para análise multivariada) com o objetivo de concentrar a discussão na comparação dos métodos de seleção.

A estatística que confere uma mensuração da probabilidade de melhor seleção (deduzida por maximização) é a correlação do valor genético com o valor predito [que é Máxima tanto para BLP quanto para BLUP, ver WHITE e HODGE (1989)]. Tal correlação é chamada acurácia seletiva, sendo apresentada por estes autores como:

$$r[\hat{g}, g] = \sqrt{\frac{V(\hat{g})}{\sigma_A^2}}, \text{ com } V(\hat{g}) = C'V^{-1}C,$$

na suposição de σ_A^2 conhecida (ver estimativas de componentes da variância).

3.4.1. Comparações entre métodos através dos BLP

A partir da forma explícita dos BLP e com o auxílio da acurácia acima descrita, compararam-se teoricamente as possíveis vantagens da utilização dos BLP tomando médias de progênies por local ou médias de parcelas ao invés da média marginal para a seleção. Para as comparações práticas a média dos BLP das progênies selecionadas foi sempre a medida do progresso por seleção (WHITE e HODGE, 1989).

3.4.2. Caso A: Seleção com base em médias marginais (gerais) de progênies:

Uma adaptação deste caso para os BLP pode ser dada por:

$$C' = \text{Cov}(g_i, \bar{Y}_{..k}) = \sigma_s^2/2 = 2\sigma_p^2; \text{ e}$$

$$V = \text{Var}(\bar{Y}_{..k}) = \sigma_p^2 + \sigma_{pe}^2/a + \sigma_{ab}^2 + \sigma_{abn}^2.$$

Para este caso, notar que $C'V^{-1} = h_s^2$, ou seja, a própria herdabilidade para a seleção entre médias marginais de progênes, como apresentada por MORAES (1987).

3.4.3. Caso B: Seleção com base em médias de progênes por local:

Neste caso, $C' = \text{Cov}(g_i, \bar{Y}_{i.k}) = \sigma_s^2/2 = 2\sigma_p^2$; e V é uma matriz simétrica cujas diagonais valem $\text{Var}(\bar{Y}_{i.k}) = \sigma_p^2 + \sigma_{pe}^2 + \sigma_b^2 + \sigma_{bn}^2$ e cujos elementos fora da diagonal valem $\text{Cov}(\bar{Y}_{i.k}, \bar{Y}_{i'.k})$ (covariância entre médias da mesma progênie em diferentes ambientes $= \sigma_p^2$). Este caso será a base dos desenvolvimentos e da discussão teórica.

3.4.4. Caso C: Seleção com base em médias de parcelas:

Neste caso, para a seleção com vistas a um ambiente intermediário entre os diferentes ambientes, $C = \text{Cov}(g_i, \bar{Y}_{ijk}) = \sigma_s^2/2 = 2\sigma_p^2$, e para a seleção com vistas a um ambiente alvo específico "E_i", $\text{Cov}(g_i, \bar{Y}_{ijk}) = 2(\sigma_p^2 + \sigma_{pei}^2)$. Numa representação matricial do conjunto dos possíveis vetores C para os diferentes locais-alvo considerados, teríamos:

Local-alvo da seleção:	Interme-diário	Anhembi	Brotas	Lençóis Paulista
$C =$	$\begin{bmatrix} 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) & 2\sigma_p^2 & 2\sigma_p^2 \\ 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) & 2\sigma_p^2 & 2\sigma_p^2 \\ 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) & 2\sigma_p^2 & 2\sigma_p^2 \\ 2\sigma_p^2 & 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) & 2\sigma_p^2 \\ 2\sigma_p^2 & 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) & 2\sigma_p^2 \\ 2\sigma_p^2 & 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) & 2\sigma_p^2 \\ 2\sigma_p^2 & 2\sigma_p^2 & 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) \\ 2\sigma_p^2 & 2\sigma_p^2 & 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) \\ 2\sigma_p^2 & 2\sigma_p^2 & 2\sigma_p^2 & 2(\sigma_p^2 + \sigma_{pei}^2) \end{bmatrix}$			

A matriz V é a matriz bloco-diagonal, simétrica, cujos blocos não nulos são dados, para cada progênie, pelas seguintes relações entre as componentes de variância (supostas conhecidas) fornecidas por MORAES (1987):

$$\begin{bmatrix} V(\bar{Y}_{ijk}) & \sigma_p^2 + \sigma_{pe1}^2 & \sigma_p^2 + \sigma_{pe1}^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 + \sigma_{pe1}^2 & V(\bar{Y}_{ijk}) & \sigma_p^2 + \sigma_{pe1}^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 + \sigma_{pe1}^2 & \sigma_p^2 + \sigma_{pe1}^2 & V(\bar{Y}_{ijk}) & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & V(\bar{Y}_{jk}) & \sigma_p^2 + \sigma_{pe2}^2 & \sigma_p^2 + \sigma_{pe2}^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 + \sigma_{pe2}^2 & V(\bar{Y}_{jk}) & \sigma_p^2 + \sigma_{pe2}^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 + \sigma_{pe2}^2 & \sigma_p^2 + \sigma_{pe2}^2 & V(\bar{Y}_{jk}) & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & V(\bar{Y}_{jk}) & \sigma_p^2 + \sigma_{pe3}^2 & \sigma_p^2 + \sigma_{pe3}^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 + \sigma_{pe3}^2 & V(\bar{Y}_{jk}) & \sigma_p^2 + \sigma_{pe3}^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 + \sigma_{pe3}^2 & \sigma_p^2 + \sigma_{pe3}^2 & V(\bar{Y}_{jk}) \end{bmatrix}$$

Note-se que nesta matriz os elementos da diagonal valem $V(\bar{Y}_{ijk}) = \sigma_p^2 + \sigma_{pe1}^2 + \sigma_p^2 + \sigma_{pe2}^2 / n$, ou seja, leva-se em conta as diferenças nas variâncias fenotípicas por local. A validade do procedimento foi analisada na discussão teórica do item anterior (Caso B). Este modelo não é válido para a análise BLUP, em que as variâncias de locais foram supostas não homogêneas (não se pôde verificar sua influência na determinação dos "pesos" de local).

A esperanças dos efeitos fixos foram estimadas num modelo fixo reduzido, qual seja: $\bar{Y}_{ijk} = m + E_i + B_{j(i)} + \varepsilon_{ijk}$, em que ε_{ijk} representa o efeitos médios (confundidos) da progênie k , e do erro (entre e dentro) da parcela ijk , sendo os demais efeitos os mesmos do modelo misto completo acima apresentado. Tal modelo foi analisado pela rotina GLM do SAS (SAS, 1992) de forma a obter as estimativas aproximadas dos BLUE($m + E_i + B_{j(i)}$), conforme metodologia indicada por WHITE e HODGE (1989). Os comandos do SAS empregados na análise podem ser vistos no apêndice 3.

3.5. Índice de seleção

O cálculo da predição através do índice pode ser descrito pelo ajuste de um índice do tipo:

$$\hat{g} = b [y - E(y)],$$

em que b é o vetor de "pesos" e $E(y)$ é a esperança matemática do vetor de observações (y) .

Assim, para o modelo de seleção considerado, ter-se-ia:

$$\begin{aligned} \hat{g}_{ijkl} = I_{ijkl} = & b_1 [Y_{ijkl} - E(Y_{ijkl})] + b_2 [Y_{ijk.} - E(Y_{ijk.})] + b_3 [Y_{i.k.} - E(Y_{i.k.})] \\ & + b_4 [Y_{..k.} - E(Y_{..k.})], \end{aligned}$$

em que b_1 é o peso para a árvore individual, b_2 é o peso para a parcela, b_3 é o peso para a média da progênie no local e b_4 é o peso para a média marginal da progênie.

A obtenção de tais pesos na abordagem BLP resume-se então a obter o vetor $b' = C'V^{-1}$, para o qual C constitui um vetor (por progênie) de covariâncias entre o valor genético individual a ser predito e cada nível de observação considerado para o índice e V a matriz das respectivas variâncias e covariâncias das observações, ou seja:

$$C = \begin{bmatrix} \text{Cov}(g_{ijkl}, Y_{ijkl}) \\ \text{Cov}(g_{ijkl}, \bar{Y}_{ijk.}) \\ \text{Cov}(g_{ijkl}, \bar{Y}_{i.k.}) \\ \text{Cov}(g_{ijkl}, \bar{Y}_{..k.}) \end{bmatrix} = \begin{bmatrix} \left(\frac{\sigma_A^2}{4}\right) \left(1 + \frac{3}{n}\right) \\ \left(\frac{\sigma_A^2}{4}\right) \left(1 + \frac{3}{bn}\right) \\ \left(\frac{\sigma_A^2}{4}\right) \left(1 + \frac{3}{abn}\right) \end{bmatrix} \quad e$$

$$V = \begin{bmatrix} \text{Var}(Y_{ijkl}, Y_{ijkl}) & \text{Var}(Y_{ijkl}, \bar{Y}_{ijk.}) & \text{Var}(Y_{ijkl}, \bar{Y}_{i.k.}) & \text{Var}(Y_{ijkl}, \bar{Y}_{..k.}) \\ \text{Var}(Y_{ijkl}, \bar{Y}_{ijk.}) & \text{Var}(Y_{ijk.}, \bar{Y}_{ijk.}) & \text{Var}(Y_{ijk.}, \bar{Y}_{i.k.}) & \text{Var}(Y_{ijk.}, \bar{Y}_{..k.}) \\ \text{Var}(Y_{ijkl}, \bar{Y}_{i.k.}) & \text{Var}(Y_{ijk.}, \bar{Y}_{i.k.}) & \text{Var}(Y_{i.k.}, \bar{Y}_{i.k.}) & \text{Var}(Y_{i.k.}, \bar{Y}_{..k.}) \\ \text{Var}(Y_{ijkl}, \bar{Y}_{..k.}) & \text{Var}(Y_{ijk.}, \bar{Y}_{..k.}) & \text{Var}(Y_{i.k.}, \bar{Y}_{..k.}) & \text{Var}(Y_{..k.}, \bar{Y}_{..k.}) \end{bmatrix}.$$

3.5.1. Comparações entre métodos pelo Índice

O índice obtido foi distribuído às árvores individuais através do programa SAS ("data step"). Procedeu-se à seleção, verificação de coincidências e comparações quanto a progressos esperados pela seleção e tamanhos efetivos resultantes dos procedimentos da seleção pelo índice e pela seleção sequencial entre e dentro de progênies. Os comandos do SAS empregados na análise podem ser vistos no apêndice 3.

Para as comparações práticas a média do índice para as árvores selecionadas foi sempre a medida do progresso por seleção (WHITE e HODGE, 1989).

3.6. Estimação de BLUEs e Predição de BLUPs

O modelo misto para a predição dos valores genéticos aditivos das matrizes no teste de progênie pode ser representado matricialmente por $y = Xb + Zu + e$, em que Xb é a parte fixa e Zu a parte referente aos efeitos aleatórios. A equação do BLUP das variáveis aleatórias (efeitos de progênies) é então:

$$\tilde{u} = E(u) + \text{cov}(u, y') [\text{var}(y)]^{-1} [y - E(y)] \quad (\text{SEARLE, 1987}),$$

$$\text{ou ainda, } \text{BLUP}(u) = \tilde{u} = GZV^{-1}(y - X\hat{\beta}).$$

A estimação da parte fixa é dada pela equação:

$$\text{BLUE}(X\beta) = X\hat{\beta} = X(X'V^{-1}X)^{-1}X'V^{-1}y,$$

que, com V conhecida, é uma aplicação direta de um modelo linear generalizado.

Para a seleção de progênies através do BLUP foram utilizadas as chamadas "equações de modelos mistos" (MME), equivalentes às equações normais para modelos fixos particionados, com uma devida correção para os efeitos aleatórios, qual seja, a soma da relação entre as componentes da variância do erro e dos demais efeitos aleatórios (matriz G) na diagonal correspondente a estes efeitos:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \sigma_e^2 G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}, \text{ SEARLE (1987).}$$

Procedeu-se à predição através da rotina computacional SAS-MIXED, que trabalha desta forma a inversão de matrizes de ordem $p+q$ (número de efeitos fixos e aleatórios, respectivamente).

Na rotina SAS-Mixed, $V = ZGZ' + R$, e as equações de estimação e predição são as que se seguem:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix},$$

esta igualdade foi primeiro sugerida por Henderson (1950, citado por HENDERSON, 1975). Segundo SEARLE (1987), foi primeiro demonstrada em HENDERSON et alii (1959), para a completa generalidade das matrizes R e G . A aparente diferença entre os dois grupos de equações é devido às matrizes (G e R) apresentadas por Searle (1987) incorporarem os valores das variâncias e as do SAS referirem-se apenas aos coeficientes das relações entre variâncias. No caso de $R=I\sigma_e^2$, para o contexto de OLSE e sem levar em conta diferenças ambientais na σ_e^2 , as equações se equivalem.

Como as componentes da variância (e portanto G e V) não são conhecidas, a rotina utiliza a cada passo de iteração suas estimativas REML como componentes conhecidas, de forma a obter $BLUE_V$ e $BLUP_V$ (SEARLE, 1987; SAS, 1993; WOLFINGER, 1993). Os comandos SAS da análise podem ser vistas no apêndice 3.

3.7. Componentes de variância e BLUP iterativo

Para o cálculo dos BLP e dos Índices de seleção foram empregadas as estimativas de componentes da variância apresentadas por MORAES (1987) produzidas pelo método dos momentos, na suposição de um delineamento de blocos ao acaso,

em ensaios de progênies no modelo aleatório em três locais (representados na análise por efeitos fixos). Tais estimativas entraram na composição das matrizes **V** e **C** do BLP e dos Índices, supostas como variâncias conhecidas (sem erro) e estão sumariadas no apêndice 2.

Para o BLUP iterativo foi empregada a "Proc-MIXED" (SAS), com a opção padrão de produção de estimativas REML (derivativas), com iterações sucessivas a partir de uma estimativa inicial MIVQUE.

A cada estimativa das componentes da variância está associado um erro padrão da estimativa, calculado na suposição de normalidade. Tais estimativas foram comparadas às de MORAES (1987).

Os comandos do SAS empregados na análise podem ser vistos no apêndice 3.

4. RESULTADOS E DISCUSSÃO

4.1. Comparações Teóricas

Um primeiro resultado importante na comparação entre os métodos é a forma explícita do BV e do BLP para comparar a seleção de progênies através da média marginal com o BLP (a partir de médias de progênies por local). Analisar-se-á o caso de variâncias fenotípicas homogêneas ao longo dos locais e o caso de diferenças nas variâncias fenotípicas.

A ocorrência de possíveis diferenças é atribuída inicialmente aos pesos de cada ambiente (ou nível do efeito fixo) e em seguida, às diferenças nos diferenciais de seleção $[y-E(y)]$, oriundas de se supor conhecida $E(y)$ nos BLP.

4.1.1. Comparações entre os pesos

BV: $C' = \sigma_a^2/2 = 2\sigma_p^2$ e $V = \text{Var}(\bar{y}_{..k}) = \sigma_p^2 + \sigma_{pe}^2/I + \sigma_e^2/Ib + \sigma_a^2/Ibn$, portanto, os vetores de pesos tornam-se: $b' = C'V^{-1} = \frac{2\sigma_p^2}{\sigma_{\bar{y}..k}^2} = h_{\bar{y}..k}^2$, o que representa a herdabilidade como um peso uniforme para todos os resíduos, aplicado ao diferencial de seleção marginal.

A correlação $r[\hat{g}, g] = \sqrt{\frac{V(\hat{g})}{\sigma_a^2}}$, com $V(\hat{g}) = C'V^{-1}C = \frac{(2\sigma_p^2)^2}{\sigma_{\bar{y}..k}^2}$, torna-se então: $r[\hat{g}, g] = \sqrt{\frac{\sigma_p^2}{\sigma_{\bar{y}..k}^2}}$, resultados já bastante demonstrados, consistentes com HODGE e WHITE, (1992).

BLP: C' é um vetor de covariâncias uniforme para a seleção com vistas a um ambiente médio entre os ambientes estudados, sendo cada um de seus elementos dado pela covariância

$\text{Cov}(g_i, \bar{y}_{i.k.}) = \sigma_p^2 / 2 = 2\sigma_p^2$; neste caso (um ambiente intermediário), $\mathbf{V}$ é uma matriz de variâncias e covariâncias cuja forma depende do número de ambientes considerado. Sua diagonal é dada por $\text{Var}(\bar{y}_{i.k.}) = \sigma_p^2 + \sigma_{pe}^2 + \sigma_b^2 / b + \sigma_{bn}^2 / bn$ e fora da diagonal ela vale $\text{Cov}(\bar{y}_{i.k.}, \bar{y}_{i'.k.}) = \sigma_p^2$, então $\mathbf{C}'\mathbf{V}^{-1}$ tem sua forma dependente do número de locais considerado.

Foi possível demonstrar (para até cinco ambientes) ou verificar (para cinco ou mais ambientes) que, para um número qualquer de ambientes, a expressão para $\mathbf{C}'\mathbf{V}^{-1}$ assume a forma do vetor:

$$\mathbf{C}'\mathbf{V}^{-1} = \frac{2\sigma_p^2}{|\mathbf{V}_a|} \begin{bmatrix} \prod_{i \neq 1} (\sigma_{\bar{y}_{i.k.}}^2 - \sigma_p^2) & \prod_{i \neq 1} (\sigma_{\bar{y}_{i'.k.}}^2 - \sigma_p^2) & \dots \end{bmatrix},$$

em que $|\mathbf{V}_a|$ é o determinante da matriz $\mathbf{V}$ para a ambientes. Uma tal expressão pode ser melhor entendida se particularizada para os diferentes números de ambientes considerados, ou seja:

A) Dois ambientes (forma geral):

$$\mathbf{C}'\mathbf{V}^{-1} = \frac{2\sigma_p^2}{(\sigma_{\bar{y}_{1.k.}}^2 - \sigma_p^2)^2} \begin{bmatrix} (\sigma_{\bar{y}_{2.k.}}^2 - \sigma_p^2) & (\sigma_{\bar{y}_{1'.k.}}^2 - \sigma_p^2) \end{bmatrix}, \text{ ou seja,}$$

o peso do ambiente 1 assemelha-se desta forma a uma herdabilidade, apenas que multiplicada por um fator de correção que depende da variância fenotípica (livre da variância de progênies) do ambiente 2, e vice-versa. Heuristicamente, um maior peso para o ambiente de menor variância fenotípica.

Supondo-se agora que $\sigma_{\bar{y}_{2.k.}}^2 = \sigma_{\bar{y}_{1'.k.}}^2$ (variâncias homogêneas ao longo dos ambientes), a expressão para os pesos torna-se:

$$\mathbf{C}'\mathbf{V}^{-1} = \frac{2\sigma_p^2}{\sigma_{\bar{y}_{1.k.}}^2 + \sigma_p^2} \begin{bmatrix} 1 & 1 \end{bmatrix} = \frac{\sigma_p^2}{\sigma_{\bar{y}_{..k.}}^2} \begin{bmatrix} 1 & 1 \end{bmatrix},$$

ou seja, a mesma herdabilidade aplicada a cada porção do resíduo aleatório em cada ambiente, o que é idêntico à seleção com base nas médias marginais [a menos de alterações no diferencial de seleção, supondo conhecidas $E(y)$].

Quanto à variância das predições, no caso geral, admitindo a presença de interação do tipo progênes por ambientes, tem-se:

$$V(\hat{g}) = C'V^{-1}C = \frac{(2\sigma_p^2)^2}{\sigma_{Y2.k}^2 \sigma_{Y1.k}^2 - (\sigma_p^2)^2} [\sigma_{Y2.k}^2 + \sigma_{Y1.k}^2 - 2\sigma_p^2],$$

expressão que simplifica-se para

$$V(\hat{g}) = \frac{(2\sigma_p^2)^2}{\sigma_{Y..k}^2},$$

(a mesma que para a seleção baseada nas médias fenotípicas gerais). Isto ocorre tanto no caso de ausência de interação quanto no caso desta ser estimada através das médias de progênes nos dois locais (o que é usual), produzindo a estimativa:

$$\hat{\sigma}_{Y..k}^2 = \frac{(\hat{\sigma}_{Y1.k}^2 + \hat{\sigma}_{Y2.k}^2)}{2},$$

que na prática fornece o mesmo resultado que $\sigma_{Y2k}^2 = \sigma_{Y1k}^2$, pois não há como estimar a heterogeneidade das variâncias do erro nos diferentes ambientes.

B) Três ambientes (forma geral):

$$C'V^{-1} = \frac{2\sigma_p^2}{|V_3|} [(\sigma_{Y2.k}^2 - \sigma_p^2)(\sigma_{Y3.k}^2 - \sigma_p^2) \quad (\sigma_{Y3.k}^2 - \sigma_p^2)(\sigma_{Y1.k}^2 - \sigma_p^2) \quad (\sigma_{Y2.k}^2 - \sigma_p^2)(\sigma_{Y1.k}^2 - \sigma_p^2)],$$

$$\text{em que } |V_3| = \sigma_{Y1.k}^2 \sigma_{Y2.k}^2 \sigma_{Y3.k}^2 + (\sigma_p^2)^2 (\sigma_p^2 - \sigma_{Y1.k}^2 - \sigma_{Y2.k}^2 - \sigma_{Y3.k}^2).$$

Desta forma, a variância das predições é dada por:

$$V(\hat{g}) = \frac{(2\sigma_p^2)^2 [\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2]}{\sigma_p^2 [\sigma_1^2 \sigma_2^2 + \sigma_1^2 \sigma_3^2 + \sigma_2^2 \sigma_3^2] + \sigma_1^2 \sigma_2^2 \sigma_3^2},$$

em que $\sigma_i^2 = \sigma_{Yi.k}^2 - \sigma_p^2$, ou seja, a variância fenotípica livre da variância de progênes para cada ambiente.

Para determinar quando (e quanto) o BLP será mais preciso que BV, pode-se comparar diretamente as expressões das variâncias das predições, pois estas formam o numerador

da acurácia, sendo que o seu denominador é comum (variância genética, suposta sempre como conhecida).

Um quociente entre as duas variâncias de predições assume, para três locais, a seguinte forma:

$$Q = \frac{V(\hat{g}_{BLP})}{V(\hat{g}_{BV})} = \frac{(\sigma_p^2 + \sigma_i^2/3)[\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2]}{\sigma_p^2[\sigma_1^2\sigma_2^2 + \sigma_1^2\sigma_3^2 + \sigma_2^2\sigma_3^2] + \sigma_i^2\sigma_1^2\sigma_2^2\sigma_3^2},$$

em que $\sigma_i^2 = \frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}{3}$, ou seja, a média das variâncias livres de σ_p^2 para os três ambientes.

Pode-se mostrar que este valor de Q é sempre maior ou igual a um, indicando ser o BV sempre um caso particular do BLP e também menos preciso. Isto também indica que a seleção com base em médias gerais é mais sujeita a erros de seleção do que o BLP calculado a partir de médias de progênes por ambientes.

No caso particular em que as variâncias são homogêneas, tem-se que:

$$Q = \frac{V(\hat{g}_{BLP})}{V(\hat{g}_{BV})} = \frac{(\sigma_p^2 + \sigma_i^2/3)[3(\sigma_i^2)^2]}{\sigma_p^2[3(\sigma_i^2)^2] + (\sigma_i^2)^3} = 1,$$

com σ_i^2 sendo esta variância homogênea. Esta expressão indica que, neste caso, o BLP proporciona variância das predições iguais ao BV. Nestas circunstâncias, os pesos por ambientes são uniformes e qualquer diferença entre BLP e o ordenamento por médias marginais convencional deve ser atribuída aos primeiros momentos, que se alteram conforme a estimativa de E(y) empregada.

Tome-se dois exemplos de relações entre componentes da variância possíveis de serem encontradas na prática dos testes de progênes florestais:

B1) Supondo $\sigma_p^2=1$; $\sigma_1^2=1,5$; $\sigma_2^2=2,0$; $\sigma_3^2=1,3$; tem-se:

$$Q = \frac{(1+1,6 \div 3)[7,55]}{[7,55] + 3,9} = 1,01,$$

com o vetor $C' = [2,0 \quad 2,0 \quad 2,0]$ e os elementos da matriz V : diagonal= $[2,5 \quad 3,0 \quad 2,3]$ e fora da diagonal= $[1,0]$;

B2) Supondo $\sigma_p^2=0,5$; $\sigma_1^2=6,1$; $\sigma_2^2=3,7$; $\sigma_3^2=1,4$; tem-se:

$$Q = \frac{(0,5 + 3,73 \div 3)[35,25]}{(0,5)[35,25] + 31,598} = 1,25,$$

com o vetor $C' = [2,0 \quad 2,0 \quad 2,0]$ e os elementos da matriz V : diagonal= $[2,5 \quad 3,0 \quad 2,3]$ e fora da diagonal= $[1,0]$;

A tabela 1 sumaria os resultados acima em termos de valores para a relação Q , as variâncias $V(\hat{g})$ e a acurácia seletiva $r(\hat{g}, g)$. Estes resultados indicam que a superioridade do valor explicativo dos BLP sobressai-se com a diminuição da herdabilidade, ou seja, reforçam a noção corrente de que para uma alta herdabilidade a importância do refinamento do método diminui e mesmo médias menos precisas de valores fenotípicos passam a ser bons indicadores do valor genético aditivo.

Tabela 1: Valores das variâncias das predições $[V(\hat{g})]$, acurácia seletiva $[r(\hat{g}, g)]$ e quocientes teóricos entre as variâncias das predições (Q) para dois casos hipotéticos de seleção a três ambientes pela média marginal (BV) e pelo BLP a partir de médias por ambiente:

CASO	Q	$V(\hat{g})$		$r(\hat{g}, g)$	
		BV	BLP	BV	BLP
B1: $\sigma_p^2=1,0$ $\sigma_1^2=1,5; \sigma_2^2=2,0; \sigma_3^2=1,3$	1,01	2,61	2,64	0,808	0,812
B2: $\sigma_p^2=0,5$ $\sigma_1^2=6,1; \sigma_2^2=3,7; \sigma_3^2=1,4$	1,25	0,58	0,73	0,53	0,60

É interessante ressaltar que o segundo exemplo (B2) traz relações entre as componentes da variância compatíveis com resultados comumente encontrados para caracteres de crescimento no melhoramento de eucaliptos no Brasil. E além disso, são relações que em geral permitiriam a análise

conjunta sem que se detectasse sua heterocedasticidade pelo teste de $F_{\text{máximo}}$, como indicado por GOMES (1987).

A relação proposta é, desta forma, um indicador prático da superioridade do BLP sobre o BV para casos em que já se possua estimativas das componentes da variância (por exemplo, para avaliar a necessidade de se reavaliar a seleção feita pelas médias fenotípicas).

Como a abordagem BLP apresentada implica uma decomposição da interação progênes por ambientes em suas componentes individuais por ambiente, espera-se que os casos mais críticos em que se devem registrar ocorrências de alterações de ordem de progênes selecionadas são os que apresentam a magnitude das variâncias fenotípicas correlacionadas às médias ambientais.

Pode-se esperar, portanto, uma considerável superioridade, mesmo para esta primeira abordagem, dos BLP sobre a seleção fenotípica baseada em médias gerais de progênes.

C) Quatro ambientes (forma geral):

$$C'V^{-1} = \frac{2\sigma_p^2}{|V_4|} [(\text{I}) \quad (\text{II}) \quad (\text{III}) \quad (\text{IV})] \quad , \quad \text{em que:}$$

$$|V_4| = \sigma_{Y1.k}^2 \cdot \sigma_{Y2.k}^2 \cdot \sigma_{Y3.k}^2 \cdot \sigma_{Y4.k}^2 + 2(\sigma_p^2)^3 (\sigma_{Y1.k}^2 + \sigma_{Y2.k}^2 + \sigma_{Y3.k}^2 + \sigma_{Y4.k}^2) + \\ - (\sigma_p^2)^2 (\sigma_{Y1.k}^2 \cdot \sigma_{Y2.k}^2 + \dots + \sigma_{Y3.k}^2 \cdot \sigma_{Y4.k}^2) - 3(\sigma_p^2)^4 \quad ,$$

$$(\text{I}) = (\sigma_{Y2.k}^2 - \sigma_p^2)(\sigma_{Y3.k}^2 - \sigma_p^2)(\sigma_{Y4.k}^2 - \sigma_p^2) \quad ,$$

$$(\text{II}) = (\sigma_{Y1.k}^2 - \sigma_p^2)(\sigma_{Y3.k}^2 - \sigma_p^2)(\sigma_{Y4.k}^2 - \sigma_p^2) \quad ,$$

$$(\text{III}) = (\sigma_{Y1.k}^2 - \sigma_p^2)(\sigma_{Y2.k}^2 - \sigma_p^2)(\sigma_{Y4.k}^2 - \sigma_p^2) \quad \text{e}$$

$$(\text{IV}) = (\sigma_{Y1.k}^2 - \sigma_p^2)(\sigma_{Y2.k}^2 - \sigma_p^2)(\sigma_{Y3.k}^2 - \sigma_p^2) \quad .$$

D) Cinco ambientes (forma geral):

$$C'V^{-1} = \frac{2\sigma_p^2}{|V_s|} [(I) \quad (II) \quad (III) \quad (IV) \quad (V)], \text{ em que:}$$

$|V_s|$ é o determinante da matriz V para cinco ambientes,

$$(I) = (\sigma_{\bar{Y}_{2,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{3,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{4,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{5,k}}^2 - \sigma_p^2),$$

$$(II) = (\sigma_{\bar{Y}_{1,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{3,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{4,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{5,k}}^2 - \sigma_p^2),$$

$$(III) = (\sigma_{\bar{Y}_{1,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{2,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{4,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{5,k}}^2 - \sigma_p^2),$$

$$(IV) = (\sigma_{\bar{Y}_{1,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{2,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{3,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{5,k}}^2 - \sigma_p^2) \text{ e}$$

$$(V) = (\sigma_{\bar{Y}_{1,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{2,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{3,k}}^2 - \sigma_p^2)(\sigma_{\bar{Y}_{4,k}}^2 - \sigma_p^2).$$

Pode-se assim estender para um maior número de ambientes a forma geral de $C'V^{-1}$, e por conseguinte as demais conclusões sobre a variância das predições e a correlação entre os valores genéticos aditivos preditos e esperados.

Para os casos com número de ambientes maior que três, é ainda mais pronunciada a superioridade da abordagem BLP, pois um maior número de variâncias, potencialmente de diferentes magnitudes, interferem na determinação dos pesos $C'V^{-1}$, além das covariâncias dos efeitos de progênie (fora da diagonal da matriz V).

É comum na prática florestal a ocorrência de ensaios que apresentam a magnitude das variâncias fenotípicas correlacionadas às médias ambientais com pronunciadas diferenças de desenvolvimento das progênie entre os ambientes considerados, as variâncias maiores associadas aos ambientes de maior média (o chamado efeito de escala). A tendência da média marginal nestes casos é de superestimar o peso relativo das progênie nos melhores ambientes e minimizar a superioridade de progênie em ambientes ruins, exatamente o inverso do que ocorre com o BLP, que atribuirá pesos baixos à superioridade de progênie em ambientes ótimos e pesos altos à superioridade em ambientes ruins.

Pode-se objetar que este procedimento é excessivamente conservador: pode-se estar interessado em selecionar exatamente para estas respostas extremas em ambientes ótimos. Para isto, basta que a seleção seja orientada para estes ambientes ótimos, lançando mão de vetores C' adequados para ambientes-alvo ao invés do vetor uniforme apresentado, adequado para um ambiente intermediário. O mesmo se pode dizer da seleção orientada para ambientes extremamente ruins, (sujeitos a algum estresse, por exemplo): basta que se disponha de estimativas de covariâncias genéticas entre os diferentes ambientes e o ambiente-alvo e desta forma, produzam-se os vetores C' adequados.

É interessante notar que o vetor de "pesos" para ambientes ótimos só se altera pela capitalização da σ_{pe}^2 para o ambiente considerado, os demais "pesos" permanecendo os mesmos! Pode-se verificar que, nestes casos, a acurácia será sempre maior que para a seleção considerando apenas um ambiente, segundo a idéia de que informações adicionais apenas reforçam a acurácia.

4.1.2. Comparações entre os efeitos do modelo

Os "diferenciais de seleção", ou a porção aleatória sobre a qual incidem os pesos no procedimento BLP podem ser descritos em termos do modelo misto como se segue:

BV: O peso único (herdabilidade na seleção entre médias marginais de progênie) aplica-se à seguinte diferença:

$$\begin{aligned}\bar{y}_{..k} - \bar{y}_{....} &= (p_k - \bar{p}_{..}) + (pE)_{..k} - (pE)_{...} + \bar{e}_{..k} - \bar{e}_{...} + \bar{d}_{..k} - \bar{d}_{...} \\ &= p_k + pE_{..k} + \bar{e}_{..k} + \bar{d}_{..k} - [(pE)_{...} + \bar{e}_{...} + \bar{d}_{...}].\end{aligned}$$

BLP: Cada um dos pesos aplica-se à seguinte diferença (para observações ao nível de médias de progênie por ambiente):

$$\begin{aligned}\bar{y}_{i.k} - E(\bar{y}_{i.k}) &= (p_k - \bar{p}_{..}) + (pE)_{ik} - (pE)_{i..} + \bar{e}_{ik} - \bar{e}_{i..} + \bar{d}_{ik} - \bar{d}_{i..} \\ &= p_k + (pE)_{ik} + \bar{e}_{ik} + \bar{d}_{ik},\end{aligned}$$

ou ainda, (para observações ao nível de médias de parcelas):

$$\bar{y}_{ijk} - E(\bar{y}_{ijk}) = p_k + (pE)_{ik} + e_{ijk} + \bar{d}_{ijk}.$$

$$\bar{y}_{ijk} - E(\bar{y}_{ijk}) = p_k + (pE)_{ik} + e_{ijk} + \bar{d}_{ijk}.$$

Observando estas expressões notamos que p_k é o mesmo para todas e corresponde à metade do valor g_i , a ser predito; para o BLP, $[(pE)_{ik} + \bar{e}_{ik} + \bar{d}_{ik}]$ no caso de médias por ambiente ou $[(pE)_{ik} + \bar{e}_{ik} + \bar{d}_{ijk}]$ no caso de médias de parcelas tendem a um valor de progênie, pois estão sujeitos aos efeitos genéticos (especialmente dentro de parcelas); o fato destes efeitos aparecerem "livres" de fatores perturbadores é uma decorrência de se supor conhecidos $E(\bar{Y}_{i.k.})$ como $E(\bar{Y}_{ijk.})$.

Para o BV, mesmo considerando nulos os fatores perturbadores $(\bar{e}_{ik} + \bar{d}_{ijk})$, os efeitos da interação progênie por ambientes não são aproveitados, pois não tendem a valores de progênie, ou seja, $[(p\bar{E})_{ik} - (p\bar{E})_{..}] = 0$.

O rigor com que se admite conhecidos os efeitos fixos depende, além da interação, dos fatores que influenciam diretamente a precisão das suas estimativas que serão empregadas no BLP, quais sejam:

- a) O delineamento experimental, estratificação ambiental, presença, adequação e eficiência de blocos, forma e tamanho de parcelas, etc.;
- b) A qualidade amostral dos ensaios, o número de progênie e de árvores por progênie que ocupa cada nível dos efeitos fixos estimados;
- c) A presença ou ausência de análises que levem à melhoria das estimativas de efeitos fixos, como análises de vizinhança e superfícies de resposta; ou ainda,
- d) A realização da estimação dos efeitos fixos num modelo generalizado ("Generalized Least Squares Estimation", ou GLSE, ao invés de OLSE ou de WLSE). Neste processo leva-se em conta as estimativas das c.v., o que mais efetivamente aproxima as estimativas (BLUE) dos valores paramétricos e as predições BLP do método BLUP.

4.2. Médias fenotípicas gerais (BV) versus BLP

(a partir de médias de parcelas)

Serão primeiro apresentados os vetores **C**, as matrizes **V**, os respectivos "pesos" $C'V^{-1}$ (herdabilidades efetivas), as variâncias das predições, dos erros associados às predições e a correlação entre os valores preditos e observados, para cada ambiente e cada um dos caracteres considerados.

A seguir, serão apresentadas as comparações entre os procedimentos de seleção, as progênies selecionadas, os progressos esperados pela seleção em cada caráter e as respostas correlacionadas pela seleção nos demais caracteres.

4.2.1. Vetores **C**, Matrizes **V** e vetores de pesos $C'V^{-1}$:

Para a seleção entre médias marginais de progênies, o peso de cada ambiente é uniforme e corresponde à herdabilidade para a seleção entre médias. Estes resultados são derivados das estimativas de componentes da variância obtidas por MORAES (1987), reunidas no apêndice 3, aplicados à seleção entre médias marginais.

Para o BLP, obteve-se os seguintes vetores e matrizes (matriz bloco-diagonal com valores correspondentes a cada progênie, valores de variâncias e covariâncias entre parcelas de progênies diferentes é zero):

DAP: $C'_{1 \times 9} = [0,2446 \quad 0,2446 \quad \dots \quad 0,2446];$

V_{9x9}: a diagonal de **V** tem elementos iguais, por ambiente:

diagonal= {0,8034 0,8034 0,8034
 1,0219 1,0219 1,0219
 0,8525 0,8525 0,8525};

fora da diagonal={0,1223} .

Altura: $C'_{1 \times 9} = [0,5478 \quad 0,5478 \quad \dots \quad 0,5478];$

$V_{9 \times 9}$: diagonal= {1,5468 1,5468 1,5468
 1,3243 1,3243 1,3243
 1,1271 1,1271 1,1271};

fora da diagonal={0,2739} .

Volume: $C'_{1 \times 9} = [9,30 \times 10^{-4} \quad 9,30 \times 10^{-4} \quad \dots \quad 9,30 \times 10^{-4}];$

$V_{9 \times 9}$: diagonal={ $7,134 \times 10^{-2}$ $7,134 \times 10^{-2}$ $7,134 \times 10^{-2}$
 $4,745 \times 10^{-2}$ $4,745 \times 10^{-2}$ $4,745 \times 10^{-2}$
 $2,394 \times 10^{-2}$ $2,394 \times 10^{-2}$ $2,394 \times 10^{-2}$ };

fora da diagonal={ $4,65 \times 10^{-4}$ }.

Densidade: $C'_{1 \times 9} = [4,06 \times 10^{-4} \quad 4,06 \times 10^{-4} \quad \dots \quad 4,06 \times 10^{-4}];$

$V_{9 \times 9}$: diagonal={ $2,50 \times 10^{-2}$ $2,50 \times 10^{-2}$ $2,50 \times 10^{-2}$
 $3,03 \times 10^{-2}$ $3,03 \times 10^{-2}$ $3,03 \times 10^{-2}$
 $2,90 \times 10^{-2}$ $2,90 \times 10^{-2}$ $2,90 \times 10^{-2}$ };

fora da diagonal={ $2,03 \times 10^{-4}$ }.

Os pesos relativos a cada local e as herdabilidades estão reunidos na tabela 2 a seguir (herdabilidades como apresentadas por MORAES, 1987).

Os resultados indicam que para cada local-alvo da seleção ocorre uma "redistribuição" da herdabilidade efetiva de forma a incorporar as informações dos demais locais. Estas informações novas são ponderadas com peso tanto maior quanto maiores sejam as porções das variâncias fenotípicas livres da variância de progênie para os demais ambientes. O mesmo se pode dizer da seleção para um ambiente intermediário.

Tabela 2: Pesos para cada ambiente-alvo de seleção e para cada caráter e local, aplicáveis a cada porção aleatória (no nível de repetições dentro de locais) e estimativas de herdabilidades na seleção entre médias de progênie por ambiente e entre médias gerais de progênie para os quatro caracteres ($h_{v_{1k}}^2$, $h_{v_{2k}}^2$, $h_{v_{3k}}^2$ e $h_{v_{4k}}^2$). Teste de progênie *E. grandis*, sete anos.

AMBIENTE ALVO	PESO NO AMBIENTE	DAP (cm ²)	ALTURA (m ²)	VOLUME (m ⁶ /arv ²)	DENSIDADE (g ² /cm ⁶)
ANHEMBI:	ANHEMBI	0,2954	0,4096	0,1715	0,3134
	BROTAS	0,0712	0,0046	0,0710	0,1943
	LENÇÓIS	0,0567	0,0792	0,1143	0,1194
$h_{v_{1k}}^2$ (%)	ANHEMBI	26,81	40,82	13,53	83,07
BROTAS:	ANHEMBI	0,0842	0,0021	0,0371	0,1945
	BROTAS	0,3187	0,4142	0,3255	0,3325
	LENÇÓIS	0,0498	0,0785	0,0819	0,1050
$h_{v_{2k}}^2$ (%)	BROTAS	27,68	40,24	27,42	70,36
LENÇÓIS:	ANHEMBI	0,0309	0,0218	0,0196	0,0425
	BROTAS	0,0230	0,0304	0,0269	0,0373
	LENÇÓIS	0,5342	0,5470	0,5026	0,5793
$h_{v_{3k}}^2$ (%)	LENÇÓIS	62,00	65,45	55,73	89,43
GERAL:	ANHEMBI	0,1367	0,0963	0,0631	0,2568
	BROTAS	0,1016	0,1346	0,0865	0,2255
	LENÇÓIS	0,0809	0,1367	0,1393	0,1386
$h_{v_{4k}}^2$ (%)	GERAL	34,39	45,49	27,72	85,50

As variâncias das predições e dos erros das predições, bem como as correlações entre valores aditivos e as predições estão reunidos na tabela 3 a seguir.

Comparando as tabelas 2 e 3 (pp. 59 e 60), nota-se que tal "redistribuição" da herdabilidade efetiva não equivale a dizer que o ambiente de maior peso é o de maior herdabilidade. Apenas para o caso do volume cilíndrico, por exemplo, este princípio intuitivo se aplica, ainda assim sem correlação perfeita entre pesos e herdabilidades.

Tabela 3: Variâncias das predições [$V(\hat{g})$, na unidade correspondente], dos erros das predições [$V(\hat{e})$, como porcentagem da variância aditiva], e acurácia seletiva [$r(\hat{g},g)$, em porcentagem] para cada caráter, conforme o método considerado (BLP ou BV) e o ambiente-alvo da seleção. Teste de Progenies de *E. grandis*, aos sete anos.

AMBIENTE/ PARÂMETRO	MÉTODO DE PREDIÇÃO	DAP (cm ²)	ALTURA (m ²)	VOLUME (m ⁶ /arv ²)	DENSIDADE (g ² /cm ⁶)
Anhembi:					
$V(\hat{g})$	BLP	0,4756	1,6894	0,001510	0,000773
	BV	0,2310	1,0309	0,000522	0,000691
$V(\hat{e})$ (%)	BLP	44,80	33,11	60,89	7,06
	BV	73,19	59,18	86,48	16,94
$r(\hat{g},g)$ (%)	BLP	74,30	91,79	62,55	96,39
	BV	51,78	63,89	36,78	91,14
Brotas:					
$V(\hat{g})$	BLP	0,6393	1,4570	0,002873	0,000790
	BV	0,3132	0,8577	0,001427	0,000599
$V(\hat{e})$ (%)	BLP	43,50	31,65	44,80	7,31
	BV	72,32	59,76	72,588	29,69
$r(\hat{g},g)$ (%)	BLP	75,16	82,67	74,30	96,29
	BV	52,61	63,44	52,36	83,88
Lençóis:					
$V(\hat{g})$	BLP	1,7338	2,5068	0,004153	0,000997
	BV	1,3109	1,9313	0,002974	0,000926
$V(\hat{e})$ (%)	BLP	18,00	15,05	22,18	3,73
	BV	38,00	34,55	44,27	10,62
$r(\hat{g},g)$ (%)	BLP	90,55	92,17	88,22	98,10
	BV	78,74	80,90	74,65	94,57
Geral:					
$V(\hat{g})$	BLP	0,2343	0,6042	0,000806	0,000756
	BV	0,1682	0,4983	0,000516	0,000694
$V(\hat{e})$ (%)	BLP	52,11	44,86	56,67	6,87
	BV	65,62	54,52	72,26	14,53
$r(\hat{g},g)$ (%)	BLP	69,20	74,26	65,83	96,49
	BV	58,64	67,45	52,65	92,47

Pode-se atribuir tal expectativa frustrada à suposição corrente de que a herdabilidade tomada apenas num ambiente pode medir o progresso pela seleção das progênes naquele ambiente, o que equivale a se supor o "aproveitamento" da interação progênes por local no local considerado.

Nos modelos de seleção, a interação não deve ser vista como a falta de consistência da resposta de um genótipo (enfoque adequado para o modelo fixo). Uma progênie de meios-irmãos é uma população com ampla variabilidade interna de genótipos. A natureza dos efeitos da interação deve ser assim abrandada e enfatizados os efeitos das demais amostras da progênie nos diferentes ambientes. Certamente isto é mais efetivo para a seleção, uma vez que os efeitos fixos de ambientes sejam bem estimados.

Conclui-se que medidas de progresso genético esperado através das herdabilidades gerais são apenas aproximações, bem menos eficazes que as predições BLP, seja qual for o tipo de resposta esperada, ou seja: para um ambiente qualquer, BLP melhora as predições pelas informações dos demais ambientes devidamente ponderadas (informações em geral desprezadas); para um ambiente intermediário, BLP pondera melhor que a média geral o valor de cada progênie.

4.2.2. Progênes Seleccionadas Através do BLP

O ordenamento de progênes para a seleção e os seus respectivos valores genéticos preditos pelos métodos BV e BLP é apresentado nas tabelas 4 a 7, nas páginas 62 e 63. Destes resultados percebem-se certas diferenças de ordenamento, as quais implicariam em diferenças de seleção qualquer que fosse o número de progênes seleccionadas. Tais diferenças são atribuídas a erros de ordenamento pelo método BV, pois este apresentou menor correlação entre predições e valor genético real.

Tabela 4: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLP. T.p. *E. grandis*, 7 anos. Caráter: DAP.

Progenies	BV	BLP	n.o.
M024	0,4227	0,2952	1
S061	0,2719	0,1912	9
S031	0,2595	0,2833	2
S076	0,2572	0,2809	3
S038	0,2527	0,2246	6
S005	0,2436	0,2316	5
S019	0,2334	0,2057	7
S074	0,2243	0,1323	12*
S080	0,2130	0,2338	4
S012	0,2107	0,0458	19**
M002	0,1960	0,1401	11
M013	0,1892	0,0417	20**
S052	0,1722	0,1738	10
S057	0,1631	0,0946	16**
S023	0,1614	0,1247	14
S028	0,1461	0,1294	13
M057	0,1314	0,2048	8
M017	0,0832	0,0298	23
S053	0,0725	0,0883	17
M055	0,0555	0,0953	15
S002	0,0385	0,0261	25
M039	0,0283	0,0319	21
S040	0,0124	0,0827	18
M044	0,0090	-0,0416	32
S056	0,0022	-0,0593	35
S029	-0,0012	0,0162	27
S024	-0,0114	0,0134	29
S051	-0,0250	0,0274	24
S062	-0,0363	0,0256	26
S017	-0,0465	-0,0361	31
M058	-0,0545	-0,0010	30
S004	-0,0567	-0,0681	37
S034	-0,0715	-0,0460	33
S045	-0,0828	-0,1218	40
S006	-0,0987	0,0147	28
M005	-0,0987	-0,0520	34
S078	-0,1213	-0,1287	42
S008	-0,1270	-0,1284	41
S016	-0,1315	-0,1090	39
S036	-0,1361	-0,0683	38
S003	-0,1429	-0,1741	46
S058	-0,1610	0,0312	22
S079	-0,1655	-0,1843	47
M048	-0,1712	-0,0656	36
I324	-0,1780	-0,1295	43
S039	-0,1893	-0,1654	45
M047	-0,2143	-0,1376	44
S042	-0,2335	-0,2141	49
S075	-0,3004	-0,2113	48
S071	-0,3684	-0,3332	52
I327	-0,4126	-0,2419	50
M046	-0,4137	-0,2861	51

Tabela 5: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLP. T.p. *E. grandis*, 7 anos. Caráter: altura.

Progenies	BV	BLP	n.o.
S076	0,5557	0,5933	1
S019	0,4571	0,4288	2
M013	0,4118	0,3179	5
S012	0,4072	0,3221	4
M024	0,4027	0,3105	7
S061	0,3109	0,1833	9
S031	0,3075	0,3240	3
S074	0,2531	0,2458	8
S053	0,2146	0,3142	6
S052	0,2078	0,1470	15*
S038	0,2055	0,1818	10
S008	0,2044	0,0959	21**
M017	0,1721	0,0725	25**
S004	0,1500	0,1057	20**
M044	0,1477	0,1739	13
S080	0,1239	0,1356	17
S034	0,1012	0,0800	22
S024	0,0956	0,1510	14
M002	0,0956	0,0632	27
M055	0,0944	0,0767	23
S023	0,0842	0,0762	24
S040	0,0786	0,1763	12
S002	0,0616	0,1392	16
S017	0,0287	0,0378	30
S028	0,0253	0,1804	11
S005	0,0185	0,1191	19
M057	0,0151	0,1269	18
M005	0,0128	0,0382	29
M039	0,0004	-0,0410	33
M058	-0,0132	0,0245	31
S057	-0,0257	-0,1253	40
S016	-0,0416	0,0578	28
M047	-0,0438	-0,0675	35
S051	-0,0461	0,0648	26
S056	-0,0540	-0,0566	34
S029	-0,0846	-0,0285	32
S039	-0,1413	-0,1625	43
M048	-0,1674	-0,1180	39
S062	-0,1730	-0,0889	38
S071	-0,1776	-0,0804	37
S045	-0,1923	-0,0786	36
S003	-0,2252	-0,1963	44
S058	-0,2263	-0,1486	42
S075	-0,2354	-0,2372	46
S006	-0,2637	-0,1410	41
S036	-0,2796	-0,2758	47
S042	-0,3419	-0,2356	45
M046	-0,3442	-0,2857	48
S078	-0,3884	-0,3364	49
I324	-0,4303	-0,3574	50
S079	-0,6309	-0,4079	51
I327	-0,7170	-0,5464	52

nota: as linhas horizontais referem-se à seleção truncada e os asteriscos aos erros do método (BV), para a intensidade de dez (*) ou quinze (**) progênies selecionadas.

Tabela 6: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLP. T.p. *E. grandis*, 7 anos. Caráter: volume.

Progenies	BV	BLP	n.o.
M024	0,03644	0,02131	1
S076	0,02568	0,01856	2
S080	0,01950	0,00626	15*
S061	0,01821	0,00845	12*
S052	0,01712	0,01269	4
S031	0,01690	0,01514	3
S038	0,01513	0,01219	5
S012	0,01388	0,00824	13*
S023	0,01275	0,01009	9
S053	0,01246	0,00875	11*
S019	0,01227	0,01001	10
M013	0,00942	0,00237	20**
S074	0,00940	0,01076	6
S028	0,00856	0,01016	8
S005	0,00778	0,01042	7
M017	0,00748	0,00432	17
M055	0,00635	0,00352	18
M002	0,00573	0,00227	22
S029	0,00544	0,00489	16
S036	0,00418	-0,00288	35
S057	0,00416	0,00231	21
S056	0,00317	-0,00218	33
M057	0,00287	0,00695	14
S040	0,00203	0,00269	19
S024	0,00001	0,00076	25
S051	-0,00120	0,00094	24
M048	-0,00158	0,00036	26
M039	-0,00293	-0,00292	36
S002	-0,00298	0,00111	23
M044	-0,00315	-0,00215	32
S008	-0,00613	-0,00915	45
S004	-0,00620	-0,00344	39
S034	-0,00646	-0,00332	38
S045	-0,00646	-0,00378	40
S062	-0,00670	-0,00115	30
S058	-0,00674	-0,00278	34
S078	-0,00723	-0,00113	29
M058	-0,00774	-0,00323	37
S017	-0,00779	-0,00201	31
S079	-0,00859	-0,00084	28
S016	-0,00915	-0,00041	27
M005	-0,01063	-0,00723	42
S003	-0,01134	-0,00750	43
S039	-0,01281	-0,00938	46
M047	-0,01331	-0,01281	48
S006	-0,01576	-0,00559	41
S071	-0,01817	-0,00815	44
I324	-0,01967	-0,00968	47
M046	-0,01970	-0,01552	51
S075	-0,02105	-0,01623	52
S042	-0,02129	-0,01299	49
I327	-0,02219	-0,01313	50

Tabela 7: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLP. T.p. *E. grandis*, 7 anos. Caráter: densidade.

Progenies	BV	BLP	n.o.
M055	0,0139	0,0241	1
I327	0,0125	0,0197	2
S053	0,0099	0,0151	3
S045	0,0078	0,0127	5
S042	0,0070	0,0120	6
M044	0,0070	0,0113	7
M046	0,0068	0,0127	4
S040	0,0054	0,0055	14*
S057	0,0050	0,0100	8
S078	0,0049	0,0055	13*
S079	0,0046	0,0093	9
M039	0,0046	0,0081	10
S051	0,0037	0,0066	11
S016	0,0034	0,0065	12
I324	0,0020	0,0032	19**
S076	0,0017	0,0040	17
S062	0,0016	0,0033	18
S061	0,0014	0,0047	15
S029	0,0014	0,0020	25
M005	0,0009	0,0026	20
S075	0,0007	-0,0008	27
S080	0,0006	-0,0002	26
S071	0,0002	0,0023	23
S038	0,0000	0,0021	24
M058	-0,0002	-0,0010	28
S019	-0,0003	0,0041	16
S008	-0,0004	0,0024	22
M013	-0,0006	0,0025	21
S039	-0,0010	-0,0017	29
S023	-0,0013	-0,0022	31
S052	-0,0019	-0,0027	33
S031	-0,0020	-0,0019	30
S074	-0,0021	-0,0035	37
S003	-0,0022	-0,0023	32
S028	-0,0024	-0,0072	45
S058	-0,0025	-0,0057	41
S034	-0,0026	-0,0039	38
S006	-0,0033	-0,0031	35
M024	-0,0034	-0,0031	36
M057	-0,0034	-0,0059	42
M047	-0,0040	-0,0042	39
M002	-0,0040	-0,0048	40
S012	-0,0040	-0,0059	43
S004	-0,0043	-0,0029	34
S056	-0,0047	-0,0073	46
S017	-0,0049	-0,0060	44
S024	-0,0061	-0,0106	49
M017	-0,0065	-0,0096	47
S002	-0,0075	-0,0104	48
S036	-0,0091	-0,0158	50
M048	-0,0103	-0,0172	51
S005	-0,0123	-0,0182	52

nota: as linhas horizontais referem-se à seleção truncada e os asteriscos aos erros do método (BV), para a intensidade de dez (*) ou quinze (**) progênies selecionadas.

Para a seleção de 10 dentre as 52 progêniez comuns, por exemplo, são selecionadas (ou descartadas) de forma errônea 2 progêniez para o caráter DAP, 1 para a altura, 4 para o volume e 2 para a densidade. Na seleção de 15 das 52, o mesmo ocorre com 3 progêniez para o caráter DAP, 3 para a altura, 1 para o volume e 1 para a densidade.

Grosso modo, observam-se diferenças entre 10% e 20% das seleções, para uma faixa situada entre 10% e 30% de progêniez selecionadas. Não obstante, a correlação entre os valores aditivos preditos pelos dois métodos apresentou-se em geral bastante alta, tal como: 93,8% para o caráter DAP, 96,2% para a altura, 93,9% para o volume e 98,2% para a densidade. Estes fatos indicam que embora revelem-se erros de ordenamento, pouca diferença deve ser encontrada quanto às predições de ganho.

As tabelas 15 e 16 da página 73 mostram as predições de ganhos e respostas correlacionadas para cada um dos métodos. Pode-se notar que o BLP nem sempre apresentou progresso genético superior ao BV. Para o DAP e a altura de árvores, por exemplo, o BV produziu ganhos maiores! A precisão de tais predições é menor, pois representam a média das predições de valores genéticos de progêniez calculadas por um método cujas propriedades estatístico-genéticas são piores, como discutido anteriormente.

É claro que o progresso medido como função dos BLP resultaria sempre maior pelo método BLP... Entrementes, isto não teria qualquer valor ilustrativo sobre a superioridade do BLP, sendo apenas uma tautologia. Deve-se ressaltar, ao contrário, razões objetivas para a observação de tais valores superiores para BV, tais como:

a) As estimativas das componentes da variância empregadas em ambos os métodos são as oriundas do processo de cálculo associado ao BV (método dos momentos aplicado às variâncias

de médias marginais das médias ajustadas de progênes, obtidas através dos látices);

b) As esperanças das médias de parcelas ($m+E_i+B_j(i)$) foram calculadas pelo método dos quadrados mínimos ordinários, num modelo fixo (através da "proc GLM" do SAS) ou seja, não houve qualquer influência das componentes da variância neste cálculo, o que prejudica as estimativas das referidas esperanças.

c) O número de ambientes analisados (apenas três) é o menor para o qual a superioridade do BLP se revela (a menos de estimativas independentes das componentes da variância para a análise de dois ambientes, como já discutido no item 4.1).

Pode-se concluir que o BLP apresentou superioridade em relação ao BV, mais flagrante em termos de ordenamento de progênes para a seleção do que em termos de predição de valores aditivos, ou seja, na seleção pelo método usual deve-se incorrer em maiores erros de ordenamento, embora isto possa não significar perda importante no progresso genético.

4.3. BLUP Iterativo

Serão primeiro apresentadas as estimativas de componentes da variância associadas à estimação-predição REML. A seguir apresentar-se-ão os resultados da seleção de progênes pelo método BLUP.

4.3.1. Estimativas REML de Componentes da Variância

A tabela 8 da página 66 apresenta as estimativas das componentes da variância e os respectivos erros-padrão a elas associados, além das estimativas de médias de locais.

Tais estimativas podem ser comparadas ao conjunto de estimativas correspondentes apresentado por MORAES (1987) e relacionadas no apêndice 2.

Tabela 8: Estimativas de componentes da variância e respectivos erros padrão obtidas através do método REML para cada um dos quatro caracteres em estudo, por locais e conjuntas, no modelo misto. Estimativas de médias gerais por locais e conjuntas. Teste de progenies de *E. grandis*, aos sete anos.

CARÁTER	PARÂMETRO	ANHEMBI		BROTAS		LENÇÓIS		CONJUNTA	
		Estimativa	Erro Padrão	Estimativa	Erro Padrão	Estimativa	Erro Padrão	Estimativa	Erro Padrão
DAP	σ_p^2	0,000000	—	0,238997	0,287807	0,522625	0,173870	0,130263	0,083314
	σ_{pe}^2							0,153674	0,119532
	σ_e^2	0,000000	—	2,367019	0,436282	0,146510	0,145743	0,628462	0,150602
	σ_d^2	20,785407	0,691695	9,440134	0,329866	9,354134	0,331708	13,395905	0,271628
	Média do DAP	16,55		14,24		12,15		14,36	
ALTURA	σ_p^2	0,521600	0,267121	0,763255	0,361173	1,381823	0,370922	0,566846	0,184150
	σ_{pe}^2							0,369494	0,163420
	σ_e^2	0,000000	—	2,493117	0,410684	1,143596	0,222762	1,022050	0,178447
	σ_d^2	27,040725	0,909781	6,223028	0,217453	5,093877	0,180948	13,020797	0,263694
	Média de Altura	25,67		21,43		19,88		22,33	
VOLUME	σ_p^2	0,000000	—	0,000944	0,001422	0,001386	0,000487	0,000712	0,000424
	σ_{pe}^2							0,000449	0,000591
	σ_e^2	0,000000	—	0,012999	0,002218	0,000829	0,000436	0,002247	0,000802
	σ_d^2	0,180645	0,006011	0,038219	0,001336	0,024260	0,000860	0,083405	0,001690
	Média do Volume	0,6451		0,3792		0,2581		0,4275	
DENSIDADE	σ_p^2	0,000210	0,000057	0,000222	0,000061	0,000280	0,000059	0,000077	0,000031
	σ_{pe}^2							0,000159	0,000031
	σ_e^2	0,000110	0,000034	0,000166	0,000038	0,000028	0,000013	0,000101	0,000016
	σ_d^2	0,001494	0,000052	0,001185	0,000042	0,000744	0,000026	0,001144	0,000023
	Média da Densidade	0,420		0,469		0,471		0,453	

Desta comparação entre o método REML para o modelo misto e o método dos momentos para o modelo aleatório, observam-se algumas regularidades, tais como:

a) As estimativas de σ_e^2 , especialmente as produzidas pelo método da ANAVA, foram bastante baixas.

Este resultado é relativamente comum na literatura de análises de testes de progênies de meios-irmãos de eucaliptos e pode estar associado à forma e tamanho de parcelas, bem como à inadequação do modelo, que produz mesmo estimativas negativas para as componentes da variância.

b) Para o DAP, na análise por local, às estimativas nulas da σ_e^2 da ANAVA, correspondem estimativas nulas de σ_e^2 e σ_p^2 do método REML.

Para a ANAVA, estimativas nulas são apresentadas em substituição às suas correspondentes negativas. Para REML, sempre que surgir uma estimativa negativa em qualquer iteração. REML é, portanto, mais sensível na anulação das estimativas.

c) Para a ANAVA, o fato (estimativas nulas) se repetiu para o caráter volume e o local "Lençóis Paulista", enquanto que para o REML nenhuma outra estimativa nula foi obtida.

Este é um indício claro da maior adequação do modelo misto e não uma influência do método de estimação (como vimos, REML é mais sensível a estimativas nulas).

d) Houve alterações importantes nas relações entre as componentes da variância, acarretando aumento nas estimativas de herdabilidade para os caracteres DAP, altura, mas uma importante redução nesta estimativa para o caráter densidade. Para o REML, a relação proposta por Shelbourne (σ_{pe}^2/σ_p^2) foi menor do que pelo método da ANAVA, para os caracteres DAP, altura e volume (respectivamente 1,18; 0,65; 0,63) porém foi maior para o caráter densidade (2,06).

Estes fatos indicam que, na comparação do BLUP com BV, pode-se esperar uma maior superioridade que a registrada na comparação do BLP com o BV, pois tanto as estimativas de efeitos de ambientes e blocos quanto as predições de valores aditivos estarão sujeitas aos efeitos das componentes da variância, das quais têm-se novas estimativas.

Embora os erros associados às estimativas da ANAVA não tenham sido calculados para a comparação, os erros das estimativas REML são apresentados. Pode-se dizer que as componentes da variância em geral não foram calculadas com boa precisão pelo método REML. Especialmente imprecisas são as estimativas da σ_e^2 . As melhores estimativas foram as da análise conjunta e as relações entre as estimativas da análise conjunta são as empregadas nas matrizes da predição BLUP do SAS. A tais estimativas associam-se em geral intervalos de confiança estreitos, exceto para a componente σ_{pe}^2 do caráter volume, cujo erro é maior que a estimativa. Desta forma, espera-se a obtenção de predições BLUP mais precisas que BLP, exceto pela não utilização pelos BLUP da decomposição da interação σ_{pe}^2 por locais.

4.3.2. Progênie Selecionadas Através do BLUP

O ordenamento de progênie para a seleção e os seus respectivos valores genéticos preditos pelos métodos BV e BLP é apresentado nas tabelas 9 a 12, nas páginas 69 e 70. Registraram-se também aqui diferenças de ordenamento que implicariam em diferenças de seleção, qualquer que fosse o número de progênie selecionadas. Tais diferenças são atribuídas a erros de ordenamento pelo método BV, como já discutido para o BLP.

A tabela 13 (página 71), indica os erros na seleção de 10 progênie, dentre as 52 comuns, na seleção através dos látices individuais e do BV (análise conjunta) quando comparados aos respectivos BLUP.

Tabela 9: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLUP. T.p. E. *grandis*, 7 anos. Caráter: DAP.

Progênes	BV	BLUP	n.o.
M024	0,4227	0,4634	1
S061	0,2719	0,2402	12*
S031	0,2595	0,1904	14*
S076	0,2572	0,4044	3
S038	0,2527	0,4094	2
S005	0,2436	0,3532	6
S019	0,2334	0,2942	7
S074	0,2243	0,2270	13*
S080	0,2130	0,2648	10
S012	0,2107	0,2700	8
M002	0,1960	0,3548	4
M013	0,1892	0,0840	24**
S052	0,1722	0,3540	5
S057	0,1631	0,1534	17**
S023	0,1614	0,1314	19**
S028	0,1461	0,1430	18
M057	0,1314	0,2606	11
M017	0,0832	0,0172	29
S053	0,0725	0,0806	25
M055	0,0555	0,1280	20
S002	0,0385	0,0770	26
M039	0,0283	0,0304	28
S040	0,0124	0,0906	23
M044	0,0090	-0,0458	35
S056	0,0022	0,0678	27
S029	-0,0012	0,0000	31
S024	-0,0114	-0,0810	39
S051	-0,0250	0,1018	22
S062	-0,0363	-0,0122	33
S017	-0,0465	-0,0394	34
M058	-0,0545	-0,1212	45
S004	-0,0567	-0,0842	41
S034	-0,0715	-0,1816	52
S045	-0,0828	-0,0820	40
S006	-0,0987	-0,0028	32
M005	-0,0987	-0,1146	44
S078	-0,1213	-0,1110	43
S008	-0,1270	-0,2286	55
S016	-0,1315	-0,1018	42
S036	-0,1361	-0,1686	50
S003	-0,1429	-0,2024	54
S058	-0,1610	-0,0708	37
S079	-0,1655	-0,1388	48
M048	-0,1712	-0,1278	46
I324	-0,1780	-0,2454	58
S039	-0,1893	-0,2302	56
M047	-0,2143	-0,1760	51
S042	-0,2335	-0,2862	59
S075	-0,3004	-0,2892	60
S071	-0,3684	-0,4126	63
I327	-0,4126	-0,4186	64
M046	-0,4137	-0,4778	65

Tabela 10: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLUP. T.p. E. *grandis*, 7 anos. Caráter: altura.

Progênes	BV	BLUP	n.o.
S076	0,5557	2,0090	1
S019	0,4571	1,3822	2
M013	0,4118	1,0790	6
S012	0,4072	1,2384	5
M024	0,4027	1,2488	4
S061	0,3109	0,7392	10
S031	0,3075	0,7180	11*
S074	0,2531	0,6264	14*
S053	0,2146	0,9556	7
S052	0,2078	0,7940	9
S038	0,2055	0,8242	8
S008	0,2044	0,4382	18**
M017	0,1721	0,3880	21**
S004	0,1500	0,2982	25**
M044	0,1477	0,3188	24**
S080	0,1239	0,3352	23
S034	0,1012	0,2240	29
S024	0,0956	0,4230	19
M002	0,0956	0,2550	27
M055	0,0944	0,3922	20
S023	0,0842	0,2158	30
S040	0,0786	0,6496	12
S002	0,0616	0,4602	17
S017	0,0287	0,1778	32
S028	0,0253	0,5622	16
S005	0,0185	0,2384	28
M057	0,0151	0,3702	22
M005	0,0128	0,0162	34
M039	0,0004	-0,0942	38
M058	-0,0132	-0,0210	36
S057	-0,0257	-0,1144	40
S016	-0,0416	0,1870	31
M047	-0,0438	-0,1160	41
S051	-0,0461	0,2700	26
S056	-0,0540	0,0570	33
S029	-0,0846	-0,1088	39
S039	-0,1413	-0,5918	51
M048	-0,1674	-0,2054	42
S062	-0,1730	-0,3640	46
S071	-0,1776	-0,0782	37
S045	-0,1923	-0,3000	45
S003	-0,2252	-0,7396	55
S058	-0,2263	-0,4980	50
S075	-0,2354	-0,6440	53
S006	-0,2637	-0,3800	47
S036	-0,2796	-0,6124	52
S042	-0,3419	-0,8878	57
M046	-0,3442	-1,0174	58
S078	-0,3884	-1,0946	60
I324	-0,4303	-1,4118	61
S079	-0,6309	-1,8566	65
I327	-0,7170	-1,7544	63

nota: as linhas horizontais referem-se à seleção truncada e os asteriscos aos erros do método (BV), para a intensidade de dez (*) ou quinze (**) progênes selecionadas.

Tabela 11: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLUP. T.p. E. *grandis*, 7 anos. Caráter: volume.

Progenies	BV	BLUP	n.o.
M024	0,03644	0,04300	1
S076	0,02568	0,03800	2
S031	0,01950	0,02280	5
S052	0,01821	0,01760	8
S038	0,01712	0,03320	3
S074	0,01690	0,01380	13*
S005	0,01513	0,02680	4
S028	0,01388	0,01900	6
S023	0,01275	0,01320	15*
S019	0,01246	0,01540	12*
S053	0,01227	0,01740	9
S061	0,00942	0,00120	31**
S012	0,00940	0,00880	21**
M057	0,00856	0,01180	19**
S080	0,00778	0,01660	10
S029	0,00748	0,00400	28
M017	0,00635	0,01240	17
M055	0,00573	0,01540	11
S040	0,00544	0,00560	25
M013	0,00418	0,00480	27
S057	0,00416	0,00360	29
M002	0,00317	0,00840	22
S002	0,00287	0,01200	18
S051	0,00203	0,00920	20
S024	0,00001	-0,00440	36
M048	-0,00120	0,00640	24
S016	-0,00158	0,00480	26
S079	-0,00293	-0,00700	41
S078	-0,00298	0,00340	30
S062	-0,00315	-0,00920	48
S017	-0,00613	-0,01080	51
M044	-0,00620	-0,00720	44
S056	-0,00646	-0,00620	38
S058	-0,00646	-0,01580	54
S036	-0,00670	-0,00700	43
M039	-0,00674	-0,00040	33
M058	-0,00723	-0,00980	50
S034	-0,00774	-0,01600	55
S004	-0,00779	-0,00480	37
S045	-0,00859	-0,00720	45
S006	-0,00915	-0,00640	39
M005	-0,01063	-0,01380	53
S003	-0,01134	-0,01740	57
S071	-0,01281	-0,01620	56
S008	-0,01331	-0,00940	49
S039	-0,01576	-0,00840	47
I324	-0,01817	-0,02080	59
M047	-0,01967	-0,03040	64
S042	-0,01970	-0,02460	61
I327	-0,02105	-0,02480	62
M046	-0,02129	-0,02800	63
S075	-0,02219	-0,02220	60

Tabela 12: Predições de valores aditivos em ordem decrescente para BV e segundo o número de ordem do BLUP. T.p. E. *grandis*, 7 anos. Caráter: densidade.

Progenies	BV	BLUP	n.o.
M055	0,0139	0,0732	1
I327	0,0125	0,0658	2
S053	0,0099	0,0464	3
S045	0,0078	0,0336	7
S042	0,0070	0,0398	4
M044	0,0070	0,0376	6
M046	0,0068	0,0388	5
S040	0,0054	0,0222	12*
S057	0,0050	0,0286	8
S078	0,0049	0,0234	11*
S079	0,0046	0,0272	9
M039	0,0046	0,0244	10
S051	0,0037	0,0190	13**
S016	0,0034	0,0180	14**
I324	0,0020	0,0118	17**
S076	0,0017	0,0084	21
S062	0,0016	0,0086	19
S061	0,0014	0,0014	29
S029	0,0014	0,0078	23
M005	0,0009	0,0072	24
S075	0,0007	0,0024	28
S080	0,0006	0,0044	25
S071	0,0002	-0,0004	35
S038	0,0000	0,0002	32
M058	-0,0002	0,0004	31
S019	-0,0003	-0,0002	34
S008	-0,0004	0,0000	33
M013	-0,0006	0,0030	27
S039	-0,0010	-0,0022	37
S023	-0,0013	-0,0078	39
S052	-0,0019	-0,0100	43
S031	-0,0020	-0,0080	40
S074	-0,0021	-0,0062	38
S003	-0,0022	-0,0098	42
S028	-0,0024	-0,0212	55
S058	-0,0025	-0,0126	45
S034	-0,0026	-0,0090	41
S006	-0,0033	-0,0128	46
M024	-0,0034	-0,0168	49
M057	-0,0034	-0,0152	47
M047	-0,0040	-0,0170	50
M002	-0,0040	-0,0172	51
S012	-0,0040	-0,0178	53
S004	-0,0043	-0,0164	48
S056	-0,0047	-0,0220	57
S017	-0,0049	-0,0214	56
S024	-0,0061	-0,0296	58
M017	-0,0065	-0,0326	59
S002	-0,0075	-0,0332	60
S036	-0,0091	-0,0476	63
M048	-0,0103	-0,0508	64
S005	-0,0123	-0,0578	65

nota: as linhas horizontais referem-se à seleção truncada e os asteriscos aos erros do método (BV), para a intensidade de dez (*) ou quinze (**) progenies selecionadas.

Tabela 13: Diferenças de ordenamento entre os látices individuais e os BV (análise conjunta dos látices), em relação aos BLUP. Seleção de 10 progênes dentro 52. Teste de progênes de *E. grandis*, sete anos.

CARÁTER	ANHEMBI	BROTAS	LENÇÓIS	GERAL
DAP	4	1	2	3
Altura	2	2	1	2
Volume	1	2	1	3
Densidade	1	0	1	2

Na seleção de 15 dentre as 52 progênes comuns, são incluídas (ou descartadas) de forma errônea 3 progênes para o caráter DAP, 4 para a altura, 3 para o volume e 3 para a densidade. Observam-se, portanto, diferenças maiores que para o BLP (entre 15% e 25% das seleções), para uma faixa situada entre 10% e 30% de progênes selecionadas.

Os valores das correlações entre as predições dos métodos BV e BLUP (na análise conjunta) foram: 94,2% para o caráter DAP, 97,6% para a altura, 95,0% para o volume e 98,8% para a densidade. Valores ainda maiores que para as correlações entre BLP e BV. Para tais valores não se esperaria maiores erros de ordenamento, tampouco maiores diferenças nas predições de ganho, como as reveladas.

Quanto à medição das propriedades estatísticas do BLUP, os parâmetros usados na comparação BLP-BV estão na tabela 14 (abaixo), oriundos das saídas de erros padrão da predição por progênie produzidas pela "proc MIXED" do SAS.

Tabela 14: Variância das predições e variância (média) do erro das predições para o BLUP de progênes como porcentagens da variância aditiva. Teste de progênes de *E. grandis*, sete anos.

PARÂMETRO	DAP	ALTURA	VOLUME	DENSIDADE
$V(\hat{g})\%$	62,37%	71,60%	62,53%	60,39%
$V(\hat{e})\%$	37,63%	28,40%	37,47%	39,61%

O BLUP foi mais preciso para os caracteres de crescimento, ao contrário da densidade básica da madeira (comparar com a tabela 3 da página 60). A principal razão para esta tendência é a diferença nas componentes da variância empregadas pelo BLUP das empregadas nos métodos

BLP e BV, com a forte redução da estimativa da variância de progênie (livre da interação) para este caráter.

A importância das componentes da variância e dos desbalanceamentos seria melhor demonstrada num ensaio de simulação (e.g. POLLAK e QUAAS, 1981a e 1981b). Para fins da comparação em questão é importante notar que o método da ANAVA com o emprego das médias gerais ignora a influência das referidas componentes no ordenamento.

A tabela 17 da página 73 relaciona ganhos e respostas correlacionadas para cada caráter sob seleção pelos BLUP de progênie.

O BLUP apresentou progressos genéticos sempre superiores ao BV. Este fato, somado às melhores propriedades do método, faz com que o BLUP seja preferível ao BV. Quanto ao BLP, a superioridade do BLUP depende do conjunto de estimativas das c.v. considerado. No caso estudado, BLUP foi superior, exceto para a densidade da madeira, em que a forma BLP estudada propiciaria maiores ganhos e melhor valor explicativo do comportamento diferencial das progênie ao longo dos ambientes.

Ambas as abordagens (BLP e BLUP) permitem o cálculo de valores genéticos relativamente precisos para as progênie não comuns a todos os ambientes. Os resultados deste tipo de seleção podem ser encontrados nas tabelas 18 e 19 (pág. 74) e indicam a possibilidade de progênie não comuns serem selecionadas (entre as 10 ou 15 melhores, por exemplo) por qualquer dos métodos de predição.

Tal possibilidade não é sequer cogitada na teoria usual, embora não seja raro que progênie falhadas ou não plantadas em determinado ambiente chamem a atenção do melhorista, que deve então confirmar sua observação com novos experimentos (sem conexão com os anteriores).

Tabela 15: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de 10 progênies das 52, baseada em suas médias marginais (BV). Teste de Progênies de *Eucalyptus grandis* aos sete anos.

CARÁTER SELECIONADO	RESPOSTA CORRELACIONADA NO CARÁTER			
	DAP	ALTURA	VOLUME	DENSIDADE
DAP	0,26 1,80%	0,30 1,36%	0,0121 2,80%	-0,0020 -0,45%
Altura	0,23 1,61%	0,35 1,57%	0,0116 2,68%	-0,0001 -0,03%
Volume	0,24 1,65%	0,25 1,12%	0,0131 3,03%	-0,0024 -0,53%
Densidade	-0,10 -0,66%	-0,15 -0,66%	-0,0031 -0,73%	0,0080 1,77%

Tabela 16: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de 10 progênies entre as 52, baseada nos seus BLP. Teste de Progênies de *Eucalyptus grandis* aos sete anos.

CARÁTER SELECIONADO	RESPOSTA CORRELACIONADA NO CARÁTER			
	DAP	ALTURA	VOLUME	DENSIDADE
DAP	0,23 1,62%	0,26 1,14%	0,0172 3,97%	-0,0017 -0,37%
Altura	0,18 1,24%	0,32 1,44%	0,0170 3,92%	0,0018 0,40%
Volume	0,19 1,35%	0,26 1,15%	0,0188 4,34%	0,0010 0,22%
Densidade	-0,08 -0,54%	-0,12 -0,52%	-0,0061 -1,42%	0,0135 2,97%

Tabela 17: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de 10 progênies entre as 62, baseada nos seus BLUP. Teste de Progênies de *Eucalyptus grandis* aos sete anos.

CARÁTER SELECIONADO	RESPOSTA CORRELACIONADA NO CARÁTER			
	DAP	ALTURA	VOLUME	DENSIDADE
DAP	0,34 2,38%	0,89 3,95%	0,02 5,64%	-0,01 -2,69%
Altura	0,28 1,94%	1,16 5,17%	0,02 5,20%	0,00 0,53%
Volume	0,31 2,18%	0,98 4,35%	0,02 5,77%	0,00 -0,74%
Densidade	-0,11 -0,73%	-0,44 -1,94%	-0,01 -1,69%	0,04 9,14%

Tabela 18: Predições de valores genéticos aditivos para a seleção para um ambiente intermediário pelo método BLP, para os caracteres DAP, Altura, Volume e Densidade, para as progênies não comuns e testemunhas

PROGÊNIES NÃO COMUNS	DAP	ALTURA	VOLUME	DENSIDADE
I041	-0,0819	-0,144	-0,005	-0,0011
I047	-0,0135	-0,071	-0,0007	-0,0072
I060	0,00872	-0,0503	6,8E-05	-0,0063
I323	-0,0078	0,03418	0,00046	0,00074
I325	-0,2436	-0,1824	-0,0083	0,0009
I326	-0,0974	-0,1828	-0,0036	0,00236
I328	-0,1828	-0,5164	-0,0109	0,00388
S037	-0,0625	-0,1867	-0,0085	0,00152
S055	0,03892	0,09981	0,0025	0,0026
S066	0,05256	0,26354	0,00295	0,00023
S068	-0,027	0,10281	-0,0021	-0,0022
<u>Testemunhas</u>				
T_B	0,03533	-0,0773	-0,001	-0,0028
T_C	-0,1805	-0,4746	-0,0163	-0,0099

Tabela 19: Predições de valores genéticos aditivos para a seleção para um ambiente intermediário pelo método BLUP, para os caracteres DAP, Altura, Volume e Densidade, para as progênies não comuns e testemunhas

PROGÊNIES NÃO COMUNS	DAP	ALTURA	VOLUME	DENSIDADE
I041	-0,20	-0,69	-0,0070	-0,0174
I047	0,01	-0,27	0,0006	-0,0370
I060	-0,05	-0,45	-0,0042	-0,0368
I323	-0,07	-0,26	-0,0032	0,0084
I325	-0,24	-1,09	-0,0206	0,0082
I326	-0,16	-0,76	-0,0078	0,0092
I328	-0,30	-1,81	-0,0136	0,0156
S037	-0,14	-0,47	-0,0070	0,0040
S055	0,19	0,63	0,0132	0,0168
S066	0,19	1,32	0,0134	0,0014
S068	0,11	0,58	0,0070	-0,0012
<u>Testemunhas</u>				
T_B	0,27	0,00	0,0182	-0,0108
T_C	-0,36	-1,74	-0,0496	-0,0184

4.4. SELEÇÃO SEQUENCIAL VERSUS ÍNDICE DE SELEÇÃO

Inicialmente apresentaremos os índices de seleção obtidos e suas propriedades para duas formas de cálculo (componentes da variância fornecidas por MORAES, 1987 e estimativas para o modelo misto) e a seguir as comparações entre o índice empregado (com as estimativas de MORAES, 1987) e a seleção sequencial (SS).

4.4.1. Índices de seleção calculados

Os pesos calculados para os índices de seleção dos quatro caracteres pelo modelo misto e a partir das componentes da variância estimadas por MORAES (1987) são apresentados nas tabelas 20 e 21 , na página seguinte.

Como podemos observar, os pesos relativos a cada nível de observação funcionam como herdabilidades para a observação considerada e é comum que alguns autores representem índices-ótimos de seleção combinada (do tipo Smith-Hazel, para modelos fixos) com os pesos sendo as herdabilidades (e.g., FALCONER, 1981; NAMKOONG, 1979).

À diferença dos índices ótimos, nos quais os contrastes que compõem o índice são independentes, as diferenças $y-E(y)$ em cada nível de observação do índice proposto não são independentes entre si. Isto porque cada nível (média marginal de progênie, média de progênie por local, média de parcela, e valor fenotípico individual) têm uma componente aleatória, e todas estas componentes são correlacionadas entre si para as árvores de uma mesma progênie. Um reflexo disto é a presença de pesos negativos e muito pouco expressivos para médias de progênie por local e médias de parcela (b_2 e b_3 , respectivamente) ao se usar as componentes da variância dadas por MORAES (1987).

Tabela 20: Pesos calculados para o índice de seleção adotado, segundo os caracteres e níveis de observação; variância das predições (valor real e como porcentagem da variância aditiva) e correlação entre valor genético e suas predições: a partir das componentes da variância apresentadas por MORAES (1987). Teste de progênies de *E. grandis*, aos sete anos.

CARÁTER	PARÂMETRO	b ₁	b ₂	b ₃	b ₄
DAP	C'v ⁻¹	0,33793	-0,0087	-0,00504	0,03121
	% de b1	100,00%	-2,57%	-1,49%	9,24%
	$V(\hat{g}) = 0,056$	% da $\sigma_A^2 = 11,45\%$		$r(\hat{g}, g) = 33,84\%$	
ALTURA	C'v ⁻¹	0,44226	-0,01481	-0,02304	0,06567
	% de b1	100,00%	-3,35%	-5,21%	14,85%
	$V(\hat{g}) = 0,18446$	% da $\sigma_A^2 = 16,84\%$		$r(\hat{g}, g) = 41,04\%$	
VOLUME	C'v ⁻¹	0,27374	-0,00371	-1,6E-06	0,0165
	% de b1	100,00%	-1,36%	0,00%	6,03%
	$V(\hat{g}) = 0,00016$	% da $\sigma_A^2 = 8,62\%$		$r(\hat{g}, g) = 29,36\%$	
DENSIDADE	C'v ⁻¹	0,68971	-0,09727	-0,2571	0,54975
	% de b1	100,00%	-14,10%	-37,28%	79,71%
	$V(\hat{g}) = 0,0005$	% da $\sigma_A^2 = 61,76\%$		$r(\hat{g}, g) = 7859\%$	

Tabela 21: Pesos calculados para um índice de seleção alternativo, segundo os caracteres e níveis de observação, variância das predições (valor real e como porcentagem da variância aditiva) e correlação entre valor genético e suas predições: a partir das componentes da variância estimadas sob o modelo misto. Teste de progênies de *E. grandis*, sete anos.

CARÁTER	PARÂMETRO	b ₁	b ₂	b ₃	b ₄
DAP	C'v ⁻¹	0,32029	-0,00175	-0,0117	0,02954
	% de b1	100,00%	-0,54%	-3,65%	9,22%
	$V(\hat{g}) = 0,05627$	% da $\sigma_A^2 = 10,80\%$		$r(\hat{g}, g) = 32,86\%$	
ALTURA	C'v ⁻¹	0,51508	0,04431	-0,07228	0,1306
	% de b1	100,00%	8,60%	-14,03%	25,36%
	$V(\hat{g}) = 0,57219$	% da $\sigma_A^2 = 25,28\%$		$r(\hat{g}, g) = 52,28\%$	
VOLUME	C'v ⁻¹	0,3603	0,00314	-0,02805	0,02561
	% de b1	100,00%	0,87%	-7,79%	7,11%
	$V(\hat{g}) = 0,00031$	% da $\sigma_A^2 = 11,04\%$		$r(\hat{g}, g) = 33,23\%$	
DENSIDADE	C'v ⁻¹	0,48353	-0,07388	-0,09468	0,20192
	% de b1	100,00%	-15,28%	-19,58%	41,76%
	$V(\hat{g}) = 8,5E-05$	% da $\sigma_A^2 = 27,57\%$		$r(\hat{g}, g) = 52,51\%$	

Para as estimativas de componentes da variância do modelo misto os caracteres altura e volume apresentaram inversão na tendência dos pesos b_2 (foram positivos, mas sempre muito próximos de zero) e b_3 (foram maiores que os do índice empregado).

A seleção sequencial não envolve predição de valores aditivos, mas pode-se pensar numa analogia de predição ao se considerar as intensidades de seleção entre e dentro de progênies, (que são no entanto aplicadas independentemente). Isto permite antecipar uma maior probabilidade de selecionar maior número de progênies quando a relação b_4/b_1 for maior que a relação entre as respectivas intensidades de seleção.

Com as intensidades consideradas da SS, isto se daria apenas com $(b_4/b_1) > [(1/10):(14/52)] = 37,14\%$. Apenas o caractere densidade básica da madeira apresenta um índice com tal relação (qualquer que seja o conjunto das c.v. considerado), ou seja, apenas para este caractere espera-se que o índice aumente o número de progênies ao invés do número de árvores nas melhores progênies.

O índice mostrou-se bastante sensível em seus pesos quanto às alterações nas estimativas das componentes da variância. Uma comparação entre os pesos calculados para os dois conjuntos de estimativas destas componentes revelou uma tendência de dar maior importância relativa à seleção de progênies do que dentro de progênies pelo índice empregado de fato, exceto para o caráter densidade. O índice com estimativas do modelo misto valoriza mais os pesos dos efeitos intermediários (médias de progênies por local e de parcelas) em seus valores absolutos. O que também ocorreu com as próprias estimativas das componentes, no item 4.3. (maiores componentes para a variância do erro entre parcelas).

O índice com as estimativas de modelos mistos apresentou melhores propriedades, com valores para $V(\hat{g})$ maiores para os

caracteres de crescimento e menores apenas para a densidade. Este resultado está associado a maiores acurácias seletivas pelo emprego do índice com tais estimativas, ou seja, novamente, as estimativas do modelo misto parecem se adequar melhor à situação de seleção. Para o teste de progênies analisado isto significaria incluir ainda mais árvores das progênies melhores ao invés de ampliar o número de famílias incluídas na seleção, acarretando ainda maior restrição do tamanho efetivo populacional.

Tanto os índices quanto a seleção sequencial sofrem o problema de serem critérios únicos que não levam em conta desbalanceamentos. O número de desbalanceamentos foi em geral concentrado no nível de parcelas, porém foi mascarado pela adoção de parcelas grandes com árvores-bordadura também avaliadas em caso de falhas.

Diversos autores enfatizam que BLP e BLUP são superiores ao índice por ajustarem um índice a cada árvore (ou parcela), de forma a combater eventualmente todos os tipos de desbalanceamentos. Isto evidencia uma fraqueza do índice, que só é BLP no caso de homogeneidade das variâncias e do número de árvores por progênie, além de boas estimativas dos efeitos fixos.

Quanto à seleção sequencial, é uma forma ainda menos robusta que o índice, pois ignora completamente o desbalanceamento, seja no nível de progênies (ordenamento por valores decrescentes de médias gerais de progênies, como discutido para o BV apresentado nos itens 4.1 a 4.3) seja na seleção dentro (seleção da melhor árvore de cada parcela, ou mesmo no caso da seleção fenotípica das melhores árvores das progênies selecionadas).

Como ressaltado na discussão teórica do BLP, também para o índice um fenótipo muito superior ao seu valor de bloco evidencia as boas árvores. Tal fato não é levado em conta na

seleção sequencial, nem mesmo em índices ótimos de modelo fixo (e.g. BUENO FILHO, 1992).

Nestes procedimentos está presente ou implícito o contraste entre a árvore e sua média de parcela como avaliador do resíduo aleatório sobre o qual a herdabilidade incide. Nos índices do contexto de modelos mistos, tal resíduo é maior $[Y_{ijkl} - E(Y_{ijkl})]$, em que $E(Y_{ijkl})$ é dada pela estimativa do BLUE de bloco (ou mais simplesmente, pela média de bloco).

Este resíduo estará correlacionado ao resíduo aleatório que a própria parcela apresentar frente a sua esperança (também a média de bloco) e também ao resíduo que a média da progênie no local apresentar de sua esperança (média do local). Todas estas ponderações revelam a superioridade genética dos índices de modelos mistos não só sobre a SS, mas também sobre os índices do contexto de modelos fixos, em que a presença de tais covariâncias é simplificada.

4.4.2. Árvores selecionadas através do índice e da seleção sequencial.

As tabelas 22 a 25 das páginas 80 a 82 apresentam o número de árvores selecionadas de cada progênie, bem como o tamanho efetivo populacional resultante de cada método de seleção (a seleção sequencial e o índice com as componentes da variância apresentadas por MORAES, 1987).

Para a seleção sequencial de 14 progênies das 52, com uma árvore por parcela em cada progênie selecionada, as árvores selecionadas em comum através dos dois métodos foram em número de 39 (em 127) para o DAP, 26 (em 135) para a altura, 23 (em 125) para o volume e 56 (em 120) para a densidade.

Tabela 22: Progênie e número de árvores por progênie na seleção sequencial e pelo índice; cálculo dos tamanhos efetivos populacionais resultantes de cada método de seleção para o caráter DAP. Teste de Progênie de *Eucalyptus grandis*, aos sete anos.

Parâmetro	Seleção Sequencial		Índice	
	Progênie	árvores	Progênie	árvores
1	M044	10	S057	54
2	S005	10	S012	30
3	M002	9	S074	24
4	M005	9	S023	13
5	M013	9	M013	4
6	M017	9	M017	1
7	S008	9	S042	1
8	S012	9		
9	S023	9		
10	S042	9		
11	S057	9		
12	S074	9		
13	S079	9		
14	S078	8		
N_1 :		127		127
$\bar{k}$:		9,07		18,14
S_k^2 :		0,23		379,14
$S_k^2/\bar{k}$:		0,0248		20,8976
N_e :		41,99		12,08

Tabela 23: Progênie e número de árvores por progênie na seleção sequencial e pelo índice; cálculo dos tamanhos efetivos populacionais resultantes de cada método de seleção para o caráter Altura. Teste de Progênie de *Eucalyptus grandis*, aos sete anos.

Parâmetro	Seleção Sequencial		Índice	
	Progênie	árvores	Progênie	árvores
1	S023	18	S074	59
2	M002	9	S012	58
3	M005	9	S057	16
4	M013	9	M013	1
5	M017	9	S008	1
6	M024	9		
7	M044	9		
8	S012	9		
9	S024	9		
10	S042	9		
11	S057	9		
12	S068	9		
13	S074	9		
14	S079	9		
N_1 :		135		135
$\bar{k}$:		9,64		27,00
S_k^2 :		5,79		864,50
$S_k^2/\bar{k}$:		0,6000		32,0185
N_e :		40,78		8,71

Tabela 24: Progênie e número de árvores por progênie na seleção sequencial e pelo índice; cálculo dos tamanhos efetivos populacionais resultantes de cada método de seleção para o caráter Volume. Teste de Progênie de *Eucalyptus grandis*, aos sete anos.

Parâmetro	Seleção Sequencial		Índice	
	Progênie	árvores	Progênie	árvores
1	M002	9	S057	62
2	M005	9	S012	58
3	M013	9	M013	1
4	M017	9	M017	1
5	M044	9	S042	1
6	S012	9	S045	1
7	S023	9	S074	1
8	S042	9		
9	S045	9		
10	S057	9		
11	S068	9		
12	S074	9		
13	S079	9		
14	S056	8		
N_1 :		125		125
$\bar{k}$:		8,93		17,86
S_k^2 :		0,077		830,14
$S_k^2/\bar{k}$:		0,0080		46,4880
N_e :		41,89		7,42

Tabela 25: Progênie e número de árvores por progênie na seleção sequencial e pelo índice; cálculo dos tamanhos efetivos populacionais resultantes de cada método de seleção para o caráter Densidade. Teste de Progênie de *Eucalyptus grandis*, aos sete anos.

Parâmetro	Seleção Sequencial		Índice	
	Progênie	árvores	Progênie	árvores
1	I327	9	I327	16
2	M002	9	S040	10
3	M039	9	S053	10
4	M047	9	I060	6
5	M058	9	M039	6
6	S016	9	S045	6
7	S040	9	S058	5
8	S042	9	S038	4
9	S053	9	S042	4
10	S058	9	I326	3
11	S076	9	M002	3
12	S045	8	M024	3
13	S078	8	M047	3
14	S037	5	M058	3
15			S016	3
16			S061	3

Parâmetro	Seleção Sequencial		Índice	
	Progênie	árvores	Progênie	árvores
17			I328	2
18			S008	2
19			S029	2
20			S037	2
21			S052	2
22			S056	2
23			S062	2
24			S071	2
25			S076	2
26			S078	2
27			S080	2
28			I041	1
29			I324	1
30			M005	1
31			M017	1
32			M044	1
33			M048	1
34			S031	1
35			S039	1
36			S075	1
37			S079	1
<hr/>				
N_1 :		120		120
$\bar{k}$:		8,57		3,24
S_k^2 :		1,19		9,69
$s_i^2/\bar{k}$:		0,1385		2,9975
N_e :		40,99		52,00

A seleção sequencial (SS) produziu tamanhos efetivos finais em torno de 41 a 42 (em geral recomenda-se números maiores que quarenta para programas de seleção recorrente). Para a seleção através do índice os tamanhos efetivos resultantes foram: $Ne_{DAP}=12,08$; $Ne_{Altura}=8,71$; $Ne_{Volume}=7,42$ e $Ne_{Densidade}=52,00$.

Estes resultados concordam com as expectativas do item anterior: redução da base genética para os caracteres de crescimento e aumento para a densidade. Inesperada, no entanto, foi a forte intensidade da restrição do Ne para os primeiros, que decorre do fato de poucas árvores terem sido

selecionadas em comum pelos dois métodos, com a concentração das árvores selecionadas nas melhores famílias.

Quanto aos progressos genéticos, estão sumariados nas tabelas 26 e 27 a seguir (que incluem também as respostas correlacionadas nos demais caracteres).

O progresso foi sempre superior para o índice, com ganho adicional relativo de 40% para o caráter DAP, 56% para a altura, 55% para o volume e 26% para a densidade, quando comparado à seleção sequencial.

Tais resultados indicam que se pode esperar melhora considerável nos progressos ao substituir a seleção sequencial pelo índice. O aumento real no progresso ficou em torno de 1% para altura e DAP a 3% para o volume e a densidade.

Tabela 26: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção sequencial de árvores dentro das melhores progênies. Teste de Progênies de *Eucalyptus grandis* aos sete anos.

CARÁTER SELECIONADO	RESPOSTA CORRELACIONADA NO CARÁTER			
	DAP	ALTURA	VOLUME	DENSIDADE
DAP	0,45 3,16%	0,71 3,16%	0,0242 5,67%	0,0021 0,45%
Altura	0,43 2,97%	0,72 3,24%	0,0225 5,26%	0,0054 1,19%
Volume	0,46 3,21%	0,7 3,13%	0,0250 5,85%	0,0061 1,35%
Densidade	-0,06 -0,40%	-0,08 -0,35%	-0,0035 -0,81%	0,0473 10,44%

Tabela 27: Progressos genéticos pela seleção em cada um dos quatro caracteres em estudo e respostas correlacionadas nos demais caracteres: seleção de árvores baseada no índice de seleção. Teste de Progenies de *Eucalyptus grandis* aos sete anos.

CARÁTER SELECIONADO	RESPOSTA CORRELACIONADA NO CARÁTER			
	DAP	ALTURA	VOLUME	DENSIDADE
DAP	0,63 4,40%	0,98 4,40%	0,0365 8,54%	0,0052 1,14%
Altura	0,55 3,83%	1,12 5,00%	0,0322 7,52%	0,0027 0,59%
Volume	0,60 4,17%	0,98 4,40%	0,0387 9,05%	0,0037 0,81%
Densidade	-0,02 -0,17%	-0,09 -0,42%	-0,0020 -0,46%	0,0597 13,17%

4.5. Predição no Melhoramento Florestal

Neste capítulo será apresentada uma discussão conjunta dos resultados já apresentados numa visão mais aplicada às possibilidades do emprego das técnicas de predição de modelos mistos no melhoramento prático.

Discutiremos inicialmente as consequências dos diferentes métodos de predição na dispersão dos valores genéticos preditos; a seguir a seleção de matrizes e a seleção de árvores, com os métodos mais ajustados a cada uma; discutiremos ainda uma sugestão de estratégia de melhoramento por testes de que lance mão eficazmente dos métodos discutidos e finalmente os problemas para a adoção da abordagem de modelos mistos pelos melhoristas de plantas e florestais.

4.5.1. Caracteres em Questão

É comum os dados de melhoramento envolverem medidas de altura e DAP, das quais é produzida uma medida do volume cilíndrico. Medidas da densidade básica da madeira podem se tornar mais frequentes com o surgimento e o crescente barateamento de técnicas não destrutivas de mensuração. Por hora, não são frequentes.

As figuras 1 a 4 (pp. 86 e 87) mostram a dispersão das predições de valores genéticos de progênies produzidas pelos três métodos de seleção de matrizes. Pode-se observar que o BLUP foi sempre o método que apresentou maior dispersão das predições, seguido alternativamente pelo BV, (para DAP e altura) ou do BLP (para volume e densidade).

Figura 1: Dispersão das predições para o DAP pelos três métodos (BV, BLP e BLUP). Teste de progênie *E. grandis*, aos sete anos.

Figura 2: Dispersão das predições para a altura pelos três métodos (BV, BLP e BLUP). Teste de progênie *E. grandis*, aos sete anos.

Figura 3: Dispersão das predições para o volume pelos três métodos (BV, BLP e BLUP). Teste de progênes *E. grandis*, aos sete anos.

Figura 4: Dispersão das predições para a densidade básica da madeira pelos três métodos (BV, BLP e BLUP). Teste de progênes *E. grandis*, aos sete anos.

Por estes gráficos deve-se reconhecer a superioridade do BLUP em identificar materiais para a seleção, sendo no entanto questionável a superioridade do BLP sobre o BV para este quesito. Isto decorre das estimativas para a aplicação do BLP serem pouco adequadas, além da correção para a variância da interação (presente apenas no BLP) reduzir a herdabilidade efetiva do método, quando comparado ao BV. Tal redução, no entanto, não significa menor adequação.

É importante ressaltar que a explosão na variabilidade genética predita para a densidade básica da madeira deve ser vista com cautela, não se podendo assegurar a realização de ganhos tão expressivos como os preditos pelo BLUP (ou mesmo pelos outros métodos). Isto devido aos seguintes empecilhos:

- a) embora de alta herdabilidade o caráter resultou em uma distribuição de predições que leva a crer que a pressuposição de normalidade questionável e conseqüentemente as magnitudes das predições também (o ordenamento é mais confiável que os valores absolutos);
- b) As interações do tipo genótipos por ambientes são melhor medidas pelas componentes calculadas pelo modelo misto (BLUP), porém apenas o BLP estaria apto a lidar com tais estimativas de forma a produzir as devidas correções para efeitos de escala e alterações no ordenamento ao longo dos ambientes. O BLUP, como proposto, produziu correção apenas para desbalanceamento, com a componente da variância da interação comum, calculada para os três ambientes.

O valor isolado das predições não pode ser o critério de decisão entre os diferentes métodos. Para a comparação da adequação é necessário verificar a correlação entre valores genéticos e preditos (acurácia). Esta foi sempre maior tanto para o BLP quanto para o BLUP do que havia sido para o BV. Isto sim indica uma tendência: a superioridade dos métodos de modelos mistos sobre as médias marginais usuais em ordenar progênies de meios-irmãos para a seleção.

4.5.2. Seleção de Matrizes (pomar de primeira geração)

As figuras 5 a 8 (páginas 90 a 93) mostram comparações globais por caráter dos valores aditivos preditos pelos três métodos de seleção de progênies. Os valores estão ordenados por BV decrescente, simulando o ordenamento usual para a seleção (médias marginais de progênies).

Na interpretação destes gráficos deve-se observar as "quebras" de sentido nas linhas do BLP e BLUP: cada vez que a linha de um destes métodos cresce, ocorre uma inversão de ordenamento com relação ao BV. Supondo, como exemplo, que a seleção truncada para o caráter DAP seja representada por uma linha vertical na décima progênie. Traçando agora uma horizontal na altura do valor da ordenada da progênie de menor BLP (ou BLUP), todos os valores à direita do gráfico que estiverem acima desta ordenada representam progênies melhores que esta, ou seja: erros de classificação do BV.

Para o caso do BLP, isto ocorreria com 8 progênies para o dap, 4 para a altura, 4 para o volume e 4 para a densidade. Para o caso do BLUP, isto ocorreria com 3 progênies para o dap, 1 para a altura, 4 para o volume e 2 para a densidade.

Analizando o padrão de oscilação do BLP e do BLUP em torno dos BV, nota-se uma grande coincidência de cumes e vales ao longo das progênies, apesar do BLUP encontrar-se em um nível mais elevado. Isto indica a boa concordância dos dois métodos (que pode ser verificada através do ordenamento pelo BLUP decrescente, por exemplo). Indica também que os ganhos de seleção preditos para as mesmas progênies são muito maiores para o BLUP que para o BLP (ou BV), ou seja, que as predições de ganho podem estar subestimadas devido à presença de estimativas imprecisas das componentes da variância.

DAP

Figura 5: Predições de valores aditivos pelos três métodos de seleção de matrizes, para o DAP. Teste de progênies de E. grandis, aos sete anos.

Altura

Figura 6: Predições de valores aditivos pelos três métodos de seleção de matrizes, para a altura de árvores. Teste de progenies de *E. grandis*, aos sete anos.

Figura 7: Predições de valores genéticos aditivos pelos três métodos de seleção de matrizes, para o volume cilíndrico. Teste de Progenies, *E. grandis*, aos sete anos.

Figura 8: Predições de valores aditivos pelos três métodos de seleção de matrizes, para a densidade básica da madeira. Teste de progenies de *E. grandis*, aos sete anos.

O BLP como apresentado permite ganhos imediatos de concatenação de resultados de testes de progênie já realizados contendo material genético comum, especialmente na seleção para ambientes-alvo, de maneira mais simples que o BLUP, que depende de recursos computacionais mais avançados e do domínio de programas de predição-estimação.

Tanto o BLP quanto o BLUP permitem a utilização de todos os dados na determinação das predições, o que contrasta com as abordagens tradicionais, centradas nas progênie comuns. Um exemplo disto pode ser visto nas tabelas 18 e 19 (pág. 74), que trazem respectivamente os BLP e os BLUP de progênie para as progênie não comuns a todos os ambientes.

4.5.2. Seleção de Árvores (pomar de segunda geração)

As figuras 9 a 12 (pág. 95 e 96) mostram os progressos genéticos preditos na seleção para cada um dos caracteres, para todos os métodos estudados.

É interessante notar que o patamar de ganhos pela seleção de árvores é bem superior à seleção de matrizes enfocada anteriormente (exceto para os caráter altura pelo BLUP), ainda que se considere a seleção de matrizes mais intensa (10 progênie em 52 para matrizes comparada a 14 progênie em 52 para a SS, no caso analisado).

De um modo geral, a seleção para os caracteres de crescimento não teve reflexos na alteração da densidade, (exceto para a seleção de matrizes com o BLUP). Tal resultado remete a outros da literatura em que a correlação negativa entre o volume e a densidade básica da madeira foi considerada significativa. Neste caso, a construção de índices multivariados seria bastante complicada (e.g. MORAES, 1987; MAGNUSSEN, 1991). O oposto (fraca correlação negativa) foi encontrado por KLEIN (1995) para *Pinus banksiana* e, segundo o autor, é comum em coníferas.

Figura 9: Progresso genético na seleção para DAP e resposta correlacionada nos demais caracteres, para os cinco métodos de seleção/predição considerados. Teste de progênies de E. grandis, aos sete anos.

Figura 10: Progresso genético na seleção para altura e resposta correlacionada nos demais caracteres, para os cinco métodos de seleção/predição considerados. Teste de progênies de E. grandis, aos sete anos.

Figura 11: Progresso genético na seleção para volume e resposta correlacionada nos demais caracteres, para os cinco métodos de seleção/predição considerados. Teste de progênies de *E. grandis*, aos sete anos.

Figura 12: Progresso genético na seleção para densidade e resposta correlacionada nos demais caracteres, para os cinco métodos de seleção/predição considerados. Teste de progênies de *E. grandis*, aos sete anos.

O índice aparece assim como a melhor estratégia em termos de ganhos, sendo que as ressalvas para sua utilização são as que seguem:

- a) O índice não pondera para os desbalanceamentos, sendo necessário o emprego de BLP ou BLUP de árvores individuais para tanto;
- b) O índice é muito sensível a alterações nas componentes da variância, e as empregadas no presente caso são menos precisas e menos adequadas ao modelo de predição que as do modelo misto;
- c) O índice em questão não corrige para efeitos de escala devido aos diferentes níveis dos efeitos fixos (heterogeneidades nas componentes da variância) como o faz o BLP. Desta forma, o índice tende a selecionar árvores de ambientes melhores, mesmo que a variância fenotípica de tais ambientes seja relativamente maior;
- d) O índice em questão restringiu sobremaneira o tamanho efetivo final para os caracteres de crescimento. Esta tendência se agravaria com as componentes da variância do modelo misto;
- e) As estimativas de efeitos fixos que compõem o índice não são influenciadas pelas componentes da variância (estimativas OLSE), sendo portanto, imprecisas;
- f) A seleção de árvores apresenta menor correlação entre valor genético e predito que a seleção de progênies.

Portanto, o índice apresenta superioridade sobre a SS, aliada porém a um grande número de inconvenientes.

4.5.4. Perspectivas de Utilização

Comparando-se as propriedades estatísticas de cada um dos métodos através das tabelas 3 (p. 60), 14 (p. 71), 20 e 21 (p.76), conclui-se que a seleção de matrizes, embora se apresente menos responsiva que a seleção de árvores, apresenta maiores correlações entre valores preditos e estimados (maior acurácia seletiva) e, portanto, menores chances de se incorrer em erros de seleção.

Esta é a lógica que leva os melhoristas práticos a optarem pela SS, pois estes investem na segurança da seleção entre progênie, enquanto que a seleção dentro de parcelas traz um ganho adicional...

Isto é um engano, pois a seleção fenotípica dentro da progênie apresenta neste caso ainda menor acurácia: selecionar as filhas envolve menor precisão que selecionar a mãe!

Os resultados apresentados indicam que o ordenamento de progênie pode ganhar com a abordagem de modelos mistos através do BLP e do BLUP, enquanto que o ordenamento dentro de progênie pode ser melhor que a seleção fenotípica de uma árvore por parcela. Este último pode se dar pelo uso do índice apresentado ou pelo desenvolvimento de BLP e BLUP de árvores individuais.

Assim, pode-se sugerir uma mescla das estratégias avaliadas de forma a maximizar os ganhos por seleção e minimizar os efeitos da redução no tamanho efetivo para a seleção de árvores.

Já se disse que os melhoristas em geral procuram concentrar os testes de gerações avançadas em poucas progênie promissoras, o que gera conflitos de objetivos para o teste de progênie (entre estabelecer boas estimativas de componentes da variância e efetivamente melhorar a população).

Uma estratégia mesclada pode também maximizar as vantagens dos modelos mistos com uma criteriosa estimação de componentes da variância em modelos e delineamentos estatísticos que também sejam adequados à predição.

Para tanto, é preciso trocar a visão de estimação de efeitos fixos de tratamentos (progênes) e encarar o fato de que os valores genéticos de progênes são, intrinsecamente, variáveis aleatórias. Deve-se procurar ajustar o modelo aos efeitos biológicos presentes e obter respostas mais realistas. Neste sentido propõe-se os seguintes passos:

- a) O estabelecimento de testes com um número de progênes efetivamente grande, representativo de uma ampla classe de cruzamentos da população (da ordem de centenas, e não de dezenas de progênes, como é usual), de forma a obter estimativas mais precisas das componentes da variância. Esta indicação já é conhecida da genética quantitativa florestal, entre outro exemplos, RESENDE (1995) estudou delineamentos que maximizam a acurácia seletiva com conclusões semelhantes.
- b) Concatenar ensaios, de diferentes ambientes e idades com material comum, num único modelo de estimação-predição. Pode-se pensar em testes e na análise conjunta de informações de árvores com diferentes graus de parentesco presentes em dados de produção florestal (meias-irmãs, irmãs completas, irmãs de progênes autofecundadas e clones), para que amostras sucessivas das melhores progênes possam ser incorporadas à predição. SOH (1994) é um exemplo das vantagens do BLUP para a seleção entre materiais de parentesco complexo.
- c) O estabelecimento de ensaios que levem à estimação mais precisa dos efeitos fixos. Tais ensaios devem envolver parcelas de plantas individuais ou parcelas não contíguas e possivelmente a análise de vizinhança. A precisa estimação de efeitos fixos é condição para revelar os erros de ordenamento, o que ficou claro na análise dos BLUP versus látices por local (item 4.3). Deve-se também preferir

delineamentos como os em blocos incompletos por sua alta eficiência estatística, apenas que invertendo a lógica da análise dos látices (progênes aleatórias e blocos fixos). Isto é necessário se se quiser um corpo teórico robusto para a estimação-predição.

d) O procedimento de seleção no contexto experimental acima descrito pode se basear na seleção recorrente intrapopulacional de árvores, com a predição BLUP ou BLP de valores de progênes, seguida de uma predição de valores de árvores para as progênes superiores, como na seleção sequencial, apenas que com uma menor rigidez quanto a árvores selecionadas. Desta forma se preserva o tamanho efetivo final ou, alternativamente, se maximiza o ganho para um tamanho efetivo final mínimo que assegure a adequada continuidade da seleção.

e) Uma alternativa é a utilização da divergência de populações para objetivos de seleção distintos, que resulta na condução de duas ou mais populações pelo método descrito no item (d). Para uma tal estratégia, o mais indicado é utilizar o BLP de progênes para ambientes-alvo distintos, estratégia superior a ignorar a informação dos ambientes em que a progênie foi testada, mas para os quais não será avaliada a seleção.

A maior adequação e robustez estatística dos modelos mistos tem sido demonstrada nas últimas duas décadas por um crescendo de resultados de simulação e de comparações entre ganhos genéticos preditos e realizados na seleção de alguns organismos modelo (em especial para progênes de irmãos completos em ratos, coelhos, touros e outros animais, ver SORENSEN e KENEDY, 1984a, 1984b; BENIWAL et alii, 1995a, 1995b). Também têm se verificado as vantagens dos métodos de modelos mistos quanto à superioridade nos progressos esperados e à maior acurácia no melhoramento florestal (RESENDE et alii, 1995; KLEIN, 1995), ou mesmo no melhoramento de autógamias (PANTHER e ALLEN, 1995).

Os maiores empecilhos para o seu emprego como novo paradigma de seleção são de natureza bem diversa...

Um modelo explicativo de um fenômeno qualquer (não necessariamente biológico) é apenas uma abstração do que realmente ocorre e que permite intervir na realidade com consequências razoavelmente previsíveis. Sobre a relação entre modelos formalmente corretos e sua validade explicativa, Einstein² afirmou:

"(...) Os conceitos e proposições adquirem "sentido" ou "conteúdo" apenas através de suas conexões com as experiências sensoriais. A conexão desta últimas com os primeiros é puramente intuitiva, e não de natureza lógica em si mesma. (...)"

A aceitação da idéia por trás de um modelo não decorre de sua adequação formal e robustez interna, mas ao contrário: depende da experiência do observador.

No caso do melhoramento florestal, a aceitação de tais modelos, para além das demonstrações, depende do acúmulo de experiência dos pesquisadores e melhoristas no trato com os procedimentos e na crítica dos resultados de modelos mistos para de predição de valores genéticos aditivos.

² Einstein, Albert Notas Autobiográficas. 3ª ed. Ed. Nova Fronteira, 1982, p. 22.

5. CONCLUSÕES

- a) Os métodos usuais de ordenamento e seleção de progênies são casos especiais de BLP, construídos a partir de médias gerais de progênies. Em situações de desbalanceamento ou de heterogeneidade das variâncias, tais simplificações perdem sua precisão e valor explicativo;
- b) Apesar de se usar representar progênies de polinização livre por efeitos aleatórios, nos processos de estimação são consideradas como efeitos fixos e isto tem levado a erros de ordenamento para a seleção;
- c) Considerar efeitos ambientais como fixos na estimação é usual (exceto para a recuperação da informação interblocos dos látices) e acarreta vantagens para fins de seleção;
- d) A abordagem BLP foi superior à usual na seleção para um ambiente intermediário no caso de se analisar três ou mais ambientes, na presença de qualquer tipo de desbalanceamento ou heterogeneidade das variâncias fenotípicas;
- e) A abordagem BLP permite melhorar as predições para um ambiente específico através dos dados de progênies comuns de outro ambiente (seleção para ambiente-alvo);
- f) A abordagem BLUP acoplada ao método REML para estimação de componentes da variância no modelo misto, proporciona vantagens sobre o método usual e o BLP tanto em ganhos de seleção como na dispersão das predições e sua maior confiabilidade;

- g) A abordagem BLUP pode ser melhorada pela incorporação da heterocedasticidade das variâncias ao longo dos ambientes, como no procedimento BLP analisado;
- h) Na seleção de árvores, o índice de seleção propiciou predições de progressos genéticos maiores que a seleção sequencial ou qualquer outro método considerado;
- i) O uso das estimativas de componentes da variância do modelo misto na construção do índice tende a aumentar sua acurácia, mas pode agravar a restrição do N_e para os caracteres de crescimento;
- j) As predições de valores genéticos aditivos (e de progressos pela seleção) para o pomar de primeira geração foram as de maior confiabilidade;
- k) É possível mesclar as diferentes abordagens numa única estratégia de seleção, com as predições mais precisas do valor de progênies associadas a predições do valor genético aditivo de árvores dentro de progênies;
- l) As razões que levam os melhoristas práticos a não utilizar os modelos mistos de predição vão desde sua pouca divulgação até os resultados satisfatórios com a abordagem usual, embora esta seja menos precisa.
- m) As predições com modelos mistos são superiores às usuais como explicação dos fenômenos biológicos e seu emprego, em estudos e na prática do melhoramento florestal, tende a substituir o paradigma de tratar efeitos aleatórios como fixos nos processos de estimação (predição) e no ordenamento para a seleção em testes de progênies florestais.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLARD, R.W. *Princípios do Melhoramento Genético das Plantas*. São Paulo: Edgard Blucher, 1971. 381p.
- BENIWAL, B.K.; HASTINGS, I.M.; THOMPSON, R.; et al. Estimating of changes in genetic parameters in selected lines of mice using REML with na animal model. 1- Lean mass. *Heredity* v. 69, n.4, p.352-360, 1995a.
- BENIWAL, B.K.; HASTINGS, I.M.; THOMPSON, R.; et al. Estimating of changes in genetic parameters in selected lines of mice using REML with an animal model. 2-Body weight, body composition and litter size. *Heredity* v. 69, n.4, p.361-371, 1975b.
- BERNARDO, R. Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Science* v. 34, p.20-25. 1994.
- BERNARDO, R. Genetic models for predicting maize single-cross performance in unbalanced yeld trial data. *Crop Science* v. 35, p.141-147. 1995.
- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L.D. et al. A Manual for use of MTDFREML - A Set of Programs to Obtain Estimates of Variances and Covariances. USDA-ARS, 1993. ["Draft"] Revisado em abril de 1995.
- BRIDGWATER, F. Mating Designs. In: Fins, L.; Friedman, S.T.; Brotshol, J.V. (eds.) *Handbook of Quantitative Forest Genetics*. Norwell: Kluwer Acad. Publ., pp.69-95. 1992.

- BROWN, K.G. Asymptotic behavior of MINQUE-type estimators of variance components. *Ann. Stat.* V. 4, p.746-754. 1976.
- BUENO FILHO, J.S.S. *Seleção Combinada Versus Seleção Sequencial no Melhoramento de Populações Florestais*. Diss. Ms/ESALQ-USP. 1992. 96p.
- BURDON, R.D. Testing and selection: strategies and tactics for the future. *IUFRO-Conference, Breeding Tropical Trees*. Cali 1992. 10p.
- COTTEERRIL, P.P. Genetic gains expected from alternative strategies including simple low cost options. *Silvae Genetica*, v. 33, p.84-90. 1986a.
- COTTEERRIL, P.P. Breeding strategy: don't underestimate simplicity. In: *IUFRO-Joint Meet. of Working Parties on Breeding Theory*. Williamsburg, 1986b. pp 8-23.
- COTTEERRIL, P.P.; JACKSON, N. On index selection. I. Method of determining economic weights. *Silvae Genetica*, v. 34, p.56-63. 1985.
- COCHRAN, W.G. Improvement by means of selection. *Proc. 2nd Berkeley Symp. Of Math. Statistics and Production*. 1951. pp. 449-470.
- COCHRAN, W.G.; COX, G. *Diseños Experimentales*. 7^a Ed., Editorial Trillas, Mexico. 1981. 661p.
- FALCONER, D.S. *Introduction to Quantitative Genetics*. 2nd Edition. Longman Group Limited, New York. 1981, 340p.
- FEDERER, W.T. *Experimental Design: theory and application*. Macmillan Eds., New York. 1955. 591p.
- FOSTER, G.S. Estimating yield: beyond breeding values. In: Fins, L.; Friedman, S.T.; Brotshol, J.V. (eds.) *Handbook of Quantitative Forest Genetics*. Norwell: Kluwer Acad. Publ., pp.229-269. 1992.
- GIANOLLA, D. A method of sire evaluation for dichotomies. *J. An. Sci.* V. 51, n.6, p.1266-1276. 1980.
- GOMES, F.P. *Curso de Estatística Experimental*. Nobel, 1987. 446p.

- GRAYBILL, F.A. *Theory and Applications of the Linear Model*. Boston: Buxbury Ed., 1976.
- HARTLEY, H.O.; RAO, J.N.K. Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika*, v. 54, p.93-108. 1967.
- HELGADOTTIR, Á.; BJÖRNSSON, H.; KRISTJÁNSDOTTIR, T.A. Analysis of a site x year experiment with timothy polycross progeny. *Euphytica*, v. 82, p.241-251. 1995.
- HENDERSON, C.R. Estimation of variance and covariance components. *Biometrics*, v.9, p.226-252. 1953.
- HENDERSON, C.R. Selection Index and expected genetic advance. In: Hanson, W.D.; Robinson, H.F. (eds.) *Statistical Genetics and Plant Breeding*. NAS-NRC. pp. 141-163. 1963.
- HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, v.31, p.423-447. 1975.
- HENDERSON, C.R. A simple method of computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics*, v. 32, p.69-83. 1976.
- HENDERSON, C.R. *Applications of Linear Models in Animal Breeding*. University of Guelph, Guelph, 1984.
- HENDERSON, C.R. Progress in statistical methods applied to quantitative genetics since 1976. In: WEIR, B. et al. *Proceedings of the Second International Conference on Quantitative Genetics*. pp.85-90. 1988.
- HENDERSON, C.R.; KEMPTHORNE, O.; SEARLE, S.; et al. Estimation of environmental and genetic trends from records subject to culling. *Biometrics*, v.15, p.192-218. 1959.
- HOCKING, R.R. *The Analysis of Linear Models*. Monterey: Brooks/Cole. 1985.

- HODGE, G.R.; WHITE, T.L. Concepts of gain prediction. In: Fins, L.; Friedman, S.T.; Brotshol, J.V. (eds.) *Handbook of Quantitative Forest Genetics*. Norwell: Kluwer Acad. Publ., 1992.
- KAGEYAMA, P.Y. *Variação genética em progênes de uma população de Eucalyptus grandis* (Hill) Maiden. (tese Dr. ESALQ-USP), Piracicaba, 1980. 125p.
- KANG, H. Components of juvenile-mature correlations in forest trees. *Theoretical and Applied Genetics*, v. 81, p.173-184. 1991.
- KANOWSKI, P.J. Forest genetics and tree breeding. *Plant Breeding Abstracts*, v. 63, n.6, p.717-726. 1993.
- KEMPTHORNE, O. *An Introduction to Genetic Statistics*. New York: John Willey. 1957.
- KEMPTHORNE, O. Fixed and mixed models in analysis of variance. *Biometrics* v. 31, p.473-486. 1975.
- KEMPTHORNE, O. An overview of the field of quantitative genetics. In: WEIR, B. et al. *Proceedings of the Second International Conference on Quantitative Genetics*. pp. 47-56, 1988.
- KENEDY, B.W.; SORENSEN, D.A. Properties of Mixed-model methods for prediction of genetic merit In: WEIR, B. et al. *Proceedings of the Second International Conference on Quantitative Genetics*. pp. 91-103, 1988.
- KLEIN, J.I. Multiple-trait combined selection in jack pine family-test plantations using best linear prediction. *Silvae Genetica*, v. 44, n.5-6, p.362-375. 1995.
- LANDE, R. Marker assisted selection in relation to traditional methods of plant breeding. In: STALKER, H.T.; MURPHY, J.P. *Plant Breeding in The 1990's*. pp.437-460. 1992.
- LINDGREEN, D. How should breeders respond to breeding values? In: *IUFRO-Joint Meet. of Working Parties on Breeding Theory*. Williamsburg, 1986. pp. 361-372.

- LOO-DINKINS, J. Field test design. In: Fins, L.; Friedman, S.T.; Brotshol, J.V. (eds.) *Handbook of Quantitative Forest Genetics*. Norwell: Kluwer Acad. Publ., pp. 96-139. 1992.
- LOWE, W.J.; Van BUIJTENEN, J.P. The development of a sublining system in an operational tree improvement program. In: *IUFRO-Joint Meet. of Working Parties on Breeding Theory*. Williamsburg, 1986. pp.98-106.
- MAGNUSSEN, S.; YEATMAN, C.W. Predictions of genetic gain from various selection methods in open pollinated *Pinus banksiana* progeny trials. *Silvae Genetica*, v. 39, p.140-1533, 1990.
- MAGNUSSEN, S. Index selection with nonlinear profit function as a tool to achieve simultaneous genetic gain. *Theoretical and Applied Genetics*, v. 79, p.89-96. 1991.
- MARTINS, E.N. Uso de modelos mistos no melhoramento animal. *Simpósio Internacional de Produção de Não Ruminantes*. (anais) Maringá, 1993. pp. 25-51.
- MARTINS, E.N.; LOPES, P.S.; SILVA, M.A.; REGAZZI, A.J. *Modelo Linear Misto*. UFV - Imprensa Universitária, Viçosa. 1993, 46p.
- MORAES, M.L.T. de *Variação Genética da Densidade Básica da Madeira em Progenies de Eucalyptus grandis HILL ex Maiden e suas relações com as características de crescimento*. (Diss. Mestrado - ESALQ-USP) 1987. 115p.
- NAMKOONG, G. *Introduction to Quantitative Genetics in Forestry*. USDA Tec. Bull. n° 1588. 1979, 342p.
- NAMKOONG, G.; KANG, H. Quantitative genetics of forest trees. *Plant Breeding Abstracts*, v.8, p.139-188. 1990.
- PACHECO, C.A.P., RIBEIRO JUNIOR, J.I.; CRUZ, C.D. Aplicação do melhor preditor linear não-viesado (BLUP) no método de seleção entre e dentro de progênies de meios-irmãos na população de milho CMS-39. (resumo) *Congresso Nacional de Milho e Sorgo*, Londrina, 1996. p.90.

- PALOMARES, G.; SANJUAN, B. BLUP evaluation of *Phaseolus vulgaris* individual subjected to a selection program. *Bean Improvement Coop.* v.39, p.235-236. 1996.
- PANTHER, D.M.; ALLEN, F.L. Using best linear predictions to enhance breeding for yield in soybean: I. Choosing parents. *Crop Science*, v.35, p.397-405. 1995.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. *Biometrika*, v. 58, p.545-554. 1971.
- POLLAK, E.J.; QUAAS, R.L. Monte Carlo study of within-herd MT evaluation of beef cattle growth traits. *J. An. Sci.* v.52, n.2, p.248-256. 1981a.
- POLLAK, E.J.; QUAAS, R.L. Monte Carlo study of genetic evaluations using sequentially selected records. *J. An. Sci.* v.52, n.2, p.257-264. 1981b.
- QUAAS, R.L.; POLLAK, E.J. Mixed model methodology for farm and ranch beef cattle testing programs. *J. An. Sci.* v. 51, n.6, p.1277-1286. 1980.
- RAO, C.R. *Linear Statistical Inference and Its Applications*. New York: John Willey. 1965. 625p.
- RESENDE, M.D.V.de Melhoria genética de essências florestais. *Simpósio de Atualização em Genética e Melhoramento de Plantas*. Lavras, 1997 (no prelo).
- RESENDE, M.D.V.de; ARAÚJO, A.J.; SAMPAIO, P.T.B.; et al. Acurácia seletiva, intervalos de confiança e variâncias de ganhos genéticos associados a 22 métodos de seleção em *Pinus caribea* var. hondurensis. *Revista Floresta*, v.25, p.25-38, 1995.
- RESENDE, M.D.V. de; PRATES, D.F.; YAMADA, C.K.; et al. Estimção de componentes de variância e predição de valores genéticos pelo método da máxima verossimilhança restrita (REML) e melhor predição linear não viciada

- (BLUP) em *Pinus*. *Boletim de Pesquisa Florestal*, Colombo, n.32-33, p.23-42. 1996.
- RESENDE, M.D.V. de; HIGA, A.R.; LAVORANTI, O.J. Predição de valores genéticos no melhoramento de *Eucalyptus* - Melhor Predição Linear. *1º Congresso Florestal Pan-Americano* (anais), p. 144-147, Curitiba, 1993.
- RESENDE, M.D.V.de; HIGA, A.R. Estimação de valores genéticos no melhoramento de *Eucalyptus*: seleção em um caráter com base em informações do indivíduo e seus parentes. *Boletim de Pesquisa Florestal*, Colombo, n.28-29, p.11-36. 1994.
- RESENDE, M.D.V.de; BERTOLUCCI, F. de L.G. Maximization of genetic gain with restriction of effective population size and inbreeding in *Eucalyptus grandis*. IUFRO Meeting, Hobart, 1995. pp.167-170.
- RESENDE, G.D.S.P.; BERTOLUCCI, F. de L.G. Uso da seleção combinada na determinação da eficiência da seleção precoce em progênies de meios-irmãos de *Eucalyptus urophylla*. *1º Congresso Florestal Pan-Americano* (anais), p. 158-160, Curitiba, 1993.
- ROBINSON, J. The distribution of a general quadratic form in normal variables. *Austr. Journal of Statistics*, v.7, p.110-114. 1965.
- SAS *SAS Technical Report: SAS/STAT Software: Changes and Enhancements, Release 6.07. (Chapter 16 - The MIXED Procedure)*. SAS Institute, Cary, 1992. pp. 323-366.
- LATOUR, D.; LATOUR, K.; WOLFINGER, R. *Getting Started with Proc Mixed*. SAS Institute, Cary. 1994. 121p.
- SCHEFFÉ, H. *The Analysis of Variance*. New York: John Willey, 1959. 477p.
- SCHWARZ, C. The mixed model ANOVA: the truth, the computer package, the books. *Journal of the Am. Stat. Ass.* v.47, n.1, p.48-59. 1993.

- SEARLE, S. Topics on variance components estimation. *Biometrics*, v.27, p.1-76. 1971.
- SEARLE, S. *Linear Models for Unbalanced Data*. New York: John Willey, 1987.
- SEARLE, S.; CASELLA, G.; McCULLOCH, C.E. *Variance Components*. New York: John Willey, 1992. 501p.
- SILVA, R.G.da *Métodos de Genética Quantitativa Aplicados ao Melhoramento Animal*. Rib. Preto: Ed. Soc. Bras. Genetica, 1982. 162p.
- SILVA, M.V.G.B.da *Utilização de Modelos Animais Uni e Bivariados no Estudo das Relações entre Eficiência Reprodutiva e produção de Leite na Raça Holandesa*. (Dissertação de Mestrado - UFMG - Escola de Veterinária). 1995.
- SMITH, S.P.; GRASER, H.V. Estimating variance components in a class of mixed models by restricted maximum likelihood. *J. Dairy Sci.*, v.69, n.4, p.1156-1165. 1986.
- SOH, A.C. Ranking parents by best linear unbiased prediction (BLUP) breeding values in oil palm. *Euphytica*, v. 76, n.13-21. 1994.
- SORENSEN, D.A.; KENEDY, B.W. Estimation of genetic variances from unselected and selected populations. *J. An. Sci.* v. 59, n.5, p.1213-1223. 1984a.
- SORENSEN, D.A.; KENEDY, B.W. Estimation of response to selection using LS & MM Methodology. *J. An. Sci.* v. 58, n.5, p.1097-1106. 1984b.
- TALBERT, J. Quantitative genetics: why bother? In: Fins, L.; Friedman, S.T.; Brotshol, J.V. (eds.) *Handbook of Quantitative Forest Genetics*. Norwell: Kluwer Acad. Publ., pp. 1-28. 1992.
- VALERIO FILHO, W.V. *Comparação de Métodos para Estimção de Componentes da Variância Através de Simulação de Dados*. Piracicaba, 1991. (Tese - ESALQ/USP). 160p.

- VAN BUIJTENEN, J.P. Quantitative Genetics in Forestry. In: WEIR, B. et al. *Proceedings of the Second International Conference on Quantitative Genetics*. 1988.
- VAN VLECK, L.D. *Selection Index and Introduction to Mixed Models Methods*. Boca Raton: CRC Press, 1993, 479p.
- VENCOVSKY, R.; BARRIGA, P. *Genética Biométrica no Fitomelhoramento*. Rio Preto: Ed. Soc. Bras. Gen., 1992.
- VENCOVSKY, R. Extensão do conceito de tamanho efetivo populacional a populações submetidas à seleção. I- Espécies monóicas. Piracicaba: *Relatório Científico do Dep. Gen Esalq*. v.10, p.223-228. 1976.
- VENCOVSKY, R. Effective size of monoecious populations submitted to artificial selection. *Revista Brasileira de Genética*, v.1, n.3, p.181-191. 1978.
- VERNEQUE, R. da S. *Procedimentos Numéricos de Estimação de Componentes de Covariância em Análise Multivariada Pelo Método da Máxima Verossimilhança Restrita - Modelos Mistos Aplicados ao Melhoramento Animal*. Piracicaba, 1994. (Tese - DS - ESALQ/USP). 157p.
- VIEIRA, S.; HOFFMAN, R. *Estatística Experimental*. São Paulo: Ed. Atlas, 1988, 179p.
- WHITE, T.L.; HODGE, G.R. *Predicting Breeding Values with Applications in Forest Tree Improvement*. Dordrecht: Kluwer Acad. Publ., 1989. 367p.
- WHITE, T.; HODGE, G.; G.L. POWELL An advanced generation tree improvement plan for slash pine in the Southeastern United States. *Silvae Genetica*, v. 42, n.6, p.359-371. 1993.
- WOLFINGER, R. *A Tutorial on Mixed Models*. Cary: SAS Inst., 1993.
- ZHU, J.; WEIR, B. Diallel Annalysis of sex-linked and maternal effects. *Theorethical and Applied Genetics*, v. 92, p.1-9. 1996.

APÊNDICE 1: Numeração dos tratamentos correspondentes às progênies para os três locais de plantio dos ensaios. Teste de progênies de *Eucalyptus grandis*, aos sete anos. (MORAES, 1987).

Material Genético	LOCAIS		
	Anhembi	Brotas	Lençóis Paulista
Progênies comuns			
I324	54	58	58
I327	56	60	60
M002	39	43	41
M005	40	44	42
M013	41	45	43
M017	42 e 43	46 e 47	44 e 45
M024	44	48	46
M039	45	49	47
M044	46	50	48
M046	47	51	49
M047	48	52	50
M048	49	53	51
M055	50	54	52
M057	51	55	53
M058	52	56	54
S002	1	1	1
S003	2	2	2
S004	3	3	3
S005	4	4	4
S006	5	5	5
S008	6	6	6
S012	20	24	22
S016	7	7	7
S017	9	9	9
S019	21	25	23
S023	9	9	9
S024	23	27	25
S028	10	10	10
S029	11	11	11
S031	24	28	26
S034	25	29	27
S036	12	12	12
S038	26	30	28
S039	27	31	29
S040	13	14	14
S042	28	32	30
S045	14	15	15
S051	29	33	31
S052	30	34	32
S053	15	16	16
S056	16	18	17
S057	17	19	18
S058	31	35	33
S061	18	20	19
S062	32	36	34
S071	33	37	35
S074	19	23	31
S075	34	38	36
S076	35	39	37
S078	36	40	38
S079	37	41	39
S080	38	42	40

Material Genético	LOCAIS		
	Anhembi	Brotas	Lençóis Paulista
Progênies não comuns			
I041	-	-	55
I047	53	-	-
I060	-	57	56
I323	-	-	57
I325	55	-	-
I326	-	59	59
I328	-	61	61
S037	-	13	13
S055	-	17	-
S066	-	21	-
S068	-	22	20
Testemunhas			
T _B	57 a 60	62 e 64	62 e 64
T _C	61 a 64	63	63

Maiores detalhes quanto à origem do material, características dos locais de ensaio, um croqui da área experimental e da casualização do látice, podem ser encontrados em MORAES (1987).

APÊNDICE 2: Estimativas de componentes da variância, herdabilidades para a seleção entre médias fenotípicas e médias por locais das 52 progênies comuns. Teste de progênies de *Eucalyptus grandis*, aos sete anos. (MORAES, 1987).

Caráter	Parâmetro	Local			Geral
		Anhembi	Brotas	Lençóis Paulista	
DAP	σ_{pE}^2				0,2326
	σ_p^2	0,2154	0,2829	0,5286	0,1223
	σ_e^2	0,0000	1,3188	0,1015	0,2263
	σ_d^2	17,6400	8,9832	8,7034	11,7543
	Relação de Shelbourne:	$(\sigma_{pE}^2 / \sigma_p^2) =$		1,90	
	h_V^2	26,81%	27,68%	62,00%	34,39%
	média	16,64	14,32	12,33	14,43
Altura	σ_{pE}^2				0,3421
	σ_p^2	0,6314	0,5329	0,7377	0,2739
	σ_e^2	0,0000	1,8004	0,7251	0,6761
	σ_d^2	27,4629	5,7381	4,4312	12,5123
	Relação de Shelbourne:	$(\sigma_{pE}^2 / \sigma_p^2) =$		1,25	
	h_V^2	40,82%	40,24%	65,45%	45,49%
	média	25,81	21,57	20,13	22,50
Volume	σ_{pE}^2				0,000818
	σ_p^2	0,000965	0,001301	0,001334	0,000465
	σ_e^2	0,000000	0,006628	0,000000	0,000000
	σ_d^2	0,185068	0,037047	0,031789	0,084565
	Relação de Shelbourne:	$(\sigma_{pE}^2 / \sigma_p^2) =$		1,76	
	h_V^2	13,53%	27,42%	55,73%	27,72%
	média	0,656	0,386	0,269	0,437
Densidade	σ_{pE}^2				0,000034
	σ_p^2	0,000208	0,000213	0,000259	0,000203
	σ_e^2	0,000112	0,000161	0,000020	0,000097
	σ_d^2	0,000152	0,001082	0,000718	0,001108
	Relação de Shelbourne:	$(\sigma_{pE}^2 / \sigma_p^2) =$		0,17	
	h_V^2	83,07%	70,36%	89,43%	85,50%
	média	0,420	0,470	0,470	0,453

APÊNDICE 3: Principais comandos e rotinas do SAS empregados na análise

/ *DADOS.SAS: cria e ordena o arquivo de dados */

/ *especificações de opções de saída */

options ps=500 ls=100 nocenter;

/ *criação do arquivo de dados */

/ *variáveis **DAP**, Altura de árvores, Volume cilíndrico e **Densidade básica da madeira** */

data arquivo;

input L\$ B\$ T P\$ R A DAP H V D;

cards;

An X1	1	S002	1	1	16.80	28.00	0.620678	0.408
-------	---	------	---	---	-------	-------	----------	-------

.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

Br Z8	18	S057	3	1	13.05	21.30	0.284899	0.482
-------	----	------	---	---	-------	-------	----------	-------

.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

LP Z8	63	T_C	3	10	11.45	18.40	0.189460	0.453
-------	----	-----	---	----	-------	-------	----------	-------

;

run;

/ *ordenamento do arquivo de dados */

proc sort;

by P T L R B;

run;

/ *EXEMPLO DE ANÁLISE DO MODELO FIXO */

/ *quadrados mínimos ordinários para estimar */

/ *médias de blocos usadas no BLP */

/ *caráter DAP*/

proc glm data=arquivo;

class P T L R B;

model DAP = L B(L) / p;

run;

```

/*EXEMPLO DE ANÁLISE CONJUNTA DOS LÁTICES */
/* quadrados mínimos generalizados para estimar */
/* médias de ambientes e progênies e predição BLUP de blocos */
/* recuperação da informação interblocos para o caráter DAP*/

```

```

proc mixed data=arquivo cl noclprint noitprint sigiter;
  class P T L R B;
  model DAP = L P P*L / noint;
  random B(L);
  repeated T*B(L) / ;
  lsmeans L P / cl;
run;

```

```

/*EXEMPLO DE ANÁLISE DO MODELO MISTO */
/* quadrados mínimos generalizados para estimar */
/* médias de blocos e predição BLUP de progênies */
/* caráter DAP*/

```

```

proc mixed data=arquivo cl noclprint noitprint sigiter;
  class P T L R B;
  model DAP = L B(L) / noint;
  random P P*L / s cl type=UN;
  repeated T*B(L) / ;
  lsmeans L B(L) / cl;
run;

```

/*INDICES.SAS: programa de distribuição dos índices de seleção */

/*a) ordenamento dos arquivos de dados geral e de parcelas */

```

proc sort data=drive.arquivo;
  by L R B T P A;
proc sort data=drive.parblpro;
  by L R B T P;

```

/*b) junção dos arquivos no arquivo de índice */

```

data drive.indices;
  merge drive.parblpro drive.arquivo;
  by L R B T P;

```

/* c) cálculo dos índices: Local Anhembi */

```
if L='An' then
  iDAP = 0.33793*(DAPmm-14.316) -0.0087*(DAPp-16.55515)
        -0.00504*(DAPm-DAPb) +0.03121*(DAP-DAPb);
  iH = 0.44226*(Hm-22.32712) -0.01481*(Hp-25.67155)
        -0.02304*(Hm-Hb) +0.06567*(H-Hb);
  iV = 0.27374*(Vmm-0.427491) -0.00371*(Vp-0.645083)
        -0.0000016*(Vm-Vb) +0.0165*(V-Vb);
  iD = 0.68971*(Dmm-0.45332) -0.09727*(Dp-0.420052)
        -0.2571*(Dm-Db) +0.54975*(D-Db);
```

/* c) cálculo dos índices: Local Brotas */

```
if L='Br' then
  iDAP = 0.33793*(DAPmm-14.316) -0.0087*(DAPp-14.24246)
        -0.00504*(DAPm-DAPb) +0.03121*(DAP-DAPb);
  iH = 0.44226*(Hm-22.32712) -0.01481*(Hp-21.42878)
        -0.02304*(Hm-Hb) +0.06567*(H-Hb);
  iV = 0.27374*(Vmm-0.427491) -0.00371*(Vp-0.379246)
        -0.0000016*(Vm-Vb) +0.0165*(V-Vb);
  iD = 0.68971*(Dmm-0.45332) -0.09727*(Dp-0.46918)
        -0.2571*(Dm-Db) +0.54975*(D-Db);
```

/* c) cálculo dos índices: Local Lençóis Paulista */

```
if L='LP' then
  iDAP = 0.33793*(DAPmm-14.316) -0.0087*(DAPp-12.15039)
        -0.00504*(DAPm-DAPb) +0.03121*(DAP-DAPb);
  iH = 0.44226*(Hm-22.32712) -0.01481*(Hp-19.88104)
        -0.02304*(Hm-Hb) +0.06567*(H-Hb);
  iV = 0.27374*(Vmm-0.427491) -0.00371*(Vp-0.258144)
        -0.0000016*(Vm-Vb) +0.0165*(V-Vb);
  iD = 0.68971*(Dmm-0.45332) -0.09727*(Dp-0.470728)
        -0.2571*(Dm-Db) +0.54975*(D-Db);
```

run;

/* d) exemplo de ordenamento para a seleção no caráter DAP */

```
proc sort data=drive.indice;
  by descending iDAP;
run;
```