

# INTRODUÇÃO

Os *delineamentos aumentados* foram propostos por Walter T. Federer na década de cinquenta para solucionar problemas experimentais inerentes ao melhoramento de plantas. Mais precisamente, buscava-se uma solução ao problema de testar um grande número de clones de cana-de-açúcar sob escassez de material de propagação, nas fases preliminares do programa de melhoramento da estação experimental HSPA (Hawaiian Sugar Planter's Association) no Havaí, EUA (Federer, 1956). O autor argumentava que o uso de uma variedade testemunha a cada três parcelas, por exemplo, não propiciava uma estimativa do erro experimental para fazer as comparações desejadas entre os novos clones. Assim, propôs que se escolhesse um delineamento padrão para as variedades testemunhas e aumentasse o número de parcelas de seus blocos (linhas e/ou colunas), para acomodar os novos clones (tratamentos adicionais ou novos).

Nas duas décadas seguintes à sua proposição, essa classe de delineamentos passou por um certo esquecimento. Mas, nos últimos anos sua utilização tem crescido significativamente. Em levantamento feito numa base de dados bibliográficos (CAB ABSTRACTS<sup>1</sup>), constatou-se, entre 1975 e 1999, a publicação de cerca de 50 trabalhos relacionados, distribuídos como: 2% até 1980; 28% entre 1981 e 1990; e 70% a partir de 1991. A maioria deles (quase 70%) refere-se, exatamente, ao seu emprego em programas de melhoramento genético vegetal, envolvendo espécies como soja, feijoeiro, grão-de-bico, leucena, cana-de-açúcar, aveia, cevada, trigo, batata e tomateiro. Outros 25% tratam de aspectos metodológicos de sua utilização, incluindo temas relacionados à análise estatística e modificações de planejamento com vistas à melhoria de sua eficiência experimental. O restante (cerca de 5%) refere-se a aplicações em outras áreas de pesquisa.

No Programa de Melhoramento da Soja desenvolvido pelo Setor de Genética Aplicada às Espécies Autógamas, do Departamento de Genética da ESALQ/USP, desde 1991 vem sendo empregado o delineamento de *blocos aumentados*. Entre as razões dessa adoção, em detrimento de delineamentos clássicos como blocos ao acaso e *lattices*, pode-se citar a maior flexibilidade quanto

---

<sup>1</sup>/ Base internacional de dados, disponível em CD-ROM, mantida pelo *Commonwealth Agricultural Bureaux International (CAB Internacional)*, com cerca de 150 mil registros bibliográficos indexados por ano, entre artigos de periódicos, livros, teses, etc., com cobertura nas áreas de agricultura e ciências afins.

ao número de tratamentos, a possibilidade de testar boa parte dos tratamentos sem repetições e a economia em área, insumos agrícolas, etc. Sua aplicação, entretanto, tem se deparado com problemas de ordem estatístico-experimental que merecem investigação. Diante disso, o processo rotineiro de seleção tem sido feito com base no desempenho das progênies ou linhagens (novos tratamentos) em relação à *performance* das testemunhas mais próximas. Este processo, embora propicie algum ajustamento para efeitos da variação local, também não fornece uma estimativa do *erro* e nem permite inferências acerca das populações de que os genótipos originaram. Ademais, tem sido demonstrado que, para se garantir um controle local efetivo, a frequência de parcelas testemunhas deve ser mantida abaixo de uma em cinco (Cullis *et al.*, 1989).

Um motivo do uso da seleção com base em testemunhas intercalares, no contexto do referido Programa, pode ser atribuído à dificuldade de se implementar uma análise estatística mais eficiente, com a rapidez operacional que o processo exige. Assim, análises mais apuradas têm-se restringido a trabalhos de tese (Farias Neto, 1995; Gomes, 1995; Laínez-Mejía, 1996; Azevedo Filho, 1997; Pinheiro, 1998; Hamawaki, 1998; Unêda-Trevisoli, 1999). E, mesmo nestes casos, o tratamento estatístico predominante foi a análise intrablocos. Logo, o estabelecimento de rotinas computacionais que possibilitem uma aplicação rotineira de critérios seletivos de maior rigor estatístico, para os ensaios assim delineados, representa uma necessidade atual do Programa. Isso justifica-se também pela debilidade natural dos delineamentos aumentados em termos de precisão, sobretudo para contrastes de grande interesse como os que envolvem os novos tratamentos.

Problemas dessa natureza são relatados, com frequência, por pesquisadores que empregam delineamentos aumentados no melhoramento de plantas. Souza (1997) afirma que o delineamento não se mostrou apropriado à estimação de parâmetros genéticos e fenotípicos devido à baixa precisão das estimativas. Rios (1997) reforça essa mesma conclusão. Gomes (1995) argumenta que o delineamento não apresentou eficiência para estimar coeficientes de correlação entre caracteres, com base nos tratamentos adicionais. A autora levanta, então, a necessidade de uma maior discussão quanto ao uso do delineamento na estimação de parâmetros genéticos. Algumas destas limitações, todavia, revelam-se por dificuldades operacionais na realização de uma análise estatística mais elaborada, seja pelo desbalanceamento inerente desses delineamentos, seja pela complexidade dos modelos adotados, em geral, mistos. Enfim, isso acaba determinando uma exploração deficiente dos dados por boa parte dos usuários.

Bearzoti (1994) sugere que uma maneira de melhorar a qualidade das estimativas em blocos aumentados seria, à semelhança dos delineamentos em blocos incompletos tradicionais, recuperar a informação interblocos (modelo assumindo blocos aleatórios). Isto, segundo o autor, parece ainda

não ter sido suficientemente investigado. Wolfinger *et al.* (1997) e Federer (1998) acrescentam que, nos estágios iniciais dos programas seletivos, a natureza aleatória dos genótipos sob teste também precisa ser levada em consideração. Em síntese, as duas sugestões apontam para a abordagem geral de análise baseada em modelos mistos. As conseqüências de sua completa aplicação em blocos aumentados, todavia, ainda são pouco estudadas, tendo se constituído, por isso, num dos objetivos centrais do presente trabalho.

Sabe-se que na abordagem de modelos mistos é fundamental dispor-se de eficientes estimadores dos componentes de variância associados aos efeitos aleatórios (Bueno Filho, 1997; Resende *et al.*, 1996a). Dada a carência de informações específicas, a diversidade de métodos de estimação e as peculiaridades dos ensaios conduzidos no Programa (ex: genótipos oriundos de diferentes cruzamentos), outro estudo que se mostrou necessário foi uma avaliação da qualidade dos principais estimadores desses componentes. É oportuno ressaltar que, para uma avaliação rigorosa de métodos estatísticos, não é suficiente um estudo baseado em conjuntos de dados reais. No contexto do melhoramento genético, entre as propriedades estatísticas desejáveis num estimador, a não tendenciosidade é provavelmente uma das de maior interesse. Tal propriedade, contudo, não pode ser avaliada a partir de observações reais. Assim, a simulação de dados mostra ser uma ferramenta indispensável para uma boa avaliação comparativa dos diferentes estimadores (Littell & McCutchan, 1987). Isso porque, nesta situação, os parâmetros têm seus valores conhecidos, o que permite avaliar o viés associado aos estimadores de cada método, bem como sua eficiência relativa.

Ainda, no que concerne à seleção de genótipos em fases preliminares dos programas de melhoramento, via delineamentos aumentados, outro aspecto preocupante é o pequeno tamanho das parcelas. Nessas fases é comum a adoção de parcelas de apenas uma ou duas fileiras de plantas em razão da pouca disponibilidade de material de propagação (sementes, tubérculos, etc.). Também é usual a alocação sistemática dos cultivares testemunhas ou mesmo de grupos dos tratamentos adicionais com alguma procedência comum. Tudo isso gera preocupações acerca da validade da adoção de uma análise clássica de variância. É de conhecimento geral que esse tipo de análise assenta-se sobre suposições bastante fortes como homocedasticidade, independência entre observações, normalidade e aditividade dos efeitos considerados no modelo. No presente contexto, a preocupação especial recai sobre a suposição de independência, haja vista a grande possibilidade de correlação espacial entre parcelas vizinhas. Diante disso, não parece razoável confiar apenas à casualização (em certo nível) a tarefa de neutralizar os efeitos desta correlação a ponto de não prejudicar, por exemplo, o ordenamento dos genótipos. Os procedimentos estatísticos que permitem lidar com esse tipo de problema são conhecidos como métodos de ajustamento com base na

informação de parcelas vizinhas ou simplesmente métodos de análise espacial de experimentos. Por esta abordagem é possível remover os efeitos da correlação espacial, de forma a obter estimativas, predições e contrastes de tratamentos mais eficientes (Cullis *et al.*, 1989; Stroup & Mulitze, 1991; Zimmerman & Harville, 1991; Cullis & Gleeson, 1991; Cressie, 1993; Es & Es, 1993; Brownie *et al.*, 1993; Scharf & Alley, 1993; Stroup *et al.*, 1994; Grondona *et al.*, 1996; Clarke & Baker, 1996; Federer, 1998; Cullis *et al.*, 1998; entre outros). Sua aplicação aos delineamentos aumentados, contudo, ainda carece de divulgação e adaptação aos *softwares* estatísticos mais amplamente difundidos, tendo se constituído, por isso, num outro objetivo deste trabalho.

Outras questões como o tamanho ideal de blocos, a proporção das parcelas do bloco com variedades testemunhas e o número de testemunhas, também não estão ainda esclarecidas nesse tipo de experimentação. Neste sentido, considerando a adoção usual de blocos relativamente grandes (com até cerca de 50 parcelas), é provável que em boa parte dos ensaios os blocos não sejam suficientemente homogêneos. Esta característica, associada à aplicação de tratamentos não repetidos e à pequena discrepância entre genótipos nos programas de seleção de espécies já bastante melhoradas como a soja, também deve contribuir para uma análise estatística de baixa eficiência. O presente trabalho certamente não traz respostas para todas essas questões. Mas, procurar-se-á situá-las no contexto dos procedimentos alternativos de análise, passíveis de aplicação nesse tipo de experimentos, e, assim, orientar o usuário no sentido de evitar decisões equivocadas.

Enfim, a presente pesquisa buscou soluções que possibilitem uma melhor utilização dos delineamentos aumentados (*blocos de Federer*) em programas de melhoramento genético. As prioridades eleitas estiveram, portanto, ligadas diretamente às características atuais do referido Programa. Especificamente o trabalho está dividido em cinco capítulos, cujos objetivos principais são: *i*) revisar os princípios teóricos relacionados aos delineamentos aumentados, à abordagem de modelos mistos e à análise espacial de experimentos (Capítulo 1); *ii*) investigar as implicações da suposição dos efeitos genotípicos, se fixos ou aleatórios, sobre as respectivas médias e seu ordenamento para fins de seleção (Capítulo 2); *iii*) apresentar e comparar os principais modelos de análise estatística para blocos aumentados, incluindo-se as análises intrablocos e com recuperação de informação interblocos, combinadas com as suposições de tratamentos adicionais fixos ou aleatórios (Capítulo 3); *iv*) comparar numericamente os principais métodos de estimação de componentes de variância, num modelo de blocos aumentados que acomoda tratamentos adicionais de diferentes origens (Capítulo 4); e, *v*) aplicar e avaliar as potencialidades da análise estatística espacial no contexto dos ensaios em blocos aumentados (Capítulo 5).