

**SOBRE O EMPREGO E A ANÁLISE ESTATÍSTICA  
DO DELINEAMENTO EM BLOCOS AUMENTADOS NO  
MELHORAMENTO GENÉTICO VEGETAL**

**JOÃO BATISTA DUARTE**  
**Engenheiro Agrônomo**

**Orientador: Prof. Dr. ROLAND VENCOVSKY**

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

**PIRACICABA**  
**Estado de São Paulo – Brasil**  
**Março – 2000**

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO – Campus “Luiz de Queiroz”/USP**

Duarte, João Batista

Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal / João Batista Duarte. - - Piracicaba, 2000.  
293 p.

Tese (doutorado) – Escola Superior de Agricultura Luiz de Queiroz, 2000.  
Bibliografia.

1. Análise estatística 2. Delineamento de experimento 3. Melhoramento genético vegetal 4. Modelo matemático 5. Seleção genética 6. Soja 7. Variabilidade espacial I. Título

CDD 631.522

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

*Ofereço este trabalho ao mais humilde dos brasileiros que, através de seus impostos ajudou a patrocinar os meus estudos, sem, todavia, poder optar em oferecê-los a seus filhos!*

### *Oração de um aluno*

*Deus, nosso pai, quero falar contigo sobre a minha vida.  
Há tantas coisas por fazer e aprender!  
Eu quero estar preparado para as minhas  
responsabilidades na construção de um mundo melhor.  
Senhor, faça com que os meus estudos me ajudem a descobrir  
a natureza das coisas, mas, sobretudo a arte  
de viver e de conviver,  
Ajude-me, então, a usar a inteligência  
para compreender principalmente as pessoas.  
Fortifica-me a vontade para que eu assumo  
a minha formação com persistência e honestidade,  
sem temer os obstáculos e os sacrifícios.  
Não permitas que a busca de diplomas,  
promoções e lucros tire do meu coração  
a alegria de servir com desinteresse e de me doar por amor.  
Cristo, que o teu Evangelho seja a minha luz  
e terei a certeza de que não vou me enganar  
na conquista da verdadeira realização e felicidade.  
Amém!*

*Mãe, obrigado por sonhar que teus meninos podiam  
aprender um pouco mais e, também, por esta oração.  
Pai, com saudades, minha eterna gratidão por sua  
luta e, também, pelas lições do trabalho e do esmero.*

*A vocês, com amor, eu Dedico.*



## AGRADECIMENTOS

- À Universidade Federal de Goiás, particularmente à Escola de Agronomia e à Pró-Reitoria de Pesquisa e Pós-graduação / PICD-CAPES, pela oportunidade desta qualificação.
- À Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (ESALQ/USP), pelo completo apoio institucional.
- Ao Centro de Informática na Agricultura / USP - Campus de Piracicaba (CIAGRI / USP), pela disponibilização de recursos computacionais.
- Aos professores Dr. Lázaro José Chaves, Dr. Edward Madureira Brasil e Dr. José Baldin Pinheiro, pela amizade e por assumirem as minhas tarefas na Escola de Agronomia / UFG, possibilitando o meu afastamento.
- Ao professor Dr. Roland Vencovsky, por oportunizar-me uma convivência fecunda e amistosa com o seu admirável espírito científico e humano, e sobretudo por auxiliar-me e acompanhar este meu “vôo”, reservando-me a prazerosa escolha do destino.
- Ao professor Dr. José Branco Miranda Filho, pela amizade, estímulo, disponibilidade e grandes lições de vida.
- Ao professor Dr. Natal Antônio Vello, pela confiança em ceder-me o material básico para o desenvolvimento deste tema; e também aos funcionários do Setor de Genética Aplicada às Espécies Autógamas (Marcos, Roberto e Claudinei), pelo trabalho dedicado que resultou em parte dos dados utilizados nesta pesquisa.
- Aos demais professores do Departamento de Genética da ESALQ/USP, pelos muitos ensinamentos, em especial, o Dr. Cláudio L. Souza Jr., Dr. Isaías O. Geraldi, Dr<sup>a</sup>. Margarida L. R. Aguiar-Perecin e Dr<sup>a</sup>. Maria Lúcia C. Vieira.
- Aos professores do Departamento de Ciências Exatas da ESALQ/USP, Dr. Antônio Francisco Iemma, Dr. Décio Barbin, Dr. Carlos Tadeu S. Dias e Dr<sup>a</sup> Maria Izalina F. Alves, pela amizade e contribuição valiosa à minha formação.
- Ao pesquisador Dr. Sidney Rosa Vieira, do Instituto Agrônomo (IAC), pelo incentivo e valiosas sugestões.
- À Léia, Adriana, Carmem, Elisabete, Silvana e Silvana (da biblioteca), Antônio (do café), Rosa, Solange e Luciane (Ciências Exatas), pelos atendimentos sempre solícitos.
- Aos amigos e colegas, pela interlocução científica e pela alegria de podermos compartilhar os nossos cursos - Dario, Alexandre Siqueira, Camilo, Edson, Leonardo, Rogério, Rhainer, Paulo Barroso, Maria Imaculada, Aldete, Glauce, Luciana Carlini, Marlene, Ângela, Sandra Helena, Andréia, Fábio, Maurisrael, Gilberto, Cláudio, Josué, Derli, Rosângela, Rainério, Maria Inês, Alexandre Sebben, Eduardo, Francisco, Raulindo, Jorge, Juan, ..., Adilson, Christian, Helena, Idemauro, Jeanete, Sandra Denisen e Wilson.
- Um agradecimento especial eu reservo a minha esposa Rôzi, pelo carinho, compreensão e auxílio permanentes, incluindo a cuidadosa revisão deste texto, e a minha filha Letícia, por tornar esta jornada muito mais alegre e estimulante.
- Enfim, a **Deus**, por proporcionar-me tudo isso e permitir-me oferecer esta contribuição.



## SUMÁRIO

<b>RESUMO</b> .....	<b>xi</b>
<b>SUMMARY</b> .....	<b>xiii</b>
<b>INTRODUÇÃO</b> .....	<b>1</b>
<b>Capítulo 1. REFERENCIAL TEÓRICO</b> .....	<b>5</b>
<b>1. Delineamentos em blocos</b> .....	<b>5</b>
1.1. Informações gerais .....	5
1.2. Aspectos da análise estatística de blocos incompletos .....	8
1.2.1. A análise intrablocos .....	8
1.2.2. A análise com recuperação da informação interblocos .....	13
<b>2. Os delineamentos aumentados</b> .....	<b>20</b>
2.1. Princípios e evolução histórica .....	20
2.2. Aspectos da análise estatística .....	23
2.3. Viabilidade e aplicação no melhoramento de plantas .....	27
2.4. Desenvolvimentos correlatos .....	31
<b>3. Sobre a suposição de efeitos fixos e aleatórios nos ensaios de competição de genótipos com delineamento de blocos</b> .....	<b>36</b>
<b>4. Modelos mistos e componentes de variância</b> .....	<b>41</b>
4.1. Considerações gerais .....	41
4.2. Estimação e predição dos efeitos no modelo linear misto .....	43
4.3. Componentes de variância .....	47
4.3.1. Métodos de estimação .....	47
4.3.2. Sobre a qualidade dos estimadores .....	50
<b>5. Variabilidade espacial em experimentos agrícolas</b> .....	<b>54</b>
5.1. Considerações gerais .....	54
5.2. Diagnosticando a presença de variabilidade espacial .....	59
5.3. Alguns métodos de análise espacial de experimentos .....	63
5.4. Eficiência da análise espacial de experimentos .....	69
<b>Capítulo 2. ESTIMAÇÃO E PREDIÇÃO SOB MODELO LINEAR MISTO COM ÊNFASE NA ORDENAÇÃO DE MÉDIAS DE TRATAMENTOS GENÉTICOS</b> .....	<b>75</b>
<b>Resumo</b> .....	<b>75</b>
<b>Abstract</b> .....	<b>76</b>
<b>1. Introdução</b> .....	<b>77</b>
<b>2. Um modelo de delineamento em blocos</b> .....	<b>78</b>
<b>3. Estimação e predição num modelo linear misto</b> .....	<b>79</b>
<b>4. Explicitação do <i>BLUP</i> de <math>\gamma</math></b> .....	<b>82</b>

<b>5. O efeito “shrinkage” nas médias <i>BLUP</i> .....</b>	<b>86</b>
<b>6. Ordenamento comparativo das médias <i>BLUP</i> .....</b>	<b>90</b>
<b>7. Duas outras variações no modelo .....</b>	<b>94</b>
<b>7.1. Modelo com blocos aleatórios .....</b>	<b>94</b>
<b>7.2. Modelo com tratamentos de diferentes populações .....</b>	<b>96</b>
<b>8. Conclusões .....</b>	<b>97</b>
<b>Capítulo 3. MÉTODOS DE ANÁLISE ESTATÍSTICA PARA BLOCOS AUMENTADOS COM APLICAÇÃO NO MELHORAMENTO DE PLANTAS (UMA ABORDAGEM - SAS® ORIENTADA)..</b>	<b>99</b>
<b>Resumo .....</b>	<b>99</b>
<b>Abstract .....</b>	<b>100</b>
<b>1. Introdução .....</b>	<b>101</b>
<b>2. Metodologia .....</b>	<b>103</b>
<b>2.1. Material .....</b>	<b>103</b>
<b>2.2. Modelos alternativos de análise estatística .....</b>	<b>104</b>
<b>2.3. Procedimentos estatístico-computacionais .....</b>	<b>107</b>
<b>2.3.1. Preparação e leitura do arquivo de dados .....</b>	<b>108</b>
<b>2.3.2. Instruções para a execução das análises estatísticas .....</b>	<b>109</b>
<b>3. Resultados e Discussão .....</b>	<b>118</b>
<b>3.1. Avaliação dos diferentes modelos de análise estatística .....</b>	<b>118</b>
<b>3.1.1. Influência dos modelos na precisão das médias genotípicas .....</b>	<b>118</b>
<b>3.1.2. Influência dos modelos na seleção dos genótipos .....</b>	<b>122</b>
<b>3.1.3. Um caso ilustrativo .....</b>	<b>125</b>
<b>3.2. Aspectos experimentais relacionados à análise estatística .....</b>	<b>127</b>
<b>3.3. Sobre as suposições para os efeitos de blocos e de progênies ...</b>	<b>129</b>
<b>3.4. Os modelos de análise sob progênies de diferentes origens .....</b>	<b>135</b>
<b>4. Conclusões .....</b>	<b>140</b>
<b>Capítulo 4. COMPARAÇÃO ENTRE ESTIMADORES DE COMPO- NENTES DE VARIÂNCIA NUM MODELO DE BLOCOS AUMENTADOS COM TRATAMENTOS NOVOS DE UMA OU MAIS POPULAÇÕES .....</b>	<b>141</b>
<b>Resumo .....</b>	<b>141</b>
<b>Abstract .....</b>	<b>142</b>
<b>1. Introdução .....</b>	<b>143</b>
<b>2. Metodologia .....</b>	<b>147</b>



2.1. O modelo matemático .....	147
2.2. Os parâmetros e os casos simulados .....	148
2.3. Procedimentos estatístico-computacionais de estimação .....	150
2.4. Avaliação da qualidade dos estimadores .....	153
<b>3. Resultados e Discussão .....</b>	<b>154</b>
3.1. Tratamentos novos de uma só população .....	154
3.2. Tratamentos novos de diferentes populações .....	163
<b>4. Conclusões .....</b>	<b>169</b>
<b>Capítulo 5. APLICAÇÃO DE ANÁLISE ESTATÍSTICA ESPACIAL A EXPERIMENTOS GENÉTICOS EM BLOCOS AUMENTADOS .....</b>	<b>173</b>
<b>Resumo .....</b>	<b>173</b>
<b>Abstract .....</b>	<b>174</b>
<b>1. Introdução .....</b>	<b>175</b>
<b>2. Material e métodos .....</b>	<b>177</b>
2.1. Material .....	177
2.2. Procedimentos estatístico-computacionais .....	178
2.2.1. Os modelos de análise estatística .....	178
2.2.2. Ajustando o modelo de observações espacialmente independentes .....	180
2.2.3. Diagnosticando e estimando a covariância espacial .....	182
2.2.4. Ajustando o modelo de análise espacial .....	185
2.2.5. Comparando os modelos de análise espacial e não espacial .....	187
<b>3. Resultados e Discussão .....</b>	<b>188</b>
3.1. Caracterização da covariância espacial .....	188
3.2. Comparação dos modelos de análise espacial e não espacial ....	191
<b>4. Conclusões .....</b>	<b>195</b>
<b>CONSIDERAÇÕES FINAIS .....</b>	<b>197</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>199</b>
<b>ANEXOS .....</b>	<b>211</b>
<b>APÊNDICES .....</b>	<b>243</b>

## Curriculum Vitae

**João Batista Duarte** é brasileiro, natural do município de Corumbáiba, estado de Goiás, em 26 de novembro de 1962. Graduou-se em Engenharia Agrônômica, no ano de 1984, pela Escola de Agronomia da Universidade Federal de Goiás (UFG), Goiânia-Goiás. Em dezembro de 1988, concluiu o seu curso de mestrado em Genética e Melhoramento de Plantas, também pela UFG, tendo defendido dissertação intitulada “Estudo da adaptabilidade e estabilidade fenotípica em linhagens e cultivares de feijão mulatinho (*Phaseolus vulgaris* L.)”. Desenvolveu este trabalho junto ao Centro Nacional de Pesquisa de Arroz e Feijão da EMBRAPA, entre 1985 e 1988, sob a orientação da pesquisadora Dra. Maria José de Oliveira Zimmermann. Nesta instituição, prestou serviços também nas atividades de seleção e avaliação de linhagens e cultivares de arroz. Em junho de 1987, ingressou como professor substituto na Escola de Agronomia / UFG, tendo ministrado, em nível de graduação, disciplinas como ‘Estatística e Experimentação’ e ‘Melhoramento de Plantas’. Em 1994, foi, então, nomeado em caráter efetivo como professor assistente desta mesma Escola. Nesta instituição, com a qual mantém vínculo empregatício até a presente data, além das atividades de ensino, desenvolveu trabalhos de pesquisa relacionados ao tema ‘interação de genótipos com ambientes’. Em nível de especialização, concluiu, em 1994, o ‘Curso de Metodologia do Ensino Superior’, pela Associação Brasileira de Educação Agrícola Superior / Universidade Federal de Viçosa. E, em 1996, o ‘Curso de Especialização em Estatística’, pelo Instituto de Matemática e Estatística / UFG. Em 1996, iniciou o seu curso de doutorado na área de ‘Genética e Melhoramento de Plantas’, junto à Escola Superior de Agricultura ‘Luiz de Queiroz’ / Universidade de São Paulo, sob a orientação do professor Dr. Roland Vencovsky, com quem publicou pela Sociedade Brasileira de Genética, a monografia “Interação genótipos x ambientes - Uma introdução à análise AMMI”.

# **SOBRE O EMPREGO E A ANÁLISE ESTATÍSTICA DO DELINEAMENTO EM BLOCOS AUMENTADOS NO MELHORAMENTO GENÉTICO VEGETAL**

Autor: JOÃO BATISTA DUARTE

Orientador: Prof. ROLAND VENCOVSKY

## **RESUMO**

A presente pesquisa propôs-se a investigar problemas de ordem estatístico-experimental relacionados à aplicação do delineamento de blocos aumentados, em programas de melhoramento genético vegetal. Nesses desenhos experimentais há duas categorias de tratamentos: as testemunhas (repetidas) e os tratamentos adicionais ou novos (normalmente não repetidos). As primeiras usualmente são cultivares comerciais já recomendados, e os outros, os novos materiais genéticos sob avaliação. Sua relação com os delineamentos de blocos em geral, princípios para a sua aplicação e análise estatística, bem como informações correlatas de outros temas aqui investigados (modelos mistos, componentes de variância e análise espacial de experimentos) foram objetos de revisão no Capítulo 1. O material que fundamentou este trabalho foi um conjunto de 41 ensaios de competição de linhagens de soja (genótipos), em gerações de autofecundação nunca inferiores a  $F_5$ . Os ensaios, assim delineados, fazem parte do Programa de Melhoramento da Soja desenvolvido pelo Setor de Genética Aplicada às Espécies Autógamas, do Departamento de Genética da ESALQ/USP. Foram escolhidos, preferencialmente, os ensaios que dispunham de mapa de campo completo, o que possibilita associar cada observação à sua posição espacial na área do experimento. Entre os diversos caracteres avaliados, apenas os dados de produtividade de grãos (kg/ha) foram aqui considerados. Sabe-se que esses ensaios são conduzidos em etapas preliminares, quando as linhagens podem ser estatisticamente assumidas como de efeitos aleatórios, ou em fases intermediárias e finais, quando são preferencialmente tidas como de efeitos fixos. Logo, uma curiosidade inicial foi investigar, nos delineamentos em blocos, a influência destas suposições sobre as estimativas das médias genotípicas, bem como sobre o seu ordenamento para fins de seleção (Capítulo 2). O fato também abre a possibilidade de diferentes modelagens para a análise estatística dos dados desses experimentos, incluindo-se o modelo fixo e modelos mistos (análise intrablocos, análise com recuperação de informação interblocos e análises recuperando informação intergenotípica). A apresentação destas alternativas e suas implicações na seleção de genótipos constitui o objetivo principal do Capítulo 3. Acrescenta-se que, na abordagem de modelos mistos,

uma etapa fundamental é a de estimação dos componentes de variância. Isto é passível de realização por meio de vários métodos estatísticos, cujos resultados, todavia, podem ser bastante conflitantes; sobretudo em caso de desbalanceamento. Dada a escassez de informações específicas e o fato de esses delineamentos serem naturalmente desbalanceados, avaliaram-se através de simulação em computador as propriedades dos principais estimadores disponíveis: *ANOVA*, *MIVQUE(0)*, *ML* e *REML* (Capítulo 4). Outra característica marcante desses ensaios, particularmente nas etapas iniciais, é a adoção de parcelas de pequeno tamanho, haja vista a pouca disponibilidade de material de propagação. Isto, associado à costumeira alocação sistemática de testemunhas e/ou grupos de linhagens aparentadas, suscitou a avaliação de um procedimento de análise estatística que não ficasse sujeito à clássica suposição de independência espacial entre observações (Capítulo 5). Entre os resultados e conclusões obtidas pode-se destacar: *i*) a análise intrablocos (modelo fixo) pode fornecer ordenamentos inadequados das médias dos genótipos se estes forem de efeitos aleatórios e, sobretudo, se estiverem relacionados a diferentes populações de referência (Capítulo 2); *ii*) a classificação das novas linhagens em relação aos cultivares testemunhas pode mudar sensivelmente de um modelo de análise para outro, principalmente quando se passa da análise intrablocos para uma análise que recupera informação interlinhagens (Capítulo 3); *iii*) o método *MIVQUE(0)* comparativamente fornece estimativas de melhor qualidade para os componentes de variância, em particular se os genótipos vierem de população(ões) com baixa(s) variância(s) genotípica(s) e os experimentos forem relativamente pequenos (Capítulo 4); e, *iv*) em experimentos com observações correlacionadas espacialmente, a discriminação genotípica e o ordenamento dos genótipos para fins de seleção podem ser melhorados consideravelmente através da análise estatística espacial (Capítulo 5).

# **APPLICATION AND STATISTICAL ANALYSIS OF AUGMENTED BLOCK DESIGN IN PLANT BREEDING**

Author: JOÃO BATISTA DUARTE  
Adviser: Prof. ROLAND VENCOVSKY

## **SUMMARY**

This research investigates experimental and statistical problems related with the use of the augmented block design in plant breeding programs. A characteristic of these designs is that treatments are of two categories, namely: common checks (generally commercial varieties replicated over blocks) and the additional or new treatments (usually not replicated), the latter being the ones under evaluating for selection purposes. The relation of this design with block designs in general and fundamentals required for its applications and analysis, as well as information about other topics (mixed models, components of variance and spatial analysis) were reviewed in the Chapter 1. For analysis and discussion a group of 41 trials, set up at the Department of Genetics (ESALQ, USP) for evaluating soybean inbred lines, was used. Only trials having a complete layout of the spatial position of plots in the field were taken. Analyses and discussions given here refer only to grain yield (kg/ha). It is known that these trials are usually conducted at preliminary stages, when genotypes can be statistically assumed as having random effects, or at intermediate and final stages when they are preferentially considered with fixed effects. The first interesting point here investigated was the influence of these assumptions on the estimates of genotypic means, as well as on the ranking of lines for selection purposes (Chapter 2). The assumptions open the possibility of considering either fixed or mixed models for analysis (intrablock analysis, or others with recovering of interblock or intergenotypic information). Chapter 3 shows the corresponding analytical procedures and their consequences on the ranking and selection of genotypes. Under mixed models a fundamental step is the estimation of variance components, for which several procedures are available. It is known that these procedures may lead to different estimates of the same parameter, specially when experiments are unbalanced. Due to the lack of specific information about its point, and since these designs are naturally unbalanced, computer simulations were made here for evaluating the properties of the different available estimators, specifically: *ANOVA*, *MIVQUE(0)*, *ML* and *REML* (Chapter 4). An additional characteristic of these design, at initial stages of breeding programs, is the use of small-sized plots, necessary to accommodate the lack of seeds. This fact in

addition to the common practice of systematic arrangement of check plots in the field and/or the arrangement of genetically related treatments in sets, motivated the application of spatial statistics procedures (Chapter 5). The following main results and conclusions can be pointed out: *i*) the intrablock analysis (fixed model) can provide inadequate ranking of the genotypic means if genotypes have random effects and, specially if they are stem from different populations (Chapter 2); *ii*) the classification of the new genetic treatments in relation to check varieties may change considerably, depending on the model adopted, specially when the intrablock and the analyses with recovery of intergenotypic information are compared (Chapter 3); *iii*) the *MIVQUE(0)* method, in general, furnished more efficient estimates of variance components, particularly if genotypes are derived from population(s) with low genotypic variance(s) and experiments are of small size (Chapter 4); and, *iv*) if the experimental observations are spatially correlated the discrimination among genotypes and the ranking of genetic treatments can be substantially improved through spatial statistical analysis (Chapter 5).