

**Universidade de São Paulo  
Escola Superior de Agricultura “Luiz de Queiroz”**

**Delineamentos ótimos para experimentos com cana-de-açúcar**

**Alessandra de Lima Goes**

Dissertação apresentada para obtenção do título de  
Mestra em Ciências. Área de concentração: Estatística e Experimentação Agronômica

**Piracicaba  
2020**

**Alessandra de Lima Goes**  
**Matemática**

**Delineamentos ótimos para experimentos com cana-de-açúcar**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:

Profa. Dra. **RENATA ALCARDE SERMARINI**

Dissertação apresentada para obtenção do título de  
Mestra em Ciências. Área de concentração: Estatística e Experimentação Agrônômica

**Piracicaba**  
**2020**

**Dados Internacionais de Catalogação na Publicação  
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Goes, Alessandra de Lima

Delineamentos ótimos para experimentos com cana-de-açúcar / Alessandra de Lima Goes. -- versão revisada de acordo de acordo com a resolução CoPGr 6018 de 2011 - - Piracicaba 2020 .

43 p.

Dissertação (Mestrado) -- USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Delineamento parcialmente replicado 2. Delineamento *grid-plot* 3. Estudo de simulação 4. Ganho genético 5. Correlação 6. Assertividade . I. Título.

## DEDICATÓRIA

Dedico este trabalho ao meu amor  
Maria Flor,  
É por você,  
Para você,  
E com você.  
Luz do meu caminho.

## AGRADECIMENTOS

Obrigada meu Deus por me dar a honra de poder ter participado desse mestrado, e por não me deixar desistir mesmo nas horas em que tudo que eu queria era chutar o balde, obrigada meu Deus por me manter firme mesmo em meio a cansaços, mesmo em meio a uma gravidez inesperada mesmo em meio a tantas frustrações, obrigada meu Deus porque antes de tudo e de todos o Senhor não desistiu de mim, o Senhor me ajudou a ser perseverante e obrigada por colocar em meu caminho todas as pessoas as quais irei igualmente agradecer a baixo.

Em primeiro lugar, não poderia deixar de agradecer a minha orientadora Renata Alcarde Ser-marini, por toda a paciência, tempo disposto e orientação durante a construção desse trabalho. Obrigada por entender e não desistir de mim, sem você com certeza eu não teria chegado ao fim desse mestrado. Você é e esta sendo muito mais que uma orientadora.

Desejo igualmente agradecer também a todos os meus colegas de mestrado em Estatística e Experimentação Agronômica pelo apoio, companheirismo em especial Suelen Gasparetto e sua mãe Arlete Gasparetto por sempre me receberem bem e ouvirem minhas frustrações, eu não tenho palavras para agradecer todo carinho e amor por mim e pela Maria Flor.

A minha pequena Maria Flor, que no último ano esteve tão próxima de mim (literalmente), foi tão presente no desenvolvimento desse trabalho, me ensinando a ser paciente e perseverante. É quem me inspira a querer ser mais do que fui até hoje. A sua existência é o reflexo mais perfeito da existência de Deus.

Aos meus pais, irmãos e marido por me encorajarem cada um a sua maneira a continuar mesmo em dias de tanto cansaço. Por cuidarem da Maria Flor para que eu pudesse estudar, por ouvirem as minhas lamentações.

Aos professores do programa de pós-graduação em Estatística e Experimentação Agronômica, e todos aqueles que de alguma forma contribuíram para o meu trabalho em especial os professores que se dispuseram a compor a minha banca.

A todos que de alguma forma contribuíram para a realização deste trabalho, o meu muito obrigada!

**EPÍGRAFE**

E, se bem que seja obscuro  
Tudo pela estrada fora,  
E falso, ele vem seguro,  
E, vencendo estrada e muro,  
Chega onde em sono ela mora.  
E, inda tonto do que houvera,  
A cabeça, em maresia,  
Ergue a mão, e encontra hera,  
E vê que ele mesmo era A Princesa que dormia.  
*Fernando Pessoa*

## SUMÁRIO

Resumo . . . . .	7
Abstract . . . . .	8
Lista de Figuras . . . . .	9
Lista de Tabelas . . . . .	10
1 Introdução . . . . .	11
2 Revisão de Literatura . . . . .	13
2.1 Delineamentos Experimentais . . . . .	13
2.1.1 Delineamentos Grid-plot . . . . .	13
2.1.1.1 Delineamentos Aumentados . . . . .	14
2.1.2 Delineamentos Parcialmente Replicados . . . . .	17
2.2 Delineamentos Ótimos . . . . .	19
2.3 Comparando Delineamentos . . . . .	20
3 Metodologia . . . . .	23
3.1 Gerando delineamentos . . . . .	23
3.2 Medidas das comparação . . . . .	24
4 Resultados e Discussão . . . . .	27
4.1 Delineamentos obtidos . . . . .	27
4.2 Medidas de comparação . . . . .	27
4.3 Estimativas dos componentes de variância . . . . .	33
5 Conclusão . . . . .	39
Referências . . . . .	41

## RESUMO

### Delineamentos ótimos para experimentos com cana-de-açúcar

O principal objetivo dos programas de melhoramento genético de cana-de-açúcar é selecionar novas variedades, com boas características de mercado, em um curto período de tempo. As fases iniciais do melhoramento envolvem um grande número de tratamentos (*test lines*) resultando em limitações como disponibilidade de material genético para a realização de várias repetições e o tamanho da área experimental. Desse modo, são necessários delineamentos eficientes que permitam a obtenção da maior informação possível a respeito dos tratamentos. O presente trabalho avalia, por meio de estudos de simulação, o ganho genético e a qualidade da seleção de 24 delineamentos distintos se *grid-plot* ou parcialmente replicado; ao número de variedades padrões para uma fixada porcentagem de *test lines* repetidas; a porcentagem de *test lines* repetidas e critérios de otimização,  $C$  ou  $D$ , para diferentes cenários de conjuntos de dados e uma dada área experimental. A partir dos resultados concluiu-se que as diferenças entre os delineamentos são pequenas, sendo mais evidente em termos do ganho genético, mas que na prática implica na seleção de uma *test line* adicional, estando a porcentagem média de acertos na seleção entre 25 e 40% para os casos estudados.

**Palavras-chave:** Delineamento parcialmente replicado, Delineamento *grid-plot*, Estudo de simulação, Ganho genético, Correlação, Assertividade



## ABSTRACT

### Optimum designs in sugarcane experiments

The main goal of the sugarcane breeding programs is to select new varieties with desirable characteristics in a short time period. In the early stage of the plant-breeding program there is a large number of treatments (test lines), resulting in limitations such as availability of genetic material for several replications and the size of the experimental area. In this sense, efficient designs are necessary in order to obtain as much information as possible about the treatments. The present study evaluates, using simulation studies, the genetic gain and the selection quality of 24 designs, distinct in relation to the type, whether grid-plot or partially replicated; the number of standard varieties for a fixed percentage of replicated test lines; the percentage of replicated test lines and optimality criteria,  $C$  or  $D$ , for different data scenarios and a given experimental area. From the results we concluded that the differences between the designs are small, being more evident in terms of genetic gain, but that in practice implies the selection of an additional test line, being the average percentage of correct test lines in the selection between 25 and 40% for the studied cases.

**Keywords:** Partially replicated design, Grid-plot design, Simulation study, Genetic gain, Correlation, Assertiveness

## LISTA DE FIGURAS

2.1	Delineamentos <i>grid-plot</i> em um arranjo com 10 linhas por 6 colunas e quatro variedades padrões (A, B, C e D). Em (a) as padrões são repetidas quatro vezes e alocadas em diagonal e tem-se 44 <i>test lines</i> (1, ..., 44) não repetidas; (b) as padrões são repetidas três vezes e ocupam as parcelas de acordo com o movimento do cavaleiro, tem-se 48 <i>test lines</i> (1, ..., 48) não repetidas . . . . .	14
2.2	Delineamento em blocos aumentados com quatro variedades padrões (A, B, C e D) repetidas três vezes e 48 <i>test lines</i> não repetidas . . . . .	15
2.3	(a) Delineamento quadrado latino para três tratamentos e (b) Delineamento quadrado de Youden com cinco tratamentos em uma área experimental com quatro linhas e cinco colunas	16
2.4	(a) Delineamento quadrado latino aumentado para três variedades padrões e 18 <i>test lines</i> em uma área experimental com três linhas e nove colunas e (b) Delineamento quadrado de Youden aumentado com quatro variedades padrões e 24 <i>test lines</i> em uma área experimental com três linhas e 12 colunas . . . . .	16
2.5	Delineamento quadrado de Youden com três linhas e sete tratamentos . . . . .	17
2.6	Delineamento quadrado de Youden modificado (aumentado) com sete linhas, sete colunas, três variedades padrões repetidas sete vezes e 28 <i>test lines</i> . . . . .	17
2.7	Delineamento aumentado tipo 1 de acordo com Lin e Poushinsky (1983), com 12 linhas, 12 colunas, quatro variedades padrões (A, B, C e D) repetidas seis vezes e 120 <i>test lines</i> (1, ..., 120) não repetidas . . . . .	18
2.8	Delineamento aumentado tipo 2 de acordo com LIN e POUSHINSKY (1985), com 25 linhas, 6 colunas, três variedades padrões, A, B e C, repetidas 30, 10 e 10 vezes, respectivamente, e 100 <i>test lines</i> (1, ..., 100) não repetidas . . . . .	18
2.9	Exemplo de um delineamento <i>p</i> -rep com 115 <i>test lines</i> , sendo 29 duplicadas (em cinza) e 86 não repetidas (em branco) em uma área com 24 linhas e seis colunas, em que as 12 primeiras linhas correspondem a um bloco e as demais a outro bloco, caracterizando um delineamento resolvível . . . . .	19
4.1	Delineamentos <i>C</i> -ótimo e <i>D</i> -ótimo para o delineamento 3, no qual 22% das <i>test lines</i> são duplicadas e tem-se quatro variedades padrões repetidas cinco vezes cada, considerando-se a presença de variâncias globais para colunas e linhas . . . . .	27
4.2	Estimativas do termo relacionado à variância genética ( $\hat{\gamma}_g$ ) para os diferentes delineamentos e cenários. A reta vertical representa o valor esperado para o parâmetro e o ponto representa a média das estimativas do parâmetro. . . . .	34
4.3	Estimativas do termo relacionado à variância genética ( $\hat{\gamma}_g$ ) para os diferentes delineamentos e cenários. A reta vertical representa o valor esperado para o parâmetro e o ponto representa a média das estimativas do parâmetro. . . . .	35
4.4	Estimativas do componente de variância globais para colunas e linhas, $\hat{\gamma}_c$ e $\hat{\gamma}_l$ , respectivamente, para os diferentes delineamentos e cenários avaliados . . . . .	36
4.5	Estimativas da correlação residual entre medidas observadas na mesma coluna, $\hat{\phi}_c$ para os diferentes delineamentos e cenários avaliados . . . . .	37
4.6	Estimativas da correlação residual entre medidas observadas na mesma linha, $\hat{\phi}_l$ para os diferentes delineamentos e cenários avaliados . . . . .	38

## LISTA DE TABELAS

3.1	Caracterização das propostas de delineamentos avaliados para 180 <i>test lines</i> em uma área experimental de 10 colunas por 24 linhas . . . . .	24
4.1	Número de casos de não convergência do algoritmo para os diferentes delineamentos, critérios de otimização e cenários para os dados, em 1000 conjuntos de dados . . . . .	28
4.2	Relação entre fatores de tratamentos e correspondentes fontes de variação para a obtenção das eficiências . . . . .	28
4.3	Anatomia do delineamento 1 ( $p=0\%$ , com quatro padrões repetidas 15 vezes) com efeito global de colunas e linhas, $C$ -ótimo, para 180 <i>test lines</i> em uma área experimental de 10 colunas por 24 linhas . . . . .	29
4.4	Anatomia do delineamento 2 ( $p=11\%$ , com quatro padrões repetidas 10 vezes) com efeito global de colunas e linhas, $C$ -ótimo, para 180 <i>test lines</i> em uma área experimental de 10 colunas por 24 linhas . . . . .	29
4.5	Anatomia do delineamento 6 ( $p=33\%$ , sem padrões) com efeito global de colunas e linhas, $C$ -ótimo, para 180 <i>test lines</i> em uma área experimental de 10 colunas por 24 linhas . . . . .	29
4.6	Eficiências médias no último estrato (Coluna#Linha) para os contrastes simples entre <i>test lines</i> duplicadas (Duplicada) e entre <i>test lines</i> não duplicadas (NDuplicadas) para os 24 delineamentos . . . . .	30
4.7	Médias e desvios-padrões para ganho genético percebido para cada um dos 24 delineamentos considerando-se os dois cenários para o componente de variância genética, $\gamma_g = 0,3$ e $\gamma_g = 0,8$ . . . . .	31
4.8	Médias e desvios-padrões para correlação entre os verdadeiros efeitos genéticos e os EBLUPs para cada um dos 24 delineamentos considerando-se os dois cenários para o componente de variância genética, $\gamma_g = 0,3$ e $\gamma_g = 0,8$ . . . . .	32
4.9	Médias e desvios-padrões para a assertividade (porcentagem de verdadeiras melhores <i>test lines</i> selecionadas) para cada um dos 24 delineamentos considerando-se os dois cenários para o componente de variância genética, $\gamma_g = 0,3$ e $\gamma_g = 0,8$ . . . . .	32

## 1 INTRODUÇÃO

Originária das regiões tropicais do Sul e do Sudeste da Ásia, com caules robustos e articulados que são cheios de sacarose, a cana-de-açúcar é um dos elos significativos para a economia mundial. A partir dela é produzido o açúcar, um alimento importante da dieta humana e o álcool utilizado em bebidas e combustíveis como o etanol, sendo este de alto interesse econômico e ambiental por ser uma fonte de energia renovável e poder competir com derivados do petróleo.

Trazida ao Brasil por volta do século XVI foi a base da economia do nordeste brasileiro que se estendeu até o século XVII. Hoje, porém, é no interior de São Paulo que se tem a maior parte dos canaviais, fazendo do Brasil o principal produtor de cana do mundo (CONAB, 2017). Segundo Vieira et al. (2007) o Brasil tem o menor custo de produção entre os principais competidores do mercado internacional, colhe cerca de 647,6 milhões de toneladas em aproximadamente 8,84 milhões de hectares, além de liderar o conhecimento da biotecnologia da cana, junto com a Austrália e a África do Sul.

Mesmo que o Brasil seja um dos maiores produtores, tal produtividade ainda pode ser melhorada, e é este o objetivo dos programas de melhoramento genético, obter variedades com melhor rendimento e mais resistentes às doenças. Ainda, é sabido que as variedades de cana-de-açúcar precisam ser substituídas periodicamente no mundo todo, pois acabam entrando em decadência após anos de cultivo (Cesnik e Miocque, 2004).

Para que novas variedades cheguem ao produtor é necessário o cumprimento de algumas etapas: cruzamento de variedades existentes, produção de mudas, seleção de clones, avaliação dos clones e comercialização de novas variedades.

Nas fases iniciais do melhoramento genético da cana-de-açúcar são realizados ensaios de campo para a avaliação dos clones, usualmente denominados *test lines*, com o objetivo de maximizar o ganho genético e conseqüentemente, selecionar as melhores *test lines*, o que justifica o uso de delineamentos experimentais eficientes.

É importante ressaltar que tais experimentos das fases iniciais do melhoramento consideram em um único ensaio centenas de *test lines*, o que implica em limitações relacionadas à disponibilidade de material genético para a realização de várias repetições de um mesmo tratamento (*test line*) e também à área experimental.

Historicamente, os experimentos eram instalados seguindo delineamentos *grid-plot*, que podem ser sistemáticos ou delineamentos aumentados (Federer, 1956), que permitem um certo controle da heterogeneidade ambiental e conseqüentemente a diminuição do erro experimental Kempton (1984).

Uma característica desses ensaios é a presença de uma única parcela por experimento associada a cada *test line*, ou seja, os tratamentos de maior interesse não apresentam repetições. A variabilidade ambiental é estimada a partir de repetições de variedades padrões, que são variedades consolidadas no mercado, e os efeitos dos tratamentos preditos a partir de uma única observação.

Buscando contornar o problema da ausência de repetições para as *test lines*, Cullis et al. (2006) propuseram os delineamentos parcialmente replicados (*p-rep*), em que *p* corresponde a proporção de *test lines* repetidas no experimento. Os autores sugerem a substituição das parcelas ocupadas por repetições das variedades padrões por uma repetição adicional de parte das *test lines*, sendo os delineamentos obtidos espacialmente otimizados e o critério *C* de otimização adotado, ou seja, o delineamento com menor variância média para os contrastes simples das *test lines*.

Tendo como referência o estudo realizado por Santos (2017), o objetivo do presente trabalho consiste em comparar delineamentos espacialmente otimizados, por meio de estudos de simulação, quanto ao ganho genético e qualidade de seleção de acordo com (i) tipo de delineamento, se *grid plot* ou *p-rep*; (ii) número de variedades padrões para uma fixada porcentagem de *test lines* repetidas; (iii) porcentagem de

*test lines* repetidas; (iv) critérios de otimização, se  $C$  ou  $D$ , considerando-se dois valores para a variância genética e a presença ou ausência de efeitos aleatórios globais, para uma área experimental de 10 colunas por 24 linhas e 180 *test lines*.

## 2 REVISÃO DE LITERATURA

Nesta seção serão apresentados os principais delineamentos experimentos utilizados em fases iniciais da experimentação dos programas de melhoramento genético de cana-de-açúcar, sendo estes os delineamentos *grid-plot*, que podem ser sistemáticos ou delineamentos aumentados e os delineamentos *p-rep*, como bem resume Clarke e Stefanova (2011).

Como tais delineamentos apresentam confundimento entre os efeitos de interesse, *test lines*, e os efeitos de controle local, uma vez que nenhuma ou apenas parte destas são repetidas, deve-se trabalhar com delineamentos ótimos, que em grande parte dos casos, buscam maximizar a informação relacionada aos tratamentos. Desse modo, serão apresentados alguns dos critérios de otimização.

Ainda, com o objetivo de comparar diferentes delineamentos a seção 2.3 traz diferentes medidas de comparação.

### 2.1 Delineamentos Experimentais

#### 2.1.1 Delineamentos Grid-plot

Buscando métodos de controle local para a fertilidade do solo e também uma solução para o problema de escassez de material genético, que impossibilita a repetição de todas as *test lines*, Kempton (1984) cita o uso de parcelas contendo uma ou mais variedades padrões, denominadas *checks*. As variedades padrões são distribuídas em toda a área experimental entre as parcelas que contêm as variedades em teste, como forma de fornecer ferramentas para quantificar a variância global, que é aquela que atinge todo o campo, como descrito por Gilmour et al. (1997). Cullis et al. (2006) chamam esses delineamentos de *grid-plot*.

A vantagem do *grid-plot* é que não é essencial aleatorizar os tratamentos, podendo até mesmo usar um arranjo sistemático, como por exemplo alocar as variedades padrões às parcelas em diagonal da área experimental considerando-se uma distância fixada ou um arranjo chamado movimento do cavaleiro (*knight's move*). Exemplos para os dois delineamentos sistemáticos são apresentados nas Figuras 2.1 (a) e 2.1 (b), respectivamente.

Segundo Kempton (1984) a escolha das variedades padrões deve ser feita de forma cuidadosa a fim de diminuir a variância do erro e não aumentá-la. O mesmo autor sugere a utilização de ao menos duas variedades com fundo genético semelhante ao das variedades em teste. Geralmente, quão maior a quantidade de parcelas ocupadas pelas variedades padrões maior a eficácia do controle local. Entretanto, o tamanho do experimento tem um papel importante, em experimentos com o número de linhas e colunas fixados, o aumento no número de parcelas ocupadas por padrões diminuiria o número de parcelas disponíveis para as *test lines*.

Clarke e Stefanova (2011) concluíram, a partir de estudos de simulação em que os erros experimentais foram considerados de um ensaio de uniformidade, que aumentando o número de parcelas ocupadas por padrões resultará em um aumento no número de graus de liberdade do resíduo, permitindo assim o ajuste de modelos espaciais mais sofisticados. Entretanto, a posição de tais padrões no experimento pouco influencia a qualidade de seleção do material genético, assim como consideração de mais variedades padrões com um número menor de repetições ou menos variedades padrões repetidas um maior número de vezes, para um dado número de graus de liberdade do resíduo.

(a)		(b)																																																																																																																																																										
Linhas	<table border="1"> <tr><td>10</td><td>7</td><td>C</td><td>24</td><td>20</td><td>31</td><td>D</td></tr> <tr><td>9</td><td>A</td><td>5</td><td>33</td><td>11</td><td>B</td><td>22</td></tr> <tr><td>8</td><td>35</td><td>16</td><td>27</td><td>A</td><td>21</td><td>13</td></tr> <tr><td>7</td><td>36</td><td>34</td><td>C</td><td>29</td><td>3</td><td>28</td></tr> <tr><td>6</td><td>10</td><td>D</td><td>2</td><td>12</td><td>14</td><td>C</td></tr> <tr><td>5</td><td>B</td><td>39</td><td>19</td><td>32</td><td>A</td><td>23</td></tr> <tr><td>4</td><td>18</td><td>26</td><td>15</td><td>B</td><td>44</td><td>8</td></tr> <tr><td>3</td><td>37</td><td>17</td><td>D</td><td>6</td><td>25</td><td>43</td></tr> <tr><td>2</td><td>4</td><td>C</td><td>41</td><td>38</td><td>40</td><td>B</td></tr> <tr><td>1</td><td>A</td><td>42</td><td>9</td><td>1</td><td>D</td><td>30</td></tr> <tr> <td></td> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td> </tr> </table>	10	7	C	24	20	31	D	9	A	5	33	11	B	22	8	35	16	27	A	21	13	7	36	34	C	29	3	28	6	10	D	2	12	14	C	5	B	39	19	32	A	23	4	18	26	15	B	44	8	3	37	17	D	6	25	43	2	4	C	41	38	40	B	1	A	42	9	1	D	30		1	2	3	4	5	6	<table border="1"> <tr><td>10</td><td>48</td><td>B</td><td>32</td><td>37</td><td>45</td><td>3</td></tr> <tr><td>9</td><td>44</td><td>46</td><td>10</td><td>C</td><td>5</td><td>6</td></tr> <tr><td>8</td><td>A</td><td>19</td><td>9</td><td>20</td><td>13</td><td>D</td></tr> <tr><td>7</td><td>43</td><td>41</td><td>C</td><td>34</td><td>15</td><td>47</td></tr> <tr><td>6</td><td>4</td><td>26</td><td>30</td><td>25</td><td>D</td><td>42</td></tr> <tr><td>5</td><td>7</td><td>B</td><td>27</td><td>18</td><td>31</td><td>28</td></tr> <tr><td>4</td><td>29</td><td>2</td><td>11</td><td>A</td><td>22</td><td>39</td></tr> <tr><td>3</td><td>D</td><td>24</td><td>8</td><td>33</td><td>36</td><td>A</td></tr> <tr><td>2</td><td>17</td><td>16</td><td>C</td><td>1</td><td>12</td><td>40</td></tr> <tr><td>1</td><td>35</td><td>14</td><td>38</td><td>23</td><td>B</td><td>21</td></tr> <tr> <td></td> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td> </tr> </table>	10	48	B	32	37	45	3	9	44	46	10	C	5	6	8	A	19	9	20	13	D	7	43	41	C	34	15	47	6	4	26	30	25	D	42	5	7	B	27	18	31	28	4	29	2	11	A	22	39	3	D	24	8	33	36	A	2	17	16	C	1	12	40	1	35	14	38	23	B	21		1	2	3	4	5	6
10	7	C	24	20	31	D																																																																																																																																																						
9	A	5	33	11	B	22																																																																																																																																																						
8	35	16	27	A	21	13																																																																																																																																																						
7	36	34	C	29	3	28																																																																																																																																																						
6	10	D	2	12	14	C																																																																																																																																																						
5	B	39	19	32	A	23																																																																																																																																																						
4	18	26	15	B	44	8																																																																																																																																																						
3	37	17	D	6	25	43																																																																																																																																																						
2	4	C	41	38	40	B																																																																																																																																																						
1	A	42	9	1	D	30																																																																																																																																																						
	1	2	3	4	5	6																																																																																																																																																						
10	48	B	32	37	45	3																																																																																																																																																						
9	44	46	10	C	5	6																																																																																																																																																						
8	A	19	9	20	13	D																																																																																																																																																						
7	43	41	C	34	15	47																																																																																																																																																						
6	4	26	30	25	D	42																																																																																																																																																						
5	7	B	27	18	31	28																																																																																																																																																						
4	29	2	11	A	22	39																																																																																																																																																						
3	D	24	8	33	36	A																																																																																																																																																						
2	17	16	C	1	12	40																																																																																																																																																						
1	35	14	38	23	B	21																																																																																																																																																						
	1	2	3	4	5	6																																																																																																																																																						
	Colunas	Colunas																																																																																																																																																										

**Figura 2.1.** Delineamentos *grid-plot* em um arranjo com 10 linhas por 6 colunas e quatro variedades padrões (A, B, C e D). Em (a) as padrões são repetidas quatro vezes e alocadas em diagonal e tem-se 44 *test lines* (1, ..., 44) não repetidas; (b) as padrões são repetidas três vezes e ocupam as parcelas de acordo com o movimento do cavaleiro, tem-se 48 *test lines* (1, ..., 48) não repetidas

### 2.1.1.1 Delineamentos Aumentados

De acordo com Federer (1956) os delineamentos sistemáticos, como os apresentados na Figura 2.1, eram ineficientes com relação à comparação dos tratamentos em fases iniciais da experimentação em melhoramento genético de plantas. Como alternativa o mesmo propôs uma classe de delineamentos, os delineamento aumentados.

O primeiro delineamento aumentado proposto foi o delineamento em blocos aumentados ou também conhecido como delineamento de Federer. Neste, assume-se um delineamento em blocos para as variedades padrões, ou seja, cada repetição das padrões ocorre uma única vez em cada bloco e às demais parcelas em cada bloco alocam-se as *test lines* não repetidas. Um exemplo pode ser encontrado na Figura 2.2.

Uma vez que delineamentos experimentais em fases iniciais do melhoramento de plantas tendem a ocupar áreas retangulares no campo, comumente são utilizados delineamentos em linhas e colunas, os quais consideram o controle local nas duas direções. Exemplo de delineamento em linhas e colunas é o quadrado latino, que é um delineamento linha-coluna-ortogonal, uma vez que o número de tratamentos a serem avaliados no experimento é igual ao número de linhas e ao número de colunas e ainda, cada tratamento ocorre uma única vez em cada linha e uma única vez em cada coluna (ver Figura 2.3 (a)).

Outro exemplo de delineamento em linhas e colunas é o quadrado de Youden. Trata-se um delineamento linha-ortogonal em que o componente do delineamento relacionado às colunas é um delineamento em blocos incompletos balanceados (John e Williams, 1995). Uma ilustração para o quadrado de Youden pode ser encontrada na Figura 2.3 (b).

Federer (1956); Federer e Raghavarao (1975) também sugeriram o delineamento quadrado latino aumentado e o quadrado de Youden aumentado ( Figura 2.4 (a) e (b), respectivamente).

34	31	13	28	19	B	45	41	C	A
D	32	38	44	21	40	10	8	23	15
16	17	5	D	35	25	3	29	A	11
B	30	C	20	18	6	27	47	7	24
A	1	2	22	42	9	39	43	46	33
12	37	C	D	4	36	26	48	B	14

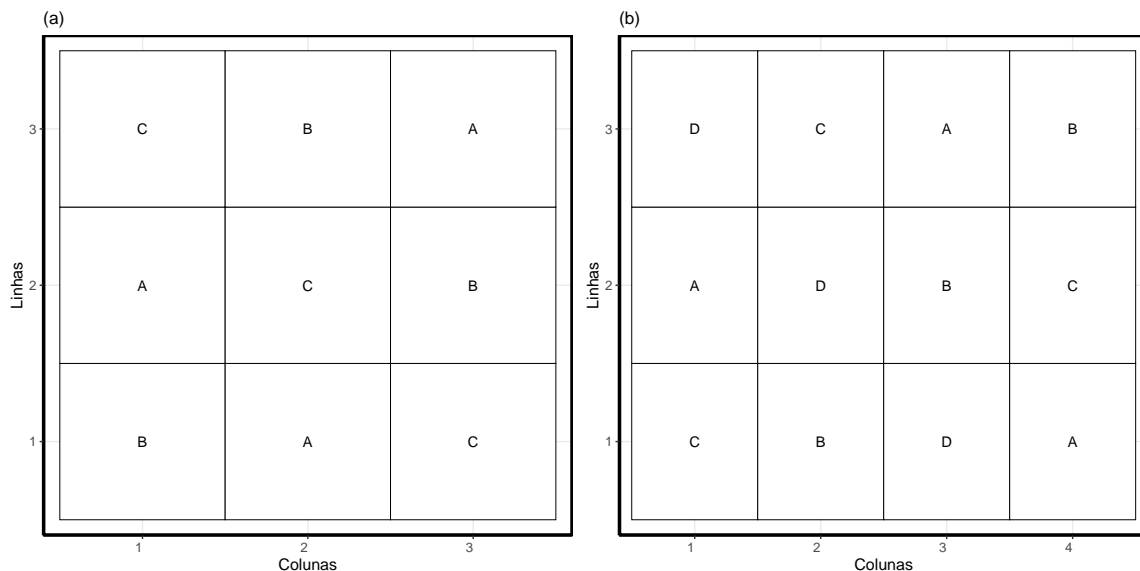
**Figura 2.2.** Delineamento em blocos aumentados com quatro variedades padrões (A, B, C e D) repetidas três vezes e 48 *test lines* não repetidas

Federer et al. (1975) modificaram o quadrado de Youden com  $k$  linhas e  $v$  tratamentos, com o objetivos de obter um arranjo quadrado  $v \times v$  em que  $k$  variedades padrões são repetidas  $v$  vezes, propondo, dessa forma, uma inversão de papéis das linhas e dos tratamentos. Com o objetivo de facilitar a compreensão do delineamento modificado observe a Figura 2.5, esta apresenta um delineamento seguindo o quadrado de Youden com sete colunas, e portanto sete tratamentos (1, ..., 7), e três linhas (A, B e C).

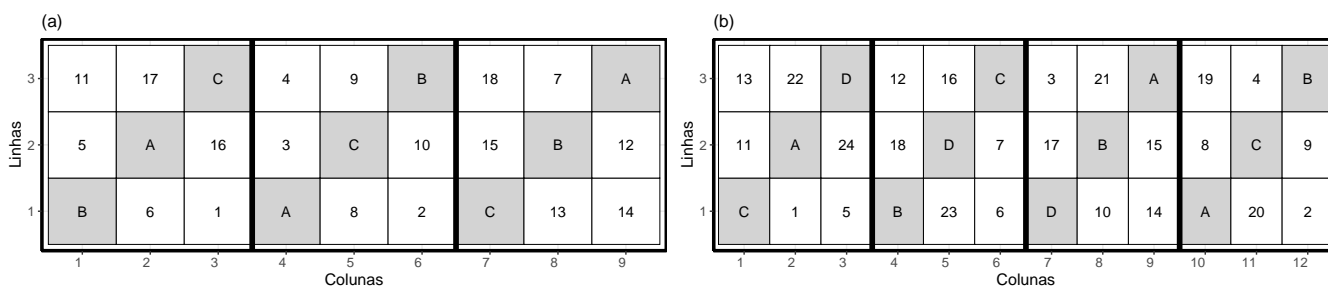
A proposta dos autores consiste em criar um arranjo  $7 \times 7$  e alocar as variedades padrões (A, B e C), que antes representavam as linhas, às linhas que ocupam as posições dos tratamentos indicados na Figura 2.5. O resultado é apresentado na Figura 2.6.

Lin e Poushinsky (1983) estudando delineamentos aumentados, sugeriram um menor comprometimento do número de parcelas ocupadas pelas variedades padrões, sendo a indicação de 13% a 17% do número total de parcelas no experimento. Para delineamentos em áreas experimentais quadradas ou aproximadamente quadradas os autores propuseram uma estrutura em parcelas subdivididas, em que qualquer delineamento poderia ser utilizado para as parcelas, porém as sub-parcelas devem ser arranjos





**Figura 2.3.** (a) Delineamento quadrado latino para três tratamentos e (b) Delineamento quadrado de Youden com cinco tratamentos em uma área experimental com quatro linhas e cinco colunas



**Figura 2.4.** (a) Delineamento quadrado latino aumentado para três variedades padrões e 18 *test lines* em uma área experimental com três linhas e nove colunas e (b) Delineamento quadrado de Youden aumentado com quatro variedades padrões e 24 *test lines* em uma área experimental com três linhas e 12 colunas

do tipo  $3 \times 3$ , sendo a parcela central ocupada pela variedade padrão (parcela), eventualmente uma parcela vizinha ocupada pela mesma variedade padrão e as sete ou oito vizinhas ocupadas por *test lines* não repetidas, conforme o exemplo na Figura 2.7. As parcelas que apresentam repetição da variedade padrão em alguma de suas sub-parcelas permitirão a estimação do erro das sub-parcelas.

Os delineamentos propostos por Lin e Poushinsky (1983) fornecem uma parametrização menos detalhada das diferenças entre linhas e colunas, além de uma estimação da variância residual menos apropriada, quando comparados com os delineamentos aumentados de Federer e Raghavarao (1975). Entretanto, parecem ser mais robustos para um número de *test lines* superior a 120.

Tratando-se de áreas experimentais retangulares, LIN e POUHINSKY (1985) apresentaram um segundo tipo de delineamento aumentado. Neste as parcelas são como no primeiro tipo e as sub-parcelas arranjadas em cinco linhas dentro de cada parcela, sendo a sub-parcela central ocupada por um tratamento controle (ver exemplo, Figura 2.8). Do mesmo modo que no tipo 1, algumas parcelas são aleatoriamente selecionadas para receber variedades padrões adicionais, de tal modo a permitir a estimação do erro das sub-parcelas. Os autores concluíram que o uso de cinco sub-parcelas, ou um número ímpar que preserve a forma quadrada para parcela, são suficientes para manter a homogeneidade das sub-parcelas e ainda permitir a avaliação de um padrão de variação no solo.

C	7	2	5	6	3	1	4
B	2	6	1	3	4	7	5
A	1	7	4	2	6	5	3
	1	2	3	4	5	6	7

**Figura 2.5.** Delineamento quadrado de Youden com três linhas e sete tratamentos

7	C	A	11	4	13	B	28
6	5	B	2	C	A	8	22
5	20	7	C	26	24	A	B
4	6	18	A	25	B	16	C
3	3	27	14	B	C	1	A
2	B	C	9	A	15	12	10
1	A	19	B	17	23	C	21
	1	2	3	4	5	6	7

**Figura 2.6.** Delineamento quadrado de Youden modificado (aumentado) com sete linhas, sete colunas, três variedades padrões repetidas sete vezes e 28 *test lines*

Delineamentos aumentados continuaram sendo objeto de estudo de vários outros pesquisadores. Federer (2002) avaliou latices quadrados aumentados e Piepho e Williams (2016) publicaram uma estratégia para gerar delineamentos em linhas e colunas aumentados com a adição de fatores de controle local, às linhas e colunas, seguindo a proposta dos delineamentos do tipo SuDoKu propostos por Sarkar e Sinha (2015). Mais trabalhos relacionados são: Federer (1961), Bhagyalakshmi e Somarajan (1999), Federer e Crossa (2001) e Federer (2005), entre outros.

### 2.1.2 Delineamentos Parcialmente Replicados

Como alternativa aos delineamentos *grid-plot* e aos delineamentos aumentados, Cullis et al. (2006) propuseram os delineamentos parcialmente replicados, ou comumente denominado, delineamentos *p-rep*. A ideia central consistiu da substituição das parcelas ocupadas pelas variedades padrões por *test lines*. Desse modo, parte dos tratamentos ( $p\%$ ) apresentam repetição e parte é não repetida  $(100-p)\%$ . De acordo com os autores a escolha das variedades a serem replicadas pode ser feita ao acaso, pela disponibilidade do material genético ou pela escolha do produtor.

12	51	99	37	A	91	81	62	90	33	7	55	11
11	66	B	4	104	A	75	27	C	41	D	D	5
10	72	120	20	17	50	52	C	92	32	113	2	57
9	22	101	87	111	103	18	58	86	48	35	70	25
8	31	C	54	95	B	B	24	D	13	80	A	89
7	36	12	C	94	107	16	117	47	67	19	85	49
6	D	44	96	39	112	14	10	61	64	59	60	115
5	106	D	46	28	C	105	A	A	119	6	B	3
4	109	98	71	34	110	40	83	77	38	15	88	76
3	97	79	100	53	56	69	93	21	29	43	30	73
2	84	A	68	63	D	42	114	B	116	82	C	9
1	8	23	26	1	74	65	108	45	B	102	118	78
	1	2	3	4	5	6	7	8	9	10	11	12

Figura 2.7. Delineamento aumentado tipo 1 de acordo com Lin e Poushinsky (1983), com 12 linhas, 12 colunas, quatro variedades padrões (A, B, C e D) repetidas seis vezes e 120 *test lines* (1, ..., 120) não repetidas

25	15	49	C	5	83	81
24	24	69	73	35	45	61
23	A	A	A	A	A	A
22	94	13	B	34	47	21
21	41	89	53	44	70	65
20	93	36	9	38	96	50
19	C	98	C	26	7	43
18	A	A	A	A	A	A
17	17	66	B	85	59	91
16	B	88	84	19	95	39
15	37	77	48	76	B	C
14	82	C	51	67	C	100
13	A	A	A	A	A	A
12	42	27	40	62	14	60
11	22	B	4	32	3	B
10	46	10	B	23	55	30
9	78	6	90	56	1	92
8	A	A	A	A	A	A
7	2	C	57	25	28	16
6	63	B	C	8	72	97
5	B	52	79	29	B	87
4	75	31	74	33	99	18
3	A	A	A	A	A	A
2	C	12	86	11	64	54
1	80	58	68	20	C	71
	1	2	3	4	5	6

Figura 2.8. Delineamento aumentado tipo 2 de acordo com LIN e POUSHINSKY (1985), com 25 linhas, 6 colunas, três variedades padrões, A, B e C, repetidas 30, 10 e 10 vezes, respectivamente, e 100 *test lines* (1, ..., 100) não repetidas

Inicialmente os *p*-rep tratavam de delineamentos em linhas e colunas resolvíveis, ou seja, a área experimental é dividida em dois blocos e cada bloco recebe uma única repetição das *test lines*. Desse modo, cada bloco corresponde a uma repetição completa dos tratamentos duplicados. Exemplo de um delineamento *p*-rep pode ser observado na Figura 2.9.

Vários autores avaliaram o desempenho dos delineamentos *p*-rep quando comparados com *grid*-

24	3	22	108	101	94	19
23	5	7	85	79	100	83
22	77	109	9	14	13	17
21	8	86	25	102	2	96
20	23	28	20	113	6	10
19	73	12	107	1	87	89
18	110	95	84	80	15	75
17	112	106	82	92	18	114
16	76	4	97	104	105	90
15	99	103	111	78	24	115
14	81	74	26	27	29	16
13	91	11	98	93	21	88
12	12	69	68	32	54	24
11	65	26	28	31	51	48
10	35	60	71	5	43	67
9	58	18	25	39	34	41
8	33	66	4	23	72	19
7	40	1	64	49	21	70
6	11	10	44	15	56	13
5	47	63	22	38	42	2
4	62	30	61	50	52	37
3	9	59	3	46	6	20
2	14	45	17	7	16	27
1	29	8	55	53	36	57

**Figura 2.9.** Exemplo de um delineamento  $p$ -rep com 115 *test lines*, sendo 29 duplicadas (em cinza) e 86 não repetidas (em branco) em uma área com 24 linhas e seis colunas, em que as 12 primeiras linhas correspondem a um bloco e as demais a outro bloco, caracterizando um delineamento resolvível

*plots* e delineamentos aumentados. Estes afirmam que os delineamentos  $p$ -rep mostram-se superiores ao selecionar o material genético (Cullis et al., 2006; Smith, A. and Thompson, R. and Butler, D. and Cullis, 2011; Clarke e Stefanova, 2011; Stringer et al., 2011; Williams et al., 2011; Hunt et al., 2013; Moehring et al., 2014; Williams et al., 2014; Smith et al., 2015; Piepho e Williams, 2016; Santos, 2017).

É importante observar que os delineamentos  $p$ -rep são obtidos por meio de delineamentos ótimos ou aproximadamente ótimos, pois não é um delineamento balanceado, que apresenta os mesmo número de repetições para todos os tratamentos, além de considerar valores para os parâmetros de interesse.

## 2.2 Delineamentos Ótimos

Uma forma de se obter um delineamento eficiente é a utilização de delineamentos ótimos. A teoria de delineamentos experimentais ótimos está centrada na caracterização e na construção de delineamentos que são ótimos em algum sentido, ou seja, no modelo adotado para o delineamento e no critério de busca (Shah e Sinha, 1989).

Grande parte dos pesquisadores faz uso de critérios relacionados à matriz de informação dos tratamentos, que por sua vez depende do modelo do delineamento. Seja um modelo genérico para uma variável resposta  $Y$ , dado por

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_g\mathbf{u}_g + \mathbf{Z}_0\mathbf{u}_0 + \mathbf{e}, \quad (2.1)$$

em que  $\mathbf{y}$  é o vetor de observações ( $n \times 1$ ),  $\boldsymbol{\beta}$  é o vetor ( $h \times 1$ ) de efeitos fixos com matriz de delineamento  $\mathbf{X}$  ( $n \times h$ ),  $\mathbf{u}_g$  é o vetor de efeitos de tratamentos ( $m \times 1$ ) com matriz de delineamento  $\mathbf{Z}_g$  ( $n \times m$ ),  $\mathbf{u}_0$  é o vetor de efeitos aleatórios ( $b \times 1$ ) com matriz de delineamento  $\mathbf{Z}_0$  ( $n \times b$ ) e  $\mathbf{e}$  é o vetor de erros aleatórios ( $n \times 1$ ). Assume-se, geralmente,  $\mathbf{u}_0 \sim N(\mathbf{0}, \mathbf{G}_0)$  e  $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$  e caso  $\mathbf{u}_g$  seja aleatório,  $\mathbf{u}_g \sim N(\mathbf{0}, \mathbf{G}_g)$ .

A matriz de informação de tratamentos,  $\mathbf{C}$ , de acordo com Hooks et al. (2009), é dada por:

$$\mathbf{C} = \mathbf{Z}'_g(\mathbf{Z}_0\mathbf{G}_0\mathbf{G}'_0 + \mathbf{R})^{-1}\mathbf{Z}_g + \mathbf{G}_g^{-1} - \mathbf{Z}'_g(\mathbf{Z}_0\mathbf{G}_0\mathbf{Z}'_0 + \mathbf{R})^{-1}\mathbf{X}(\mathbf{X}'(\mathbf{Z}_0\mathbf{G}_0\mathbf{Z}'_0 + \mathbf{R})^{-1}\mathbf{X})^{-1}\mathbf{X}'(\mathbf{Z}_0\mathbf{G}_0\mathbf{Z}'_0 + \mathbf{R})^{-1}\mathbf{Z}_0. \quad (2.2)$$

Para os casos em que os efeitos de tratamentos são fixos, assume-se  $\mathbf{G}_g^{-1} = \mathbf{0}$ .

Um critério comumente utilizado é aquele que minimiza a expressão 2.2,

$$\text{AVPD} = \frac{2}{t-1} \left( \text{tr}(\mathbf{C}^{-1}) - \frac{1}{t} \mathbf{1}'\mathbf{C}^{-1}\mathbf{1} \right), \quad (2.3)$$

em que  $t$  corresponde ao número de tratamentos,  $\mathbf{C}$  é a matriz de informação apresentada na equação 2.2 e  $\text{tr}()$  denota o traço de uma matriz. Tais delineamentos apresentam a média (ou soma) da variância dos erros de predição dos contrastes simples para efeitos de tratamentos, ou seja, diferença entre pares de efeitos, mínima.

Para ensaios contendo *test lines* e variedades padrões Martin et al. (2006) adaptou o critério que minimiza o AVPD (*average variance of pairwise differences* - Equação 2.3), denotando por  $A_{nn}$  o critério que toma o delineamento com mínima variância dos erros de predição entre pares de *test lines* somente,  $A_{ss}$  o critério que toma o delineamento com mínima variância dos erros de predição entre pares de variedades padrões e  $A_{ns}$  o critério que toma o delineamento com mínima a variância dos erros de predição entre pares de *test lines* e variedades padrões. Os autores concluíram que os delineamentos apresentam comportamentos distintos em termos de eficiência relativa dependendo do número de *test lines* e da correlação residual, sendo a eficiência pouco afetada para quando ambos são altos.

Outro critério usado com frequência é o que minimiza a média (ou a soma) da variância dos erros de predição (APEV - *average prediction error variance*), ou seja, busca por delineamentos com valor mínimo para

$$\text{APEV} = \text{tr}(\mathbf{C}^{-1}). \quad (2.4)$$

Na literatura pode-se encontrar os dois critérios apresentados como critério  $A$ . Isso se justifica pelo fato de que quando os efeitos de tratamentos e os efeitos de controle local presentes no modelo são todos de efeito fixo, os critérios são equivalentes. Neste trabalho será adotada a seguinte notação, o critério  $A$  será o que minimiza a APEV (Equação 2.4) e o critério  $C$  o que minimiza a AVPD (Equação 2.3).

Ainda, considerando-se funções da matriz de informação dos tratamentos, tem-se o critério  $D$ , que busca por delineamentos com mínima variância generalizada dos erros de predição (GPEV - *generalized prediction error variance*). Assim, delineamentos  $D$ -ótimos são o que apresentam mínimo valor para

$$\text{GPEV} = |\mathbf{C}^{-1}|, \quad (2.5)$$

em que  $|\cdot|$  denota o determinante de uma matriz.

### 2.3 Comparando Delineamentos

É evidente que as medidas utilizadas no critério para busca do delineamento ótimo podem ser usadas para a comparação dos delineamentos obtidos. Porém, medidas adicionais podem ser utilizadas.

Cullis et al. (2006) compararam delineamentos por meio do ganho genético percebido (RGG - *realized genetic gain*), sendo quão maior melhor, o que, de acordo com Smith et al. (2006), equivale a um menor APEV. O RGG corresponde à média dos  $s\%$  superiores EBLUPs (*empirical best linear unbiased prediction*) relativo ao verdadeiro ganho genético, ou seja, à média dos  $s\%$  superiores verdadeiros efeitos de *test lines*, em que  $s\%$  corresponde a porcentagem de seleção do material genético. É importante ressaltar

que o cálculo do RGG é possível apenas com estudos de simulação, pois os verdadeiros efeitos de *test lines* são desconhecidos na prática. Tal medida também foi utilizada por Santos (2017), que comparou delineamentos *grid-plot*, *p*-rep e uma combinação destes, e por Sermarini et al. (2018) que avaliaram os delineamentos *p*-rep sob diferentes condições.

Müller et al. (2010) utilizaram a média da variância empírica e Clarke e Stefanova (2011) uma medida do erro padrão das comparações entre *test lines* e variedades padrões como uma proporção do erro padrão em delineamento inteiramente casualizado.

Alternativamente, Williams e Piepho (2013) fizeram uso da eficiência relativa, que consiste da média harmônica dos autovalores não nulos da matriz de informação dos tratamentos, sob a suposição de que o único termo aleatório no modelo é referente ao erro.

Também com estudos de simulação, medida como a correlação entre os EBLUPs e os verdadeiros efeitos de *test lines* foi utilizada por Fouilloux e Lalölë (2001) Smith, A. and Thompson, R. and Butler, D. and Cullis (2011) e Moehring et al. (2014), em que Moehring et al. (2014) calcularam a correlação apenas para a porcentagem de seleção. Santos (2017) e Sermarini et al. (2018) também utilizaram como medida secundária de comparação a assertividade, que é definida pela porcentagem das verdadeiras melhores *test lines* selecionadas.



### 3 METODOLOGIA

Na presente seção serão apresentados os modelos para a obtenção dos delineamentos experimentais, bem como os critérios utilizados, sendo avaliados delineamentos *grid-plot* e *p-rep*, para os quais as diferenças consistem no número de padrões e suas repetições, para o primeiro tipo e na proporção de *test lines* repetidas para o segundo tipo de delineamento, caracterizando os cenários de estudo. Para ser possível a comparação entre os delineamentos por meio do ganho genético e da assertividade, serão simulados dados. Logo, esta seção traz também o modelo para tal simulação e as medidas utilizadas para a comparação.

#### 3.1 Gerando delineamentos

Tratando-se de delineamentos ótimos deve-se trabalhar com duas suposições: (i) o modelo para a obtenção dos mesmos e (ii) o critério a ser utilizado, como visto no capítulo anterior. De acordo com Butler et al. (2014) o modelo utilizado na busca pelo delineamento ótimo ou quase-ótimo deve ser tão próximo da realidade quanto possível. Entretanto, esta proximidade envolve suposições como assumir efeitos fixos ou aleatórios e quando aleatórios sua correspondente distribuição e componentes de variância.

Para obter resultados que se aproximem da realidade usou-se, assim como Gilmour et al. (1997), Cullis et al. (2006), Santos (2017) e Sermarini et al. (2018), modelos mistos que incorporam termos para avaliação local, global e externa. Os efeitos de tratamentos, sejam as variedades padrões para o caso dos *grid-plots* ou as *test lines* foram assumidas de efeitos fixo, como sugerido por Sermarini et al. (2018).

Desse modo, uma generalização do modelo para a obtenção dos delineamentos é apresentada na Equação 2.1, repetida a seguir, com as suposições que seguem:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_g\mathbf{u}_g + \mathbf{Z}_0\mathbf{u}_0 + \mathbf{e},$$

em que,

$$\mathbb{E} \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \text{ e } \text{Var} \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{e} \end{bmatrix} = \sigma^2 \begin{bmatrix} \mathbf{G}_o(\boldsymbol{\gamma}_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{R}(\boldsymbol{\phi}) \end{bmatrix},$$

com  $\boldsymbol{\gamma}_0 = (\gamma_c, \gamma_l)'$ , sendo  $\gamma_c = \sigma_c^2/\sigma^2$  e  $\gamma_l = \sigma_l^2/\sigma^2$ ,  $\sigma_c^2$  e  $\sigma_l^2$  são respectivamente, as variâncias entre colunas e entre linhas. Ainda,  $\boldsymbol{\phi} = (\phi_c, \phi_l)'$ , em que  $\phi_c$  corresponde à correlação residual entre colunas e  $\phi_l$  à correlação residual entre linhas. Adicionalmente, assumiu-se  $\mathbf{R}(\boldsymbol{\phi}) = \text{AR1}(\phi_c) \otimes \text{AR1}(\phi_l)$ , ou seja, processos separáveis autorregressivos de primeira ordem.

No presente estudo  $\phi_c = 0,6$  e  $\phi_l = 0,4$ , fixou-se valores iguais para variâncias entre colunas e entre linhas, sendo dois valores avaliados,  $\gamma_c = \gamma_l = 0$  (ausência de variação global entre colunas e entre linhas) e  $\gamma_c = \gamma_l = 0,1$  (presença de variação global entre colunas e entre linhas baixa). O conjunto de valores adotados foi baseado em Santos (2017). Sem perda de generalidade, assumiu-se  $\sigma^2 = 1$ .

Os delineamentos avaliados formam uma generalização do trabalho de Santos (2017), para os quais foram consideradas áreas experimentais de 10 colunas por 24 linhas e 180 *test lines*. A Tabela 3.1 exhibe uma caracterização das propostas de delineamentos. Estes variaram com relação ao número de variedades padrões (tratamentos padrões) e seus correspondentes número de repetições e também com relação à porcentagem de *test lines* com repetição.



**Tabela 3.1.** Caracterização das propostas de delineamentos avaliados para 180 *test lines* em uma área experimental de 10 colunas por 24 linhas

	número de tratamentos			repetições para
	não duplicados ( <i>test lines</i> )	duplicados ( <i>test lines</i> )	padrões	cada padrão
Delineamento 1 ( $p = 0\%$ )	180	0	4	15
Delineamento 2 ( $p = 11\%$ )	160	20	4	10
Delineamento 3 ( $p = 22\%$ )	140	40	4	5
Delineamento 4 ( $p = 22\%$ )	140	40	2	10
Delineamento 5 ( $p = 22\%$ )	140	40	1	20
Delineamento 6 ( $p = 33\%$ )	120	60	0	0

Os delineamentos ótimos foram obtidos por meio de dois critério de otimização, o  $C$  e o  $D$ . Para os delineamentos  $C$ -ótimos fez-se uso da biblioteca `od` (Butler, 2014) para o *software* R (R Core Team, 2017) com 20.000 movimentos para a busca pelo ótimo. Os delineamentos  $D$ -ótimos foram obtidos por meio do procedimento PROC OPTEX do *software* SAS (SAS Institute Inc., 2015) com 1.000 movimentos. O número deferente de movimento para a busca pelo delineamento ótimo ou quase-ótimo para os dois critérios se justifica pela diferença no esforço computacional. A obtenção do delineamento  $D$ -ótimo pelo PROC OPTEX exige mais tempo.

Logo, considerando-se os dois conjuntos de parâmetros não genéticos, as seis caracterizações e os dois critérios de otimização, foram estudados 24 delineamentos.

### 3.2 Medidas das comparação

As principais medidas de comparação foram RGG (Cullis et al., 2006; Santos, 2017; Sermarini et al., 2018), isto é, o ganho genético percebido e a assertividade, ambos para uma seleção de 10% das melhores *test lines*.

Como mencionado anteriormente, tais medidas dependem de simulação de dados, pois os verdadeiros efeitos genéticos são desconhecidos. Assim, o modelo utilizado para gerar os dados (Equação 3.1) foi similar ao usado para obter os delineamentos (Equação 2.1), exceto pela suposição dos efeitos aleatórios de *test lines*, dado por

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_g\mathbf{u}_g + \mathbf{Z}_0\mathbf{u}_0 + \mathbf{e}, \quad (3.1)$$

em que  $\boldsymbol{\beta}$  é o vetor de efeitos fixos (neste caso uma constante comum a todas as observações), com matriz de delineamento  $\mathbf{X}$ ,  $\mathbf{u}_g$  é o vetor de efeitos aleatórios de *test lines*, com matriz de delineamento  $\mathbf{Z}_g$ ,  $\mathbf{u}_0$  é o vetor de efeitos aleatórios de colunas e de linhas ( $\mathbf{u}_0 = (\mathbf{u}'_c, \mathbf{u}'_l)'$ ), com matriz de delineamento  $\mathbf{Z}_0$  e  $\mathbf{e}$  é o vetor de erros aleatórios. Ainda, assumiu-se que os efeitos aleatórios são normalmente distribuídos, tais que

$$E \begin{bmatrix} \mathbf{u}_g \\ \mathbf{u}_o \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \text{ e } \text{Var} \begin{bmatrix} \mathbf{u}_g \\ \mathbf{u}_o \\ \mathbf{e} \end{bmatrix} = \sigma^2 \begin{bmatrix} \mathbf{G}_g(\gamma_g) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_o(\gamma_o) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}(\phi) \end{bmatrix},$$

com  $\gamma_g = \sigma_g^2/\sigma^2$  e todos os demais termos são descritos conforme o modelo para a obtenção dos delineamentos ótimos (Equação 2.1).

Para cada um dos 24 delineamentos foram gerados dados considerando-se dois valores para a variância genética,  $\gamma_g = 0,3$  e  $\gamma_g = 0,8$ . Os valores para os parâmetros não genéticos foram os mesmos assumidos para a obtenção do modelo, ou seja,  $\gamma_c = \gamma_l = 0$  ou  $\gamma_c = \gamma_l = 0,1$ ,  $\phi_c = 0,6$ ,  $\phi_l = 0,4$  e  $\sigma^2 = 1$ . É importante observar que os dados para delineamentos ótimos considerando efeitos aleatórios de

colunas e linhas apresentaram  $\gamma_c = \gamma_l = 0, 1$  e dados para os delineamentos sem tais efeitos apresentaram  $\gamma_c = \gamma_l = 0$ .

Para diminuir o viés na comparação dos delineamentos, foram gerados três vetores de tamanho 1000 da distribuição normal padrão, um para cada um dos efeitos aleatórios,  $\mathbf{u}_g$ ,  $\mathbf{u}_0$  e  $\mathbf{e}$ . Estes foram utilizados em todas as simulações. Assim, para cada um dos cenários os vetores foram transformados, de tal modo a serem amostras aleatórias de uma distribuição normal com os parâmetros de variâncias conforme o cenário, e combinados para a geração dos dados.

Para cada um dos 1000 conjuntos de dados, para cada cenário e delineamento ajustou-se o modelo conforme a Equação 3.1 de tal modo que, para os casos com a presença de variedades padrões, seus correspondentes efeitos foram assumidos fixos. Dos modelos ajustados foram registrados os EBLUPs, para o cálculo das medidas de comparação, as estimativas dos componentes de variância e a condição de convergência do algoritmo. Todas as análises foram realizadas no PROC MIXED do *software* SAS (SAS Institute Inc., 2015).

Adicionalmente, foram calculadas as eficiências médias no último estrato da análise para os contrastes entre *test lines* para cada um dos 24 delineamentos, assumindo-se que o único efeito aleatório presente no modelo é o erro, fazendo uso da biblioteca *dae* (Brien, 2017a) para o *software* R (R Core Team, 2017) como proposto por Brien (2017b). Embora as suposições para a obtenção de tais eficiências sejam fortes, estas medidas fornecem uma ideia do quanto da informação dos tratamentos estará confundida com efeitos de linhas e de colunas, uma vez que os delineamentos estudados no presente trabalho não são ortogonais.



## 4 RESULTADOS E DISCUSSÃO

### 4.1 Delineamentos obtidos

Em todos os delineamentos as *test lines* estão distribuídas aleatoriamente, como os exemplos apresentados na Figura 4.1. Nos 24 delineamentos obtidos não foram observados agrupamentos das *test lines* repetidas próximo às fronteiras da área experimental ou um arranjo sistemático das mesmas.

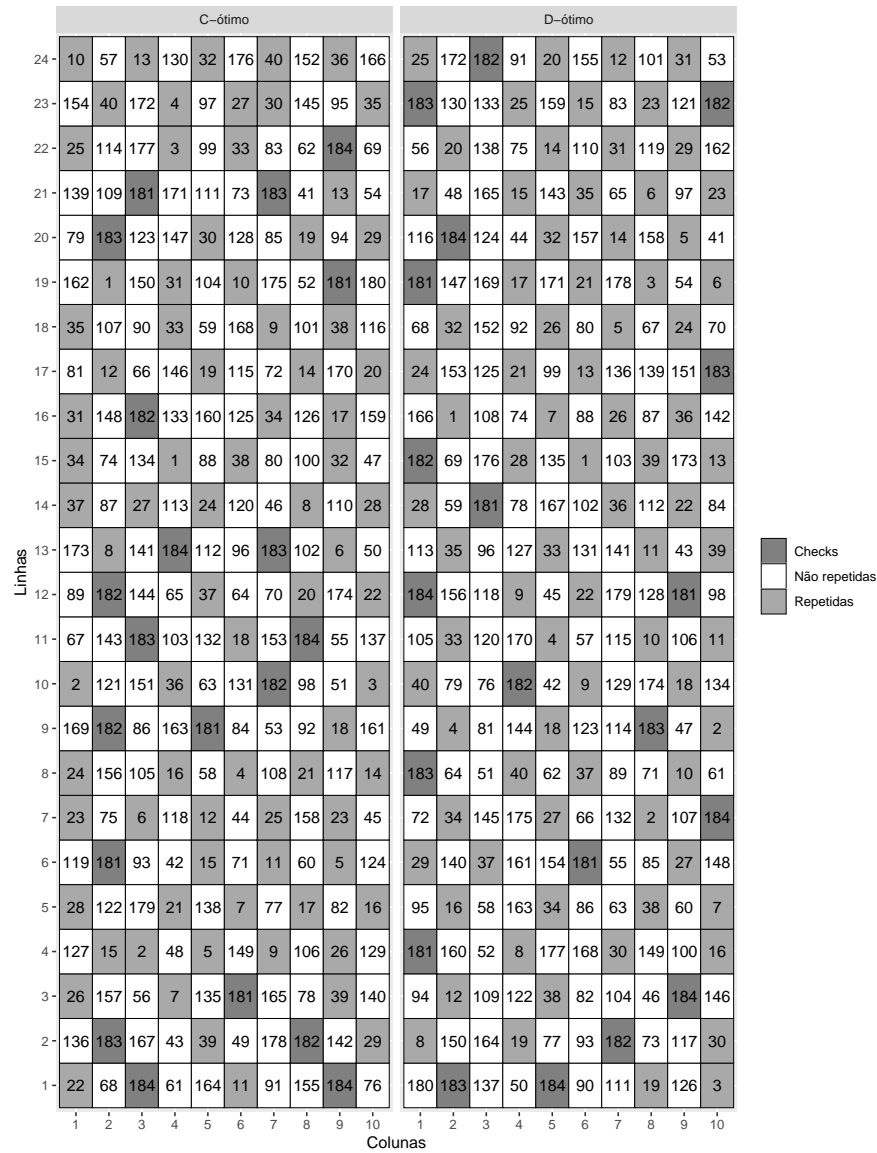


Figura 4.1. Delineamentos C-ótimo e D-ótimo para o delineamento 3, no qual 22% das *test lines* são duplicadas e tem-se quatro variedades padrões repetidas cinco vezes cada, considerando-se a presença de variâncias globais para colunas e linhas

### 4.2 Medidas de comparação

Embora sejam poucos os casos de não convergência do algoritmo, observa-se na Tabela 4.1 que tais casos ocorreram na presença de baixa variância genética,  $\gamma_g = 0,3$ , sendo que as maiores frequências

são relativas ao delineamento 1, aquele que não apresenta repetição das *test lines*.

**Tabela 4.1.** Número de casos de não convergência do algoritmo para os diferentes delineamentos, critérios de otimização e cenários para os dados, em 1000 conjuntos de dados

	$\gamma_g = 0,3$				$\gamma_g = 0,8$			
	$\gamma_l = \gamma_c = 0$		$\gamma_l = \gamma_c = 0,1$		$\gamma_l = \gamma_c = 0$		$\gamma_l = \gamma_c = 0,1$	
	C	D	C	D	C	D	C	D
Delineamento 1	14	11	5	8	0	0	0	0
Delineamento 2	2	5	5	7	0	0	0	0
Delineamento 3	2	0	3	1	0	0	0	0
Delineamento 4	5	4	3	1	0	0	0	0
Delineamento 5	5	2	1	0	0	0	0	0
Delineamento 6	1	1	5	2	0	0	0	0

Na Tabela 4.2 são apresentadas as relações entre os fatores de tratamentos, bem como as fontes de variação originadas. Tem-se que: Variedade representa os contrastes entre variedades padrões (*checks*) e *test lines*, Checks representa contrastes entre variedades padrões, Categoria representa contrastes entre todas as *test lines* duplicadas e todas as não duplicadas, Duplicada representa contrastes entre *test lines* repetidas e NDuplicada representa contrastes entre *test lines* que não apresentam repetição.

**Tabela 4.2.** Relação entre fatores de tratamentos e correspondentes fontes de variação para a obtenção das eficiências

Delineamento	Relação	Fontes de variação
1	Variedade/(Checks + + NDuplicada)	Variedade + Checks[Variedade] + + NDuplicada[Variedade]
2-4	Variedade/(Checks + + Categoria/(Duplicada + NDuplicada))	Variedade + Checks[Variedade] + + Categoria[Variedade] + + Duplicada[Variedade^Categoria] + + NDuplicada[Variedade^Categoria]
5	Variedade/Categoria/(Duplicada + + NDuplicada)	Variedade + Categoria[Variedade] + + Duplicada[Variedade^Categoria] + + NDuplicada[Variedade^Categoria]
6	Categoria/(Duplicada + NDuplicada)	Categoria + Duplicada[Categoria] + + NDuplicada[Categoria]

Os fatores de unidades, Coluna e Linha, são cruzados, tais que suas correspondentes fontes de variação são dadas por Coluna + Linha + Coluna#Linha.

As Tabelas 4.3, 4.4 e 4.5 são exemplos da anatomia dos delineamentos 1, 2 e 6 seguindo o critério *C* de otimização, em que o modelo do delineamento considerou  $\gamma_c = \gamma_r = 0,1$ . A partir dessas tabelas é possível verificar o confundimento presente entre as fontes de variação relacionadas às unidades e aos tratamentos, seus correspondentes número de graus de liberdade ortogonais e eficiências médias.

**Tabela 4.3.** Anatomia do delineamento 1 ( $p=0\%$ , com quatro padrões repetidas 15 vezes) com efeito global de colunas e linhas,  $C$ -ótimo, para 180 *test lines* em uma área experimental de 10 colunas por 24 linhas

unidades		tratamentos		eficiência	gl
fonte	gl	fonte	gl	média	ortogonais
Coluna	9	Variedade	1	0,0111	1
		Checks[Variedade]	3	0,0080	3
		NDuplicada[Variedade]	5	0,7576	5
Linha	23	Variedade	1	0,0133	1
		Checks[Variedade]	3	0,0485	3
		NDuplicada[Variedade]	19	0,7465	5
Coluna#Linha	207	Variedade	1	0,9756	1
		Checks[Variedade]	3	0,9174	3
		NDuplicada[Variedade]	179	0,5250	32
		Resíduo	24		

**Tabela 4.4.** Anatomia do delineamento 2 ( $p=11\%$ , com quatro padrões repetidas 10 vezes) com efeito global de colunas e linhas,  $C$ -ótimo, para 180 *test lines* em uma área experimental de 10 colunas por 24 linhas

unidades		tratamentos		eficiência	gl
fonte	gl	fonte	gl	média	ortogonais
Coluna	9	Variedade	1	0,0550	1
		Checks[Variedade]	3	0,0162	3
		Categoria[Variedade]	1	0,0065	1
		Duplicada[Variedade^Categoria]	4	0,0655	4
Linha	23	Variedade	1	0,0340	1
		Checks[Variedade]	3	0,0615	3
		Categoria[Variedade]	1	0,0129	1
		Duplicada[Variedade^Categoria]	17	0,0009	17
Coluna#Linha	207	NDuplicada[Variedade^Categoria]	1	0,7000	1
		Variedade	1	0,9110	1
		Checks[Variedade]	3	0,8919	3
		Categoria[Variedade]	1	0,8962	1
		Duplicada[Variedade^Categoria]	19	0,8628	19
		NDuplicada[Variedade^Categoria]	159	0,5224	33
		Resíduo	24		

**Tabela 4.5.** Anatomia do delineamento 6 ( $p=33\%$ , sem padrões) com efeito global de colunas e linhas,  $C$ -ótimo, para 180 *test lines* em uma área experimental de 10 colunas por 24 linhas

unidades		tratamentos		eficiência	gl
fonte	gl	fonte	gl	média	ortogonais
Coluna	9	Categoria	1	0,0056	1
		Duplicada[Categoria]	8	0,2124	8
Linha	23	Categoria	1	0,0167	1
		Duplicada[Categoria]	22	0,1840	22
Coluna#Linha	207	Categoria	1	0,9778	1
		Duplicada[Categoria]	59	0,8367	33
		NDuplicada[Categoria]	119	0,5816	33
		Residual	28		

Observa-se na Tabela 4.6 que no mínimo 81% da informação dos contrastes entre *test lines* duplicadas encontra-se no último estrato da análise, ou seja, Coluna#Linha. O mesmo não é observado para

a informação dos contrastes entre *test lines* não duplicadas, cujo valor máximo é 60%; uma consequência da não repetição dos tratamentos.

A diferença entre as eficiências médias para os diferentes critérios é pequena quando relacionadas aos contrastes entre *test lines* duplicadas. Entretanto, a diferença tende a ser maior quando relacionada às *test lines* não duplicadas, sendo mais evidente para delineamentos otimizados na presença de efeitos globais de colunas e linhas. Nestes casos os delineamentos *C*-ótimos foram melhores.

Delineamentos otimizados com a presença de efeitos globais de colunas ou linhas tendem a apresentar maiores valores para eficiência, especialmente quando relacionadas aos contrastes entre *test lines* não duplicadas. Observa-se, ainda nesses casos, que os delineamento parcialmente replicados tendem a ter um melhor desempenho e quão maior a porcentagem de *test lines* replicadas maior a eficiência.

Com relação a informação dos contrastes entre *test lines* duplicadas, os menores valores estão associados ao delineamento 6, o que não apresenta variedades padrões no experimento. Tal resultado contraria o que foi encontrado por Santos (2017), que encontrou maiores valores de eficiência relativa associados ao delineamento sem a presença de padrões quando considerou 414 *test lines*, das quais 86 foram duplicadas, em uma área experimental de 25 colunas por 20 linhas.

Quando comparados delineamentos com a mesma porcentagem de *test lines* repetidas, há um aparente melhor desempenho dos delineamento com menor número de variedades padrões distintas.

**Tabela 4.6.** Eficiências médias no último estrato (Coluna#Linha) para os contrastes simples entre *test lines* duplicadas (Duplicada) e entre *test lines* não duplicadas (NDuplicadas) para os 24 delineamentos

	$\gamma_c = \gamma_l = 0$				$\gamma_c = \gamma_l = 0, 1$			
	Duplicada		NDuplicada		Duplicada		NDuplicada	
	C	D	C	D	C	D	C	D
Delineamento 1			0,5059	0,3191			0,5250	0,3689
Delineamento 2	0,8593	0,8589	0,5073	0,4044	0,8628	0,8627	0,5224	0,5366
Delineamento 3	0,8399	0,8413	0,4573	0,4816	0,8475	0,8484	0,5443	0,5274
Delineamento 4	0,8448	0,8465	0,4974	0,4972	0,8501	0,8515	0,5673	0,5478
Delineamento 5	0,8521	0,8493	0,5048	0,5268	0,8476	0,8518	0,5829	0,5373
Delineamento 6	0,8326	0,8143	0,5034	0,3369	0,8367	0,8371	0,5816	0,5503

Na Tabela 4.7 são apresentadas as médias e os desvios-padrões para a medida RGG (ganho genético percebido) para todos os delineamentos. Pode-se observar que maiores valores para a variância genética ( $\gamma_g$ ) resultam em maiores RGGs, como observado por Santos (2017). Assim como em Cullis et al. (2006), Clarke e Stefanova (2011) e Santos (2017), os delineamentos parcialmente duplicados, ou seja, delineamentos 2-6, apresentaram melhores resultados, quando comparados ao delineamento 1, no qual apenas as variedades padrões são repetidas no mesmo experimento.

O melhor desempenho observado refere-se ao delineamento 6 ( $p = 33\%$  sem a presença de variedades padrões), independentemente do valor adotada para a variância genética e da presença ou ausência da variância global entre linhas e colunas, sendo de 7,0% a 8,1% maior quando  $\gamma_g = 0, 3$  e de 4,2% a 4,8% quando  $\gamma_g = 0, 8$ , quando comparado ao delineamento 1.

Outra característica importante dos delineamento *p*-rep é uma maior precisão relacionada ao RGG, uma vez que seus respectivos desvios-padrões são sempre menores.

Comparando os delineamentos 1, 2 e 3, nota-se uma tendência de crescimento do ganho genético conforme aumenta o número de *test lines* repetidas. Entretanto, quando se comparam os delineamentos 3, 4 e 5 não há evidências de que o número de variedades padrões afete o ganho genético para um fixado número de *test lines* repetidas. Tais resultados também estão em concordância com os obtidos por Clarke

e Stefanova (2011), que afirmam que o que irá influenciar o ganho genético é o número de graus de liberdade do resíduo; se o experimento apresenta mais repetições e poucas padrões ou mais padrões e poucas repetições, tal fato pouco afeta o ganho genético.

Com relação aos critérios de otimização  $C$  e  $D$ , não há diferença aparente, indicando que qualquer critério poderia ser utilizado.

**Tabela 4.7.** Médias e desvios-padrões para ganho genético percebido para cada um dos 24 delineamentos considerando-se os dois cenários para o componente de variância genética,  $\gamma_g = 0,3$  e  $\gamma_g = 0,8$

	$\gamma_c = \gamma_l = 0$		$\gamma_c = \gamma_l = 0,1$	
	Critério C	Critério D	Critério C	Critério D
$\gamma_g = 0,3$				
Delineamento 1	0,6142 (0,2159)	0,6175 (0,2195)	0,6087 (0,2273)	0,6101 (0,2126)
Delineamento 2	0,6286 (0,1968)	0,6285 (0,2046)	0,6230 (0,2122)	0,6147 (0,2152)
Delineamento 3	0,6523 (0,1896)	0,6503 (0,1928)	0,6383 (0,1917)	0,6272 (0,2056)
Delineamento 4	0,6534 (0,1839)	0,6492 (0,1876)	0,6409 (0,1913)	0,6404 (0,2055)
Delineamento 5	0,6548 (0,1924)	0,6404 (0,1839)	0,6473 (0,1882)	0,6462 (0,1981)
Delineamento 6	0,6586 (0,1785)	0,6675 (0,1845)	0,6548 (0,1789)	0,6531 (0,1865)
$\gamma_g = 0,8$				
Delineamento 1	0,7705 (0,1399)	0,7713 (0,1446)	0,7617 (0,1526)	0,7629 (0,1485)
Delineamento 2	0,7810 (0,1317)	0,7804 (0,1393)	0,7799 (0,1404)	0,7728 (0,1426)
Delineamento 3	0,7967 (0,1335)	0,7961 (0,1328)	0,7869 (0,1294)	0,7768 (0,1386)
Delineamento 4	0,7987 (0,1213)	0,7958 (0,1282)	0,7854 (0,1316)	0,7888 (0,1399)
Delineamento 5	0,8005 (0,1287)	0,7891 (0,1253)	0,7912 (0,1286)	0,7892 (0,1366)
Delineamento 6	0,8048 (0,1192)	0,8086 (0,1263)	0,7965 (0,1233)	0,7951 (0,1263)

As correlações entre os EBLUPs e os verdadeiros efeitos genéticos estão entre 0,62 e 0,66 para  $\gamma_g = 0,3$  e 0,77 e 0,81 para  $\gamma_g = 0,8$  (Tabela 4.8), o que indica que para uma baixa variância genética os valores preditos não estão tão próximos dos verdadeiros efeitos. Assim como para o RGG, quão maior a variância genética, maior a correlação.

O delineamento 6 apresentou o melhor desempenho e o delineamento 1 o pior, sendo os valores para o delineamento 6 de 6,1% a 6,7% maiores do que os observados para o delineamento 1 quando  $\gamma_g = 0,3$  e de 3,7% a 3,9% quando  $\gamma_g = 0,8$ . Os desvios-padrões foram similares entre os delineamentos. Os resultados para as demais comparações entre os delineamentos e critérios utilizados foram similares aos observados para o RGG.



**Tabela 4.8.** Médias e desvios-padrões para correlação entre os verdadeiros efeitos genéticos e os EBLUPs para cada um dos 24 delineamentos considerando-se os dois cenários para o componente de variância genética,  $\gamma_g = 0,3$  e  $\gamma_g = 0,8$

	$\gamma_c = \gamma_l = 0$		$\gamma_c = \gamma_l = 0,1$	
	Critério C	Critério D	Critério C	Critério D
$\gamma_g = 0,3$				
Delineamento 1	0,6234 (0,0456)	0,6269 (0,0443)	0,6188 (0,0452)	0,6195 (0,0453)
Delineamento 2	0,6386 (0,0449)	0,6407 (0,0299)	0,6312 (0,0450)	0,6323 (0,0438)
Delineamento 3	0,6503 (0,0439)	0,6528 (0,0430)	0,6423 (0,0437)	0,6433 (0,0420)
Delineamento 4	0,6526 (0,0423)	0,6527 (0,0423)	0,6428 (0,0422)	0,6454 (0,0425)
Delineamento 5	0,6527 (0,0425)	0,6518 (0,0420)	0,6439 (0,0443)	0,6448 (0,0450)
Delineamento 6	0,6651 (0,0418)	0,6655 (0,0415)	0,6565 (0,0418)	0,6584 (0,0412)
$\gamma_g = 0,8$				
Delineamento 1	0,7749 (0,0310)	0,7763 (0,0302)	0,7671 (0,0314)	0,7676 (0,0315)
Delineamento 2	0,7857 (0,0431)	0,7861 (0,0290)	0,7759 (0,0306)	0,7774 (0,0301)
Delineamento 3	0,7932 (0,0291)	0,7956 (0,0285)	0,7846 (0,0297)	0,7856 (0,0286)
Delineamento 4	0,7953 (0,0280)	0,7958 (0,0281)	0,7847 (0,0288)	0,7872 (0,0286)
Delineamento 5	0,7958 (0,0282)	0,8045 (0,0277)	0,7859 (0,0302)	0,7869 (0,0304)
Delineamento 6	0,8045 (0,0272)	0,8057 (0,0271)	0,7954 (0,0280)	0,7972 (0,0273)

**Tabela 4.9.** Médias e desvios-padrões para a assertividade (porcentagem de verdadeiras melhores *test lines* selecionadas) para cada um dos 24 delineamentos considerando-se os dois cenários para o componente de variância genética,  $\gamma_g = 0,3$  e  $\gamma_g = 0,8$

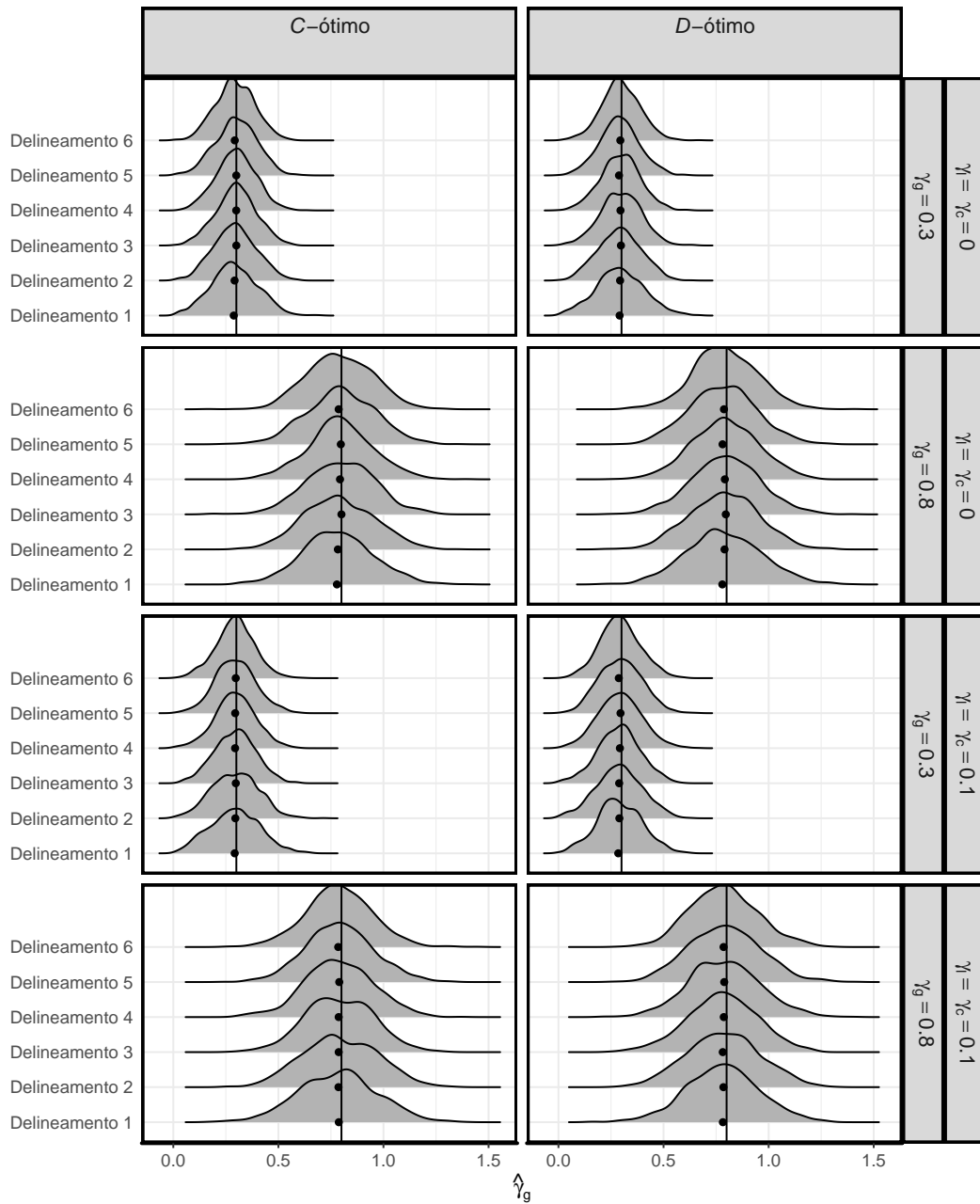
	$\gamma_c = \gamma_l = 0$		$\gamma_c = \gamma_l = 0,1$	
	Critério C	Critério D	Critério C	Critério D
$\gamma_g = 0,3$				
Delineamento 1	26,5940 (8,3003)	26,5961 (7,9515)	26,7635 (8,3009)	26,7385 (8,0979)
Delineamento 2	27,5153 (8,4301)	27,5372 (8,4319)	27,3213 (8,3106)	26,7057 (8,3500)
Delineamento 3	28,0737 (8,1823)	28,6286 (8,1571)	28,0449 (8,5720)	27,8833 (8,0942)
Delineamento 4	28,8365 (8,1543)	28,2782 (7,7694)	28,2914 (8,0864)	27,5825 (8,3547)
Delineamento 5	28,4464 (7,9810)	28,2000 (8,0667)	28,0011 (8,5127)	27,3658 (8,1205)
Delineamento 6	28,9901 (8,6661)	28,8020 (8,0974)	28,8455 (8,3027)	28,4840 (8,3089)
$\gamma_g = 0,8$				
Delineamento 1	35,9388 (7,7399)	35,6500 (7,5909)	35,6833 (7,6537)	35,5444 (7,6930)
Delineamento 2	36,6444 (7,7871)	36,8222 (8,0540)	35,9555 (7,8512)	35,6055 (7,9500)
Delineamento 3	37,1388 (7,5596)	37,6833 (7,8722)	36,8055 (8,3010)	36,5888 (7,7947)
Delineamento 4	37,7222 (7,8052)	37,2111 (7,5249)	36,7888 (7,7878)	36,5833 (8,0447)
Delineamento 5	37,3777 (7,3698)	37,6222 (7,6657)	36,7111 (7,6985)	36,3111 (7,7152)
Delineamento 6	38,0777 (7,8423)	37,8777 (7,6189)	37,4444 (7,7306)	37,3111 (8,0304)

Os resultados se repetem para a assertividade, sendo que esta sempre assume baixos valores: entre 26,6% e 29,0% quando  $\gamma_g = 0,3$  e entre 35,5% e 38,1% quando  $\gamma_g = 0,8$ . A maior diferença entre os delineamentos para um mesmo cenário é de aproximadamente 10%, entretanto, tal valor não representa mais do que uma verdadeira *test line* adicional selecionada. Embora os resultados de Santos (2017) apresentem assertividade mais próxima a 50%, em termos práticos, as diferenças entre os delineamentos também correspondem a seleção de uma *test line* adicional, em média.

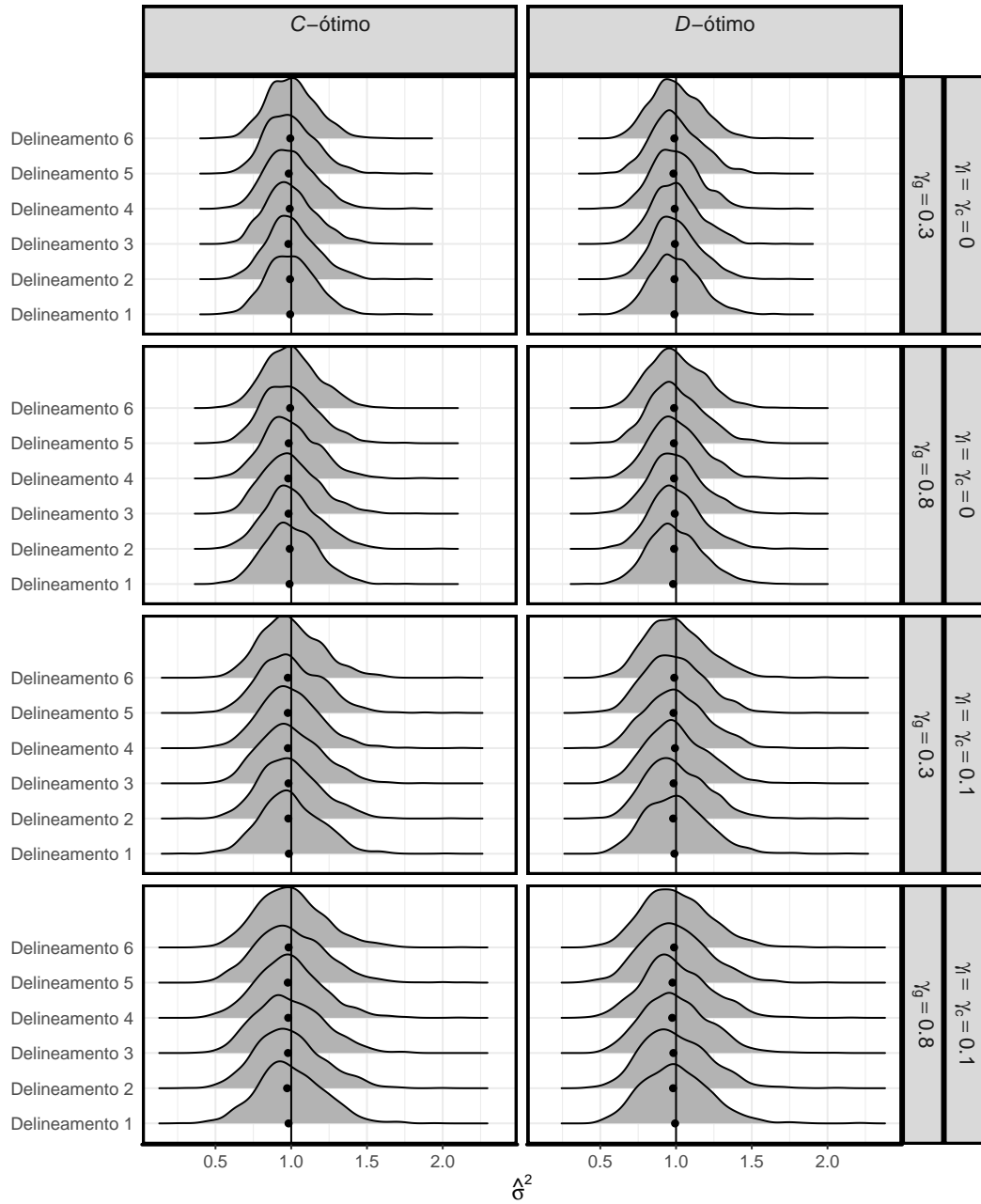
### 4.3 Estimativas dos componentes de variância

Como é possível observar nas Figuras 4.2 e 4.3, de modo geral, para os diferentes delineamentos e cenários há uma leve tendência à subestimação dos componentes de variância  $\gamma_g$  e  $\sigma^2$ , de modo mais acentuado quando  $\gamma_g = 0,8$ . Para cenários em que  $\gamma_g = 0,8$ , a dispersão de suas correspondentes estimas tende a ser maior, assim como a dispersão de  $\hat{\sigma}^2$  tende a ser maior na presença de variações globais de colunas e linhas.

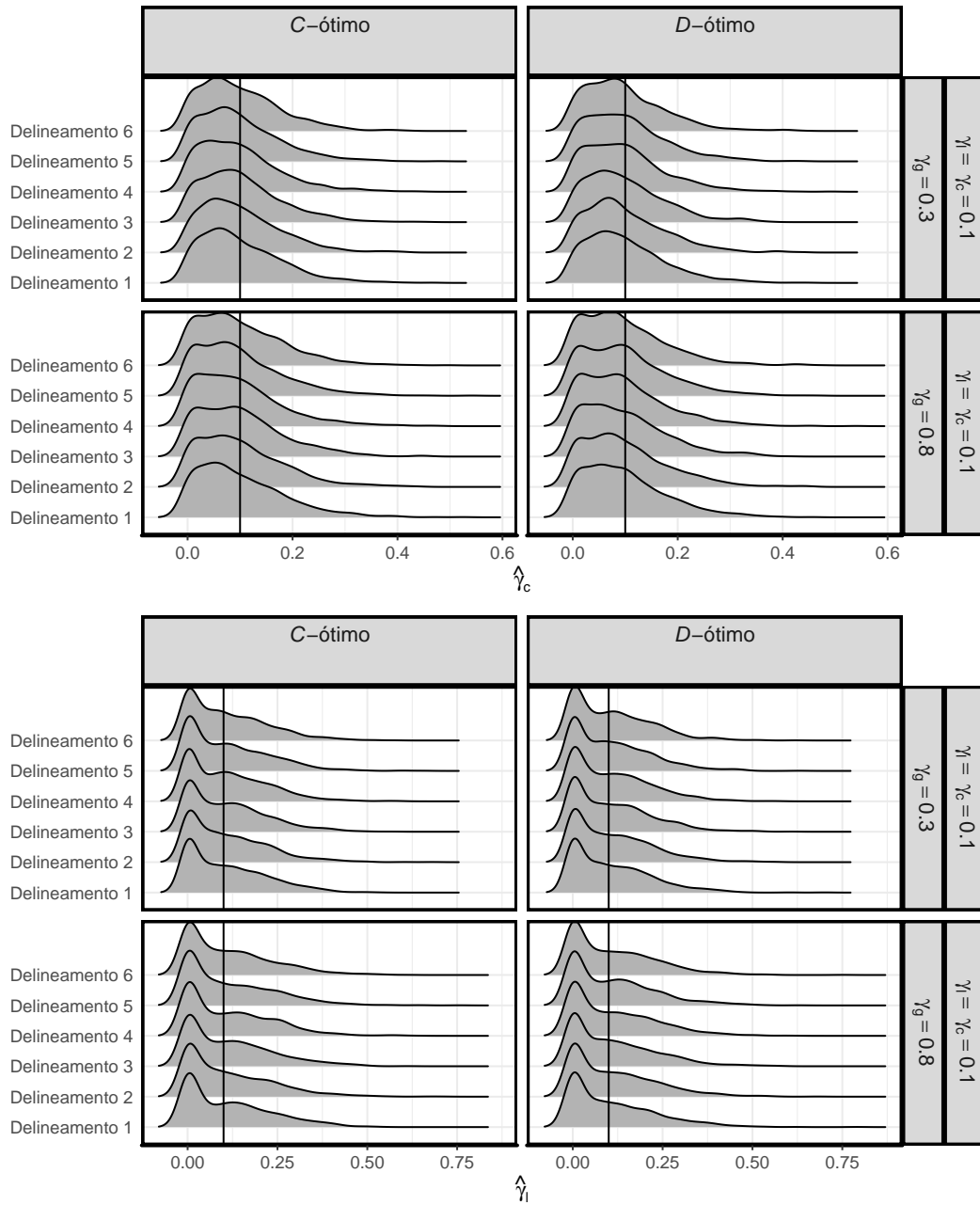
Com relação às estimativas dos componentes de variância de colunas e linhas (Figura 4.4), também há uma tendência à subestimação, independentemente do delineamento ou variância genética assumida para os dados. As correlações locais entre medidas tomadas na mesma coluna,  $\phi_c$ , e entre medidas tomadas na mesma linha,  $\phi_l$ , foram bem estimadas, como pode-se notar nas Figuras 4.5 e 4.6.



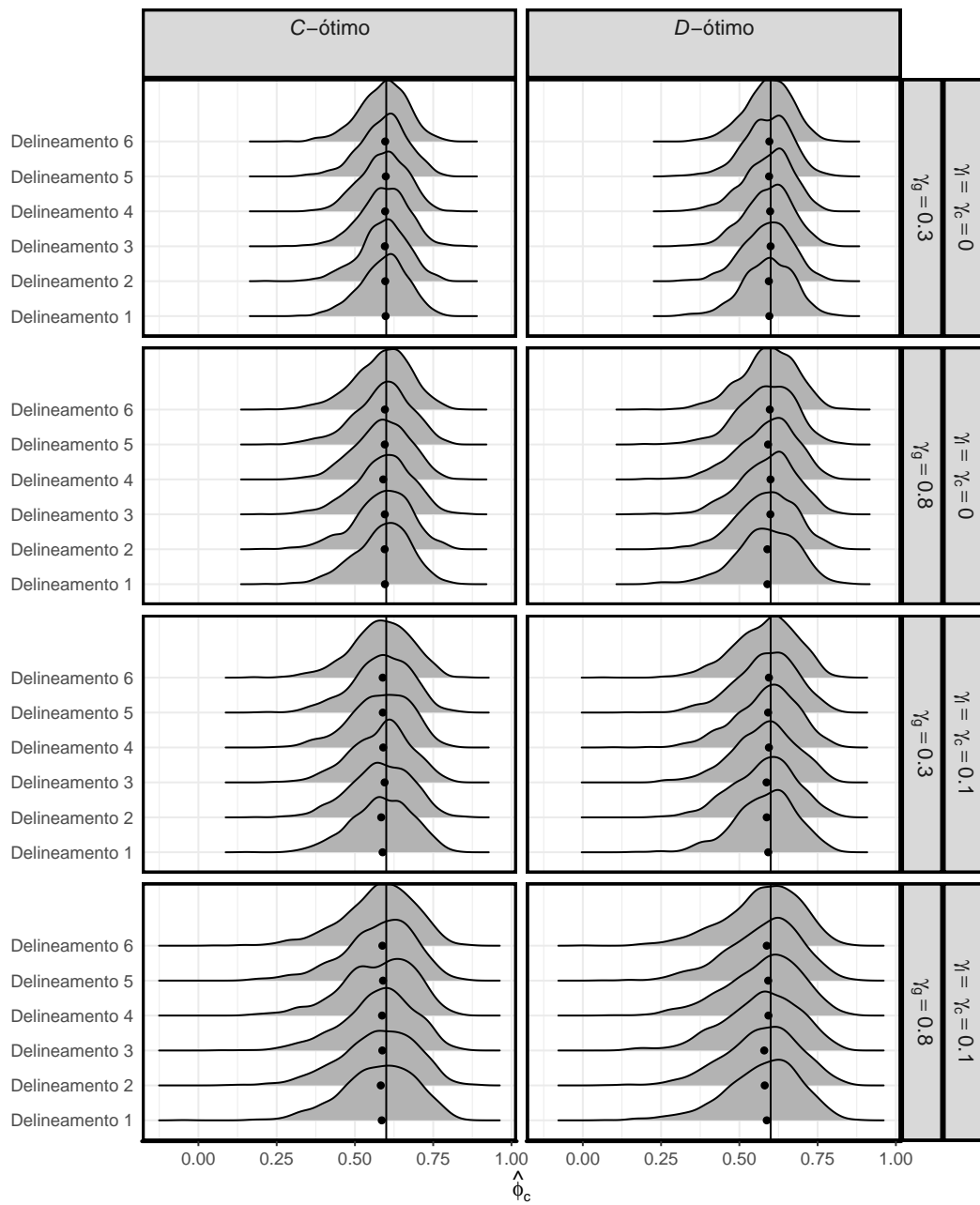
**Figura 4.2.** Estimativas do termo relacionado à variância genética ( $\hat{\gamma}_g$ ) para os diferentes delineamentos e cenários. A reta vertical representa o valor esperado para o parâmetro e o ponto representa a média das estimativas do parâmetro.



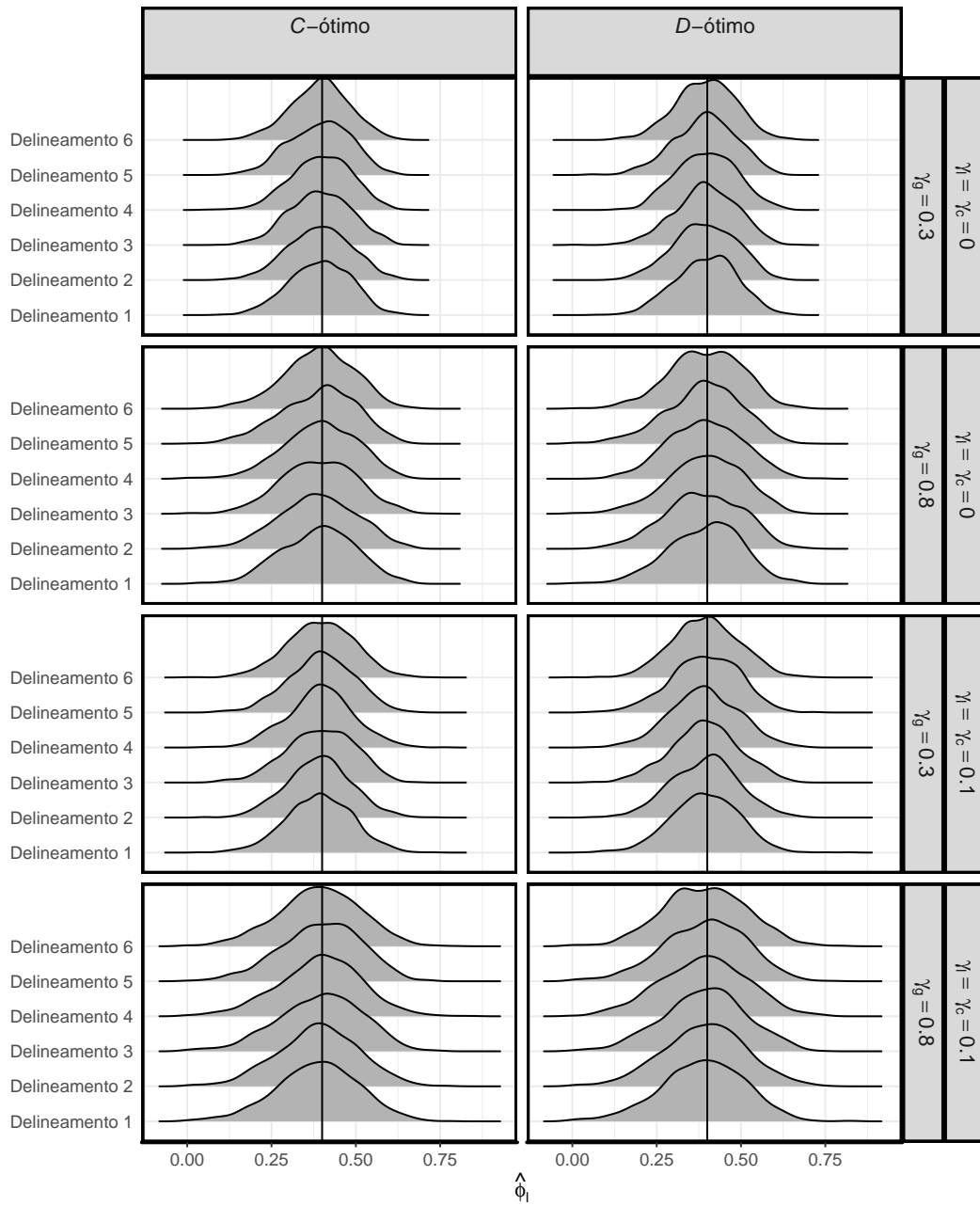
**Figura 4.3.** Estimativas do termo relacionado à variância genética ( $\hat{\gamma}_g$ ) para os diferentes delineamentos e cenários. A reta vertical representa o valor esperado para o parâmetro e o ponto representa a média das estimativas do parâmetro.



**Figura 4.4.** Estimativas do componente de variância globais para colunas e linhas,  $\hat{\gamma}_c$  e  $\hat{\gamma}_l$ , respectivamente, para os diferentes delineamentos e cenários avaliados



**Figura 4.5.** Estimativas da correlação residual entre medidas observadas na mesma coluna,  $\hat{\phi}_c$  para os diferentes delineamentos e cenários avaliados



**Figura 4.6.** Estimativas da correlação residual entre medidas observadas na mesma linha,  $\hat{\phi}_l$  para os diferentes delineamentos e cenários avaliados

## 5 CONCLUSÃO

Para os casos estudados, ou seja, experimentos correspondentes a uma área experimental com 10 colunas por 24 linhas e 180 *test lines*, o delineamento *p*-rep denotado por delineamento 6 ( $p = 33\%$  sem variedades padrões) apresentou os menores valores de eficiência para contrastes de médias entre *test lines* duplicadas. Entretanto, o mesmo delineamento foi o que resultou em maiores valores médios para o ganho genético com menor desvio-padrão. Desse modo, entende-se que delineamentos que apresentam valores altos de eficiência média não necessariamente apresentarão maiores valores em termos do ganho genético.

Entretanto, assim como ressaltado por Moehring et al. (2014) e Santos (2017), caso o pesquisador deseje comparar *test lines* com as variedades presentes no mercado, padrões, um delineamento *p*-rep com a presença das padrões seria recomendado. O presente estudo mostrou que o número de variedades padrões distintas presentes no experimento para uma fixada proporção de *test lines* repetidas, pouco afeta o ganho genético, a correlação entre os verdadeiros efeitos genéticos e o EBLUPs, a assertividade ou mesmo as estimativas dos componentes de variância.

No presente estudo não houve diferença entre os critérios de otimização, sendo o recomendado aquele que for mais acessível ao pesquisador. Cabe salientar que em todos os casos o modelo para a obtenção do delineamento ótimo assumiu os efeitos de *test lines* fixos, diferentemente do modelo utilizado para a geração dos dados e sua respectiva análise.

Ainda, mesmo havendo uma aparente diferença entre os delineamentos considerando-se a proporção de melhores *test lines* selecionadas, em termos práticos tais diferenças não representam mais do que uma *test line* adicional, sendo as porcentagens de seleção bastante baixas. Entende-se que maior informação genética se faz necessária para uma melhor qualidade de seleção.

Pretende-se, em trabalhos futuros, avaliar a qualidade de seleção dos delineamentos *p*-rep aumentados, como proposto por Williams et al. (2011), nos quais diferentes *test lines* são repetidas para grupos de experimentos instalados em locais distintos. Desse modo, todas as *test lines* apresentarão número de repetições no mínimo igual ao número de locais mais um.





## REFERÊNCIAS

- Bhagyalakshmi, K. V. e Somarajan, K. G. (1999). A modified augmented design for early selection stages in sugarcane and its limitation. *Sugar Tech*, 1(3):63–66.
- Brien, C. J. (2017a). *dae: Functions Useful in the Design and ANOVA of Experiments*. R package version 3.0-13.
- Brien, C. J. (2017b). Multiphase experiments in practice: A look back. *Australian & New Zealand Journal of Statistics*, 59(4):327–352.
- Butler, D., Cullis, B., e Taylor, J. (2014). Extensions in Linear Mixed Models and Design of Experiments. pages 32–61.
- Butler, D. G. (2014). *od: generate optimal designs*. R package version 0.75.
- Cesnik, R. e Miocque, J. (2004). *Melhoramento da cana-de-açúcar*. Embrapa Informação Tecnológica Brasília.
- Clarke, G. P. Y. e Stefanova, K. T. (2011). Optimal design for early-generation plant-breeding trials with unreplicated or partially replicated test lines. *Australian and New Zealand Journal of Statistics*, 53(4):461–480.
- CONAB, C. d. A. (2017). Acompanhamento da safra brasileira de café safra 2017. *Primeiro levantamento*, 4(1):98.
- Cullis, B. R., Smith, a. B., e Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(4):381–393.
- Federer, W. T. (1956). Augmented (or hoonuiaku) designs. *Biometrics Unit Technical Reports*, 55:191–208.
- Federer, W. T. (1961). Augmented designs with one-way elimination of heterogeneity. *Biometrics*, 17(3):447–473.
- Federer, W. T. (2002). Construction and analysis of an augmented lattice square design. *Biometrical journal*, 44:251–257.
- Federer, W. T. (2005). Augmented split block experiment design. *Agronomy Journal*, 97:578–586.
- Federer, W. T. e Crossa, J. (2001). On the design and analysis of field experiments.
- Federer, W. T., Nair, R. C., e Raghavarao, D. (1975). Some augmented row-column designs. *Biometrics*, 31(2):361–373.
- Federer, W. T. e Raghavarao, D. (1975). On augmented designs. *Biometrics*, 31(1):29–35.
- Fouilloux, M.-N. e Lalölë, D. (2001). A sampling method for estimating the accuracy of predicted breeding values in genetic evaluation. *Genetics Selection Evolution*, 33:473–486.
- Gilmour, A. R., Cullis, B. R., e Verbyla, A. P. (1997). Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. *Journal of Agricultural, Biological, and Environmental Statistics*, 2(3):269–293.
- Hooks, T., Marx, D., Kachman, S., e Pedersen, J. (2009). Optimality criteria for models with random effects. *Revista Colombiana de Estadística*, 32(1):17–31.

- Hunt, C. H., Smith, A. B., Jordan, D. R., e Cullis, B. R. (2013). Predicting additive and non-additive genetic effects from trials where traits are affected by interplot competition. *Journal of Agricultural, Biological, and Environmental Statistics*, 18(1):53–63.
- John, J. A. e Williams, E. R. (1995). *Cyclic and computer generated designs*. Chapman & Hall, 2<sup>nd</sup> edition.
- Kempton, R. A. (1984). The design and analysis of unreplicated field trials. *Vortrafe für Pflanzenzuchtung*, 7:219–242.
- Lin, C. S. e Poushinsky, G. (1983). A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. *Biometrics*, 39(3):553–561.
- LIN, C.-S. e POUSHINSKY, G. (1985). A modified augmented design (type 2) for rectangular plots. *Canadian Journal of Plant Science*, 65(3):743–749.
- Martin, R. J., Chauhan, N., Eccleston, J. A., e Chan, B. S. (2006). Efficient experimental designs when most treatments are unreplicated. *Linear Algebra and Its Applications*, 417:163–182.
- Moehring, J., Williams, E. R., e Piepho, H. P. (2014). Efficiency of augmented p-rep designs in multi-environmental trials. *Theoretical and Applied Genetics*, 127(5):1049–1060.
- Müller, B. U., Schützenmeister, A., e Piepho, H. P. (2010). Arrangement of check plots in augmented block designs when spatial analysis is used. *Plant Breeding*, 129(6):581–589.
- Piepho, H. P. e Williams, E. R. (2016). Augmented row–column designs for a small number of checks. *Agronomy Journal*, 108(6):2256–2262.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Santos, A. (2017). *Design and analysis of sugarcane breeding experiments : a case study*. PhD thesis, Universidade de São Paulo.
- Sarkar, J. e Sinha, B. K. (2015). Sudoku squares as experimental designs. *Resonance*, 20(9):788–802.
- SAS Institute Inc. (2015). *SAS/QC 14.1 User's Guide*. SAS Institute Inc., Cary, NC. accessed June 29, 2018.
- Sermarini, R., Brien, C., Demétrio, C., e dos Santos, A. (2018). Obtaining p – rep designs for plant breeding experiments.
- Shah, K. e Sinha, B. K. (1989). *Theory of Optimal Designs*. Springer Berlin Heidelberg.
- Smith, A. B., Ganesalingam, A., Kuchel, H., e Cullis, B. R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics*, 128(1):55–72.
- Smith, A. B., Lim, P., e Cullis, B. R. (2006). The design and analysis of multi-phase plant breeding experiments. *The Journal of Agricultural Science*, 144(05):393.
- Smith, A. and Thompson, R. and Butler, D. and Cullis, B. (2011). The design and analysis of variety trials using mixtures of composite and individual plot samples. *Journal of the Royal Statistical Society*, 60(3):437–455.

- Stringer, J. K., Cullis, B. R., e Thompson, R. (2011). Joint modeling of spatial variability and within-row interplot competition to increase the efficiency of plant improvement. *Journal of Agricultural, Biological, and Environmental Statistics*, 16(2):269–281.
- Vieira, M. C. A., Lima, J. F., e Braga, N. M. (2007). Setor sucroalcooleiro brasileiro: evolução e perspectivas. *Perspectivas do investimento*, 2010:207–245.
- Williams, E., Piepho, H. P., e Whitaker, D. (2011). Augmented p-rep designs. *Biometrical Journal*, 53(1):19–27.
- Williams, E. R., John, J. A., e Whitaker, D. (2014). Construction of more flexible and efficient p-rep designs. *Australian and New Zealand Journal of Statistics*, 56(1):89–96.
- Williams, E. R. e Piepho, H. P. (2013). A comparison of spatial designs for field variety trials. *Australian and New Zealand Journal of Statistics*, 55(3):253–258.