



Universidade de Brasília

2023

Experimentação 9 – Experimentos com Plantas



User

Concepta McManus, Sonia Nair Bão,
Daniel Pimentel, Felipe Pimentel, Andrea
Queiroz Maranhão, José Jivago Rolo

1/1/2023

Sumário

Desenho do Estudo: conceitos básicos	3
Fatores	5
Tratamentos	6
Escolha de tratamentos	6
Unidades experimentais	7
Tamanho da amostra e Repetições.....	9
Repetições.....	9
Escala e Forma de Replicação	10
Exemplo 1: falha em usar a escala adequada de replicação leva à identificação de diferenças médias de tratamento que não são reais.....	11
Diretrizes Gerais para Estudos Agronômicos.....	11
Exemplo 2: fatores de confusão - impacto do estrume e época de colheita na qualidade e produtividade da biomassa do milho.....	12
Solução 1: um desenho experimental com tratamentos.	14
Solução 2: replicação em vários anos ou locais.	16
Considerações específicas para experimentos com plantas perenes.....	17
Exemplo 3: Pseudo-replicação em escala de produção	18
Diretrizes para pesquisa em estufa e câmara de crescimento.....	20
Exemplo 4: estudo em estufa sobre o efeito da variação macroambiental.	23
Diretrizes para pesquisa em escala de bancada e reação	23
Diretrizes para Pesquisa Genética.....	26
Exemplo 5: um projeto de plotagem de controle para um grande número de genótipos não replicados.	28
Tipos de Experimentos de Melhoramento no Campo	29
Diretrizes para Experimentos em Bioquímica e Biologia Molecular	31
Determinando o Número Adequado ou Necessário de réplicas.....	33
Exemplo 6: determine o número de réplicas necessárias para atingir a potência desejada	34
Randomização	35
Bloqueio	37
Medições de variáveis de resposta	42
Experimentos de campo.....	42
Experimento de uniformidade:	47

Layout para experimentos de campo.....	51
Disposição do RBD.....	52
Experimentos não replicadas.....	54
Análise de dados e apresentação de resultados	57
Análise de conjuntos de dados grandes e complexos.....	59
Referências.....	61
Abbreviations	67
APPENDIX 1	68
APPENDIX 2	68
APPENDIX 3	69
APPENDIX 4.....	70

Desenho do Estudo: conceitos básicos

Por sua própria natureza, os experimentos biológicos contêm inúmeras fontes de variabilidade, algumas das quais são desejáveis e estão sob o controle do pesquisador, enquanto outras são indesejáveis e apenas parcialmente (ou nada) sob o controle do pesquisador. Por pouco, todo pesquisador que conduz experimentos comparativos críticos deseja fazer comparações biológicas entre os tratamentos que podem impactar alguma variável de medição biológica importante. Devido às duas principais fontes de variabilidade mencionada acima, a chave para o sucesso em fazer comparações biológicas é ter um certo grau de confiança de que a comparação biológica é significativa. Dentro outras palavras, a diferença entre dois tratamentos significa deve ser suficientemente grande para superar a variabilidade aleatória, incontável ou indesejável que existe dentro de um sistema biológico, ou seja, que a relação sinal-ruído exceda algum nível mínimo crítico.

Normalmente, usamos teste padrão estatísticas, por exemplo, testes t , testes F e testes χ^2 , para determinar se as relações sinal-ruído são suficientemente grandes para garantir a formação de declarações inferenciais. Costumamos usar um nível de significância de 0,05 ou 0,01, mas isso nem sempre é o ideal.

Em vez disso, a escolha de um nível de significância apropriado deve, idealmente, ser escolhido usando algum tipo de avaliação de risco e abordagem de gerenciamento de risco, que pode ser aplicada a qualquer situação experimental.

Normalmente, o objetivo de um estudo é descobrir as relações entre certos fatores explicativos e as variáveis de resposta. O desenho de um estudo consiste, portanto, em tomar decisões sobre o seguinte:

- O conjunto de fatores explicativos.
- O conjunto de variáveis de resposta.
- O conjunto de tratamentos.
- O conjunto de unidades experimentais.
- O método de randomização e bloqueio.
- Tamanho da amostra e número de repetições.
- As medições de resultados nas unidades experimentais - as variáveis de resposta.

Uma definição cuidadosa dos materiais e instalações experimentais a serem incluídos no experimento exige que cada um seja amostrado adequadamente para garantir que seja representado adequadamente. Se os tratamentos derivam de sacos de sementes, gado, patógenos ou locais de campo destinados a representar uma ou mais variáveis ambientais, a população deve ser definida e uma amostra aleatória ou representativa deve ser escolhida para representar a população. As populações podem ser definidas de forma muito ampla com uma intenção clara de escolher uma amostra aleatória. Este é frequentemente o caso quando a inferência desejada é representar uma população maior que a amostra, levando à escolha de um efeito aleatório (Fig.). Por outro lado, quando uma inferência é desejada apenas para um pequeno número de níveis, cada um dos quais pode ser incluído no experimento, a escolha geralmente é tratá-la como um efeito fixo. A seleção de uma amostra aleatória continua sendo um componente crítico para definir e aplicar os tratamentos, mesmo para efeitos fixos, para garantir que a inferência corresponda à hipótese.

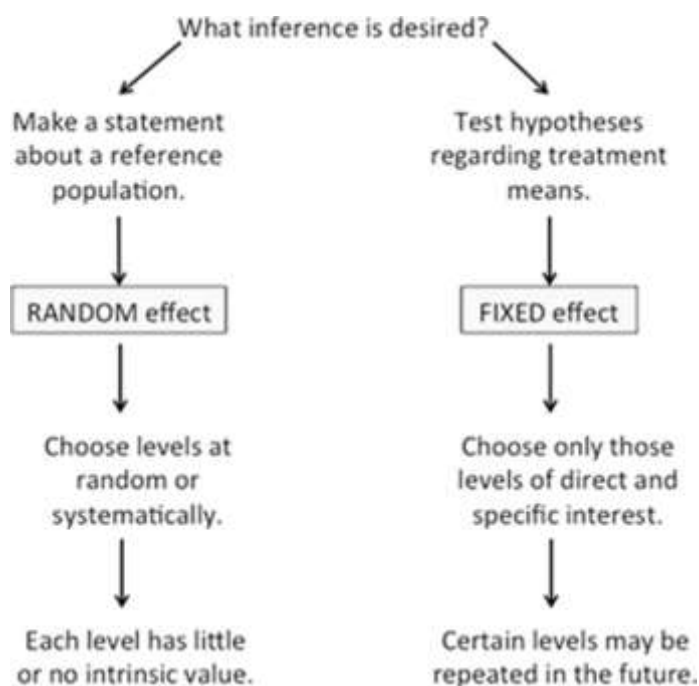


Figura . Diagrama de fluxo que ilustra a regra de decisão para efeitos fixos versus aleatórios para cada fator em um experimento. O primeiro passo é fazer a pergunta sobre a inferência desejada, o que leva à decisão de fixo versus aleatório. Após essa decisão, os tratamentos experimentais podem ser escolhidos conforme desejado para atender à inferência escolhida.

Fatores

Fatores são variáveis explicativas a serem estudadas em uma investigação.

Exemplos:

1. Em um estudo dos efeitos das cores e dos preços na venda de carros, os fatores estudados são a cor (variável qualitativa) e o preço (variável quantitativa).
2. Em uma investigação dos efeitos da educação sobre a renda, o fator estudado é o nível de educação (qualitativo, mas ordinal).

Níveis de fator

Níveis de fator são os "valores" desse fator em um experimento. Por exemplo, no estudo envolvendo a cor dos carros, o fator cor do carro poderia ter quatro níveis: vermelho, preto, azul e cinza. Em um desenho envolvendo vacinação, o tratamento poderia ter dois níveis: vacina e placebo.

Tipos de fatores

- Fatores experimentais: os níveis do fator são atribuídos aleatoriamente às unidades experimentais.
- Fatores observacionais: os níveis do fator são característicos das unidades experimentais e não estão sob o controle dos investigadores.
- Pode haver fatores observacionais em um estudo experimental.

Exemplo: no "estudo de nova droga" (consulte o Folheto 1), se também estivermos interessados nos efeitos da idade e do sexo na taxa de recuperação, então esses fatores observacionais; enquanto o tratamento (droga nova ou droga antiga) é um fator experimental.

Tratamentos

- Em um estudo de fator único, um tratamento corresponde a um nível de fator; assim, o número de tratamentos é igual ao número de diferentes níveis de fator daquele fator.
- Em um estudo multifatorial, um tratamento corresponde a uma combinação de níveis de fatores em diferentes fatores; assim, o número de todos os tratamentos possíveis é o produto do número de níveis de fatores de diferentes fatores.

Exemplos:

- No estudo dos efeitos da educação sobre a renda, cada nível de educação é um tratamento (ensino médio, faculdade, nível superior, etc).
- No estudo dos efeitos de raça e gênero na renda, cada combinação de raça e gênero é um tratamento (mulher asiática; homem hispânico, etc).

Exercício: Quantos tratamentos diferentes existem para os exemplos acima?

Escolha de tratamentos

A escolha dos tratamentos depende da escolha:

(i) dos fatores (que são os fatores importantes);

(ii) níveis de cada fator.

- Para fatores qualitativos, os níveis geralmente são indicados pela natureza do fator.

Exemplo: o gênero tem dois níveis: feminino e masculino

- Para fatores quantitativos, a escolha dos níveis reflete o tipo de tendência esperada pelo investigador.

Exemplo: tendência linear implica dois níveis; tendência quadrática implica três níveis. Normalmente, 3 a 4 níveis igualmente espaçados são suficientes.

- O alcance dos níveis também é crucial. Normalmente, é necessário conhecimento prévio para uma escolha eficaz de fatores e tratamentos.

Unidades experimentais

- Uma unidade experimental é a menor unidade de material experimental à qual um tratamento pode ser atribuído.

Exemplo: Em um estudo de dois sistemas de aposentadoria envolvendo as 10 escolas da UC, poderíamos perguntar se a unidade básica deveria ser um funcionário individual, um departamento ou uma universidade.

Resposta: A unidade básica deve ser uma Universidade inteira para viabilidade prática.

- Representatividade: as unidades experimentais devem ser representativas da população sobre a qual se pretende tirar uma conclusão.

Exemplo: Um estudo realizou pesquisas entre 5.000 estudantes universitários dos EUA e descobriu que cerca de 20% deles usaram maconha pelo menos uma vez. Se o objetivo do estudo é o uso de drogas entre americanos de 18 a 22 anos, esse é um bom projeto?

- É importante escolher um conjunto representativo de unidades experimentais que se adequem ao propósito do seu estudo.

Todo experimento tem uma unidade experimental, seja o pesquisador reconhece ou não esse conceito. Em experimentos agrônômicos, a unidade experimental é tipicamente uma única planta ou um aglomerado de plantas em uma “parcela”. Em experimentos em estufa, a unidade experimental é geralmente um único pote, mas às vezes é definido como um plano ou bandeja contendo vários potes que são todos tratados iguais. Na pesquisa de bancada, o conteúdo de um único recipiente de reator geralmente compreende a unidade experimental, se esse recipiente é um recipiente de

fermentação, uma placa de Petri ou um frasco que contém um microrganismo ou enzima interagindo com um substrato. Unidades experimentais menores são impossíveis para pesquisa de bancada, a menos que um navio possa de alguma forma ser subdivididos em unidades que podem ser tratadas e medidas independentemente umas das outras.

Finalmente, na pesquisa genética, há uma grande variedade de unidades experimentais, dependendo dos objetivos o experimento. Por exemplo, estudos de campo de diferentes cultivares ou culturas de biomassa terão uma “parcela” ou agrupamento de plantas como a unidade experimental. Estudos de variabilidade genética podem reduzir o tamanho da unidade experimental para o genótipo individual ou clone, desde que os genótipos individuais possam ser replicados adequadamente. Estudos genômicos, focados na descoberta de loci de traço quantitativo (QTL) pode ser projetado com o alelo como o nível de tratamento, de modo que grandes tamanhos populacionais em que muitos indivíduos irão representar cada nível de tratamento, pode substituir a replicação tradicional do indivíduo genótipos.

Uma vez definida a unidade experimental, o próximo passo é para o pesquisador é determinar como replicar tratamentos que formam o núcleo do experimento. A replicação adequada requer que, para cada tratamento, as seguintes condições sejam ser atendido:

- (1) deve haver várias unidades experimentais, repetidas no tempo ou no espaço ou ambos,
- (2) cada unidade experimental deve receber o tratamento, poder se expressar e ser medido independentemente de todas as outras unidades experimentais, durante todo o curso do experimento, e
- (3) os tratamentos devem ser randomizados, não organizados de forma sistemática ou maneira ordenada. O requisito de independência é fundamental e absoluto e não está aberto ao debate. Falha ao replicar tratamentos na escala adequada, às vezes denominado pseudoreplicação,

impacta significativamente a inferência que é possível a partir de um experimento.

Em muitas situações experimentais, a unidade experimental é muito grande para ser totalmente utilizada na fase de coleta de dados e requer análise de amostras que representam a unidade experimental.

Tamanho da amostra e Repetições

Falando livremente, o tamanho da amostra é o número de unidades experimentais no estudo.¹

- O tamanho da amostra geralmente é determinado pela compensação entre considerações estatísticas, como poder dos testes, precisão das estimativas e a disponibilidade de recursos, como dinheiro, tempo, mão de obra, tecnologia etc.
- Em geral, quanto maior o tamanho da amostra, melhor para inferência estatística; no entanto, o mais caro é o estudo.
- Uma consideração importante em um projeto experimental é como avaliar o poder ou a precisão em função do tamanho da amostra (planejamento do tamanho da amostra/cálculo do poder)?

Repetições

Para muitos estudos desenhados, o tamanho da amostra é um múltiplo inteiro do número total de tratamentos. Este número inteiro é o número de vezes que cada tratamento é repetido e uma repetição completa de todos os tratamentos (sob condições experimentais semelhantes) é chamada de Repetição completa do experimento.

- Exemplo: Em um estudo de temperatura de cozimento no volume de pão rápido preparado a partir de uma mistura de embalagem, quatro

¹ Casler MD (2013) Fundamentals of experimental design: guidelines for designing successful experiments. Agron J 105. doi:10.2134/ agronj2013.0114

temperaturas de forno: baixa, média, alta e muito alta foram testadas atribuindo aleatoriamente cada temperatura a 5 misturas de embalagem (todas da mesma marca). Assim, o tamanho da amostra é $20 (= 4 \times 5)$, o número de tratamentos é 4 (4 níveis de temperaturas) e há 5 Repetições completas do experimento.

Por que replicar?

Quando um tratamento é repetido sob as mesmas condições experimentais, qualquer diferença na resposta das respostas anteriores para o mesmo tratamento é devido a erros aleatórios. Assim, a replicação nos fornece algumas informações sobre erros aleatórios. Se a variação nos erros aleatórios for relativamente pequena em comparação com a variação total na resposta, teríamos evidências do efeito do tratamento.

Escala e Forma de Replicação²

Experimentos biológicos podem ser replicados em muitas escalas e, de fato, muitos experimentos contêm vários níveis diferentes de replicação. A escala é importante, porque determina o nível de inferências que são possíveis em relação aos tratamentos.

Replicação na escala exata da unidade experimental, ou seja, múltiplas unidades experimentais independentes por tratamento, fornece uma estimativa do erro experimental. É apenas esta estimativa de erro experimental, calculado na escala apropriada, que permite que o pesquisador julgue adequadamente se a variabilidade de tratamento para tratamento excede a variabilidade aleatória ou incontrolável em um grau suficiente para julgar a tratamento significa ser diferente.

Como princípio geral, a replicação em escalas maiores que a unidade experimental é conduzida para permitir que o pesquisador desenvolva inferências maiores ou mais amplas em relação a uma gama de condições experimentais sob as quais se espera que os tratamentos causem efeitos diferentes. Por exemplo, experimentos agrônômicos

² Casler, M.D., Vermerris, W. and Dixon, R.A., 2015. Replication concepts for bioenergy research experiments. *Bioenergy Research*, 8(1), pp.1-16.

podem ser replicados em vários locais, permitindo inferências potencialmente mais amplas em relação ao solo tipo, clima e/ou outras condições ambientais sob quais os efeitos do tratamento ocorrem. Da mesma forma, reator ou bancada experimentos podem ser conduzidos sob uma variedade de condições físicas, condições químicas e/ou biológicas ou com diferentes matérias-primas.

Por outro lado, a replicação em escalas menores que a unidade experimental ocorre quando a unidade observacional é menor que a unidade experimental. Em muitos tipos de pesquisa, a unidade experimental é muito grande para que as medições ocorram nessa escala, por exemplo, grandes parcelas de campo (superiores a 0,1 ha) ou grande reator navios como plantas piloto. Nessas situações, o representante amostragem da unidade experimental é um grande problema.

Frequentemente, várias observações são feitas em cada unidade experimental, por exemplo, amostras aleatórias de plantas ou alíquotas aleatórias de uma solução de reação. Normalmente, todas as medições feitas em um ponto no tempo de uma unidade experimental podem ser calculadas juntos para entregar a observação para essa unidade experimental em particular. Neste caso, a análise estatística dos dados de tais um experimento pode ser feito com base nessas médias valores por unidades experimentais como os dados de entrada. Replicação em escalas menores que a unidade experimental não é válida substituto para a replicação na escala da unidade experimental, porque as variâncias residuais tendem a se tornar menores, tanto na expectativa quanto na prática, à medida que a escala se torna menor. Por isso, usar uma escala muito pequena para replicação resulta em testes de hipóteses tendenciosos, ou seja, valores de p que são irrealisticamente baixos (superotimista) e inflação da verdadeira taxa de erro tipo I.

Exemplo 1: falha em usar a escala adequada de replicação leva à identificação de diferenças médias de tratamento que não são reais.

Diretrizes Gerais para Estudos Agronômicos

Os estudos agronômicos geralmente envolvem experimentos de campo onde a variação na composição do solo, fertilidade, temperatura, e precipitação variam de local para local e de ano para ano. Para que os dados relativos à produção de bioenergia

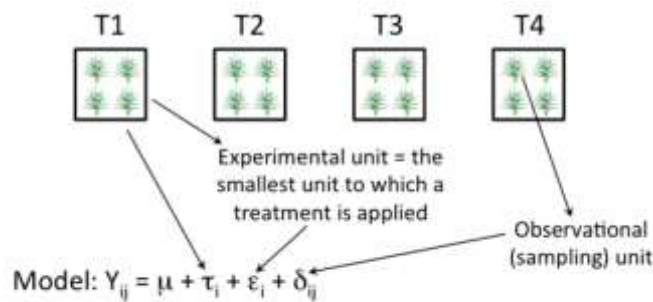
sejam extremamente relevantes, portanto, é importante não apenas adquirir dados de vários locais e/ou vários anos, mas também para replicar em um determinado site.

Exemplo 2: fatores de confusão - impacto do estrume e época de colheita na qualidade e produtividade da biomassa do milho.

Para ilustrar a importância da replicação, considere um experimento destinado a quantificar o efeito dos tratamentos com estrume e época de colheita na qualidade e produtividade da biomassa do milho. O pesquisador planeja quatro tratamentos (duas fontes de estrume em combinação com duas datas de colheita, com a colheita a ser realizada com equipamento à escala agrícola). Devido a as dificuldades logísticas de criar múltiplos experimentos unidades desses quatro tratamentos, cada tratamento é aplicado a apenas uma seção de um campo de milho uniforme e cuidadosamente definido (tratamentos T1 a T4 na Fig. 1, onde cada planta na figura representa uma unidade de medida, por exemplo, o campo específico área na qual o tratamento é aplicado, dividido em quatro seções ou amostras). A unidade experimental e o tratamento são completamente confundidos entre si, ou seja, os tratamentos são replicados, mas não na escala adequada para eliminar o problema de confusão. Como o pesquisador pode eliminar esse efeito de confusão?

Dois mecanismos de replicação estão disponíveis:

- coletar várias amostras dentro de cada unidade experimental ou
- repetir todo o experimento por dois ou mais anos.



ANOVA Source of variation	df
Treatments + Experimental error	3 (fixed)
Observational error	12

Fig. 1 Design example 1: Sixteen plants are assigned to four experimental units, each of which is assigned to one treatment level (T1 through T4, symbolized by the treatment effect, τ_i in the linear model). Treatments are confounded with experimental units because replication is conducted at a scale that does not match the treatment application

Ambas as alternativas representam a replicação em uma escala diferente dos tratamentos ou unidades experimentais. A opção (1) não resolve o problema fundamental problema que tratamentos e unidades experimentais são completamente confundidos entre si, resultando nos problemas apontados no Exemplo 1 (Tabela 1). A opção (2) resolve este problema por usando anos como fator de bloqueio ou replicação, apenas fornecido que o experimento seja conduzido em parcelas novas e ao acaso no segundo ano. Embora o campo de milho possa parecer uniforme, não há garantia de que seja suficientemente uniforme para implementar o projeto na Fig. 1, que é baseado na suposição fundamental de que a única diferença entre as quatro unidades experimentais se deve à aplicação de tratamentos.

Embora essa suposição possa estar correta, não há como saber com certeza. Por definição, essa é a natureza da confusão—dois ou mais itens estão inextricavelmente interligados, de modo que nenhuma quantidade de análise estatística, pensamento positivo ou discussão pode separá-los. O problema de confundir os tratamentos com unidades experimentais é ilustrado na análise de variância (ANOVA) fonte de variação que reflete o fato de que os efeitos do tratamento não podem ser estimados como um resultado puro e inconfundível fator (Fig. 1). Imagine o impacto potencial nos dados coletados se houver um gradiente subterrâneo de fertilidade, tipo de solo, capacidade de retenção de umidade, pH ou algum outro fator importante que não pode ser

observado ou não foi medido e que a direção deste gradiente é horizontal em relação ao mapa de campo na Fig. 1.

Table 1 Mixed models analysis of variance for an experiment with three levels of replication: three blocks with one experimental unit per treatment inside each block, two independent samples per experimental unit, and two independent laboratory determinations per sample

Source of variation	df	Mean square	Variance component associated with error term (random effect)	F ratio for treatments	p value for treatments	LSD (0.05) for comparing treatment means
Blocks	2	2485				
Treatments	4	781				
Experimental Error	8	147	20	5.31 ^a	0.0219 ^a	11 ^a
Samples within Experimental Units	15	69	23	11.32 ^b	0.0002 ^b	7 ^b
Determinations within Samples	30	23	23	33.96 ^c	<0.0001 ^c	4 ^c

Data were in vitro dry matter digestibility (mg/g) of switchgrass biomass (Casler, 2000, unpublished data). The total number of laboratory determinations was 12 per treatment for each of three hypothesis testing scenarios (a, b, c)

^a Correct (unbiased) values of F ratio and LSD calculated using experimental error mean square in this experiment with three experimental units per treatment, two samples per experimental unit, and two determinations per sample. The treatment means were 496, 502, 505, 514, and 515 mg/g

^b Values of F ratio and LSD calculated using the mean square representing variation among samples within experimental units in this experiment

^c Values of F ratio and LSD calculated using the mean square representing variation among determinations within samples in this experiment

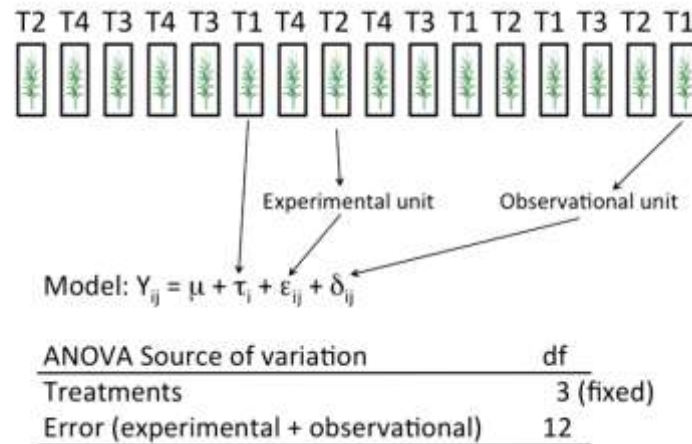
Como os pacotes de software de análise estatística não reconhecem as limitações do desenho experimental, qualquer pessoa pode conceber testes de hipóteses para este cenário e aplicá-los aos dados. Uso de várias amostras dentro de cada unidade experimental, representado pelas unidades observacionais (ou amostragem) na Fig. 1, permite a estimativa de um termo de erro e testes F subsequentes ou t testes, mas este termo de erro é enviesado para baixo devido à escala.

As unidades observacionais ocorrem em uma área menor e mais uniforme. área do que unidades experimentais, e eles estão correlacionados um com outro, resultando em uma variância menor do que seria esperado para várias unidades experimentais independentes de cada tratamento. Tal teste resulta em valores p superotimistas e um grau desconhecido de inflação da taxa de erro tipo I.

Solução 1: um desenho experimental com tratamentos.

Então, quais opções estão disponíveis para resolver esse problema? Nós começamos pelo mais simples e menos problemático, do ponto de vista da maioria dos pesquisadores. Replicação na escala adequada, resultando em estimativas imparciais de médias de tratamento e erro variâncias, requer experimentos múltiplos e independentes unidades para cada tratamento. A Figura 2 ilustra a solução mais simples para este problema, usando quatro tratamentos, cada um replicado quatro vezes.

Fig. 2 Design example 2: 16 plants are assigned to 16 experimental units, 4 of which are independently assigned to each treatment level (T1 through T4, symbolized by the treatment effect, τ_i in the linear model). This is an example of the completely randomized design



O número de observações feitas é idêntico ao projeto na Fig. 1, mas cada uma das quatro observações feitas em um tratamento único era mutuamente independente, garantindo que ambos os problemas, confusão e escala imprópria de replicação (Fig. 1), foram resolvidos. Neste caso específico, apenas uma amostra ou ponto de dados é coletado dentro de cada experimento unidade, confundindo as unidades experimentais e observacionais, mas isso geralmente não é um problema. Se os pesquisadores estão preocupados sobre este problema, por exemplo em grandes unidades que estão sujeitas a altos níveis de variabilidade, então o desenho ilustrado na Fig. 3 oferece uma solução. Neste caso, o design permanece o mesmo, com várias unidades experimentais independentes por tratamento, mas as múltiplas amostras ou observações dentro de cada unidade experimental ajudam o pesquisador para controlar a variabilidade que pode estar presente nas parcelas.

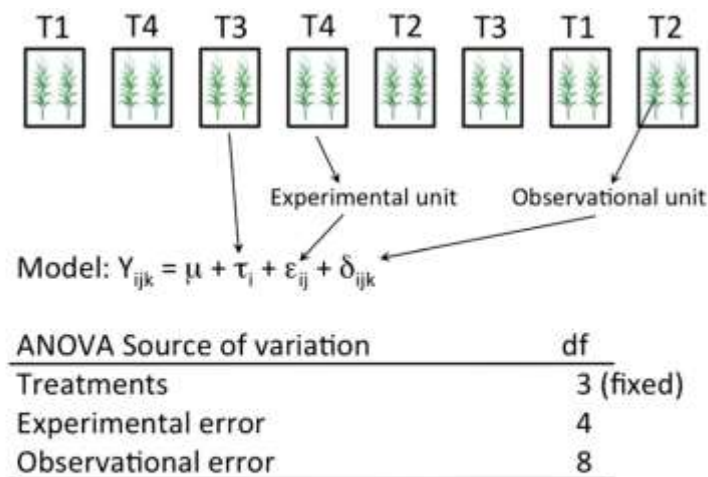


Fig. 3 Design example 3: Sixteen plants are assigned to eight experimental units, two of which are independently assigned to each treatment level (T1 through T4, symbolized by the treatment effect, τ_i in the linear model). Each experimental unit contains two observational, or sampling, units. This is another example of the completely randomized

design. In this design, the researcher is using two independent samples from each experimental unit to provide a better estimate of the performance of each experimental unit, i.e., representative sampling, as compared with design example 2

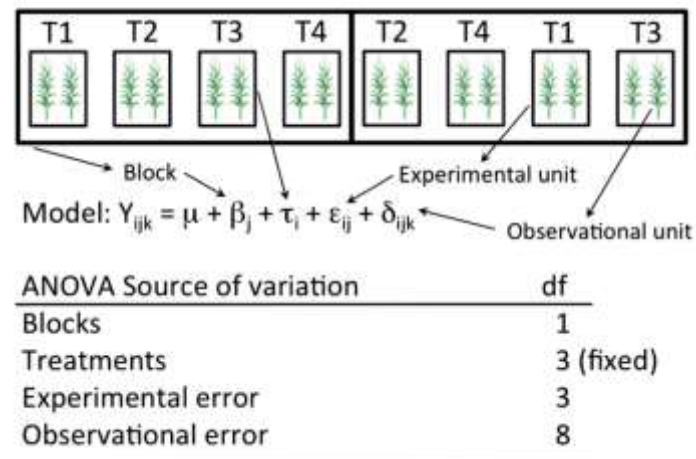
O ANOVA na Fig. 3 mostra estimativas de dois efeitos aleatórios, representando variabilidade em duas escalas (a unidade observacional e a escala da unidade experimental), mas é igualmente válido para calcular as médias de todas as amostras dentro de cada experimental unidade, em seguida, para conduzir a ANOVA valores médios da unidade experimental. As unidades experimentais na maioria das pesquisas agrônomicas são estabelecidas a partir de sementes plantadas em fileiras, semeadura ou de plantas transplantadas. Observações feitas em unidades menores do que a unidade experimental inteira podem estar em plantas individuais, em grupos de plantas dentro de uma quadra ou amostragem estrutura, ou de fardos individuais de biomassa. Figuras 2 e 3 ambos representam variações do delineamento inteiramente casualizado, o mais simples de todos os delineamentos experimentais.

Solução 2: replicação em vários anos ou locais.

Outra opção seria repetir o experimento vários anos. A maneira fácil de fazer isso não resolve o problema de confusão dentro de um ano individual ou estação de crescimento. Se os tratamentos mostrados na Fig. 1 são repetidos nas mesmas unidades experimentais por vários anos, como seria o caso com observações repetidas de uma planta, quaisquer fatores ambientais ou desconhecidos que sejam confundidos com os tratamentos permanecem como tal durante a duração do experimento. Em alguns casos, pode piorar com tempo, se os seus efeitos forem cumulativos. Mesmo que várias observações sejam feitas em cada tratamento, essas observações são não

são independentes uns dos outros, por exemplo, eles estão correlacionados com uns aos outros. Por estarem correlacionados entre si, não fornecem uma estimativa imparcial do erro experimental, pois com o projeto na Fig. 1. No entanto, vários anos podem ser usados para resolver este problema, como mostrado na Fig. 4.

Fig. 4 Design example 4: Sixteen plants are assigned to eight experimental units. Experimental units are grouped into two blocks, each of which contains a number of experimental units exactly equal to the number of treatment levels (T1 through T4, symbolized by the treatment effect, τ_i in the linear model). Each experimental unit contains two observational, or sampling, units. This is an example of the randomized complete block design, with sampling



Neste caso, o bloco à esquerda, que contém quatro unidades experimentais e quatro tratamentos é aplicado em 1 ano, enquanto o bloco no direito é aplicado a um novo campo de milho, independente do campo dos anos anteriores, no segundo ano. Isso representa uma aplicação do delineamento em blocos ao acaso, em que a cada ano funciona como um "bloco" de unidades experimentais. De É claro que tal projeto não exigiria o uso de múltiplos anos, desde que haja terra e mão de obra suficientes para aplicar várias unidades experimentais de cada tratamento em um único ano. A escolha do delineamento inteiramente casualizado vs. blocos completos ao acaso (bloquear ou não bloquear) pode ser complexo e está além do escopo deste trabalho, tendo sido efetivamente discutido em várias outras fontes.

Considerações específicas para experimentos com plantas perenes

Muitos dos obstáculos e impedimentos para criar inferências válidas para a pesquisa de culturas perenes, incluindo espécies herbáceas e culturas lenhosas de curta rotação (SRWC) são semelhantes aos discutidos acima. Apesar disso, existem algumas diferenças óbvias entre anuais e perenes, criadas em grande parte por diferenças de escala e longevidade de experimentos parcelas, com perenes muitas vezes exigindo pesquisas em maior escala física e conduzidas por um período de tempo mais longo

do que para culturas herbáceas. Muitas plantas perenes são geneticamente heterogêneas e podem exigir um tempo considerável para se tornarem estabelecido. Além disso, a mortalidade de plantas individuais cria espaços abertos que podem ser manuseados de muitas maneiras diferentes ou simplesmente ignorado e considerado parte do “tratamento”.

Às vezes, material experimental, como sementes ou rizomas, estão em falta, levando à necessidade de não replicados ou projetos parcialmente replicados, como discutido posteriormente no exemplo 5. Além disso, a idade do povoamento é frequentemente um fator importante a ser estudado, para se obter uma compreensão das mudanças na produção de biomassa ao longo do tempo. Criando um projeto experimental em que o efeito fixo da idade do povoamento não se confunde com o efeito aleatório efeitos climáticos é particularmente difícil.

Devido a esses requisitos, os cientistas que trabalham com plantas perenes não podem responder tão rápida e prontamente quanto aqueles que trabalham em anuais a novas iniciativas ou criando novos experimentos todos os anos à medida que as necessidades surgem. Como tal, muitas comparações críticas experimentos destinados a avaliar a produção de bioenergia e características de sustentabilidade das plantas perenes são conduzidas em estandes existentes que foram criados estritamente para fins de produção, criando potenciais limitações às inferências da pesquisa.

Pseudo-replicação, replicação em escala imprópria, que é geralmente menor do que a unidade experimental real, pode ser um problema comum na pesquisa sobre os povoamentos de plantas perenes existentes.

Exemplo 3: Pseudo-replicação em escala de produção

Como exemplo, considere um novo experimento a ser conduzido em uma plantação existente de SRWC ou em uma área que inclui numerosos campos de uma grama de energia perene. Uma área de pesquisa existente normalmente pode incluir várias datas de plantio e múltiplos genótipos, clones ou cultivares plantados em vários anos, com uma pequena quantidade de novas terras convertidas para a cultura de biomassa em

cada ano. Deve-se tomar muito cuidado quando utilizando este tipo de terra de produção para pesquisa em bioenergia.

A Figura 1 representa o esquema básico da área experimental potencial, em que as unidades experimentais são representadas por as terras utilizadas para o estabelecimento de culturas energéticas em qualquer ano e os tratamentos tornam-se diferentes anos de estabelecimento e genótipos. Simplesmente usar esta área experimental para tirar inferências estatísticas e biológicas sobre a idade ou genótipo do povoamento é completamente insustentável, conforme descrito acima no Exemplo 2. Tal uma área experimental teria idade, genótipo e talvez uma série de outros fatores ambientais, todos confundidos entre si em graus desconhecidos. Os pesquisadores puderam convencer-se de que a idade e o genótipo do povoamento são as fontes dominantes de variação, mas muitos revisores irão justificadamente não estar convencido de que esta suposição é válida. É muito melhor para a equipe de pesquisa enfrentar essa questão "de frente" no projeto fase do que depois de receber avaliações desfavoráveis de uma revista científica internacional.

Idealmente, um novo experimento seria desenhado, com tratamentos replicados e randomizados, exclusivamente dentro de cada indivíduo plantio homogêneo para data de estabelecimento e genótipo. Tal como acontece com o experimento agrônomo projetado acima, à randomização adequada e a replicação de tratamentos para várias unidades experimentais independentes dentro de um local experimental homogêneo permitem estatísticas e biológicas limpas e válidas comparações e testes de hipóteses. Repetição do experimento em diferentes datas de estabelecimento e/ou genótipos permitir que os pesquisadores ampliem as inferências a serem derivadas de tal experimento. Em alguns casos, isso pode exigir alguns comprometer o tamanho final da unidade experimental, resultando em "parcelas" menores do que o desejado. Na maioria situações biológicas em que um pesquisador está avaliando os custos e benefícios de aumentar o tamanho da unidade experimental vs. o número de réplicas, e não há dados específicos para ajudar orientar uma escolha informada, esta última geralmente é melhor para ambos precisão estatística e espaço de inferência.

Finalmente, mais uma solução alternativa envolveria a uso de projetos de plotagem de controle, que podem ser projetados e adaptados para atender a quase todas as

necessidades de pesquisadores. As parcelas de controle são frequentemente usadas em pesquisas agronômicas na fazenda ou em escala de fazenda como um mecanismo para ajustar observações sobre tratamentos não replicados para variação espacial que normalmente existe dentro de áreas experimentais extremamente grandes. Um projeto altamente eficiente é atribuir um dos tratamentos como controle e emparelhar o tratamento de controle com o outro tratamento em um desenho de parcela pareada.

Diretrizes para pesquisa em estufa e câmara de crescimento³

A pesquisa em estufa é geralmente conduzida em uma ou ambas as duas escalas: macro ou microambiental. Pesquisa na escala microambiental é simples e direta, envolvendo um conjunto de condições ambientais e provavelmente apresentando poucos problemas para a maioria dos pesquisadores.

As necessárias condições ambientais para a estufa devem ser pré-determinadas pela equipe de pesquisa e organizado para ser o mais uniforme possível em todo o espaço físico a ser utilizado. A unidade experimental deve ser criteriosamente definida, geralmente um único vaso, contendo uma ou mais plantas, ou às vezes uma bandeja ou apartamento de células. Uma vez que esta definição é feita, cada tratamento deve ser atribuído a unidades experimentais múltiplas e independentes (réplicas) e todas as unidades experimentais devem ser randomizado em um desenho ou padrão experimental que é capaz de lidar com a variação espacial que pode existir dentro do ambiente de estufa. Exceto em casos muito incomuns, várias plantas dentro de um vaso não podem servir como réplicas independentes do tratamento, pois não são independentes de cada outro e sua variação será uma subestimação do verdadeiro variância entre unidades experimentais tratadas da mesma forma.

O delineamento inteiramente casualizado, que envolve a randomização simples de todas as unidades experimentais em um único bloco, é uma boa idéia apenas em uma de duas circunstâncias:

³ Brien CJ, Berger B, Rabie H, Tester M (2013) Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems. Plant Methods 9:5 <http://www.plantmethods.com/content/9/1/5>

- (a) há pouca probabilidade de variação espacial, ou efeitos microclimáticos, dentro da estufa ou
- (b) os pesquisadores planejam reordenar ou reorganizar as unidades experimentais periodicamente durante o período experimental⁴.

Nós desencorajamos fortemente a opção (b) por quatro razões⁵:

- (1) o rearranjo das unidades experimentais aumenta a carga de trabalho e cria mais oportunidades para erros na randomização ou danos às plantas;
- (2) a aplicação bem sucedida desta abordagem requer que as unidades experimentais gastem quantias semelhantes de tempo dentro de cada microclima;
- (3) não há evidência que esta abordagem levará a uma maior precisão do que o uso de projetos formalmente randomizados, porque o rearranjo elimina ou cria desafios significativos na contabilidade para variação espacial ou microclimática; e
- (4) existem inúmeras opções alternativas de desenho experimental que variam em complexidade e eficácia esperada. Efeitos do microclima são comuns em ambientes de estufa e esses efeitos são geralmente previsíveis por longos períodos de tempo.

Como tal, delineamentos em bloco sistemáticos geralmente serão mais eficazes na contabilização da variação do microclima do que as abordagens de rearranjo⁶.

Blocos geralmente fornecem a solução ideal para estufas ou câmaras de crescimento que possuem alguns efeitos microclimáticos. Os projetos de bloqueio podem ser agrupados em duas categorias:

- bloqueio unidirecional (bloqueio completo aleatório, treliça e projetos de blocos incompletos balanceados) ou

⁴ Hardy EM, Blumenthal DM (2008) An efficient and inexpensive system for greenhouse pot rotation. HortSci 43:965–966

⁵ Kempthorne O (1957) 126. Query: arrangements of pots in greenhouse experiments. Biometrics 13:235–237

⁶ Edmondson RN (1989) Glasshouse design for repeatedly harvested crops. Biometrics 45:301–307

- de duas vias bloqueio (quadrados latinos, quadrados de Youden, quadrados de Tróia, quadrados de treliça e desenhos de linha-coluna).

Nem todos os designs e padrões de bloqueio são eficazes; assim, experimentos com muitos blocos pequenos, como projetos de bloqueio incompletos deve ser usado em combinação com métodos estatísticos espaciais identificar padrões de efeitos microclimáticos que podem ser usados como uma base para projetar experimentos futuros. Isto é um processo altamente empírico que pode envolver várias iterações ou repetições para identificar padrões consistentes.

Experimentos em estufas macroambientais ou câmaras de crescimento criam projetos experimentais e análises estatísticas problemas, porque a replicação raramente é realizada na escala adequada. Várias estufas ou câmaras de crescimento são frequentemente usados para criar condições ambientais altamente controladas e diferenciadas, para as quais o teste de hipóteses e/ou diferenças quantitativas são de interesse. A maioria desses experimentos são projetados como mostrado na Fig. 1, com a estufa ou câmara como unidade experimental e replicação conduzida apenas dentro da unidade experimental, em algum tipo de unidade amostral nível. Essas condições podem formar todo o experimento ou servir como um fator em um experimento multifatorial. Dentro do estojo onde eles formam um experimento de fator único, as opções viáveis são:

- (1) dobrar o número de câmaras, permitindo a verdadeira replicação dos tratamentos macroambientais, ou
- (2) repetir o experimento ao longo do tempo, re-randomizando as condições ambientais (tratamentos) para as várias estufas ou câmaras de crescimento para garantir que o termo de erro é estimado na escala apropriada, ou seja, que cada tratamento é aplicado a múltiplos e independentes unidades (estufas ou câmaras).

Uso de várias plantas, vasos, bandejas, ou apartamentos dentro de cada casa ou câmara não permite a pesquisador para estimar o termo de erro adequado, subestimando a verdadeira variação entre unidades experimentais tratadas da mesma forma (Fig. 1; Tabela 1).

Exemplo 4: estudo em estufa sobre o efeito da variação macroambiental.

Arnoult et al.⁷ utilizaram uma abordagem nova e muito econômica para resolver este problema. Eles projetaram um experimento de dois fatores com quatro tratamentos macroambientais, definidos por todas as combinações de dois fotoperíodos e dois regimes de temperatura, cada um atribuído para uma estufa. Isso é quase idêntico ao design mostrado na Fig. 1, diferindo apenas no número de tratamentos e repetições dentro de cada macroambiente. Dentro de cada estufa, o segundo fator consistiu em oito genótipos diferentes de *Miscanthus*. Os genótipos foram estabelecidos de forma completamente delineamento casualizado com dez repetições dentro de cada estufa (80 plantas e vasos dentro de cada estufa). Sem replicação dos tratamentos ambientais ao longo do tempo, as comparações entre os genótipos fornecem as únicas testes de hipótese estatística válidas. Os autores estavam parcialmente interessados em fazer comparações entre genótipos para desempenho médio em ambientes, mas seu principal objetivo era determinar se os genótipos diferiram em sua resposta aos fatores ambientais (luz e fotoperíodo). Os autores projetaram três testes t de amostra independente para avaliar a sensibilidade ambiental de cada genótipo: (1) dias longos versus dias curtos, (2) regime de temperatura quente versus fria e (3) a interação fotoperíodo x temperatura.

Cada um deles representa um contraste de interação dentro da maior interação genótipo x ambiente. Cada teste t foi aplicado aos dados brutos de cada genótipo, totalizando 24 testes de hipóteses que focaram nos efeitos simples de cada resposta ambiental dentro de cada genótipo. Comparação desses efeitos nos oito genótipos forneceram uma avaliação de suas respostas ambientais relativas.

Diretrizes para pesquisa em escala de bancada e reação

A pesquisa em escala de reação consiste em estudos de conversão de biomassa em recipientes de reação de fluxo contínuo ou descontínuo. A escala pode variar de reatores pequenos e relativamente baratos presentes em múltiplos, para reatores grandes e caros em escala piloto ou comercial, disponíveis como uma única unidade.

⁷ Arnoult S, Quillet MC, Brancourt-Hulmel M (2014) *Miscanthus* clones display large variation in floral biology and different environmental sensitivities useful for breeding. *BioEnergy Res* 7:430–441

Pesquisa de bancada pode assumir uma ampla gama de atividades disciplinares, por exemplo, testes de germinação, crescimento de patógenos fúngicos ou bacterianos, insetos, estudos de preferência ou sobrevivência, etc. Os princípios discutidos abaixo aplicam-se igualmente a qualquer tipo de experimento comparativo crítico conduzido em material biológico que está sujeito a fontes de variação aleatórias e inexplicáveis.

É comum para descrições de bancada ou escala de reator pesquisa para conter escassa ou nenhuma referência a experiências design, geralmente referenciando o uso de duplicatas de laboratório ou triplica como a única referência à replicação. Nossa recomendação aos pesquisadores dessas disciplinas é incluir uma descrição completa do desenho experimental, de preferência usando terminologia padrão classicamente usada em livros de estatística. Idealmente, os tratamentos devem ser randomizados e replicados para que qualquer variação que seja introduzida no experimento, associado a mudanças nas condições de reação ao longo do tempo ou com a incapacidade de parar cada reação exatamente ao mesmo tempo, pode ser controlado ou removido do fixo efeitos dos tratamentos. Por exemplo, se houver algum distúrbio ou variação que está associada ao tempo de parada ou amostragem de produtos de reação, o pior cenário seria têm amostras triplicadas de cada tratamento que são agrupadas juntos, amostrados consecutivamente e analisados consecutivamente (por exemplo, Fig. 1).

Neste caso, a variação indesejável da escala de tempo é completamente confundida com a variação de tratamento desejável. Se tal situação tem o potencial de ocorrer, a Fig. 4 representa o melhor cenário, permitindo a estimativa e remoção de um efeito "bloco" ou "tempo" na ANOVA. Tal desenho pode ser aplicado mesmo no caso de um único reator de grande escala, simplesmente aplicando os tratamentos na Fig. 4 em uma sequência aleatória hora extra. Neste caso, os autores devem descrever suas duplicatas de laboratório como tendo sido organizadas de forma aleatória. bloco completo no tempo. As Figuras 2 e 3 representam soluções de nível intermediário para este problema potencial, usando o desenho completamente aleatório para organizar duplicatas ou triplicados dos tratamentos de uma maneira que reduz o potencial de confusão e viés para os efeitos do tratamento.

Independentemente do desenho específico, a randomização é uma apólice de seguro essencial contra consequências indesejáveis de uma perturbação ou fonte inesperada de variação.

Outra omissão comum nos experimentos em escala de reator é a descrição da matéria-prima e sua preparação. Isso decorre do fato de que esses tipos de experimentos geralmente dependem de amostras muito grandes em escala industrial, e é muitas vezes assumiu que a composição da matéria-prima para uma determinada fonte (por exemplo, palha de trigo, palha de milho, lascas de madeira de pinho, etc.) pode ser tratado como uma constante. Isso raramente é uma suposição precisa.

Como consequência, a falta de detalhes sobre a origem da matéria-prima torna a replicação independente muito desafiadora, se não impossível. Portanto, encorajamos os autores a fornecer como muitos detalhes sobre suas amostras na seção experimental de manuscrito, incluindo o local onde a matéria-prima foi coletado, a idade do material (em unidades de tempo ou estágios fisiológicos), qualquer informação genotípica (especialmente quando as matérias-primas produzidas comercialmente foram usadas), e o maneira como as amostras foram manuseadas após a coleta (secagem, moagem, condições de armazenamento). Por último, especialmente quando uma única amostra de matéria-prima é usada, ou quando as comparações são entre duas espécies, cada uma representada por uma única fonte, os autores precisam ter cuidado para não exagerar a conclusões e deixar espaço para a possibilidade de que observados diferenças podem não ocorrer da mesma maneira quando diferentes amostras dessas matérias-primas são usadas. Por exemplo, se dois matérias-primas foram usadas em um experimento, cuidados devem ser tomados nas conclusões para não atribuir diferenças de matéria-prima apenas a um dos muitos fatores potencialmente confundidores, ou seja, os numerosos fatores que podem ou não ter diferido entre os dois matérias-primas e seus ambientes, como espécies, data e estágio, tipo de solo, clima, etc.

A falta de replicação também é um desafio comum na pesquisa em escala de reatores envolvendo digestão anaeróbica. Esses estudos normalmente dependem de uma amostra coletada de uma estação de tratamento de esgoto instalação. Experimentos são conduzidos com esta amostra para avaliar, por exemplo, a produção de metano

após digestão de uma série de matérias-primas diferentes, ou um matéria-prima que foi preparada após diferentes pré-tratamentos. É comum relatar os rendimentos de metano durante um período de dias ou semanas, e então comparar os méritos relativos de certas matérias-primas ou tratamentos. Quando as conclusões são baseadas em um único experimento de digestão anaeróbica, as conclusões aplicam-se apenas às condições que envolvem esta amostra específica de lodo. Dado que o lodo de esgoto é um microcosmo de bactérias, é improvável que seja estático. A comunidade microbiana do dia em que a amostra foi coletada pode ser diferente de 2 dias ou 2 semanas depois, quando a composição do esgoto muda, por exemplo, devido a flutuações sazonais. Além disso, realizar experimentos de laboratório baseados em uma pequena amostra de lodo o risco de mudanças aleatórias na comunidade microbiana, análogo ao conceito de deriva genética. Como consequência, seria preferível, no mínimo, replicar experimentos envolvendo lodo de esgoto incluindo incubações paralelas com subamostras do lodo. Ainda melhor, porém, seria coletar amostras de lodo em dias diferentes ou de diferentes estações de tratamento de águas residuais, a fim de alargar as conclusões para além de uma amostra específica de lamas. Amostras de lodo de esgoto coletado em datas diferentes ou de as estações de tratamento de esgoto não devem ser a granel ou homogeneizadas entre si, mas deve ser aplicado a matérias-primas em vasos de reatores independentes como um mecanismo para avaliar a matéria-prima conversão sob condições de reação potencialmente diferentes.

Cuidados consideráveis devem ser tomados para coincidir com a replicação e designação da unidade experimental à questão a ser respondida ou à hipótese principal a ser testada. Por exemplo, se os "tratamentos" incluem vários genótipos de uma matéria-prima, réplicas de campo individuais dos genótipos não devem ser a granel, mas processado independentemente como um fator de "tratamento" adicional para as fermentações. Informações críticas adicionais para estudos de fermentação incluem a duração do período de "inanição" e o substrato padrão (com rendimento de metano) para ser usado no experimento de fermentação.

Diretrizes para Pesquisa Genética

A pesquisa genética é altamente diversificada, em escopo e propósito, bem como na metodologia estatística que é usada para desenvolver inferências genéticas. Muitos

estudos de campo genético são baseados em culturas agronômicas ou SRWC, então os princípios fundamentais que se aplicam a esses experimentos foram discutidos nessas duas seções deste artigo. Talvez as distinções mais fundamentais são: (1) experimentos genéticos tendem a ser grandes com muitos tratamentos e (2) há dois cenários em que um caso pode ser feito para não replicar experimentos genéticos ou para replicá-los de forma menos tradicional. Em relação ao tamanho do experimento, experimentos genéticos geralmente exigem projetos experimentais mais sofisticados e métodos de análise espacial para desenvolver estimativas precisas do genótipo ou das médias familiares. Delineamento em blocos incompletos tendem a ser a regra, e não a exceção, muitas vezes combinado com métodos de análise espacial. Isto tópico é discutido em grande detalhe em vários artigos de revisão e livros.

Experimentos de campo tradicionais projetados para avaliar grandes número de clones, genótipos, linhagens ou famílias estão cada vez mais se movendo em direção a estruturas de replicação modificadas que dão a pesquisa genética mais flexibilidade, inferência mais ampla e maior precisão para dados fenotípicos. A maioria dos estudos genéticos, particularmente aqueles que envolvem tentativas de desenvolver associações genótipo-fenótipo (marcador-traço) geralmente têm três características comuns: (1) centenas ou milhares de genótipos ou famílias, (2) uma necessidade ou desejo de avaliar genótipos em vários ambientes e (3) quantidades extremamente limitadas de terra, sementes ou propágulos vegetativos para replicação de genótipos.

Como tal, muitos melhoristas e geneticistas estão utilizando a teoria de alocação de recursos para minimizar ou eliminar a replicação tradicional em favor da avaliação de todos os genótipos ou famílias ao mesmo tempo e vários locais. Por exemplo, se houver sementes suficientes ou tecido vegetativo para replicar 500 genótipos quatro vezes cada, teoria de alocação de recursos e resultados empíricos, ignorando custos, quase sempre apontam para a utilização de quatro locais com uma réplica em cada local como a melhor escolha.

Para estudos puramente baseados no fenótipo, como viveiros de seleção ou avaliações de cultivares-candidatas, falta de a replicação tradicional para genótipos individuais pode ser difícil para alguns revisores aceitarem. No entanto, pode-se fazer um caso para este cenário em estudos de associação genótipo-fenótipo que são projetados para

detectar loci de características quantitativas (QTL) para características de biomassa. Nesses estudos, o genótipo individual ou família não é mais a unidade de avaliação, ou seja, geralmente não importa qual genótipo ou família é melhor ou pior. Em vez, a unidade de avaliação é o alelo, ou seja, para cada locus, precisamos para saber qual alelo é melhor e precisamos estimar que efeito do alelo no fenótipo. Assim, uma simples randomização de 500 genótipos dentro de uma área de campo bastante uniforme resultará em uma randomização válida dos dois alelos em qualquer locus. Mesmo se há algum erro envolvido na medição do fenótipo de cada planta, o número de plantas que possuem cada alelo torna-se a forma real de replicação para os alelos, ou seja, o número de genótipos ou famílias (replicação de alelos) é agora o valor crítico em relação ao poder do experimento.

Exemplo 5: um projeto de plotagem de controle para um grande número de genótipos não replicados.

Em qualquer um dos casos acima, recomendamos fortemente que qualquer experimento desenhado para avaliar um grande número de genótipos e famílias sem benefício da replicação tradicional em o nível da unidade experimental incluem um mecanismo para controlar e ajuste para variação espacial, que é inevitável e inexorável na maioria dos grandes estudos de campo. A Figura 5 fornece um exemplo de um design de grade de 421 genótipos, dos quais apenas um é replicado, ou seja, o genótipo "controle". Este desenho poderia ser aplicado a qualquer tipo de genótipo ou estrutura familiar, desde que um genótipo, linhagem ou família pode ser escolhido para uma quantidade de replicação, neste caso 80 parcelas em cada local. O projeto mostrado na Fig. 5 seria replicado no desejado número de locais e a análise de dados envolveria o uso de métodos estatísticos espaciais para ajustar a observação fenotípica de cada planta para variação espacial em linhas, colunas e quadrados.

Este desenho pertence à família de delineamentos parcialmente replicados ou aumentados, que são altamente flexíveis em relação ao tamanho, dimensão e proporção de parcelas de teste para parcelas de controle.

Tal projeto fornece dois níveis de erro controle para estimar os efeitos dos alelos: (1) ajuste espacial para aumentar a precisão e exatidão dos dados fenotípicos e (2) grande

tamanho populacional para fornecer precisão adequada para a estimativa dos efeitos dos alelos.

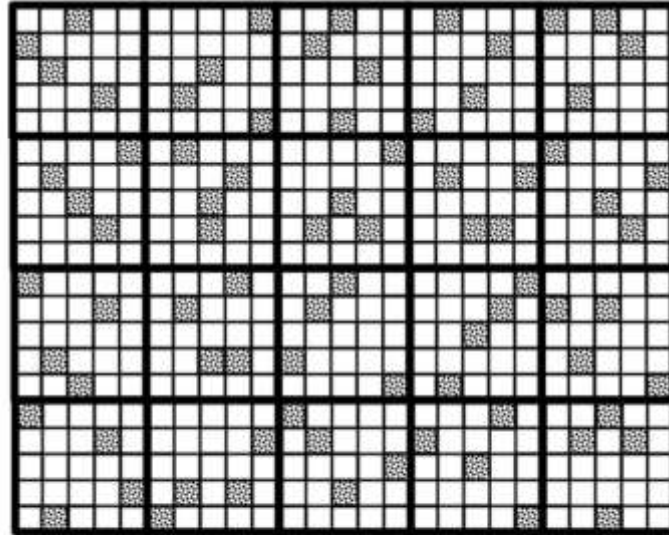


Fig. 5 Design example 5: 420 genotypes, lines, or families are evaluated in a grid design with one replicate or experimental unit per genotype at each location. In this example, 21 genotypes are randomly assigned to each of 20 groups or grids (*open squares*) and each grid contains four independent replicate plots of a single check genotype (*filled squares*). The entire design is repeated at the desired number of locations, with re-

randomization at each location. Statistical data analysis would involve the use of spatial statistical methods to adjust each plant's phenotypic observation for spatial variation in *rows*, *columns*, and *squares*. This type of design is highly flexible with regard to size, dimension, and ratio of test plots to control plots

Tipos de Experimentos de Melhoramento no Campo

RT: teste de linha (Row Trial)

O ensaio de linha é geralmente conduzido em F₃ e F₄, quando as sementes não são suficientes para a replicação com linhas de progênie de planta individual. Cada linha consiste em cerca de 20 ou mais plantas. Plantas individuais com características desejáveis são selecionadas a partir de filas de progênie superiores. Eliminam-se pragas, doenças e progênies suscetíveis com características indesejáveis.

RRT - Teste de Linha Replicada (Repeated Row Trial)

Geralmente é conduzido da geração F₃ em diante. Dependendo da disponibilidade de sementes, 3-4 mais linhas são cultivadas para cada progênie para facilitar a comparação entre progênies adotando replicações adequadas. As famílias, que se tornaram razoavelmente homozigotas, podem ser colhidas a granel. A partir dessas famílias que

mostram segregação, plantas individuais são selecionadas para os caracteres em estudo. O criador deve avaliar visualmente o potencial de produção das progênes e rejeitar as inferiores no campo e o potencial de produção deve ser avaliado em laboratório para confirmação.

PYT - Teste Preliminar de Rendimento ou Teste de Avaliação de Rendimento Inicial (IYET) - Preliminary Yield Trial or Initial Yield Evaluation Trial

É conduzido da geração F₅ em diante. Ensaio de produção preliminares com três ou mais repetições são conduzidos para avaliar o desempenho comparativo da cultura e identificar as culturas superiores entre elas. As culturas são avaliadas quanto à altura da planta, acamamento, resistência a pragas e doenças, tempo de floração, duração e rendimento, etc. Também podem ser realizados testes de qualidade. As variedades comerciais padrão devem ser incluídas como verificações para comparação. Dez a quinze culturas pendentes, se superiores às verificações, seriam avançadas para as trilhas de rendimento avançadas.

AYT - Teste de Rendimento Avançado (Advanced Yield Trial)

O Teste de Rendimento Avançado é conduzido da geração F₈ nas enfermarias. As culturas superiores identificadas no Teste de Rendimento Preliminar são testadas no Teste de Produção Replicado. Neste ensaio, as culturas são avaliadas quanto ao rendimento, pragas, resistência a doenças e acamamento, duração, qualidade, etc.

MLT - Multi Location Trial

O teste em vários locais é conduzido a partir de F₁₃ por 3 anos pelos cientistas da estação de pesquisa. Os testes de vários locais são úteis para estudos de adequação, ou seja, se uma cultura específica é capaz de ter um bom desempenho em todos os locais ou não e se a cultura específica produz todas as outras culturas desenvolvidas por estações de pesquisa e a variedade de verificação avaliada simultaneamente. Com base na avaliação, a cultura de desempenho superior e estável será promovida a ART.

ART - Teste de Pesquisa Adaptativa (Adaptive Research Trial)

É conduzido após MLT por 3 anos pelo Departamento de Agricultura. Quase 3-4 culturas são testadas e com base no desempenho de 3 anos no campo dos fazendeiros, a melhor cultura além da seleção pode ser proposta ao MAPA (Comitê Estadual de Liberação de Variedades) para liberação.

Diretrizes para Experimentos em Bioquímica e Biologia Molecular

As análises bioquímicas no tubo de ensaio geralmente estão de acordo com o estudo de bancada ou reator descrito acima. Um bom exemplo é a determinação de constantes cinéticas para duas enzimas recombinantes, sendo uma delas a enzima do tipo selvagem e a outra forma mutante (por exemplo, uma celulase projetada para atividade). Conforme descrito acima, o pior cenário seria ser ter amostras triplicadas de cada tratamento que são agrupados, amostrados consecutivamente e analisados consecutivamente. Se o ensaio enzimático foi realizado, por exemplo, usando cromatografia líquida de alta eficiência (HPLC) para quantificar a formação do produto, cada amostra deve ser analisada mais de uma vez no HPLC para determinar a variação analítica. Na maioria dos casos, isso é menor que a variação biológica (mas veja abaixo). Há também a necessidade de tal experimento para replicação adicional ao nível das preparações enzimáticas mesmos, para garantir que as diferenças aparentes na cinética entre o tipo selvagem e a enzima mutante não são o resultado de diferentes eficiências de extração enzimática, tempos diferentes desde a extração enzimática até o ensaio, ou quaisquer outras diferenças no ambiente físico em que as amostras foram manuseadas.

O método usado para determinação da atividade enzimática pode também influenciar o desenho experimental. Por exemplo, a análise por HPLC muitas vezes envolve a extração da fase aquosa em qual o ensaio enzimático foi realizado para uma fase orgânica que é injetado na coluna. A eficiência deste processo de extração (determinada pelo coeficiente de partição do produto da reação entre as fases aquosa e orgânica) pode afetar o tamanho da variação analítica, com baixas eficiências levando a variações mais altas. Tais eficiências devem ser medido e, se possível, controlado, por exemplo, por inclusão de uma norma interna que atue como proxy para o produto de interesse (aceitável, mas não ideal), ou, se o produto da reação enzimática for radiomarcado ou isótopo pesado rotulado, spiking a amostra com produto não

rotulado, ou por cravar com produto rotulado se o produto não for rotulado, antes para separação e determinação dos níveis do produto. Semelhante princípios se aplicam à quantificação dos níveis de proteína por meio de abordagens espectrométricas de massa.

A determinação dos níveis de transcrição do gene apresenta uma série de questões no que diz respeito à replicação e à interpretação.

Três métodos são geralmente usados: RT-PCR para baixa a análises de médio rendimento e microarranjo de DNA ou RNAseq para análises de nível de genoma completo de alto rendimento.

Uma questão importante com abordagens de PCR e que muitas vezes causa problemas para os autores é se o método PCR é verdadeiramente quantitativo. Por exemplo, coloração simples de géis de DNA com brometo de etídio não dá uma resposta linear e sofre de problemas de saturação, de modo que este método nunca deve ser usado para fazer inferências quantitativas.

A coloração de DNA com SYBR green/SYBR gold fornece respostas lineares em uma baixa de concentração maior. Abordagens verdadeiramente quantitativos de PCR, que exploram a relação proporcional entre a quantidade do modelo inicial e a quantidade de os produtos de PCR na fase exponencial de amplificação (ou seja, em tempo real), apresentam os mesmos problemas de replicação que encontrados para os ensaios enzimáticos, nomeadamente a necessidade de réplicas analíticas e técnicas suficientes para assegurar significância estatística. Elementos estatísticos adicionais inerentes ao método em si também exigem consideração para análise de dados e qualidade controle de experimentos de PCR em tempo real⁸.

A análise de microarrays apresenta um tipo diferente de desafio. No projeto de microarray impresso de dois canais mais antigo, o próprio processo de impressão apresenta erros e grande variação; em alguns respeitos, um microarray manchado se assemelha ao campo descrito anteriormente em que as diferenças na resposta (aqui DNA ligação) pode ser confundido por diferenças no microambiente em diferentes

⁸ Yuan JS, Reed A, Chen F, Stewart CN Jr (2006) Statistical analysis of real-time PCR data. BMC Bioinforma 7:85

partes do campo (aqui o chip). Isto é menos um fator no design do chip Affymetrix. A análise estatística de microarrays é uma disciplina inteira em si, na qual normalização da expressão gênica e controle de falsos positivos são os principais fatores de preocupação. Qualquer tipo de experimento de microarray deve, no mínimo, aderir às convenções⁹. Os tipos de considerações necessárias para superar os erros inerentes ao microarrays manchados foram discutidos¹⁰, e um exemplo dos tipos de correções necessárias na análise do Affymetrix matrizes podem ser encontradas em Naoumkina et al. (2007)¹¹.

O sequenciamento de próxima geração (RNAseq) está se tornando o método de escolha para perfis de transcrição, particularmente para aquelas espécies com genomas sequenciados ou parcialmente sequenciados (por exemplo, *Populus* ou *Panicum virgatum*). Por causa do custo ainda significativo do sequenciamento e da enorme quantidade de dados gerados, estudos observacionais sem replicação biológica são comuns, mas são claramente propensos a interpretações errôneas. O leitor é referido Auer & Doerge (2010)¹² para uma discussão dos melhores desenhos experimentais para comparações significativas de RNAseq conjuntos de dados; os princípios são fundamentalmente semelhantes aos descritos acima para o projeto de parcelas de campo e estufa.

Determinando o Número Adequado ou Necessário de réplicas

O poder do teste estatístico é a probabilidade de rejeitar uma hipótese nula que é, de fato, falsa. Em outras palavras, o poder reflete a capacidade do pesquisador de detectar diferenças verdadeiras entre os meios de tratamento - experimentos mais poderosos resultar em uma maior probabilidade de que diferenças verdadeiras sejam detectadas.

⁹ Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans S, Schulze KS, Stewart J, Taylor R, Vilo J, Vingron M (2001) Minimum information about a microarray experiment (MIAME): toward standards for microarray data. *Nat Genet* 29:365–371

¹⁰ Brown CS, Goodwin PC, Sorger PK (2000) Image metrics in the statistical analysis of DNA microarray data. *Proc Natl Acad Sci U S A* 98:8944–8949

¹¹ Naoumkina M, Farag MA, Sumner LW, Tang Y, Liu C-J, Dixon RA (2007) Different mechanisms for phytoalexin induction by pathogen and wound signals in *Medicago truncatula*. *Proc Natl Acad Sci U S A* 104:17909–17915

¹² Auer PL, Doerge RW (2010) Statistical design and analysis of RNA sequencing data. *Genetics* 185:405–416

Aumentos no poder podem ser alcançados principalmente aumentando o número de réplicas e, secundariamente, alterando a estrutura de replicação.

Exemplo 6: determine o número de réplicas necessárias para atingir a potência desejada

Considere qualquer tipo de experimento com dois níveis de replicação: verdadeira replicação de tratamentos na escala adequada e amostragem de unidades experimentais que são grandes demais para serem medidas in toto. Tal projeto se aplica a uma ampla gama de experimentos, por exemplo, fardos de biomassa como unidade experimental e núcleos dentro de fardos como unidade de amostragem, parcelas agrônomicas como a unidade experimental e as plantas dentro das parcelas como amostra unidade, ou reatores como unidade experimental e alíquotas de produtos de reação como unidade de amostragem. A Figura 6 mostra o poder estimado dos testes de hipóteses para 50 replicações potenciais cenários com uma gama de unidades experimentais de $r=3$ a 12 réplicas e unidades de amostragem de $s=3$ a 20 amostras por unidade experimental. Esses cálculos requerem dados de entrada em a forma de estimativas de variâncias de erro obtidas de experimentos e um limite de detecção desejado, ou seja, a diferença entre o tratamento significa que o pesquisador deseja detectar.

A partir deste cenário específico (variações de 5 e 10 % do média e um limite de detecção de 5 % da média), é claro que existem várias opções que podem alcançar o resultado desejado: $r=4$ $es=20$, $r=5$ $es=10$, $r=6$ $es=5$, $r=7$ $es=4$, ou $r=10$ e $s=3$. Normalmente, essas atividades levam a maioria dos pesquisadores à conclusão de que a maioria dos experimentos são insuficientes em relação aos níveis de detecção desejados. Mesmo assim, eles fornecem um mecanismo crítico para ajudar os pesquisadores a projetar experimentos melhores e mais poderosos, sem utilizar nada mais do que informações já obtidas de experimentos anteriores¹³¹⁴.

¹³ Gbur EE, Stroup WW, McCarter KS, Durham S, Young LJ, Christman M, West M, Kramer M (2012) Analysis of generalized linear mixed models in the agricultural and natural resources sciences. ASA-CSSA-SSSA, Madison

¹⁴ Casler MD (2013) Fundamentals of experimental design: guidelines for designing successful experiments. Agron J 105. doi:10.2134/agronj2013.0114

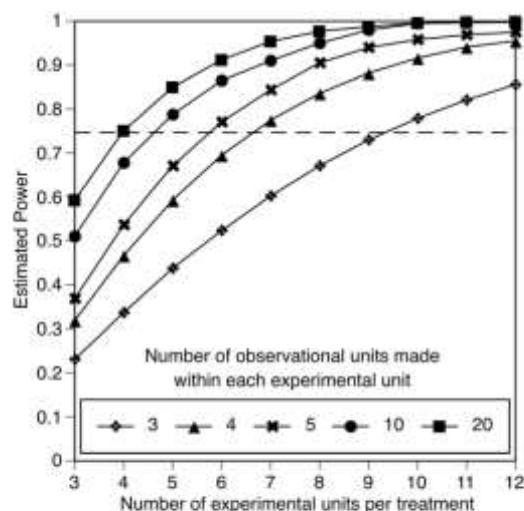


Fig. 6 Estimated power of a hypothesis test designed to detect a treatment difference of 5 % of the mean with a type I error rate of 0.05, experimental and sampling errors, respectively, of 5 and 10 % of the mean, and varying numbers of experimental units ($r=3$ to 12) and observational units ($s=3$ to 20). The dashed line represents power=0.75 and illustrates that different replication and sampling scenarios can be created to achieve the same result

Randomização

A randomização diz respeito à atribuição de tratamentos às unidades experimentais. Por definição, em um experimento adequadamente randomizado, todo tratamento tem a mesma probabilidade de ser aplicado a todas as unidades experimentais. A aplicação mais simples dessa definição é aplicar aleatoriamente r réplicas de t tratamentos a rt unidades experimentais, sem levar em conta ordem, estrutura ou prioridade. Este é o desenho completamente aleatório (CRD), que consiste em um bloco, representado nas Fig. 3 e 4. Desenhos mais complicados, envolvendo randomização hierárquica ou multietapas, são discutidos abaixo.

A randomização fornece dois serviços-chave em projetos experimentais: (i) estimativa imparcial de médias de tratamento e erros experimentais; e (ii) uma precaução contra uma perturbação que pode ou não ocorrer. A randomização é nossa apólice de seguro; como o seguro, devemos pagá-lo, não com fundos, mas com algum tempo gasto realizando randomizações e algum inconveniente na condução do experimento; no entanto, os benefícios quase sempre superam os custos. Se não houvesse gradientes de campo, variação espacial oculta e observações independentes de nossos tratamentos fornecessem resultados uniformes, não precisaríamos randomizar ou replicar nossos tratamentos; seríamos químicos ou físicos. É claro que o campo da

biologia está repleto de fontes ocultas de variabilidade, incluindo muitas fontes desconhecidas ou inesperadas durante a condução de um experimento. A randomização ajuda a garantir que essas fontes de variação desconhecidas ou inesperadas não introduzam vieses, confusão ou eliminação de testes de hipóteses válidos durante o experimento.

Isso traz à tona a questão do que é melhor: randomização ou interspersão (Hurlbert, 1984). A randomização é um processo matemático estrito no qual um gerador de números aleatórios é usado para ordenar os tratamentos. A interspersão é um conceito subjetivo em que o pesquisador se esforça para alcançar o equilíbrio e evitar o “agrupamento” de certos tratamentos ou réplicas de tratamentos. Às vezes, a randomização pode resultar em projetos experimentais que parecem ter agrupamentos ou padrões indesejáveis. Esses padrões podem resultar em viés potencial ou confundir os efeitos do tratamento se houver padrões de variação espacial subjacentes que estejam correlacionados com o plano de randomização (Martin, 1986). Eles também podem levar a diferenças significativas no nível de precisão para comparações estatísticas: tratamentos que têm distâncias médias altas entre réplicas tendem a ter variâncias maiores do que aqueles com distâncias médias baixas entre réplicas (van Es e van Es, 1993). Projetos de blocos completos espacialmente balanceados (van Es et al., 2007) foram desenvolvidos para resolver esse problema potencial, ajudando a garantir que tanto a randomização quanto a interspersão sejam componentes-chave de qualquer projeto. Esses desenhos promovem o equilíbrio espacial entre as comparações de tratamentos e não permitem que os tratamentos ocorram na mesma posição em blocos diferentes.

- A randomização tende a fazer a média entre os tratamentos, quaisquer que sejam os efeitos sistemáticos que possam estar presentes, aparentes ou ocultos, de modo que a comparação entre os tratamentos meça apenas o efeito do tratamento puro.
- A randomização é necessária não apenas para a atribuição de tratamentos às unidades experimentais, mas também para outras etapas do experimento, onde erros sistemáticos podem estar presentes.

Exemplo: Em um estudo dos efeitos da luz na taxa de crescimento das plantas, dois tratamentos são considerados: ambiente mais claro versus ambiente mais escuro. 100 plantas são aleatoriamente designadas para cada tratamento (todas geneticamente idênticas). No entanto, há apenas uma câmara de crescimento que pode cultivar 20 plantas ao mesmo tempo. Portanto, as 200 plantas precisam ser cultivadas em 10 intervalos de tempo diferentes.

Além de randomizar os tratamentos, é importante também randomizar os intervalos de tempo. Isso ocorre porque as condições da câmara de crescimento (como umidade, temperatura) podem mudar com o tempo. Portanto, cultivar todas as plantas com tratamento de luz mais brilhante nos primeiros 5 intervalos de tempo e, em seguida, cultivar todas as plantas com tratamento de luz mais escura nos últimos 5 intervalos de tempo não é um bom projeto.

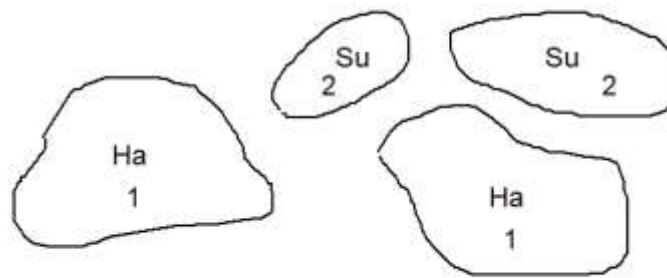
Bloqueio

Em um experimento em bloco, unidades experimentais heterogêneas (com fontes conhecidas de heterogeneidade) são divididas em subgrupos homogêneos, chamados de blocos, e experimentos aleatórios separados são conduzidos dentro de cada bloco.

- Exemplo: em um estudo de vitamina C na prevenção do resfriado, 1.000 crianças foram recrutadas. Metade deles foram escolhidos aleatoriamente e receberam vitamina C em sua dieta e a outra metade recebeu placebos. Ao final do estudo, o número de resfriados contraídos por cada criança foi registrado. (Este é um exemplo de um desenho aleatorizado completo (CRD).)
- Se sabemos (ou temos motivos suficientes para acreditar) que o gênero também pode influenciar a incidência de resfriado, então uma forma mais eficiente de conduzir o estudo é por meio do bloqueio de gênero: 500 meninas e 500 meninos foram recrutados. Entre as meninas, 250 foram escolhidas aleatoriamente e receberam vitamina C e as outras 250 receberam placebo. O mesmo é feito para os 500 meninos. (Este é um exemplo de um projeto de blocos aleatórios (RCBD).)

- Ao bloquear, remove-se a fonte de variação devido a potenciais fatores de confusão (aqui é o sexo), e assim melhora a eficiência da inferência do efeito do tratamento (aqui é a Vitamina C)
- A randomização por si só (como no CRD) não garante que o mesmo número de meninas e meninos recebam cada tratamento. Assim, se houver uma diferença na taxa de incidência de resfriado entre os sexos, as diferenças observadas entre os grupos de tratamento podem ser observadas mesmo que de fato não haja efeito do tratamento.

Neste experimento, a área disponível, além de pequena, era altamente heterogênea. Identificaram-se duas séries de solo: Suçuarana e Haras, com características físicas e químicas bem distintas. Solução - blocos não contíguos e descontínuos, de forma a garantir que cada bloco se situasse inteiramente na mesma série



Blocos descontínuos e não contíguos.

Lista de algumas estruturas de tratamento comuns utilizadas na concepção de experimentos comparativos ou manipulativos¹⁵.

Number of factors	Design name or definition	Defining characteristics
One	unstructured	there is no structure or organization to the treatments
Two or more	nested structure	factors have levels that are not repeated or have the same meaning at all levels of the other factors (Schutz and Cockerham, 1966)
Two or more	full factorial design	each factor has a specific number of levels that are repeated (have the same definition and meaning) over all levels of the other factors; the number of treatment combinations is the product of the number of levels of each factor (Cochran and Cox, 1957)
Two or more	confounding design	a full factorial in which a higher order interaction term is sacrificed as a blocking factor in order to achieve a reduction in block size (Cochran and Cox, 1957; Cox, 1958)
Two or more	composite design	a subset of a factorial designed to severely reduce the number of treatments required to evaluate the main effects and first-order interactions using regression-based modeling (Draper and John, 1988; Draper and Lin, 1990; Lucas, 1976)
Two or more	fractional factorial	a partial factorial arrangement in which only a subset of the factorial treatments is included, usually based on choosing a higher order interaction term as the defining contrast (Cochran and Cox, 1957; Cox, 1958)
Two or more	repeated measures	one or more of the treatment factors is observed over multiple time points without rerandomization of treatments to experimental units (Milliken and Johnson, 2009)

¹⁵ Casler, M.D., 2015. Fundamentals of experimental design: Guidelines for designing successful experiments. *Agronomy Journal*, 107(2), pp.692-705.

Famílias de projetos experimentais organizadas de acordo com o tipo e a complexidade dos arranjos de bloqueio.

Number of potential blocking levels	Treatment design (da Tabela em cima)	Experimental design options	Defining characteristics	References
One	any structure	randomized complete block	block size = number of treatments (t)	Steel et al. (1996)
Multiple and flexible	full factorial	split plot and variations, split block (strip plot)	design contains multiple sizes of experimental units, one for each level (or "split"); larger experimental units may be required for convenience but will be associated with increased error if $\text{Error}(a) > \text{Error}(b)$	Cochran and Cox (1957), Cox (1958), Petersen (1985), Steel et al. (1996)
Bidirectional and structured	any structure	Latin square, Graeco-Latin square, lattice square, incomplete Latin square	bidirectional blocking in perpendicular directions; number of replicates and treatments are highly restricted in some designs	Cochran and Cox (1957), Petersen (1985), Steel et al. (1996)
Multiple and flexible, bidirectional	any structure	alpha, row-column	treatments arranged in rows and columns; extremely flexible with regard to number of	John and Eccleston (1986), Patterson and Robinson (1989), Williams (1986)

Multiple and flexible	nested structure	blocks in reps (sets in reps), reps in blocks (reps in sets)	treatments, number of replicates, and block size treatments are randomly divided into sets; block size = number of treatments per set; good for inferences on random effects	Schutz and Cockerham (1966), Casler (1998)
Multiple and flexible	any structure	balanced or partially balanced incomplete blocks	potentially large reduction in block size with flexibility in both structure and field layout; block size (k) may be $t_{1/2}$ or $t_{1/3}$	Cochran and Cox (1957), Cox (1958), Petersen (1985)
Multiple and flexible	any structure with variable replication across treatments	control plot, augmented, modified augmented	designed to accommodate extremely unbalanced treatment structures and unequal replication across treatments; no restrictions on numbers of treatments and replicates	Chandra (1994), Lin and Poushinsky (1983, 1985), Wolfinger et al. (1997)

Medições de variáveis de resposta

A questão do viés de medição surge devido a diferenças irreconhecíveis no processo de avaliação.

Exemplo: O conhecimento do tratamento de um paciente pode influenciar o julgamento do médico. A fonte de viés de medição pode ser reduzida a ocultar a atribuição de tratamento tanto ao sujeito quanto ao avaliador (duplo-cego).

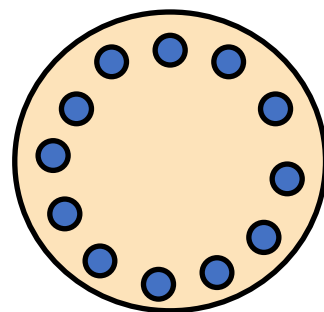
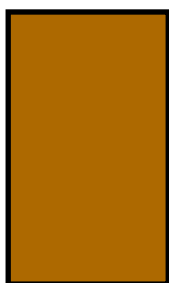
Experimentos de campo

A diferença entre um experimento de campo e um experimento de laboratório é que o primeiro ocorre em um ambiente natural em oposição a um artificial - por exemplo, uma sala de aula, um escritório, uma loja, um shopping center, uma fábrica, etc. O cenário é realista, o que tem a vantagem de não estar impondo condições artificiais, mas a desvantagem de ter menos controle.

Considerações pragmáticas podem tornar os experimentos de campo mais comuns nas ciências sociais e administrativas.

A parcela no campo

- O veículo para avaliação da resposta do material ao tratamento
- Formas
 - Retangular é o mais comum - ao longo da dimensão paralela ao qualquer gradiente
 - Leque pode ser útil quando estudando densidade
 - Forma pode ser determinada pela máquina a ou irrigação



Fatores que afetam o tamanho da parcela¹⁶:

O último pilar do desenho experimental é o menos compreendido e possui a menor quantidade de resultados teóricos para apoiar observações empíricas. Este assunto recebe pouca ou nenhuma cobertura nos livros didáticos que tratam de projeto experimental (por exemplo, menos de uma página em Steel et al., 1996), talvez devido à falta de resultados teóricos. O conceito de tamanho ótimo de parcela (unidade experimental) é muito antigo, anterior aos conceitos iniciais de Fisher de planejamento experimental e análise de variância. O conceito é baseado na Lei de Smith (Smith, 1938), que derivou da observação geral de uma relação assintótica negativa entre a variância (em uma base por unidade ou parcela única) e o tamanho da parcela (Fig. 11). A Lei de Smith parece ser um fenômeno geral que se aplica a uma ampla gama de espécies e condições ambientais (LeClerc, 1966). A Lei de Smith é mais frequentemente modelada com equações de regressão log-linear, cuja inclinação (b) pode variar amplamente entre espécies e ambientes (LeClerc, 1966), refletindo níveis diferenciais de homogeneidade espacial.

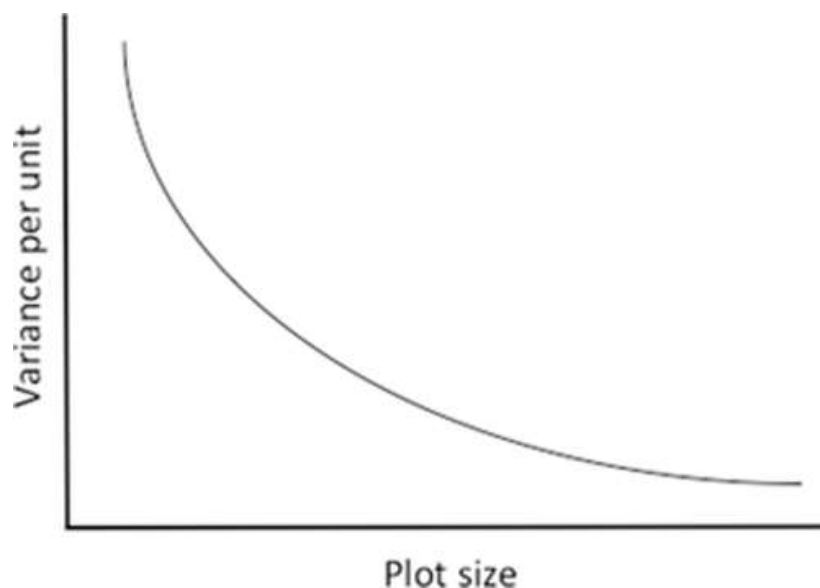


Figura . Ilustração gráfica da Lei de Smith (Smith, 1938), demonstrando a relação empírica entre variância por unidade e tamanho de parcela em experimentos agrônômicos de campo.

¹⁶ Lin CS, Binns MR (1984) Working rules for determining the plot size and numbers of plots per block in field experiments. J Agric Sci 103: 11–15

A natureza assintótica dessa relação cria um enigma para qualquer pesquisador que pretenda modificar o tamanho da parcela e prever as consequências estatísticas de tal mudança (Fig. 11). Para tamanhos de parcela iniciais pequenos, espera-se que qualquer mudança no tamanho da parcela tenha um grande efeito na variância média da unidade. Por outro lado, para pesquisadores que estão usando parcelas relativamente grandes, mudanças relativamente grandes no tamanho da parcela podem ter pouco ou nenhum impacto na variância média da unidade. É claro que a distinção entre parcelas “pequenas” e “grandes” é completamente relativa, provavelmente variando com as espécies, tipos de solo, variáveis de medição e vários outros fatores. Em meu próprio programa de pesquisa, consideramos uma parcela “pequena” como a menor área que pode ser plantada com equipamento padrão (parcelas perfuradas de cinco linhas, 0,9 por 1,2 m). Por outro lado, consideramos uma parcela “grande” como a maior distância que pode ser plantada com uma rotação da plantadora de cone sem que haja lacunas significativas na colocação de sementes (15 m) e uma largura igual a qualquer múltiplo de 0,9 m (múltiplas passagens de o plantador de cone). Para qualquer experimento, nossa escolha de tamanho de parcela depende da disponibilidade de sementes e terra e estimativas anteriores da inclinação da curva log-linear na Fig. Em cima, estimada para aquele campo em particular, conforme descrito abaixo.

A pergunta para todos que estão considerando um aumento no tamanho da parcela para obter uma diminuição na variância ou uma diminuição no tamanho da parcela para reduzir a área experimental é: “onde meus experimentos estão localizados nesta curva?” Lin e Binns (1984, 1986) usaram a Lei de Smith para desenvolver um método empírico para estimar o impacto de mudanças no tamanho de parcelas nas variâncias de erros experimentais. Eles desenvolveram uma série de cálculos simples de pós-análise a partir de qualquer projeto de bloqueio, permitindo que os pesquisadores comparassem vários projetos experimentais para experimentos futuros com base na variação do tamanho da parcela, número de tratamentos (tamanho do bloco), número de réplicas (blocos) e tamanho total do experimento. Este método é focado na estimação empírica de b , coeficiente de heterogeneidade do solo de Smith, levando a algumas generalizações. Se $b < 0,2$, aumente o número de repetições ou blocos, utilizando desenhos de blocos incompletos se o número de tratamentos for grande.

Reduções no tamanho das parcelas podem ser vantajosas, mas somente se permitirem um aumento no número de repetições. Se $b > 0,7$, aumentar o tamanho da parcela provavelmente será mais rentável do que aumentar o número de repetições. Se $0,2 < b < 0,7$, aumentar um ou ambos o tamanho da parcela e o número de repetições deve ser eficaz.

Como essas inferências são derivadas de estatísticas estimadas a partir de ANOVA ou saída de modelos mistos, elas representam variáveis aleatórias que estão sujeitas a erros de estimativa, como variação amostral e efeitos ambientais. Sua maior utilidade provavelmente derivará de muitos anos de experimentação em um local, mantendo registros históricos de estimativas de b , juntamente com dados críticos, como número de campo, orientação dentro do campo e características básicas do projeto experimental, como número de repetições, número de tratamentos, tamanho de bloco e tamanho de parcela.

- Tipo de planta
- Tipo de experimento
- Fase do programa de pesquisa
- Variabilidade do local do experimento
- Presença e natureza dos efeitos da bordadura
- Tipo de máquinas
- Número e tipos de tratamentos
- Área disponível
- Custo

	Aumenta em tamanho 				
Fator	Parcelas Pequenas				Parcelas Grandes
Variab. Solo	Uniforme				Heterogêneos
Plantio	Cereais	Plantio em linha			Árvores
Pastagem					
Fase do expto	Cedo				Tarde
Tipo de expto	Melhor.	Fertil		Prep. Solo	Irrig.
Máquina	Sem		Pesquisa		Fazenda

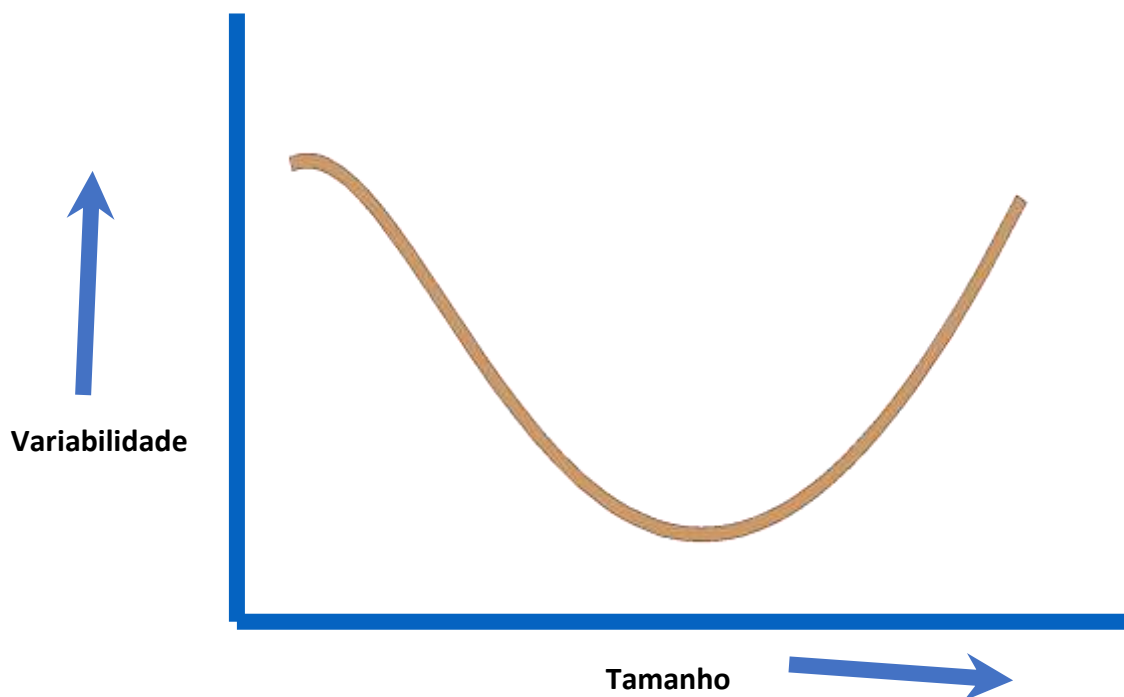
Forma e orientação de parcela

- Parcelas compridas e longas são preferências - mais econômica para trabalho de campo
- Se existe gradiente - a dimensão mais comprida deve ser na direção de maior variabilidade

Efeito na Variabilidade:

- ▶ Variabilidade por parcela diminui quando aumenta tamanho da parcela
- ▶ Parcelas grandes podem ter mais erro porque a área é mais variável
- ▶ Parcelas muito pequenas são muitas variáveis porque:
 - ▶ Perdas na coleta e erros de medição tem efeito maior
 - ▶ Número de plantas reduzidos
 - ▶ Competição e efeitos de bordadura maiores

Como tamanho de parcela afeta variabilidade



Tamanho de parcela 'Regra fácil':

- Tem limites inferiores:

- Tem que ser suficientemente grande para remover os efeitos de bordadura com suficiente sobrando para coletar e medir adequadamente
- Tem que ser suficientemente grande para as máquinas necessárias
- Quando as parcelas são suficientemente grandes para precisão é aumentado aumentando o número de repetições

Bordadura

- Plantas na beirada da parcela
 - freqüentemente desenvolve numa maneira diferente daquelas no centro
 - Diferenças maiores em parcelas menores
 - Linhas na Bordadura tem vantagem em competição
 - Máquinas pode puxar os efeitos de um tratamento dentro do próximo

Minimizando os efeitos de Bordadura

- Deixar espaços entre parcelas para minimizar efeito de máquina
- Não medir as plantas na bordadura
- Plantar parcelas de bordadura em volta do experimento

Experimento de uniformidade:

- A área é plantada com uma única cultura
- A área é dividida em áreas pequenas coletados individualmente
- produções brutas são smoothed (correndo medianos de 3 parcelas) por linha
- Smooth de novo por coluna
- Delinear áreas de produção igual
 - usado para determinar a forma e orientação das parcelas e agrupamento em blocos
 - pode ser usado para determinar o tamanho ótima de uma parcela

488	406	491	479	438	367
488	392	457	438	296	510
457	359	436	396	503	337
440	403	388	450	420	384
325	365	369	343	361	376
347	462	393	390	368	437
424	387	408	354	386	402
426	384	346	389	375	375
448	454	347	374	321	403
476	430	490	428	435	452
416	445	440	415	449	422
387	382	319	393	438	455
470	418	353	402	470	410
446	426	352	394	427	435
410	393	385	411	446	441
388	420	389	419	414	348
303	333	389	396	473	371
315	328	365	426	455	375

Smoothed correndo as medianas de 3 em linhas ➡

488	406	491	479	438	367	488	488	479	479	438	367
488	392	457	438	296	510	488	457	438	438	438	510
457	359	436	396	503	337	457	436	396	436	396	337
440	403	388	450	420	384	440	403	403	420	420	384
325	365	369	343	361	376	325	365	365	361	361	376
347	462	393	390	368	437	347	393	393	390	390	437
424	387	408	354	386	402	424	408	387	386	386	402
426	384	346	389	375	375	426	384	384	375	375	375
448	454	347	374	321	403	448	448	374	347	374	403
476	430	490	428	435	452	476	476	430	435	435	452
416	445	440	415	449	422	416	440	440	440	422	422
387	382	319	393	438	455	387	382	382	393	438	455
470	418	353	402	470	410	470	418	402	402	410	410
446	426	352	394	427	435	446	426	394	394	427	435
410	393	385	411	446	441	410	393	393	411	441	441
388	420	389	419	414	348	388	389	419	414	414	348
303	333	389	396	473	371	303	333	389	396	396	371
315	328	365	426	426	375	315	328	365	426	426	375

Smoothed correndo as medianas de 3 em colunas

488	488	479	479	438	367	488	488	479	479	438	367
488	457	438	438	438	510	488	457	438	438	438	510
457	436	396	436	396	337	457	436	403	436	420	384
440	403	403	420	420	384	440	403	396	420	396	376
325	365	365	361	361	376	347	393	393	390	390	384
347	393	393	390	390	437	347	393	387	386	386	402
424	408	387	386	386	402	424	393	387	386	386	402
426	384	384	375	375	375	426	408	384	375	375	402
448	448	374	347	374	403	448	448	384	375	375	403
476	476	430	435	435	452	448	448	430	435	422	422
416	440	440	440	422	422	416	440	430	435	435	452
387	382	382	393	438	455	416	418	402	402	422	422
470	418	402	402	410	410	446	418	394	394	427	435
446	426	394	394	427	435	446	418	394	402	427	435
410	393	393	411	441	441	410	393	394	411	427	435
388	389	419	414	414	348	388	389	393	411	414	371
303	333	389	396	396	371	315	333	389	414	414	371
315	328	365	426	426	375	315	328	365	426	426	375

Mapa de Smoothed variabilidade (Yield in dg/plot)

488	488	479	479	438	367	49	49	48	48	44	37
488	457	438	438	438	367	49	46	44	44	44	37
457	436	403	436	420	384	46	44	40	44	42	38
440	403	396	420	396	376	44	40	40	42	40	38
347	393	393	390	390	384	35	39	39	39	39	38
347	393	387	386	386	402	35	39	39	39	39	40
424	393	387	386	386	402	42	39	39	39	39	40
426	408	384	375	375	402	43	41	38	38	38	40
448	448	384	375	375	403	45	45	38	38	38	40
448	448	430	435	422	422	45	45	43	44	42	42
416	440	430	435	435	452	42	44	43	44	44	45
416	418	402	402	422	422	42	42	40	40	42	42
446	418	394	394	427	435	45	42	39	39	43	44
446	418	394	402	427	435	45	42	39	40	43	44
410	393	394	411	427	435	41	39	39	41	43	44
388	389	393	411	414	371	39	39	39	41	41	37
315	333	389	414	414	371	32	33	39	41	41	37
315	328	365	426	426	375	32	33	37	43	43	38

Layout para experimentos de campo

O objetivo básico do melhoramento de plantas é o melhoramento final da cultura. Isso resulta no desenvolvimento de variedades híbridas de alto rendimento etc., sobre as cultivares existentes e assim por diante. O desempenho das novas variedades é confirmado pelos resultados obtidos nos experimentos de campo. Para serem explicados cientificamente, os experimentos de campo são dispostos seguindo certas regras e os dados assim coletados são analisados estatisticamente. As etapas envolvidas neste processo são explicadas abaixo.

Qualquer projeto de experimentos envolve três etapas principais.

1. Seleção de unidades experimentais

Os objetos sobre os quais os tratamentos são aplicados são conhecidos como unidades experimentais. Por exemplo. Parcelas no campo, planta, etc.,

2. Fixação de tratamentos

Os objetos de comparação são conhecidos como tratamentos. Por exemplo. Variedades, espaçamentos etc.,

3. Disposição dos tratamentos nas Unidades experimentais

É composto por três princípios básicos de design

- i) Repetição: repetição de tratamentos
- ii) Randomização: alocação imparcial de tratamentos para as unidades experimentais

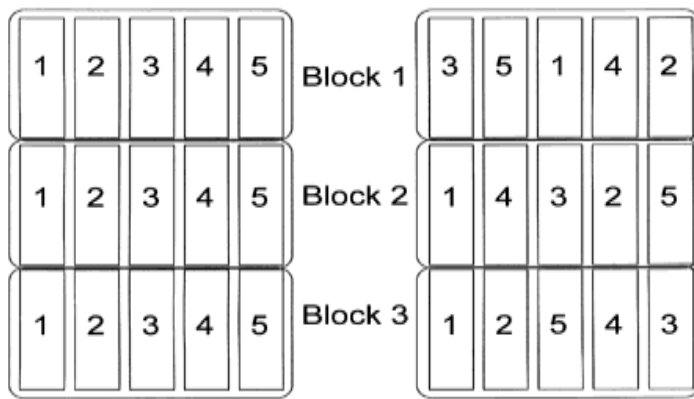


Fig. 1. Os gráficos de teste à esquerda não são aleatórios. As parcelas à direita são aleatórias. Os números (1-5) representam os cinco tratamentos neste teste

- iii) Controle local: minimizando o efeito da heterogeneidade das unidades experimentais

O objetivo da repetição, randomização e controle local é minimizar o Erro Experimental (EE). EE nada mais é do que diferenças nas respostas de unidade experimental para unidade experimental em ambientes semelhantes. Além destes, o EE pode ser reduzido ainda mais pela seleção adequada das unidades experimentais e escolha do desenho experimental mais adequado para um determinado número de tratamentos.

Tipos de desenhos experimentais básicos

1. Projeto Completamente Randomizado (CRD)
2. Projeto de Bloco Aleatório (RBD)
3. Desenho do Quadrado Latino (LSD)

Entre estes, o RBD é o design amplamente utilizado.

Disposição do RBD

A. O material experimental (campo) é dividido primeiramente em blocos constituídos por unidades experimentais homogêneas (uniformes). Cada bloco é dividido em número de tratamentos igual ao número total de tratamentos.

B. A randomização deve ser feita dentro de cada bloco e os tratamentos são aplicados seguindo a tabela de números aleatórios.

C. Coleta e análise dos dados: Após a coleta dos dados da unidade experimental individual (tratamentos) é formada a tabela ANOVA (Análise de Variância).

A significância da tabela ANOVA é que ela indica as fontes de variação exibidas pelos tratamentos, a magnitude da variação derivada de diferentes fontes e seu valor (significativo/não significativo).

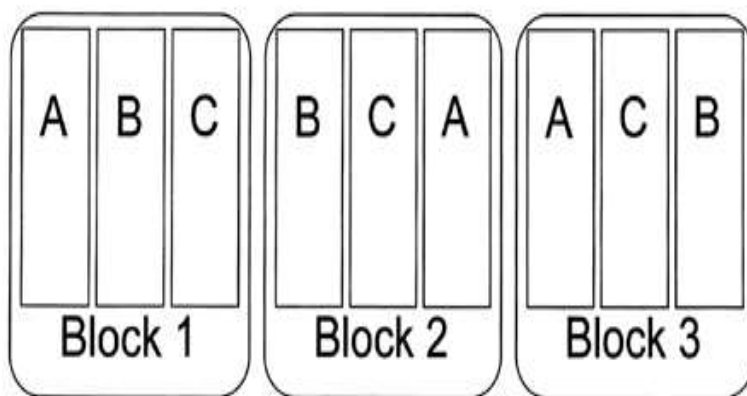


Fig. 2. Uma maneira fácil de organizar os blocos é colocá-los lado a lado no campo. As letras representam diferentes tratamentos. (Fonte: pubs.caes.uga.edu/caespubs/pubcd/B1177-w.htm)

D. Cálculo da Diferença Crítica (CD)

A Diferença Crítica é a diferença entre as médias de tratamento, o que coloca os tratamentos estatisticamente e significativamente separados. Caso contrário, se a diferença da média de dois tratamentos for menor que CD, pode-se concluir que ambos os tratamentos estão no mesmo nível.

Experimentos não replicadas

Existem inúmeras situações especiais para as quais há uma forte tentação ou necessidade de dedicar todos os recursos para vários tratamentos e nenhum para replicação ou observações independentes desses tratamentos. Muitos experimentos na fazenda representam um exemplo clássico dessa situação. Agricultores e pessoal de divulgação veem valor em tentar diferentes tratamentos em grande escala, mas veem pouco valor na replicação desses tratamentos. Os pesquisadores que organizam esses tipos de experimentos de campo têm três opções:

- (i) conduzir os experimentos em várias fazendas, usando as fazendas como blocos;
- (ii) usar desenhos de parcelas de controle em que um tratamento de controle é intercalado com os outros tratamentos, otimamente em um padrão de parcelas alternadas; ou
- (iii) usar uma combinação de ambas as abordagens.

Projetos de parcelas de controle foram desenvolvidos no início do século 20 (Pritchard, 1916; Richey, 1924) e sua popularidade durou até a década de 1970, em grande parte para uso na avaliação de números extremamente grandes de linhagens ou famílias para melhoramento de plantas (Baker e McKenzie, 1967; Mak et al., 1978). Esses projetos caíram em desuso quando ficou claro que dedicar metade das unidades experimentais a um único tratamento ou cultivar era altamente ineficiente (Melton e Finckner, 1967), e os métodos modernos de análise espacial estavam apenas surgindo (Gaeton e Guyon, 2010). No entanto, eles continuam sendo uma opção muito viável para pequenos experimentos em fazendas.

Os designs aumentados representam uma forma específica de design que pode lidar com centenas ou milhares de tratamentos, a maioria não replicada. Os tratamentos são geralmente cultivares ou linhas de reprodução que devem ser avaliadas em vários locais e para as quais os suprimentos de sementes são muitas vezes altamente limitantes e extremamente valiosos (Casler et al., 2000, 2001). Desenhos aumentados envolvem um grande número de tratamentos não replicados organizados em blocos muito pequenos, muitas vezes blocos de apenas cinco unidades experimentais (Lin e Poushinsky, 1983, 1985; Casler et al., 2000, 2001). A parcela central dentro de cada bloco consiste em uma

única cultivar que serve ao propósito de estimativa de erro e ajuste espacial em larga escala. Cultivares de verificação adicionais são adicionadas aleatoriamente a blocos individuais para aumentar a estimativa de erro e para estimar os efeitos espaciais em uma escala mais fina ou menor (Lin e Poushinsky, 1983, 1985).

Uma variação desse tema surgiu no campo da pesquisa participativa: desenhos mãe-filha (Morrone e Snapp, 2001; Snapp et al., 2002; van Eeuwijk et al., 2001). Esses projetos geralmente se concentram em um pequeno número de tratamentos práticos nos quais os agricultores ou pequenos proprietários têm um interesse genuíno e/ou um sentimento de propriedade. A aplicação mais tradicional envolve o uso de um ensaio “mãe”, realizado com a devida replicação e randomização em uma estação de pesquisa experimental, acompanhado de inúmeros ensaios “filhos” conduzidos por vários proprietários de terras (Fig. 7). Os ensaios-mãe geralmente incluem todos os tratamentos possíveis de interesse em uma região ou vila, enquanto os ensaios-filhas incluem apenas um pequeno número de tratamentos de interesse para o participante ou proprietário da terra. Cada participante é livre para projetar e/ou escolher os tratamentos de maior interesse e valor para a fazenda individual. Os ensaios de filhas estão ligados entre si e ao ensaio de mãe pelo fato de que todos os tratamentos são representados no ensaio de mãe e alguns tratamentos são repetidos em vários ensaios de filha (van Eeuwijk et al., 2001).

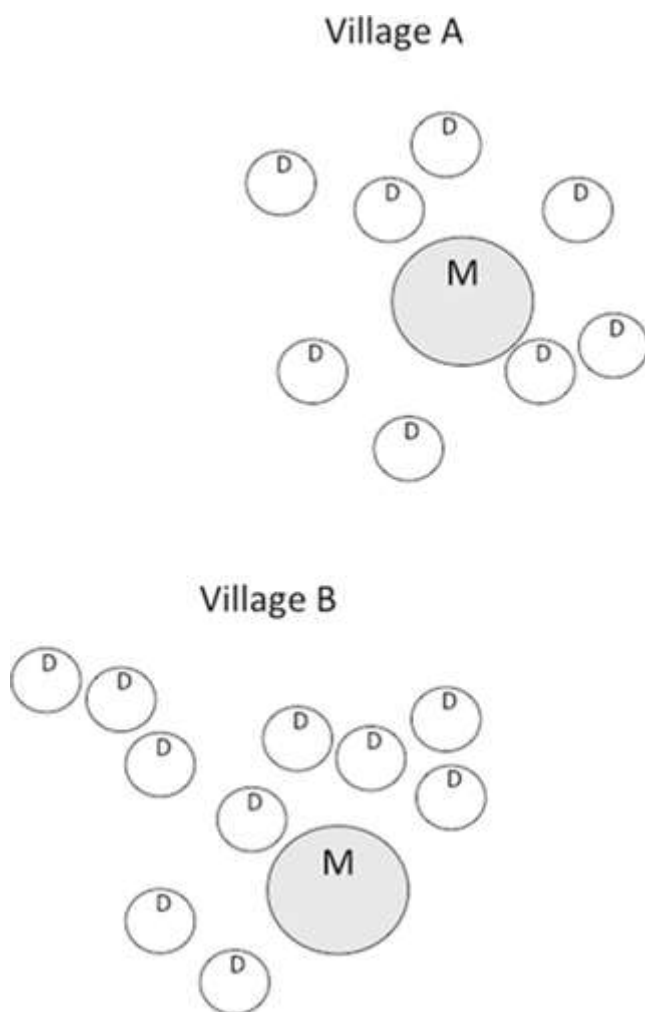


Figura. Ilustração de desenhos experimentais mãe-filha para duas aldeias de pesquisadores participativos, nos quais os ensaios-mãe (M) são conduzidos por pesquisadores em estações experimentais e os ensaios-filhas (D) são conduzidos por agricultores individuais em suas propriedades.

Por fim, ocasionalmente surgem situações em que as unidades experimentais são tão caras que a troca entre replicação e tratamentos adicionais pode favorecer o último. Como exemplo, para ser de maior valor tanto para a comunidade prática quanto para a científica, a pesquisa em pastagem deve ser realizada em vários níveis ou taxas de lotação (animais por hectare). Como as instalações de pesquisa de pastoreio são limitadas no número de piquetes discretos disponíveis para uso como unidades experimentais independentes e porque as unidades experimentais são muito caras, geralmente é impossível incluir todos os três fatores desejados: tratamentos, níveis de pastoreio e repetições. Bransby (1989) propôs delineamentos não replicados, com

análise estatística baseada em regressão de múltiplos tratamentos repetidos em vários níveis de pastejo, como uma alternativa aos desenhos replicados classicamente.

Análise de dados e apresentação de resultados

Uma reflexão significativa deve ser dada aos mais lógicos e apresentação de dados sensata para os leitores interpretarem, entenderem, e citar o trabalho. Apresentação de médias e erros padrão por si só não é aceitável para quaisquer dados em que a discussão infere diferenças de tratamento. Os leitores não podem assegurar que estas diferenças são estatisticamente significativas ou repetíveis com base em apenas médias e desvios padrão. Consequentemente, testes estatísticos apropriados precisam ser realizados e detalhes suficientes precisa ser fornecido no manuscrito para permitir replicação da análise de dados. Um comentário nos Materiais e seção de métodos que a análise estatística foi realizada com um pacote de software específico, mas sem mais detalhes é, portanto, não aceitável. O procedimento estatístico mais comum é a análise de variância (ANOVA) baseada em efeitos fixos modelos lineares ou modelos lineares de efeitos mistos padrão com matrizes de variância/covariância diagonal para efeitos aleatórios. Análise de modelos lineares mistos modernos com base em abordagens de máxima verossimilhança (REML) tornaram-se mais comum e prontamente disponível (Fig. 7).

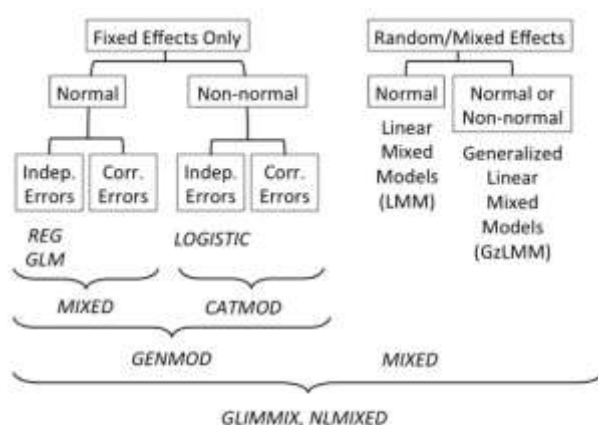


Fig. 7 Hierarchical diagram of SAS procedures (*in italics*) that are appropriate for various model assumptions and restrictions involving fixed vs. mixed or random effects at the highest level, normal vs. non-normal data at the second level, and independent vs. correlated errors at the third level. The diagram shows that generalized linear mixed models procedures that can handle the most complex cases are also the most generalizable to a wide range of model assumptions and restrictions, as well as hypothesis testing scenarios

Os autores devem relatar as seguintes informações fundamentais sobre os dados análise, conforme apropriado:

- Liste todos os efeitos que foram originalmente ajustados no modelo.
- Especificar efeitos fixos versus aleatórios, incluindo justificção.
- Identifique quaisquer efeitos de medidas repetidas no modelo.
- Identificar os critérios utilizados para selecionar o modelo final.
- Quando aplicável, identificar toda e qualquer estrutura de covariância usada no modelo final, especificamente incluindo estruturas de autocorrelação de medidas repetidas, uso de análise espacial e uso de estruturas de variância heterogêneas.
- Especifique as distribuições adequadas aos dados e as funções de link utilizadas, se apropriado.
- Especifique se e como a superdispersão foi acomodada uma abordagem generalizada de modelos lineares mistos, se necessário

Os métodos de análise de dados mudaram radicalmente nos últimos 10 anos. As suposições clássicas de ANOVA de dados normalmente distribuídos, homogeneidade de variância e erros não correlacionados não são mais necessárias devido ao desenvolvimento de métodos sofisticados de análise que podem modelar uma ampla gama de distribuições e estruturas de erro. Isso está resumido em Fig. 7 e explicado na próxima seção para o leitor interessado.

Para efeitos fixos de fatores qualitativos em que os níveis são de interesse, os autores devem considerar a seguinte opção, por ordem de preferência:

- Comparações específicas e pré-planejadas entre o tratamento significa, por exemplo, contrastes.
- Comparações de cada tratamento com um tratamento de controle, usando o teste de Dunnett.
- As comparações ad hoc de cada tratamento significam entre si média de tratamento usando um procedimento de comparação múltipla, por exemplo, LSD, HSD, etc. É mais desejável usar um teste estatística que ajusta para várias

comparações não independentes, como o teste de Tukey (HSD) ou uma correção de Bonferroni, ou uma abordagem de simulação.

Para efeitos fixos de fatores quantitativos, por exemplo, "taxa" – ou fatores relacionados ao "tempo", gráficos de dispersão ou gráficos de linha são geralmente o método de apresentação mais informativo, acompanhado por algum tipo de análise de regressão. A apresentação de resultados para fatores quantitativos pode ser efetivamente realizada em forma de tabela ou gráfico de barras se o número de níveis for muito pequeno, por exemplo, dois ou três, e se acompanhados de análise estatística para esclarecer a natureza da resposta. A legenda da tabela ou figura deve mencionar o número de réplicas usadas para calcular o desvio padrão. Dois excelentes artigos de revisão esclarecem muitos desses princípios gerais— apesar da idade desses artigos, esses são princípios atemporais¹⁷¹⁸.

Análise de conjuntos de dados grandes e complexos

A análise estatística adequada é fundamental, especialmente ao lidar com conjuntos de dados grandes e complexos que exigiam esforço e dinheiro substanciais para serem gerados. Na maioria desses casos, os estatísticos estarão envolvidos no planejamento dos experimentos e a análise dos dados, e incentiva a participação do(s) estatístico(s) na redação do manuscrito e, portanto, sua inclusão como coautor(es).

Esta seção descreve a análise desenvolvida recentemente métodos que são especialmente úteis para grandes, complexos ou conjuntos de dados não normais.

1. Modelos mistos lineares generalizados (GzLMM) oferecem ao usuário inúmeras opções para modelar uma ampla gama de distribuições comumente encontradas em dados biológicos (por exemplo, binário, binomial, binomial negativo, exponencial, Poisson, etc.). A família GzLMM fornece aos pesquisadores estruturas de ANOVA e regressão para distribuições específicas sem a necessidade de alterar os dados por transformação para atender aos pressupostos de normalidade. Além disso, tanto os modelos lineares mistos

¹⁷ Carmer SG, Walker WM (1982) Baby bear's dilemma: a statistical tale. *Agron J* 74:122–124

¹⁸ Chew V (1976) Comparing treatment means: a compendium. *HortSci* 11:348–357

quanto o GzLMM oferecem ao usuário uma ampla gama de estruturas de autocorrelação espacial e temporal, bem como a oportunidade de modelar a heterogeneidade da variância em vários níveis dentro do experimento.

2. Não paramétricos procedimentos, como o teste de Kruskal-Wallis ou Mann-Whitney também são opções válidas para projetos experimentais simples com testes de hipóteses relativamente simples e situações em que suas suposições específicas podem ser atendidas, por exemplo, simetria para Kruskal– Wallis e distribuições semelhantes para Mann-Whitney.

Usando o Sistema de Análise Estatística (SAS) como modelo, a Fig. 7 mostra oito procedimentos de software diferentes que são apropriados para situações específicas de análise estatística envolvendo efeitos fixos versus mistos ou aleatórios, normal versus não normal dados e erros independentes vs. correlacionados. Proc REG é estritamente um método baseado em regressão que é apropriado apenas em um conjunto limitado de circunstâncias. Proc GLM, Proc MIXED, Proc GENMOD e Proc GLIMMIX são todos baseados em ANOVA métodos que são projetados para resolver vários problemas e suposições falhadas nos dados, com o Proc GLIMMIX tendo a maior flexibilidade e poder de resolução de problemas.

Referências

1. Carmer SG, Walker WM (1988) Significance from the statistician's viewpoint. *J Prod Agric* 1:27–33
2. Quinn GP, Keough MJ (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge
3. Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecol Mono* 54:187–211
4. Casler MD, Vogel KP (2014) Selection for biomass yield in upland, lowland, and hybrid switchgrass. *Crop Sci* 54:626–636
5. Muttoni G, Johnson JM, Santoro N, Rhiner CJ, von Mogel KJH, Kaeppler SM, de Leon N (2012) A high-throughput core sampling device for the evaluation of maize stalk composition. *Biotech Biofuels* 5:27. doi:10.1186/1754-6834-5-27
6. Headlee WL, Zalesny RS Jr, Hall RB, Bauer EO, Bender B, Birr BA, Miller RO, Randall JA, Wiese AH (2013) Specific gravity of hybrid poplars in the north-central region, USA: within-tree variability and site×genotype effects. *Forests* 4:251–269
7. Casler MD (2013) Fundamentals of experimental design: guidelines for designing successful experiments. *Agron J* 105. doi:10.2134/agronj2013.0114
8. Steel RGC, Torrie JH, Dickey DA (1996) *Principles and procedures in statistics*, 3rd edn. McGraw-Hill, New York
9. Casler MD (1999) Repeated measures vs. repeated plantings in perennial forage grass trials: an empirical analysis of precision and accuracy. *Euphytica* 105:33–42
10. Williams ER, Matheson AC, Harwood CE (2002) *Experimental design and analysis for tree improvement*. CSIRO Publishing, Collingwood, 214pp
11. Lin CS, Binns MR (1984) Working rules for determining the plot size and numbers of plots per block in field experiments. *J Agric Sci* 103: 11–15
12. Lin CS, Binns MR (1986) Relative efficiency of two randomized block designs having different plot sizes and numbers of replications and plots per block. *Agron J* 78:531–534
13. Fisher MM, Getis A (eds) (2010) *Handbook of applied spatial analysis: Software tools, methods, and applications*. Springer, New York, 811pp
14. Stroup WW (2002) Power analysis based on spatial effects mixed models: a tool for comparing designs and analysis strategies in the presence of spatial variability. *J Agric Biol Environ Stat* 7:491–511

15. Brien CJ, Berger B, Rabie H, Tester M (2013) Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems. *Plant Methods* 9:5
<http://www.plantmethods.com/content/9/1/5>
16. Hardy EM, Blumenthal DM (2008) An efficient and inexpensive system for greenhouse pot rotation. *HortSci* 43:965–966
17. Kempthorne O (1957) 126. Query: arrangements of pots in greenhouse experiments. *Biometrics* 13:235–237
18. Edmondson RN (1989) Glasshouse design for repeatedly harvested crops. *Biometrics* 45:301–307
19. Geurtal EA, Elkins CB (1996) Spatial variability of photosynthetically active radiation in a greenhouse. *J Am Soc Hort Sci* 121:321–325
20. Wallihan EF, Garber MJ (1971) Efficiency of glasshouse pot experiments on rotating versus stationary benches. *Plant Physiol* 48:789–791
21. Arnoult S, Quillet MC, Brancourt-Hulmel M (2014) Miscanthus clones display large variation in floral biology and different environmental sensitivities useful for breeding. *BioEnergy Res* 7:430–441
22. Piepho H-P, Möhring J, Williams ER (2013) Why randomize agricultural experiments? *J Agron Crop Sci* 199:374–383
23. Rath J, Heuwinkel H, Hermann A (2013) Specific biogas yield of maize can be predicted by the interaction of four biochemical constituents. *BioEnergy Res* 6:939–952
24. Gbur EE, Stroup WW, McCarter KS, Durham S, Young LJ, Christman M, West M, Kramer M (2012) Analysis of generalized linear mixed models in the agricultural and natural resources sciences. ASA-CSSA-SSSA, Madison
25. McCann LC, Bethke PC, Casler MD, Simon PW (2012) Allocation of experimental resources used in potato breeding to minimize the variance of genotype mean chip color and tuber composition. *Crop Sci* 52:1475–1481
26. Smith AB, Lim P, Cullis BR (2006) The design and analysis of multiphase plant breeding experiments. *J Agric Sci* 144:393–409
27. Williams E, Piepho H-P, Whitaker D (2011) Augmented p-rep designs. *Biom J* 1:19–27
28. Shuford CM, Sederoff RR, Chiang VL, Muddiman DC (2012) Peptide production and decay rates affect the quantitative accuracy of protein cleavage isotope dilution mass spectrometry (PC-IDMS). *Mol Cell Proteomics* 11:814–823

29. Yuan JS, Reed A, Chen F, Stewart CN Jr (2006) Statistical analysis of real-time PCR data. *BMC Bioinforma* 7:85
30. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert CJJ, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FCP, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans S, Schulze KS, Stewart J, Taylor R, Vilo J, Vingron M (2001) Minimum information about a microarray experiment (MIAME): toward standards for microarray data. *Nat Genet* 29:365–371
31. Brown CS, Goodwin PC, Sorger PK (2000) Image metrics in the statistical analysis of DNA microarray data. *Proc Natl Acad Sci U S A* 98:8944–8949
32. Naoumkina M, Farag MA, Sumner LW, Tang Y, Liu C-J, Dixon RA (2007) Different mechanisms for phytoalexin induction by pathogen and wound signals in *Medicago truncatula*. *Proc Natl Acad Sci U S A* 104:17909–17915
33. Auer PL, Doerge RW (2010) Statistical design and analysis of RNA sequencing data. *Genetics* 185:405–416
34. Carmer SG, Walker WM (1982) Baby bear's dilemma: a statistical tale. *Agron J* 74:122–124
35. Chew V (1976) Comparing treatment means: a compendium. *HortSci* 11:348–357
36. Stroup WW (2014) Rethinking the analysis of non-normal data in plant and soil science. *Agron J* doi:10.2134/agronj2013.034237. SAS (2014) Statistical Analysis System. Cary, NC. <http://support.sas.com/documentation/>

-
- Baker, R.J., and R.I.H. McKenzie. 1967. Use of control plots in yield trials. *Crop Sci.* 7: 335– 337. doi:<https://doi.org/10.2135/cropsci1967.0011183X000700040017x>,
-
- Bransby, D.I. 1989. Compromises in the design and conduct of grazing experiments. In: G.C. Marten, editor, *Grazing research: Design, methodology, and analysis*. CSSA Spec. Publ. 16. ASA and CSSA, Madison, WI. p. 53– 67.
 - Brown, J.S., and B. Glaz. 2001. Analysis of resource allocation in final stage sugarcane clonal selection. *Crop Sci.* 41: 57– 62. doi:<https://doi.org/10.2135/cropsci2001.41157x>,
 - Casler, M.D. 1998. Genetic variation within eight populations of perennial forage grasses. *Plant Breed.* 117: 243– 249. doi:<https://doi.org/10.1111/j.1439-0523.1998.tb01933.x>,
 - Casler, M.D. 2008. Among-and-within-family selection in eight forage grass populations. *Crop Sci.* 48: 434– 442. doi:<https://doi.org/10.2135/cropsci2007.05.0267>,

- Casler, M.D., K.P. Vogel, C.M. Taliaferro, N.J. Ehlke, J.D. Berdahl, E.C. Brummer, R.L. Kallenbach, C.P. West, and R.B. Mitchell. 2007. Latitudinal and longitudinal adaptation of switchgrass populations. *Crop Sci.* **47**: 2249– 2260. doi:<https://doi.org/10.2135/cropsci2006.12.0780>,
- Casler, M.D., S.L. Fales, A.R. McElroy, M.H. Hall, L.D. Hoffman, and K.T. Leath. 2000. Genetic progress from 40 years of orchardgrass breeding in North America measured under hay management. *Crop Sci.* **40**: 1019– 1025. doi:<https://doi.org/10.2135/cropsci2000.4041019x>,
- Casler, M.D., S.L. Fales, D.J. Undersander, and A.R. McElroy. 2001. Genetic progress from 40 Years of orchardgrass breeding in North America measured under management intensive rotational grazing. *Can. J. Plant Sci.* **81**: 713– 721. doi:<https://doi.org/10.4141/P01-032>
- Chandra, S. 1994. Efficiency of check-plot designs in unreplicated field trials. *Theor. Appl. Genet.* **88**: 618– 620. doi:<https://doi.org/10.1007/BF01240927>,
- Cochran, W.G., and G.M. Cox. 1957. *Experimental designs*. John Wiley & Sons, New York.
- Cox, D.R. 1958. *Planning of experiments*. John Wiley & Sons, New York.
- Draper, N.R., and D.K.J. Lin. 1990. Small response-surface designs. *Technometrics* **32**: 187– 194. doi:<https://doi.org/10.1080/00401706.1990.10484634>,
- Draper, N.R., and J.A. John. 1988. Response-surface designs for quantitative and qualitative variables. *Technometrics* **30**: 423– 428. doi:<https://doi.org/10.1080/00401706.1988.10488437>,
- Gaeton, C., and X. Guyon. 2010. *Spatial statistics and modeling*. Springer, New York.
- Gbur, E.E., W.W. Stroup, K.S. McCarter, S. Durham, L.J. Young, M. Christman, M. West, and M. Kramer. 2012. *Analysis of generalized linear mixed models in the agricultural and natural resources sciences*. ASA, CSSA, and SSSA, Madison, WI.
- Gordillo, G.A., and H.H. Geiger. 2008. Alternative recurrent selection strategies using doubled haploid lines in hybrid maize breeding. *Crop Sci.* **48**: 911– 922. doi:<https://doi.org/10.2135/cropsci2007.04.0223>,
- Hinkelmann, K., and O. Kempthorne. 2008. Design and analysis of experiments. Vol. 1. *Introduction to experimental design*. 2nd ed. Wiley-Interscience, New York.
- Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* **54**: 187– 211. doi:<https://doi.org/10.2307/1942661>,
- John, J.A., and J.A. Eccleston. 1986. Row-column a- designs. *Biometrika* **73**: 301– 306. <https://doi.org/10.2307/2336206>

- LeClerg, E.L. 1966. Significance of experimental design in plant breeding. In: K.J. Frey, editor, *Plant breeding*. Iowa State Univ. Press, Ames. p. 243– 313.
- Legendre, P., M.R.T. Dale, M. Fortin, P. Casgrain, and J. Gurevitch. 2004. Effects of spatial structures on the results of field experiments. *Ecology* **85**: 3202– 3214. doi:<https://doi.org/10.1890/03-0677>,
- Lin, C.S., and G. Poushinsky. 1983. A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. *Biometrics* **39**: 553– 561. doi:<https://doi.org/10.2307/2531083>,
- Lin, C.S., and G. Poushinsky. 1985. A modified augmented design (Type 2) for rectangular plots. *Can. J. Plant Sci.* **65**: 743– 749. doi:<https://doi.org/10.4141/cjps85-094>,
- Lin, C.S., and H.D. Voldeng. 1989. Efficiency of Type 2 modified augmented designs in soybean variety trials. *Agron. J.* **81**: 512– 517. doi:<https://doi.org/10.2134/agronj1989.00021962008100030024x>,
- Lin, C.S., and M.R. Binns. 1984. Working rules for determining the plot size and numbers of plots per block in field experiments. *J. Agric. Sci.* **103**: 11– 15. doi:<https://doi.org/10.1017/S0021859600043276>,
- Lin, C.S., and M.R. Binns. 1986. Relative efficiency of two randomized block designs having different plot sizes and numbers of replications and plots per block. *Agron. J.* **78**: 531– 534. doi:<https://doi.org/10.2134/agronj1986.00021962007800030029x>,
- Lucas, J.M. 1976. Which response surface design is best. *Technometrics* **18**: 411– 417. doi:<https://doi.org/10.1080/00401706.1976.10489472>,
- Mak, C., B.L. Harvey, and J.D. Berdahl. 1978. An evaluation of control plots and moving means for error control in barley nurseries. *Crop Sci.* **18**: 870– 873. doi:<https://doi.org/10.2135/cropsci1978.0011183X001800050049x>,
- Martin, R.J. 1986. On the design of experiments under spatial correlation. *Biometrika* **73**: 247– 277. doi:<https://doi.org/10.1093/biomet/73.2.247>,
- McCann, L.C., P.C. Bethke, M.D. Casler, and P.W. Simon. 2012. Allocation of experimental resources to minimize the variance of genotype mean chip color and tuber composition. *Crop Sci.* **52**: 1475– 1481.
- Melton, B., and N.D. Finckner. 1967. Relative efficiency of experimental designs with systematic control plots for alfalfa yield tests. *Crop Sci.* **7**: 305– 307. doi:<https://doi.org/10.2135/cropsci1967.0011183X000700040006x>,
- Milliken, G.A., and D.E. Johnson. 2009. *Analysis of messy data. Vol. 1. Designed experiments*. 2nd ed. CRC Press, Boca Raton, FL.
- Morrone, V.L., and S.S. Snapp. 2001. Uptake of a new on-farm trial design that includes the small-holder farmer. *HortScience* **36**: 477.

- Murthy, N.S., R.V.S. Rao, and C.R. Nageswara Rao. 1994. Screening of tobacco germplasm lines using modified augmented design. *Tob. Res.* **20**: 91– 96.
- Patterson, H.D., and D.L. Robinson. 1989. Row-and-column designs with two replicates. *J. Agric. Sci.* **112**: 73– 77. doi:<https://doi.org/10.1017/S0021859600084124>,
- Patterson, H.D., and E.R. Williams. 1976. A new class of resolvable incomplete block designs. *Biometrika* **63**: 83– 92. doi:<https://doi.org/10.1093/biomet/63.1.83>,
- Peart, N. 1980. *Freewill*. Rush, Permanent Waves, Island/Mercury Records Ltd., London.
- Petersen, R.G. 1985. *Design and analysis of experiments*. Marcel Dekker, New York.
- Pritchard, F.J. 1916. The use of check plots and repeated plantings in varietal trials. *J. Am. Soc. Agron.* **8**: 65– 81.
doi:<https://doi.org/10.2134/agronj1916.00021962000800020001x>,
- Quinn, G.P., and M.J. Keough. 2002. *Experimental design and data analysis for biologists*. Cambridge Univ. Press, Cambridge, UK.
- Richey, F.D. 1924. Adjusting yields to their regression on a moving average, as a means of correcting for the heterogeneity. *J. Agric. Res.* **27**: 79– 90.
- Riesterer, J.L., M.D. Casler, D.J. Undersander, and D.K. Combs. 2000. Seasonal yield distribution of cool-season grasses following winter defoliation. *Agron. J.* **92**: 974– 980. doi:<https://doi.org/10.2134/agronj2000.925974x>,
- Sahai, H., and M.I. Ageel. 2000. *Analysis of variance: Fixed, random, and mixed models*. Springer, New York.
- Savage, A. 2009. Mythbusters. Episode 114: Demolition derby. The Discovery Channel. <http://dsc.discovery.com/tv-shows/mythbusters>
- Schutz, W.M., and C.C. Cockerham. 1966. The effect of field blocking on gain from selection. *Biometrics* **22**: 843– 863. doi:<https://doi.org/10.2307/2528078>,
- Scott, R.A., and G.A. Milliken. 1993. A SAS program for analyzing augmented randomized complete block designs. *Crop Sci.* **33**: 865– 867.
doi:<https://doi.org/10.2135/cropsci1993.0011183X003300040046x>,
- Smith, H.F. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *J. Agric. Sci.* **28**: 1– 23.
doi:<https://doi.org/10.1017/S0021859600050516>,
- Snapp, S., G. Kanyama-Phiri, B. Kamanga, R. Gilbert, and K. Wellard. 2002. Farmer and researcher partnerships in Malawi: Developing soil fertility technologies for the near-term and far-term. *Exp. Agric.* **38**: 411– 431.
doi:<https://doi.org/10.1017/S0014479702000443>
- Steel, R.G.D., J.H. Torrie, and D.A. Dickey. 1996. *Principles and procedures in statistics*. 3rd ed. McGraw-Hill, New York.

- vanEeuwijk, F.A., M. Cooper, I.H. DeLacy, S. Ceccarelli, and S. Grando. 2001. Some vocabulary and grammar for the analysis of multi-environment trials, as applied to the analysis of FPB and PPB trials. *Euphytica* **122**: 477– 490. doi:<https://doi.org/10.1023/A:1017591407285>,
 - vanEs, H.M., and C.L.vanEs. 1993. Spatial nature of randomization and its effect on the outcome of field experiments. *Agron. J.* **85**: 420– 428. doi:<https://doi.org/10.2134/agronj1993.00021962008500020046x>,
 - vanEs, H.M., C.P. Gomes, M. Sellman, and C.L.vanEs. 2007. Spatially-balanced complete block designs for field experiments. *Geoderma* **140**: 346– 352. doi:<https://doi.org/10.1016/j.geoderma.2007.04.017>,
 - Waters, R., R. Wright, and D. Gilmour. 1975. Shine on you crazy diamond. Pink Floyd, Wish You Were Here, Columbia/CBS Records, New York.
 - Williams, E.R. 1986. Row and column designs with contiguous replicates. *Aust. J. Stat.* **28**: 154– 163. doi:<https://doi.org/10.1111/j.1467-842X.1986.tb00594.x>,
 - Wolfinger, R.D., W.T. Federer, and O. Cordero-Brana. 1997. Recovering information in augmented designs, using SAS PROC GLM and PROC MIXED. *Agron. J.* **89**: 856– 859. doi:<https://doi.org/10.2134/agronj1997.00021962008900060002x>,
 - Yaklich, R.W., B. Vinyard, M. Camp, and S. Douglass. 2002. Analysis of seed protein and oil from soybean northern and southern region uniform tests. *Crop Sci.* **42**: 1504– 1515. doi:<https://doi.org/10.2135/cropsci2002.1504>,
-

Abbreviations

SRWC Short-rotation woody crops

QTL Quantitative trait loci

IVDMD In vitro dry matter digestibility

LSD Least significant difference

HSD Honestly significant difference (Tukey's LSD)

HPLC High-performance liquid chromatography

PCR Polymerase chain reaction

RT-PCR Real-time PCR

RNAseq High-throughput expression profiling (RNA sequencing)

APPENDIX 1¹⁹

SAS code to compute Tukey's test for nonadditivity in a randomized complete block design. The test is computed in two parts: (1) output the predicted values from the mixed models analysis, then (2) square the predicted values and include this term as a single-degree-of-freedom regressor variable (covariate) as a test for multiplicative block and treatment effects (Sahai and Ageel, 2000).

- data a; infile 'tukey.dat';
- input rep block trt y;
- proc mixed; class trt;
- model pcb = trt / outpred=x;
- random block;
- data x; set x;
- p2=pred*pred;
- proc mixed; class trt;
- model pcb = trt p2;
- random block;

APPENDIX 2

This SAS code to estimate the power of a hypothesis test has the following parameters: treatment means = 95 and 100, variance components = 5 and 10 (experimental error and sampling error, respectively), $r = 4$ replicates, $s = 2$ observational units per experimental unit, and the assumption of normally distributed errors (adapted from Gbur et al., 2012).

- options nocenter mprint;
- data a; input trt y;
- do rep=1 to 4 by 1;
- do samples=1 to 2 by 1;
- output;
- end;
- end;
- datalines;
- 1 95
- 2 100
- run;
- proc glimmix; class trt rep;
- model y = trt;
- random rep(trt);
- parms (5)(10) / hold=1,2;
- ods output tests3=power_terms;
- data power;
- set power_terms;
- alpha=0.05;
- ncparm=numdf*Fvalue;
- F_critical=finv(1-alpha, numdf, dendf, 0);
- power=1-probf(F_critical, numdf, dendf, ncparm);
- proc print;

¹⁹ <https://access.onlinelibrary.wiley.com/doi/full/10.2134/agronj2013.0114>

- run;

APPENDIX 3

The SAS code from [Appendix 1](#) is embedded in a macro that allows power to be estimated for a range of experimental designs with four to eight experimental units per treatment (rep, repl, repmax) and two to four observational units per experimental unit (obs, obsv, obsmax).

- options nocenter mprint;
- %macro one(obsmax,repmax);
- data a;
- %do obsv=2 %to &obsmax;
- group1=&obsv;
- %do repl=2 %to &repmax;
- group2=&repl;
- do obs=2 to &obsv by 1;
- do rep=4 to &repl by 1;
- do trt=0 to 1 by 1;
- output;
- end;
- end;
- end;
- %end;
- %end;
- %mend one;
- %one(4,8); /* change values here */
- run;
- proc sort; by group1 group2;
- data b; set a; by group1 group2;
- if trt=0 then y=95;
- if trt=1 then y=100;
- run;
- proc glimmix; class trt rep; by group1 group2;
- model y = trt;
- random rep(trt);
- parms (5)(10) / hold=1,2;
- ods output tests3=power_terms;
- data power;
- set power_terms;
- alpha=0.05;
- ncparm=numdf*Fvalue;
- F_critical=finv(1-alpha, numdf, dendif, o);
- power=1-probf(F_critical, numdf, dendif, ncparm);
- proc print;
- run;

APPENDIX 4

This SAS code to estimate the power of a hypothesis test has the following parameters: treatment means = 9 and 10, variance components = 0.02 and 0.2 (treatment $\times$ location interaction and experimental error, respectively), $r = 4$ to 20 replicates, $l = 2$ to 6 locations, and the assumption of normally distributed errors (adapted from [Gbur et al., 2012](#)).

- options nocenter mprint;
- %macro two(locmax,repmax);
- data a;
- %do locn=2 %to &locmax;
- group1=&locn;
- %do repl=4 %to &repmax;
- group2=&repl;
- do loc=2 to &locn by 1;
- do rep=4 to &repl by 1;
- do trt=0 to 1 by 1;
- output;
- end;
- end;
- end;
- %end;
- %end;
- %mend two;
- %two(6,20); /* change here */
- run;
- proc sort; by group1 group2;
- data b; set a; by group1 group2;
- if trt=0 then y=9;
- if trt=1 then y=10;
- run;
- proc glimmix; class location trt rep; by group1 group2;
- model y = trt;
- random trt*location;
- parms (0.02)(0.2) / hold=1,2;
- ods output tests3=power_terms;
- data power;
- set power_terms;
- alpha=0.05;
- ncparm=numdf*Fvalue;
- F_critical=finv(1-alpha, numdf, dendf, 0);
- power=1-probf(F_critical, numdf, dendf, ncparm);
- proc print;
- run;