



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA

ANÁLISE DE EXPERIMENTOS EM BLOCOS INCOMPLETOS

ARTHUR MOURA TREVISOL
KÁTIA FLÁVIA SANTOS GUIMARÃES COSME

Brasília

2013

ARTHUR MOURA TREVISOL
KÁTIA FLÁVIA SANTOS GUIMARÃES COSME

ANÁLISE DE EXPERIMENTOS EM BLOCOS INCOMPLETOS

Relatório apresentado à disciplina Estágio Supervisionado II do curso de graduação em Estatística, Instituto de Ciências Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Orientador: Prof. Dr. Lúcio José Vivaldi

Brasília

2013

Dedico este trabalho aos meus pais, por tudo que eles representam na minha vida, às minhas irmãs e familiares; minha grande amiga e colega de trabalho Flávia por todo o companheirismo durante a graduação.

Arthur Moura Trevisol

Dedico este trabalho aos meus pais, pelo amor, apoio e pelos valores por eles passados a mim, a minha irmã e familiares; a todos os meus verdadeiros amigos, em especial ao Arthur, grande companheiro e colega de trabalho.

Kátia Flávia Santos Guimarães Cosme

AGRADECIMENTOS

Agradecemos primeiramente a Deus, que sempre nos iluminou, tornando possível a conclusão de mais uma etapa de nossas vidas.

Ao professor Lúcio José Vivaldi por sua orientação, dedicação e inestimável auxílio durante todo o trabalho. Foi uma honra ter seus conhecimentos transmitidos a nós.

Aos demais professores, funcionários e colegas do Departamento de Estatística da Universidade de Brasília, pelo apoio e incentivo.

À Embrapa Cerrados, por ter cedido o conjunto de dados para realização das análises.

Ao SAS *Institute* Brasil por possibilitar a utilização desse software por meio de parceria acadêmica com o Departamento de Estatística da UnB.

RESUMO

Experimentos são conduzidos por pesquisadores para responder questões cujas respostas não podem ser obtidas por meio de um raciocínio dedutivo. Já o delineamento experimental trata-se do planejamento do experimento e está mais relacionado à estatística. Fisher introduziu esse assunto e estabeleceu os princípios sobre delineamento – repetição, casualização e blocagem – em seu livro, *The Design of Experiments* (1935). Dentro do princípio denominado de blocagem, existe a categoria de blocos incompletos, o foco deste trabalho.

Visando explicar as diversas técnicas de análise de blocos incompletos existentes, o estudo em questão analisa 42 experimentos com soja em blocos incompletos resolvíveis. Para cada experimento são realizadas cinco análises, sendo essas: análise de blocos ao acaso segundo Speed, Williams e Patterson (1985), análise intrabloco e análise com recuperação interbloco segundo Yates (1936, 1940), análise espacial segundo Papadakis (1937) e Bartlett (1978) e análise como um modelo de blocos incompletos e variância linear segundo Williams (1986).

A pesquisa sobre a soja é um assunto bastante pertinente, tendo em vista que, com base nos dados divulgados pela Embrapa, referentes à safra 2010/2011, o Brasil é o segundo maior produtor de soja do mundo, produzindo nesse período 75 milhões de toneladas. Ficando atrás apenas dos Estados Unidos, que no mesmo intervalo de tempo produziu 90,6 milhões de toneladas de soja. Além disso, no ano de 2012, se tratando de cereais, leguminosas e oleaginosas, 50,89% da área plantada no Brasil foi destinada a produção de soja, segundo os dados do Levantamento Sistemático da Produção Agrícola, Outubro/2012, do Instituto Brasileiro de Geografia e Estatística (IBGE).

Palavras-chave: análise, blocos ao acaso, blocos incompletos, delineamento, experimento, interbloco, intrabloco, método, recuperação.

LISTA DE QUADROS

Quadro 1 – Delineamento de blocos completos com seis tratamentos em quatro blocos de seis parcelas	5
Quadro 2 – Delineamento de blocos incompletos com quatro tratamentos em quatro blocos de três parcelas	10
Quadro 3 – Delineamento de blocos incompletos balanceados com quatro tratamentos em quatro blocos de três parcelas	11
Quadro 4 – Delineamento de blocos incompletos balanceados que podem ser agrupados em repetições com seis tratamentos em dez blocos de três parcelas	11
Quadro 5 – Delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com uma repetição com seis tratamentos em quinze blocos de duas parcelas	12
Quadro 6 – Delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com mais de uma repetição com cinco tratamentos em dez blocos de duas parcelas ..	13
Quadro 7 – Delineamento de blocos incompletos parcialmente balanceados com seis tratamentos em três blocos de quatro parcelas	14
Quadro 8 – Delineamento de blocos incompletos parcialmente balanceados e resolvíveis com seis tratamentos em quatro blocos de três parcelas	15
Quadro 9 – Delineamento de blocos incompletos parcialmente balanceados e não resolvíveis com seis tratamentos em quatro blocos de três parcelas	16
Quadro 10 – Delineamento de blocos incompletos com seis tratamentos em quatro blocos de três parcelas	19
Quadro 11 – Representação gráfica de N parcelas de um experimento num plano	31
Quadro 12 – Delineamento de blocos incompletos com dez tratamentos em quinze blocos de quatro parcelas	60

LISTA DE TABELAS

Tabela 1 – Análise de variância para delineamento completamente casualizado	7
Tabela 2 – Análise de variância para delineamento de blocos incompletos balanceados que podem ser agrupados em repetições com seis tratamentos em dez blocos de três parcelas	12
Tabela 3 – Análise de variância para delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com uma repetição com seis tratamentos em quinze blocos de duas parcelas.....	12
Tabela 4 – Análise de variância para delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com mais de uma repetição com cinco tratamentos em dez blocos de duas parcelas.....	13
Tabela 5 – Análise de variância intrabloco	18
Tabela 6 – Análise de variância interbloco	21
Tabela 7 – Análise de resultados da análise com recuperação da informação interbloco (quadrados mínimos).....	22
Tabela 8 – Análise de resultados da análise com recuperação da informação interbloco (máxima verossimilhança restrita).....	22
Tabela 9 – Teoria da análise de variância intrabloco	25
Tabela 10 – Teoria da análise de variância interbloco	27
Tabela 11 – Resultados de todas as análises.....	40
Tabela 12 – Análises recomendadas conforme estimação de η e ϕ	41

SUMÁRIO

1	INTRODUÇÃO	1
1.1	EXPERIMENTO E DELINEAMENTO	1
1.2	PRINCÍPIOS DA EXPERIMENTAÇÃO	2
1.2.1	Repetição	3
1.2.2	Casualização	3
1.2.3	Blocagem	3
1.3	ANÁLISE DE VARIÂNCIA.....	5
2	BLOCOS INCOMPLETOS	9
2.1	INTRODUÇÃO	9
2.2	TIPOS DE BLOCOS INCOMPLETOS	10
2.2.1	Blocos Incompletos Balanceados (BIB).....	10
2.2.2	Blocos Incompletos Parcialmente Balanceados (BIPB).....	13
2.2.3	Blocos Incompletos Resolvíveis.....	14
2.2.4	Blocos Incompletos Não Resolvíveis	15
3	ANÁLISE DE BLOCOS INCOMPLETOS.....	17
3.1	ANÁLISE INTRABLOCO.....	17
3.2	ANÁLISE INTERBLOCO	18
3.3	ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO	21
4	TEORIA DA ANÁLISE DE BLOCOS INCOMPLETOS	23
4.1	INTRODUÇÃO	23
4.2	ANÁLISE INTRABLOCO.....	23
4.3	ANÁLISE INTERBLOCO	25
4.4	ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO	28
5	ANÁLISE ESPACIAL.....	30
5.1	INTRODUÇÃO	30
5.2	MÉTODO DE PAPADAKIS.....	30
5.3	MÉTODO DE WILLIAMS	33

6. APLICAÇÕES E RESULTADOS.....	37
7. CONCLUSÃO	42
8. REFERÊNCIAS	43
APÊNDICE A – Programação em SAS (2003) dos dados de Federer (1963)	45
APÊNDICE B – Programação em SAS (2003) dos 42 experimentos analisados	49
ANEXO A – Dados de Federer (1963)	60

1 INTRODUÇÃO

1.1 EXPERIMENTO E DELINEAMENTO

A base de toda a ciência é a observação, é esse ato que move a vontade de pesquisar a respeito dos mais variados assuntos em diversas áreas do conhecimento. Considerando duas ou mais variáveis, um estudo pode ser completamente observacional, que investiga a associação entre essas variáveis sem controlar possíveis fatores que influenciam o estudo; experimental, quando são analisadas as relações de causa e efeito entre as variáveis explicativas e a variável resposta, isso acontece porque, nesse caso, existe controle dos fatores que influenciam na resposta observada, ou; misto, quando são encontradas no estudo características dos dois tipos (observacional e experimental).

Tem-se que, experimento é uma observação planejada em condições arranjadas pelo observador e é composto de uma série de testes em que alterações são propositalmente efetuadas nas variáveis explicativas (input) de um processo ou sistema para permitir a observação e identificação das razões de alterações ocorridas na variável resposta (output) do processo. Sendo assim, na realização de um experimento, o pesquisador não deseja controlar a resposta do processo, mas identificar fontes de variação que influenciaram em tal resposta (output) quando o processo está submetido a um conjunto de condições (input).

Porém, qualquer experimento precisa ser planejado para ser executado da melhor maneira possível, para esse planejamento dar-se o nome de delineamento. Sendo assim, o delineamento experimental é uma técnica estatística utilizada para obter, através do experimento, a maior quantidade de informação possível com os recursos disponíveis.

Estudos e metodologias fundamentam e padronizam cada vez mais os experimentos e um dos cientistas que mais contribuiu nesses estudos foi Ronald Aylmer Fisher (1890 – 1962), que trabalhou com experimentos aplicados na biologia por mais de trinta anos, descobrindo e consolidando os princípios básicos do delineamento experimental. Grande parte dessa monografia está relacionada às suas ideias.

1.2 PRINCÍPIOS DA EXPERIMENTAÇÃO

Os delineamentos de experimentos foram introduzidos por Fisher em alguns artigos na década de vinte do século passado, mas foi em seu livro de 1935, *The Design of Experiments*, que ele estabeleceu os seguintes princípios sobre delineamentos de experimentos: repetição, casualização e blocagem.

Embora as características de um experimento variem de caso para caso, um experimento na área científica ou tecnológica pode ser considerado como um sistema construído pelo pesquisador, parcialmente controlado e que vai fornecer as informações requeridas pela pesquisa desejada. Sendo assim, ele trás resultados influenciados por alguns fatores que são controlados pelo experimentador e por outros que não são. As variáveis controláveis são aquelas que fazem parte do sistema e aquelas que o pesquisador impõe ao sistema, essa última chamada de tratamentos. Já as variáveis não controláveis dividem-se em covariáveis e aquelas que causam o erro experimental, sendo a primeira, variáveis que podem ser observadas e utilizadas na análise e a segunda, aquelas que não podem ser observadas.

No delineamento experimental trabalha-se com tratamentos, unidades experimentais e erro experimental. Os tratamentos são definidos como estímulos formulados pelo pesquisador que aplicados ao sistema geram uma mudança no valor da variável resposta. A unidade experimental é a entidade física que vai dar origem às observações, sendo o local onde o tratamento será aplicado. O erro experimental reflete a variação entre as unidades experimentais definidas no sistema e independe dos tratamentos. Para um melhor experimento, procura-se encontrar as unidades experimentais mais homogêneas possíveis, pois quanto mais idênticas forem, menor e mais desprezível se torna o erro experimental.

Em experimentos, além do erro experimental, existem, pelo menos, outros três tipos de erro. O chamado erro de mensuração, associado às variáveis respostas e que se origina do instrumental ou do meio utilizado para quantificar estas variáveis; balanças e aparelhos de laboratório sensíveis utilizados na mensuração podem ocasionar esse tipo de erro. A forma incorreta de aplicação dos tratamentos ou a quantidade errada utilizada em uma unidade experimental causa o erro de aplicação. A modificação da unidade ao longo do tempo, como um dano físico não registrado ocorrido em uma delas, vai produzir o erro de unidade. Infelizmente, não há técnicas estatísticas que possam tratar estes três tipos de erro e no final

eles vão aumentar o erro experimental e viesar as comparações entre tratamentos, o melhor que se pode fazer é evitá-los.

1.2.1 Repetição

A repetição tem grande importância no delineamento experimental. Se o mesmo tratamento for alocado a duas unidades experimentais homogêneas, a diferença entre elas encerra a influência dos fatores não controláveis e assim, é possível detectar a magnitude que o erro experimental possui, ou seja, a repetição é condição necessária para se estimar o erro experimental e por isso é considerada um dos princípios da experimentação.

1.2.2 Casualização

Após alguns trabalhos, Fisher considerou a casualização como um dos princípios da experimentação. Ele defendia que a alocação de um determinado tratamento em uma determinada unidade experimental deveria ser aleatória e possuir a mesma probabilidade para cada alocação. A casualização dá sustentação à análise estatística de delineamentos de experimentos e é obtida mediante um sorteio que pode ocorrer através de urnas ou programas computacionais. As funções da casualização foram descritas no clássico livro de Cox (1958), no capítulo 5 e são as seguintes:

- i. Permite estimar o erro experimental de forma não viesada, sendo a repetição necessária, mas não suficiente;
- ii. Possibilita o uso de testes de hipóteses, e;
- iii. Possibilita a estimação não viesada da diferença entre os efeitos de tratamentos.

A casualização é, portanto, o princípio que dá sustentação à análise estatística dos delineamentos de experimentos. Uma teoria sobre a casualização em experimentos está em Hinkelmann e Kempthorne (1994).

1.2.3 Blocagem

Além de ser um dos princípios da experimentação, os experimentos em blocos formam um dos tipos mais importante de delineamento. Sabe-se que em um experimento

completamente casualizado é primordial que as unidades experimentais sejam as mais homogêneas possíveis para que o efeito do tratamento aplicado seja observado na variável resposta com um baixo erro experimental. Porém, existem situações que as unidades experimentais não são homogêneas e diferem-se em alguma característica que pode comprometer a análise dos resultados dos tratamentos, devido ao grande erro experimental ou ao viés. Nesse caso, é necessário fazer uma divisão das unidades experimentais de acordo com a característica divergente, denominada divisão em blocos.

O exemplo utilizado para demonstrar esse tipo de delineamento foi retirado do capítulo 8 de Kuehl (2001) e traduzido para a língua portuguesa.

Exemplo 1:

As recomendações atuais para a adubação de milho com nitrogênio incluem a aplicação de quantidades específicas em fases estabelecidas do crescimento da planta. As recomendações foram desenvolvidas através de uma análise periódica da quantidade de nitrogênio contida nos tecidos da espiga. A análise do tecido é um meio eficaz para supervisionar a quantidade de nitrogênio na colheita e, conseqüentemente, ter uma base para previsão do nitrogênio necessário para uma produção ótima.

Objetivo de investigação: O pesquisador quer avaliar o efeito de diferentes programas de adubação com aplicações distintas de nitrogênio para refinar as recomendações do procedimento.

Delineamento de tratamento: O delineamento de tratamento inclui seis programas diferentes de aplicação de nitrogênio que podem proporcionar as condições necessárias para avaliar o processo.

Delineamento de experimento: As respostas das plantas dependem da variabilidade da umidade disponível. Com o intuito de diminuir o erro experimental, as unidades experimentais são divididas em quatro blocos com umidades distintas, cada um com seis parcelas. Assim, qualquer diferença nas respostas das plantas causadas pela umidade pode ser associada aos blocos. O delineamento experimental resultante foi o delineamento de blocos ao acaso, com quatro blocos, cada um com seis parcelas que se alocaram aleatoriamente aos seis tratamentos de programa de aplicação de nitrogênio.

A distribuição das unidades experimentais no campo é exibida no Quadro 1, onde são apresentadas as quantidades de nitrogênio observadas ($\text{ppm} \times 10^{-2}$) para cada espiga de milho, juntamente com o número do tratamento e bloco do qual a unidade foi submetida.

Quadro 1 – Delineamento de blocos completos com seis tratamentos em quatro blocos de seis parcelas

Bloco 1	2	5	4	1	6	3
	40,89	37,99	37,18	34,98	34,89	42,07
Bloco 2	1	3	4	6	5	2
	41,22	49,42	45,85	50,15	41,99	46,69
Bloco 3	6	3	5	1	2	4
	44,57	52,68	37,61	36,94	46,65	40,23
Bloco 4	2	4	6	5	3	1
	41,90	39,20	43,29	40,45	42,91	39,97

1.3 ANÁLISE DE VARIÂNCIA

A análise de variância foi introduzida por Fisher na década de 1920 e é uma das técnicas mais difundidas entre os pesquisadores. Ela é um procedimento que visa verificar se existe diferença significativa entre os tratamentos analisados e se esses exercem influência sobre as unidades experimentais; possibilita estimar o erro experimental, testar hipóteses e comparar tratamentos.

Para melhor explicar sobre o assunto, considere Y uma variável resposta, pode-se observar que Y varia de unidade para unidade, ou seja, existe uma variabilidade dos valores de Y , denominada de variabilidade total; a análise de variância consiste em separar e quantificar as causas dessa variabilidade. Primeiramente, algumas das causas (ou fontes de variação, como são conhecidas), são devidas ao delineamento utilizado; a segunda fonte é devida a ação dos tratamentos impostos ao sistema pelo experimentador; a última fonte de variação é referente ao erro experimental, originária da desigualdade das unidades experimentais.

A forma utilizada por Fisher (1935) para concretizar a separação dos componentes da variabilidade total foi dada pela soma de quadrados em torno das médias. A análise de variância de um experimento completamente casualizado firma-se no modelo (1.1),

apresentado a seguir, que representa Y_{ij} , valor da variável resposta Y correspondente à j -ésima repetição do i -ésimo tratamento, considerando r repetições e t tratamentos.

$$Y_{ij} = \mu + \tau_i + e_{ij} \quad (1.1)$$

Onde,

μ : média geral;

τ_i : efeito do tratamento i , e;

e_{ij} : erro experimental, $e_{ij} \sim N(0, \sigma_e^2)$.

A variabilidade total da variável resposta Y , chamada de Soma de Quadrados Total (SQTotal), é dada por:

$$\text{SQTotal} = Y_{11}^2 + Y_{12}^2 + \dots + Y_{tr}^2 = \sum_{i=1}^t \sum_{j=1}^r Y_{ij}^2 \quad (1.2)$$

A variabilidade total pode ser decomposta da seguinte forma:

$$\text{SQTotal} = \sum_{i=1}^t \sum_{j=1}^r Y_{ij}^2 = \sum_{i=1}^t \sum_{j=1}^r \hat{\mu}^2 + \sum_{i=1}^t \sum_{j=1}^r \hat{\tau}_i^2 + \sum_{i=1}^t \sum_{j=1}^r \hat{e}_{ij}^2 \quad (1.3)$$

Onde $\hat{\mu}$, $\hat{\tau}_i$ e \hat{e}_{ij} são, respectivamente, as estimativas de μ , τ_i e e_{ij} . Expressas por:

$\hat{\mu} = \bar{Y}_{..}$, onde $\bar{Y}_{..}$: média geral dos dados;

$\hat{\tau}_i = \bar{Y}_{i.} - \bar{Y}_{..}$, onde $\bar{Y}_{i.}$: média do tratamento i ;

$\hat{Y}_{ij} = \hat{\mu} + \hat{\tau}_i$, onde \hat{Y}_{ij} : valor de Y_{ij} estimado pelo modelo, e;

$\hat{e}_{ij} = Y_{ij} - \bar{Y}_{i.} = Y_{ij} - \hat{\mu} - \hat{\tau}_i = Y_{ij} - \hat{Y}_{ij}$, onde $Y_{ij} - \hat{Y}_{ij}$: diferença entre o valor observado e o estimado pelo modelo ou erro estimado.

Entretanto, a forma mais frequente para esta decomposição é a seguinte:

$$\text{SQTotal} = \sum_{i=1}^t \sum_{j=1}^r (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^t \sum_{j=1}^r (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^t \sum_{j=1}^r (Y_{ij} - \bar{Y}_{i.})^2 \quad (1.4)$$

Onde,

$\sum_i^t \sum_j^r (Y_{ij} - \bar{Y}_{..})^2$: Soma de Quadrados Total (SQTotal);

$\sum_i^t \sum_j^r (\bar{Y}_{i.} - \bar{Y}_{..})^2$: Soma de Quadrados de Tratamentos (SQT), e;

$\sum_i^t \sum_j^r (Y_{ij} - \bar{Y}_{i.})^2$: Soma de Quadrados do Erro (SQE).

Assim,

$$\text{SQTotal} = \text{SQT} + \text{SQE} \quad (1.5)$$

A Tabela 1, apresentada abaixo, chamada de tabela de análise de variância (ANOVA), contém as somas de quadrados e outras informações. Considerando t o número de tratamentos e r o de repetições, tem-se:

Tabela 1 – Análise de variância para delineamento completamente casualizado

Fonte de Variação	Graus de liberdade	SQ	QM	F
Tratamentos	$t - 1$	SQT	QMT	F_0
Erro	$(tr - 1) - (t - 1)$	SQE	QME	
Total	$tr - 1$	SQTotal		

Graus de liberdade (GL) é uma constante que caracteriza a distribuição das Somas de Quadrados (SQ). O Quadrado Médio (QM) é uma estatística intermediária utilizada em testes de hipóteses. Considerando que X representa tratamentos ou erro, o valor do $\text{QM}(X)$ é encontrado a partir da seguinte equação:

$$\text{QM}(X) = \frac{\text{SQ}(X)}{\text{GL}(X)} \quad (1.6)$$

A SQT é uma medida de variação entre os tratamentos e a SQE mede a variação entre as unidades experimentais homogêneas. O Quadrado Médio do Erro (QME) é uma estimativa não viesada de σ_e^2 , parâmetro que representa a variância do erro experimental e que, conseqüentemente, faz parte do modelo normal de erros independentes ($e_{ij} \sim N(0, \sigma_e^2)$).

Uma importante propriedade da análise de variância é que se os efeitos de tratamentos não são diferentes, o Quadrado Médio de Tratamentos (QMT) também é um estimador não viesado de σ_e^2 .

O teste F ($F_0 = \frac{\text{QMT}}{\text{QME}}$) testa a hipótese da ocorrência de igualdade entre os tratamentos, como apresentado abaixo:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_t \\ H_a: \tau_i \neq \tau_j; \quad i \neq j \end{cases} \quad (1.7)$$

Sabe-se que quanto maior o valor de F_0 , maior é a evidência contra a hipótese nula. Rejeita-se H_0 se o valor de F_0 cair na região de rejeição, isto é, se o valor de F_0 for maior do que o tabelado para a distribuição F a um nível de significância especificado pelo pesquisador ($F_{\alpha, (t-1), (tr-1)}$). Outra maneira de se testar as hipóteses é utilizando o p – valor. Se o p – valor for menor que o nível de significância estabelecido pelo pesquisador, rejeita-se a hipótese nula.

2 BLOCOS INCOMPLETOS

2.1 INTRODUÇÃO

Existem situações em que é necessário utilizar blocos, mas não é possível alocar uma repetição de cada tratamento em todos os blocos, como acontece nos blocos completos. Podemos verificar essas situações quando o número de tratamentos é muito grande ou as unidades experimentais em estudo são muito heterogêneas ou, ainda, quando certas limitações restringem excessivamente o tamanho dos blocos. Quando tais acontecimentos são verificados é usado o delineamento de blocos incompletos, a sua aplicação diminui a variância do erro experimental e proporciona comparações mais precisas entre os tratamentos. Os delineamentos em blocos incompletos foram introduzidos por Yates (1936).

O exemplo 2, retirado do capítulo 9 de Kuehl (2001) e traduzido para a língua portuguesa, mostra uma situação onde o delineamento em blocos incompletos é utilizado.

Exemplo 2:

Usualmente os tomates são produzidos durante os meses de inverno nas regiões áridas tropicais, a produção se encerra ao final do verão, quando as temperaturas do solo podem exceder 40°C, passando dos 35°C, temperatura máxima sugerida para plantação.

Objetivo de investigação: Um pesquisador de plantas deseja determinar os intervalos de temperatura que inibem a germinação das sementes de tomate para um grupo de cultivo.

Delineamento de tratamento: São escolhidas quatro temperaturas – 25°C, 30°C, 35°C e 40°C – para representar um intervalo comum para uma área de cultivo. A semente de tomate é semeada a uma temperatura constante dentro de uma câmara de temperatura controlada.

Delineamento de experimento: Uma câmara é uma unidade experimental, pois a repetição verdadeira de qualquer tratamento (temperatura) requer uma plantação independente do tratamento de uma câmara. Certo número de fatores pode contribuir para a variação da resposta entre as plantações e as condições de todo o experimento devem se repetir em outras plantações, por isso a importância da divisão das plantações em blocos.

O delineamento de blocos completos requereria quatro câmaras para cada plantação, mas o pesquisador só dispõe de três. Como o bloco (plantação) tem menos câmaras (unidades

experimentais) do que tratamentos (temperaturas), é construído o delineamento de blocos incompletos.

O Quadro 2 mostra o diagrama do delineamento, em que são aplicadas três temperaturas diferentes em cada uma das quatro plantações. As plantações representam os blocos incompletos de três tratamentos de temperaturas e os tratamentos se alocaram aleatoriamente as câmeras para cada plantação.

Quadro 2 – Delineamento de blocos incompletos com quatro tratamentos em quatro blocos de três parcelas

Plantação 1	25°C	30°C	40°C
Plantação 2	40°C	25°C	35°C
Plantação 3	35°C	30°C	25°C
Plantação 4	40°C	30°C	35°C

2.2 TIPOS DE BLOCOS INCOMPLETOS

Existem duas grandes classificações feitas aos delineamentos de blocos incompletos, que são: balanceamento e resolução. Ou seja, esses delineamentos podem ser balanceados ou parcialmente balanceados e resolvíveis ou não resolvíveis.

2.2.1 Blocos Incompletos Balanceados (BIB)

Os blocos incompletos são considerados balanceados quando cada tratamento é pareado um número igual de vezes com os demais tratamentos em um mesmo bloco, considerando todos os blocos do experimento. Por exemplo, ao ser representado um delineamento com quatro tratamentos em quatro blocos de três parcelas, com a configuração apresentada no Quadro 3, percebe-se que cada tratamento aparece pareado com outro em dois blocos. A letra lambda (λ) tem sido usada pela maioria dos autores para representar esta característica. Nesse exemplo apresentado, tem-se $\lambda = 2$.

Quadro 3 – Delineamento de blocos incompletos balanceados com quatro tratamentos em quatro blocos de três parcelas

Bloco 1	1	2	3
Bloco 2	1	2	4
Bloco 3	1	3	4
Bloco 4	2	3	4

Considerando que um delineamento de blocos incompletos possui r repetições de t tratamentos em b blocos de k parcelas com $k < t$, então o número de blocos onde ocorre cada par de tratamentos é $\lambda = r(k - 1) / (t - 1)$, com $\lambda < k < b$; resultando na conhecida expressão $\lambda(t - 1) = r(k - 1)$ para os blocos incompletos balanceados (BIB), que é uma condição necessária para a existência de um delineamento de BIB, mas não suficiente (JOHN, 1971).

Como apresentado por Cochran e Cox (1957), será considerado que os experimentos em blocos incompletos balanceados podem ser enquadrados em três tipos distintos.

Tipo I – Experimentos em que os blocos podem ser agrupados em repetições

Considerando, por exemplo, um experimento com $t = 6$ tratamentos em $b = 10$ blocos de $k = 3$ parcelas, com a seguinte disposição apresentada no Quadro 4.

Quadro 4 – Delineamento de blocos incompletos balanceados que podem ser agrupados em repetições com seis tratamentos em dez blocos de três parcelas

Bloco 1	1	2	5	Bloco 6	2	3	4
Bloco 2	1	2	6	Bloco 7	2	3	5
Bloco 3	1	3	4	Bloco 8	2	4	6
Bloco 4	1	3	6	Bloco 9	3	5	6
Bloco 5	1	4	5	Bloco 10	4	5	6

Nesse caso, um tratamento qualquer aparece duas vezes no mesmo bloco com cada um dos outros. Por exemplo, o par (1,2) aparece no bloco 1 (1,2,5) e no bloco 2 (1,2,6). Sendo assim, tem-se $\lambda = 2$. Além disso, cada tratamento aparece em cinco blocos, logo o número de repetições é $r = 5$. Embora tenha cinco repetições, não é possível agrupar subconjunto de blocos que formem uma ou mais repetições, como será apresentado logo em seguida nos BIB dos Tipos II e III. A análise de variância para esse caso é a apresentada a seguir na Tabela 2.

Tabela 2 – Análise de variância para delineamento de blocos incompletos balanceados que podem ser agrupados em repetições com seis tratamentos em dez blocos de três parcelas

Fonte de Variação	Graus de liberdade
Blocos	$b - 1 = 9$
Tratamentos	$t - 1 = 5$
Erro	$(tr - 1) - (b - 1) - (t - 1) = 15$
Total	$tr - 1 = 29$

Tipo II – Experimentos em que os blocos podem ser dispostos em grupos com uma repetição

Há casos em que os blocos podem ser reunidos de modo a formar grupos com uma repetição, é o que acontece, por exemplo, no delineamento apresentado no Quadro 5, em que se tem $t = 6$ tratamentos em $b = 15$ blocos de $k = 2$ parcelas, com $r = 5$ repetições e $\lambda = 1$. Dessa maneira, o esquema de análise de variância pode ser ligeiramente modificado, como mostra a Tabela 3.

Quadro 5 – Delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com uma repetição com seis tratamentos em quinze blocos de duas parcelas

Repetição 1		Repetição 2			Repetição 3			Repetição 4			Repetição 5		
Bloco 1	1 2	Bloco 4	1 3	Bloco 7	1 4	Bloco 10	1 5	Bloco 13	1 6				
Bloco 2	3 4	Bloco 5	2 5	Bloco 8	2 6	Bloco 11	2 4	Bloco 14	2 3				
Bloco 3	5 6	Bloco 6	4 6	Bloco 9	3 5	Bloco 12	3 6	Bloco 15	4 5				

Tabela 3 – Análise de variância para delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com uma repetição com seis tratamentos em quinze blocos de duas parcelas

Fonte de Variação	Graus de liberdade
Repetições	$r - 1 = 4$
Blocos dentro das repetições	$b - r = 10$
Tratamentos	$t - 1 = 5$
Erro	$(tr - 1) - (r - 1) - (b - r) - (t - 1) = 10$
Total	$tr - 1 = 29$

Tipo III – Experimentos em que os blocos podem ser dispostos em grupos com mais de uma repetição

Há casos, ainda, em que os blocos não podem ser agrupados em repetições e nem em grupos com uma repetição. A seguir, no Quadro 6, é apresentado um exemplo com $g = 2$ grupos, $t = 5$ tratamentos em $b = 10$ blocos de $k = 2$ parcelas, com $r = 4$ repetições (duas

em cada grupo) e $\lambda = 1$. O esquema de análise de variância segue o modelo apresentado na Tabela 4.

Quadro 6 – Delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com mais de uma repetição com cinco tratamentos em dez blocos de duas parcelas

Grupo 1 (Repetições I e II)			Grupo 2 (Repetições II e III)		
Bloco 1	1	2	Bloco 6	1	4
Bloco 2	3	4	Bloco 7	2	3
Bloco 3	2	5	Bloco 8	3	5
Bloco 4	1	3	Bloco 9	1	5
Bloco 5	4	5	Bloco 10	2	4

Tabela 4 – Análise de variância para delineamento de blocos incompletos balanceados que podem ser dispostos em grupos com mais de uma repetição com cinco tratamentos em dez blocos de duas parcelas

Fonte de Variação	Graus de liberdade
Grupos	$g - 1 = 1$
Blocos dentro dos grupos	$b - g = 8$
Tratamentos	$t - 1 = 4$
Erro	$(tr - 1) - (g - 1) - (b - g) - (t - 1) = 6$
Total	$tr - 1 = 19$

2.2.2 Blocos Incompletos Parcialmente Balanceados (BIPB)

Não é possível construir delineamentos balanceados para todas as situações experimentais que requerem blocos incompletos, em alguns casos o número de repetições necessárias é muito grande e não se tem unidades experimentais suficientes para o experimento. Portanto, com frequência são construídos delineamentos parcialmente balanceados que requerem menos repetições.

Bose e Nair (1939) propuseram o desenho de blocos incompletos parcialmente balanceados (BIPB), esses delineamentos, ao contrário dos balanceados, têm alguns pares de tratamentos que aparecem em mais blocos que outros pares, assim algumas comparações de tratamentos são mais precisas que outras. É mais simples usar um delineamento de blocos balanceados, pois proporcionam a mesma precisão para todas as comparações entre tratamentos, mas se os recursos são limitados e não se pode obter repetições suficientes, o delineamento parcialmente balanceado é uma alternativa atrativa quando o delineamento balanceado requer um número excessivo de unidades experimentais.

No Quadro 7, apresentado abaixo, é possível observar um exemplo de delineamento de blocos incompletos parcialmente balanceados para seis tratamentos em três blocos de quatro parcelas.

Quadro 7 – Delineamento de blocos incompletos parcialmente balanceados com seis tratamentos em três blocos de quatro parcelas

Bloco 1	1	4	2	5
Bloco 2	2	5	3	6
Bloco 3	3	6	1	4

Pode-se perceber que os pares de tratamentos (1,4), (2,5) e (3,6) estão em dois blocos ($\lambda = 2$) e os pares restantes aparecem em apenas um ($\lambda = 1$). Sendo assim, os pares de tratamentos que se apresentam juntos em dois blocos são comparados com uma precisão um pouco maior daqueles que estão em apenas um. A diferença nas precisões das comparações de tratamentos é o ônus existente ao fazer um experimento menor, porém a diferença existente na precisão, normalmente, não é tão grande ao ponto de evitar o uso do delineamento parcialmente balanceado.

Para fim comparativo, se o delineamento de blocos incompletos balanceados fosse usado para um experimento de $t = 6$ tratamentos com blocos de $k = 4$ parcelas, seriam necessários $b = 15$ blocos e $r = 10$ repetições, conseqüentemente, $\lambda = 6$. Ou seja, seriam utilizadas $tr = bk = 60$ unidades experimentais, enquanto o delineamento parcialmente balanceado reduziu essa quantidade para doze, como mostrado acima. Teoricamente, para t tratamentos e k parcelas por bloco, sempre existe um bloco incompleto balanceado com o número de blocos $b = C_k^t$.

2.2.3 Blocos Incompletos Resolvíveis

O delineamento de blocos incompletos resolvíveis são aqueles que possuem blocos agrupados de maneira que cada grupo constitui uma ou mais repetições completa de todos os tratamentos. Esse conceito foi introduzido por Yates e sua principal característica é que ele também pode ser analisado como blocos ao acaso considerando a repetição como um bloco completo. Tanto o delineamento de BIB quanto o de BIPB podem ser planejados como resolvíveis.

Os delineamentos de BIB do Tipo II (experimentos em que os blocos podem ser dispostos em grupos com uma repetição) e do Tipo III (experimentos em que os blocos podem ser dispostos em grupos com mais de uma repetição) apresentados no tópico 2.2.1 são exemplos de BIB resolvíveis. Isso porque nos delineamento em questão, os blocos são reunidos formando grupos de uma ou mais repetições.

Como citado acima, os BIPB também podem ser classificados como resolvíveis. Por exemplo, ao considerar um delineamento com seis tratamentos em quatro blocos de três parcelas, com a configuração apresentada no Quadro 8, tem-se um delineamento de BIPB resolvíveis. Percebe-se que os pares de tratamentos (1,4), (2,5), (2,6), (3,5) e (3,6) não aparecem em nenhum bloco ($\lambda = 0$), já os pares de tratamentos (1,2), (1,3), (1,5), (1,6), (2,4), (3,4), (4,5) e (4,6) estão em um bloco ($\lambda = 1$) e, por fim, os pares de tratamentos (2,3) e (5,6) estão presentes em dois blocos ($\lambda = 2$), ou seja, tem-se um delineamento de BIPB. Além disso, o delineamento também é resolvível, porque é possível dividir os quatro blocos em dois grupos com uma repetição.

Quadro 8 – Delineamento de blocos incompletos parcialmente balanceados e resolvíveis com seis tratamentos em quatro blocos de três parcelas

Repetição 1			Repetição 2				
Bloco 1	1	2	3	Bloco 3	2	3	4
Bloco 2	4	5	6	Bloco 4	1	5	6

Historicamente, certa classe de blocos incompletos resolvíveis foi conhecida como latice e sua configuração está descrita em vários livros como o de Cochran e Cox (1957) e Federer (1963). São conhecidos os latices quadrados, os latices retangulares e os latices cúbicos. Nos latices quadrados $t = k^2$, onde t é o número de tratamentos e k o número de parcelas do bloco, com k blocos por repetição. Nos retangulares, $t = k(k + 1)$, sendo $(k + 1)$ o número de blocos por repetição. Nos cúbicos, $t = k^3$. Embora ainda sejam usados, as restrições advindas das suas formulações limitam sua utilização e atualmente existe uma classe bem maior, denominada de alfa latice, desenvolvida a partir do trabalho de Patterson e Williams (1976), cujas restrições são bem menores.

2.2.4 Blocos Incompletos Não Resolvíveis

O delineamento de blocos incompletos não resolvíveis, ao contrário dos resolvíveis, não permite o agrupamento dos blocos de maneira que cada grupo possua uma ou mais repetições completas de todos os tratamentos. Tanto o delineamento de BIB quanto o de BIPB podem ser classificados como não resolvíveis.

O delineamento de BIB do Tipo I (Experimentos em que os blocos podem ser agrupados em repetições) apresentado no tópico 2.2.1 é um exemplo de BIB não resolvíveis. Isso porque no delineamento em questão, os dez blocos não podem ser reunidos em grupos de uma ou mais repetições.

Os BIPB, como citado acima, também podem ser classificados como não resolvíveis. Por exemplo, ao considerar um delineamento com seis tratamentos em quatro blocos de três parcelas, com a configuração apresentada no Quadro 9, tem-se um delineamento de BIPB não resolvíveis. Percebe-se que os pares de tratamentos (1,4), (2,5) e (3,6) não estão presentes em nenhum bloco ($\lambda = 0$), já os pares de tratamentos (1,2), (1,3), (1,5), (1,6), (2,3), (2,4), (2,6), (3,4), (3,5), (4,5), (4,6) e (5,6) aparecem em um bloco ($\lambda = 1$), ou seja, tem-se um delineamento de BIPB. Além disso, o delineamento também é não resolvível, porque não é possível dividir os quatro blocos em grupos de repetição.

Quadro 9 – Delineamento de blocos incompletos parcialmente balanceados e não resolvíveis com seis tratamentos em quatro blocos de três parcelas

Bloco 1	1	2	3
Bloco 2	2	4	6
Bloco 3	1	5	6
Bloco 4	3	4	5

3 ANÁLISE DE BLOCOS INCOMPLETOS

A análise dos experimentos em blocos incompletos é bem mais complicada do que a dos experimentos em blocos completos. Entretanto, essa maior dificuldade de análise, como também a perda de graus de liberdade no resíduo, são frequentemente compensadas por uma redução na Soma de Quadrados do Erro (SQE) e, conseqüentemente, no Quadrado Médio do Erro (QME), ou seja, a obtenção de experimentos mais precisos.

Existem três tipos de análise de experimentos em blocos incompletos, são elas: análise intrablocos, análise interbloco e análise com recuperação da informação interbloco (análise combinada).

Os detalhes destes três tipos de análise de variância serão explicados mais adiante, para ilustrá-los serão usados os dados de um experimento relatado em Federer (1963) que tem o seguinte conteúdo: comparar métodos de preparo de ovos secos com dez tratamentos em quinze blocos de quatro parcelas, resultando em seis repetições, a variável resposta é a média de seis juizes que provaram uma comida feita com o produto de cada parcela. Os dados de Federer (1963) seguem no Anexo A.

3.1 ANÁLISE INTRABLOCO

Na análise intrablocos, idealizada por Yates (1936), são feitas comparações entre parcelas do mesmo bloco e apenas essas são usadas nas estimativas de efeito de tratamentos, sendo assim o efeito de cada bloco é considerado fixo. Essa análise pode ser usada para qualquer experimento em blocos incompletos e se baseia em métodos estatísticos exatos.

O modelo estatístico usado na análise intrablocos para representar Y_{ij} , valor da variável resposta Y que recebeu o i -ésimo tratamento e pertence ao j -ésimo bloco, é o seguinte:

$$Y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \quad (3.1)$$

Onde,

μ : média geral;

τ_i : efeito do tratamento i ;

β_j : efeito do bloco j , e;

e_{ij} : erro experimental, $e_{ij} \sim N(0, \sigma_e^2)$.

Para analisar os dados citados em Federer (1963) é encontrada a ANOVA intrablocos através da função PROC GLM do software SAS (2003), como apresentado na Tabela 5. A programação segue no Apêndice A.

Tabela 5 – Análise de variância intrablocos

Fonte de Variação	Graus de liberdade	SQ	QM	F	Pr > F
Blocos (Ajustados)	14	23,87	1,71	3,43	0,0014
Tratamentos (Ajustados)	9	321,51	35,72	71,92	0,0001
Erro	36	17,88	0,5		
Total	59	356,46			

Analisando a tabela de ANOVA, através do teste F , têm-se evidências que existe diferença entre os tratamentos, considerando o nível de significância igual a 0,05 (p – valor = 0,0001). Além disso, é observada que a estimativa da variância do erro experimental é igual a 0,5, tendo em vista que $E(QME) = \sigma_e^2$.

3.2 ANÁLISE INTERBLOCO

A análise interbloco é baseada em um modelo modificado, idealizado por Yates (1940) e citado por John e Williams (1995), que considera o efeito de bloco como aleatório. Existem dois raciocínios que defendem a natureza aleatória dos efeitos de blocos. O primeiro é que se os blocos do experimento foram selecionados aleatoriamente de um conjunto dos possíveis blocos, está-se diante de uma amostra aleatória de blocos e, portanto, o efeito de blocos é aleatório; em experimentos de campo é difícil conceber que isso ocorra, pois o pesquisador irá escolher os melhores blocos. A segunda é que após os b blocos serem escolhidos pelo pesquisador, há um sorteio entre eles para se definir qual será o bloco 1, o 2, etc., isto é, um processo aleatório que indica qual é o bloco j ; esse foi o argumento usado por Nelder (1965) e nos desenvolvimentos mais recentes, por Bailey (1991).

Sob o ponto de vista inferencial em experimentos, o efeito de blocos é secundário, isto é, não há interesse em comparar blocos; a blocagem é uma estratificação da população de parcelas que o pesquisador dispõe; os tratamentos são estímulos aplicados às parcelas cujo grau de influência vai se refletir na variável resposta; os blocos, embora possam ser construídos, muitas vezes ocorrem naturalmente e sua função é diminuir o erro experimental e o viés das comparações entre tratamentos. Acrescenta-se ainda que, sob a teoria da casualização, não existe um teste de hipótese para comparar efeitos de blocos (KEMPTHORNE, 1952). Postos esses argumentos, é factível defender o modelo para blocos incompletos com efeitos aleatórios para blocos. Assim, para representar Y_{ij} , valor da variável resposta Y que recebeu o i -ésimo tratamento e pertence ao j -ésimo bloco, considere o seguinte modelo:

$$Y_{ij} = \mu + \tau_i + b_j + e_{ij} \quad (3.2)$$

Onde,

μ : média geral;

τ_i : efeito do tratamento i ;

b_j : efeito do bloco j , $b_j \sim N(0, \sigma_b^2)$, e;

e_{ij} : erro experimental, $e_{ij} \sim N(0, \sigma_e^2)$.

Partindo do modelo (3.2), Yates criou uma nova variável resposta definida como a soma das observações de um bloco, reduzindo o número de observações para um número igual ao de blocos. Para explicar melhor como é feita essa soma, considere um experimento com seis tratamentos em quatro blocos de três parcelas, como o esquema apresentado abaixo no Quadro 10.

Quadro 10 – Delineamento de blocos incompletos com seis tratamentos em quatro blocos de três parcelas

Bloco 1	1	2	3
Bloco 2	4	5	6
Bloco 3	2	3	4
Bloco 4	1	5	6

Em cada bloco é realizado o seguinte procedimento:

$$B_1 = Y_{11} + Y_{21} + Y_{31} = 3\mu + \tau_1 + \tau_2 + \tau_3 + 3b_1 + e_{11} + e_{21} + e_{31} \quad (3.3)$$

Em relação a B_1 , têm-se:

$$E(B_1) = 3\mu + \tau_1 + \tau_2 + \tau_3 \quad (3.4)$$

$$\text{Var}(B_1) = 9\sigma_b^2 + 3\sigma_e^2 = \sigma_B^2 \quad (3.5)$$

Sendo assim,

$$B_1 \sim N(3\mu + \tau_1 + \tau_2 + \tau_3, \sigma_B^2) \quad (3.6)$$

Analogamente:

$$B_2 \sim N(3\mu + \tau_4 + \tau_5 + \tau_6, \sigma_B^2) \quad (3.7)$$

$$B_3 \sim N(3\mu + \tau_2 + \tau_3 + \tau_4, \sigma_B^2) \quad (3.8)$$

$$B_4 \sim N(3\mu + \tau_1 + \tau_5 + \tau_6, \sigma_B^2) \quad (3.9)$$

Dessa maneira, forma-se a seguinte matriz de blocos:

$$B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} \mu + \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \\ \tau_6 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} \quad (3.10)$$

Onde e_1 , e_2 , e_3 e e_4 são os erros associados a cada bloco, como por exemplo, o e_1 mostrado abaixo:

$$e_1 = 3b_1 + e_{11} + e_{21} + e_{31} \quad (3.11)$$

De maneira geral, a nova variável criada para a análise (B_j) tem $\text{Var}(B_j) = k^2\sigma_b^2 + k\sigma_e^2 = \sigma_b^2$ e são independentes. Esse modelo é mais complicado que o anterior (intrabloco) porque cada observação agora tem a presença de k tratamentos e não de somente um. Este tipo de análise só pode ser usada quando $b > t$. No caso prático em questão, tem-se quinze blocos (número de observações) e os resultados da análise com o GLM do SAS (2003) estão apresentados abaixo na Tabela 6. A programação segue no Apêndice A.

Tabela 6 – Análise de variância interbloco

Fonte de Variação	Graus de liberdade	SQ	QM	F	Pr > F
Tratamentos	9	62,764	6,9737	6,32	0,0281
Erro	5	5,5133	1,1026		
Total	14	68,2773			

Assim como na análise intrabloco, a interbloco também rejeita a hipótese nula de igualdade entre os tratamentos, considerando o nível de significância igual a 0,05 (p – valor = 0,0281). Porém, nesse caso, a estimativa da variância do erro experimental é de 1,1026, bem maior do que a gerada na análise intrabloco, que foi de 0,5.

3.3 ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO

A terceira análise é conhecida como análise com recuperação da informação interbloco, idealizada por Yates (1940) e citado por John e Williams (1995). Nessa, além das comparações realizadas dentro de cada bloco, também são consideradas as comparações entre os blocos para a estimação dos efeitos de tratamentos, isto é, com duas estimativas diferentes dos efeitos de tratamentos, obter uma terceira estimativa. Sendo assim, essa análise aproveita melhor os dados, mas usa métodos estatísticos apenas aproximados e só deve ser usada para experimentos com número de graus de liberdade relativamente grande para blocos e para os resíduos.

O resultado desta análise, apresentado na Tabela 7, de acordo com Yates, pode ser obtido pela análise dos dados originais, via PROC MIXED do SAS (2003) com a opção TYPE1 (quadrados mínimos). A programação segue no Apêndice A.

Tabela 7 – Análise de resultados da análise com recuperação da informação interbloco (quadrados mínimos)

Efeito	Graus de liberdade		<i>F</i>	Pr > <i>F</i>	Variância do erro
	Tratamentos	Erro			
Tratamentos	9	36	71,33	< 0,0001	0,4967

Outro resultado usando a teoria de Yates, mas trabalhado por Patterson e Thompson (1971), é obtido via PROC MIXED do SAS (2003) com a opção REML (máxima verossimilhança restrita), sendo ele apresentado abaixo na Tabela 8. A programação segue no Apêndice A.

Tabela 8 – Análise de resultados da análise com recuperação da informação interbloco (máxima verossimilhança restrita)

Efeito	Graus de liberdade		<i>F</i>	Pr > <i>F</i>	Variância do erro
	Tratamentos	Erro			
Tratamentos	9	36	70,89	< 0,0001	0,4998

Assim como nas duas análises realizadas anteriormente (intrabloco e interbloco), as duas maneiras de obter os resultados da análise com recuperação da informação interbloco também rejeitam a hipótese nula de igualdade entre os tratamentos, considerando o nível de significância igual a 0,05 ($p - \text{valor} = < 0,001$). É importante ressaltar que nessa análise, as duas variâncias do erro experimental são menores do que as geradas nas outras duas análises, sendo elas de 0,4967 e 0,4998, considerando os resultados utilizando quadrados mínimos e máxima verossimilhança restrita, respectivamente.

4 TEORIA DA ANÁLISE DE BLOCOS INCOMPLETOS

4.1 INTRODUÇÃO

O estudo da análise de blocos incompletos é feito usando estrutura matricial, embora possa ser desenvolvido, como foi feito anteriormente, de forma aberta, sem uso de matrizes (COCRHAN; COX, 1957). Seguindo a notação de John e Williams (1995), o modelo em forma matricial é:

$$Y = 1\mu + X\tau + Z\beta + e \quad (4.1)$$

E têm-se que,

$$E(Y) = 1\mu + X\tau + Z\beta \quad (4.2)$$

$$\text{Var}(Y) = I\sigma_e^2 \quad (4.3)$$

Num experimento com $tr = bk = n$ parcelas, t tratamentos e b blocos, o vetor $1_{(nx1)}$ tem todos os elementos iguais a 1; a matriz $X_{(nxt)}$ é chamada de matriz de incidência dos tratamentos; a matriz $Z_{(nxb)}$ é a matriz de incidência dos blocos; $e_{(nx1)}$ é o vetor dos erros; $\tau_{(tx1)}$ é o vetor de efeitos de tratamentos; $\beta_{(bx1)}$ é o vetor de efeitos de blocos.

4.2 ANÁLISE INTRABLOCO

A análise de variância de um experimento em blocos incompletos com o modelo especificado anteriormente, em (4.1), é conhecida como análise intrabloco; a denominação vem do fato de que as diferenças entre tratamentos são estimadas dentro de cada bloco. Nesse modelo aplica-se o método dos quadrados mínimos para estimação dos parâmetros. Se $W = [1 \ X \ Z]$ e $\theta' = [\mu \ \tau' \ \beta']$, então as equações normais podem ser escritas da forma $W'W\hat{\theta} = W'Y$ e com um desenvolvimento algébrico chega-se a:

$$\begin{bmatrix} n & r' & k' \\ r & R & N \\ k & N' & K \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} G \\ T \\ B \end{bmatrix} \quad (4.4)$$

Onde,

r : vetor de repetições de tratamentos e o elemento $(i) = r_i$;

k : vetor de tamanho de blocos e o elemento $(j) = k_j$;

R : matriz diagonal e o elemento $(ii) = r_i$;

K : matriz diagonal e o elemento $(jj) = k_j$;

N : matriz de incidência dos tratamentos nos blocos;

G : grande total;

T : vetor de totais de tratamentos, e;

B : vetor de totais de blocos.

O sistema de equações gerado em (4.4) é indeterminado e tem posto $(n - 2)$ desde que $1'[r \ R \ N] = [n \ r' \ k'] = 1'[k \ N' \ K]$. Uma solução para $\hat{\theta}$ em (4.4) é dada por $\hat{\theta} = (W'W)^{-1}W'Y$, sendo $(W'W)^{-1}$ uma inversa generalizada de $(W'W)$; entretanto, o interesse maior está nos tratamentos e é necessário apresentar uma solução para $\hat{\tau}$ em forma explícita. John (1971) desenvolveu a solução e chegou à seguinte forma reduzida:

$$\begin{bmatrix} 1 & \frac{1}{n}r' & \frac{1}{n}k' \\ 0 & R - NK^{-1}N' & 0 \\ 0 & 0 & K - N'R^{-1}N \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} G \\ T - NK^{-1}B \\ B - N'R^{-1}T \end{bmatrix} \quad (4.5)$$

Onde 0 é um vetor ou matriz de zeros.

Uma estimativa de μ é $\bar{Y} = \frac{G}{n}$, adotando as restrições $\sum_i r_i \tau_i = 0$ e $\sum_j k_j \beta_j = 0$; na segunda linha de (4.5) estão as equações para tratamentos e na terceira para blocos. Entretanto, o interesse maior é sobre os tratamentos, como já explicado anteriormente. Definindo $C = R - NK^{-1}N'$ e $Q = T - NK^{-1}B$, tem-se a solução $C\hat{\tau} = Q$. C é denominada matriz de informação sobre os tratamentos, tem posto incompleto $(t - 1)$ e suas linhas são contrastes, desde que $C1 = 0$, ou seja, todos os contrastes entre tratamentos são estimáveis. Assim, uma solução para $\hat{\tau}$ é dada por $\hat{\tau} = C^{-1}Q$, com $\text{Var}(\hat{\tau}) = C^{-1}CC^{-1}\sigma_e^2$.

Na forma clássica, a análise de variância é baseada na seguinte identidade dos modelos lineares $Y'Y = \hat{Y}'\hat{Y} + \hat{e}'\hat{e}$, e seguindo John (1971), a análise de variância fica como apresentada na Tabela 9.

Tabela 9 – Teoria da análise de variância intrabloco

Fonte de Variação	Graus de liberdade	SQ
Blocos (não ajustados)	$b - 1$	$B'K^{-1}B - G\hat{\mu}$
Tratamentos (ajustados)	$t - 1$	$Q'\hat{\tau}$
Erro	$(n - 1) - (b - 1) - (t - 1)$	Por diferença
Total	$n - 1$	$\sum_i \sum_j Y_{ij}^2 - \frac{G^2}{n}$

O termo tratamentos (ajustados) significa que o valor esperado da forma quadrática $Q'\hat{\tau}$ é $(t - 1)\sigma_e^2 + \tau' C \tau$, livre de efeitos de blocos. Assumindo normalidade pode-se testar

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_t \\ H_a: \tau_i \neq \tau_j; \quad i \neq j \end{cases} \quad (4.6)$$

mediante o teste F . A variância de qualquer contraste $a'\hat{\tau}$ entre tratamentos é dada por $a' C^{-1} a \sigma_e^2$ e os testes de hipóteses seguem da forma conhecida.

4.3 ANÁLISE INTERBLOCO

Como é conhecida, a análise intrabloco para blocos incompletos ocasiona confundimento parcial de alguns contrastes entre tratamentos, o que resulta em perda de informação; entretanto, se os efeitos de blocos fossem variáveis aleatórias (efeito aleatório) com média zero, essa informação seria recuperada. Essa ideia começou com Yates (1940).

Assim, para representar Y_{ij} , valor da variável resposta Y que recebeu o i -ésimo tratamento e pertence ao j -ésimo bloco, considere o seguinte modelo:

$$Y_{ij} = \mu + \tau_i + b_j + e_{ij} \quad (4.7)$$

Onde,

μ : média geral;

τ_i : efeito do tratamento i ;

b_j : efeito do bloco j , $b_j \sim N(0, \sigma_b^2)$, e;

e_{ij} : erro experimental, $e_{ij} \sim N(0, \sigma_e^2)$.

Como complemento, têm-se:

$$E(Y_{ij}) = \mu + \tau_i \quad (4.8)$$

$$\text{Var}(Y_{ij}) = \sigma_b^2 + \sigma_e^2 \quad (4.9)$$

$$\text{Cov}(Y_{ij}, Y_{hk}) = \sigma_b^2 + \sigma_e^2; \quad i \neq h, \quad j = k \quad (4.10)$$

Seguindo o desenvolvimento basicamente proposto por Yates (1940) (JOHN, 1971), com a condição $b > t$, considere a seguinte transformação no modelo:

$$B = Z'Y = Z'1\mu + Z'X\tau + Z'Z\underline{b} + Z'e \quad (4.11)$$

E assim,

$$B = k1\mu + N'\tau + K\underline{b} + Z'e \quad (4.12)$$

Onde $\underline{b}_{(bx1)}$ é o vetor de efeitos aleatórios para blocos com $E(\underline{b}) = 0$ e $\text{Var}(\underline{b}) = I\sigma_b^2$ e $B_{(bx1)}$ é o vetor coluna dos totais de blocos. Mediante suposições anteriores, têm-se que,

$$E(B) = k1\mu + N'\tau = W_1\varphi \quad (4.13)$$

$$\text{Var}(B) = k^2I\sigma_b^2 + kI\sigma_e^2 = (k^2\sigma_b^2 + k\sigma_e^2)I = I\sigma_B^2 \quad (4.14)$$

Onde,

$$W_1 = [k1 \quad N'] \quad (4.15)$$

$$\varphi = \begin{bmatrix} \mu \\ \tau \end{bmatrix} \quad (4.16)$$

A transformação em (4.12) visa concentrar a informação sobre os tratamentos em um bloco nos totais de blocos. O modelo assim representado é tipicamente de Gauss–Markov, aplicando o método dos mínimos quadrados em (4.12), têm-se as equações normais $W_1'W_1\tilde{\varphi} = W_1'B$, com um trabalho algébrico chega-se ao sistema:

$$\begin{bmatrix} bk & 1'R \\ kR1 & NN' \end{bmatrix} \begin{bmatrix} \tilde{\mu} \\ \tilde{\tau} \end{bmatrix} = \begin{bmatrix} G \\ NB \end{bmatrix} \quad (4.17)$$

Esse pode ser simplificado para:

$$\begin{bmatrix} bk & 1'R \\ 0 & NN' - \frac{1}{b}R11'R \end{bmatrix} \begin{bmatrix} \tilde{\mu} \\ \tilde{\tau} \end{bmatrix} = \begin{bmatrix} G \\ NB - \frac{G}{b}R1 \end{bmatrix} \quad (4.18)$$

Como no modelo intrabloco, se $\tilde{C} = NN' - \frac{1}{b}R11'R$ e $\tilde{Q} = NB - \frac{G}{b}R1$, então tem-se o sistema de equações $\tilde{C}\tilde{\tau} = \tilde{Q}$; como o posto de \tilde{C} é $(t - 1)$ e $\tilde{C}1 = 0$, todos os contrastes entre tratamentos são estimáveis. Assim, um contraste $a'\tau$ é estimado por $a'\tilde{\tau}$ e sua variância é dada por $a'\tilde{C}^-a\sigma_B^2$. Se \tilde{C}^- é uma inversa generalizada de \tilde{C} , então $\tilde{\tau} = \tilde{C}^-\tilde{Q}$ é uma solução para $\tilde{\tau}$.

A análise da variância do modelo interbloco de Yates, segue os procedimentos conhecidos e as informações sobre ela encontram-se a seguir, na Tabela 10. O valor de $\hat{\mu}_b$ é dado pela média dos blocos.

Tabela 10 – Teoria da análise de variância interbloco

Fonte de Variação	Graus de liberdade	SQ	$E(QM)$
Tratamentos	$t - 1$	$\tilde{Q}'\tilde{\tau}$	$\sigma_B^2 + \frac{\tau'\tilde{C}\tau}{t - 1}$
Erro	$(b - 1) - (t - 1)$	Por diferença	σ_B^2
Total	$b - 1$	$B'B - B\hat{\mu}_b^2$	

As inferências obtidas dessa análise de variância são de precisão menor e por isso não são usadas isoladamente; elas estão aqui para completar um quadro histórico. Se os tamanhos dos blocos são diferentes o método proposto por Yates não pode ser aplicado.

4.4 ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO

De acordo com tudo visto anteriormente e examinando o caso de um contraste entre tratamentos $a'\tau$, têm-se, portanto, dois estimadores não viesados, um intrabloco ($a'\hat{\tau}$) e outro interbloco ($a'\tilde{\tau}$) que são não correlacionados, como está demonstrado em John (1971). Com dois estimadores não viesados e não correlacionados, o próximo passo é combiná-los para formar um terceiro. Seja,

$$a'\tilde{\tau} = g_1 a'\hat{\tau} + g_2 a'\tilde{\tau}; \quad (g_1 + g_2) = 1 \quad (4.19)$$

Com

$$\text{Var}(a'\tilde{\tau}) = g_1^2 \text{Var}(a'\hat{\tau}) + g_2^2 \text{Var}(a'\tilde{\tau}) = g_1^2 w_1 + (1 - g_1)^2 w_2 \quad (4.20)$$

Onde,

$$w_1 = \text{Var}(a'\hat{\tau}) \quad (4.21)$$

$$w_2 = \text{Var}(a'\tilde{\tau}) \quad (4.22)$$

Aplicando derivada obtêm-se g_1 e g_2 tal que $\text{Var}(a'\tilde{\tau})$ seja mínima e eles são:

$$g_1 = \frac{w_2}{w_1 + w_2} \quad (4.23)$$

$$g_2 = \frac{w_1}{w_1 + w_2} \quad (4.24)$$

Assim,

$$a'\tilde{\tau} = \frac{w_2 a'\hat{\tau} + w_1 a'\tilde{\tau}}{w_1 + w_2} \quad (4.25)$$

$$\text{Var}(a'\tilde{\tau}) = \frac{w_1 w_2}{w_1 + w_2} \quad (4.26)$$

Certamente, esta variância é menor ou igual às fornecidas pela análise intrabloco e interbloco. O problema é que os pesos dependem das estimativas de σ_b^2 e de σ_e^2 e existem várias formas de obtê-los, mas este tópico não será tratado aqui. Os procedimentos para se testar as hipóteses relevantes usando quadrados mínimos são tratados nos livros de Cochran e Cox (1957) e John (1971).

Embora o método dos quadrados mínimos seja usado; depois do trabalho de Patterson e Thompson (1971), a análise de blocos incompletos com recuperação da informação interbloco é atualmente feita diretamente pelo método de máxima verossimilhança restrita; como está desenvolvido em Searle, Casella e McCulloch (1992). Para isso considere um modelo misto da forma:

$$Y = 1\mu + X\tau + Z\underline{b} + e \quad (4.27)$$

Onde $\underline{b}_{(bx1)}$ é o vetor de efeitos aleatórios para blocos com $E(\underline{b}) = 0$ e $\text{Var}(\underline{b}) = I\sigma_b^2$ e

$$\text{Var}(Y) = ZZ'\sigma_b^2 + I\sigma_e^2 \quad (4.28)$$

Admitindo normalidade, segue a análise. O PROC MIXED com a opção REML (máxima verossimilhança restrita) do SAS (2003) pode ser utilizado para obtenção dos resultados, mas eles não são iguais aos produzidos pelo método de Yates (quadrados mínimos), embora a diferença seja pequena; o PROC MIXED com a opção TYPE1 reproduz os resultados de Yates, pelo menos para os BIB.

5 ANÁLISE ESPACIAL

5.1 INTRODUÇÃO

O uso de delimitamento de blocos incompletos visa diminuir o erro experimental e tem sido intensivamente utilizado na prática. Em um estudo conduzido por Patterson e Hunter (1983) com 244 experimentos em blocos incompletos resolvíveis, a eficiência média foi de 1,43 comparada com blocos ao acaso. Uma crítica da análise de experimentos, mesmo em blocos com poucas parcelas, como pode ser o caso de blocos incompletos resolvíveis, é que ela não leva em conta a influência das parcelas vizinhas. Nesse contexto surgiram alguns modelos como o de Papadakis (1937), cuja eficiência foi estudada por Bartlett (1978) e o de Wilkinson, Eckert, Hancock e Mayo (1983). Especificamente para blocos incompletos, surgiu a proposta de Williams (1986). Uma síntese dos métodos de Papadakis e de Williams compõe esse capítulo.

5.2 MÉTODO DE PAPADAKIS

Devido à necessidade de se retirar as variações geográficas existentes no solo, Papadakis (1937) elaborou um método conhecido como “Méthode statistique pour des expériences sur champ”. Nesse foi sugerido o uso da análise de covariância com médias móveis, calculada através dos valores das parcelas vizinhas, com o intuito de diminuir o erro experimental. Papadakis afirmou que o método corrige os rendimentos baseando-se na correlação da produtividade entre uma parcela e toda a área maior a qual ela pertence e que a análise de covariância elimina as tendências, propiciando a diminuição do erro.

Considere um delineamento completamente casualizado para discorrer sobre o método de Papadakis. O Quadro 11 a seguir representa as N parcelas do experimento no plano.

Quadro 11 – Representação gráfica de N parcelas de um experimento num plano

		5	2	6		
	...	1		3	...	
		7	4	8		

O modelo matemático adotado é:

$$Y_{ij} = \mu + \tau_i + e_{ij} \quad (5.1)$$

Onde Y_{ij} é o valor de resposta Y na repetição j do tratamento i ; μ é a média geral; τ_i é o efeito do tratamento i , e; e_{ij} é o erro experimental ($e_{ij} \sim N(0, \sigma_e^2)$).

O primeiro passo no método de Papadakis é definir quais são os vizinhos de uma parcela; tomando a parcela rachurada no Quadro 11 como a parcela de referência, ela pode ter até oito parcelas vizinhas; para o desenvolvimento que se segue, somente as parcelas 1, 2, 3 e 4 serão as vizinhas.

O segundo passo é criar as covariáveis; se Y_{ij} é o valor da parcela de referência, sejam $Y_{(1)}$, $Y_{(2)}$, $Y_{(3)}$ e $Y_{(4)}$ os valores observados de Y nas parcelas 1, 2, 3 e 4, respectivamente; seja $\hat{\mu}$ a média geral do experimento e $\hat{\tau}_{(1)}$, $\hat{\tau}_{(2)}$, $\hat{\tau}_{(3)}$ e $\hat{\tau}_{(4)}$ os efeitos estimados dos tratamentos aplicados nas parcelas vizinhas. Considerando n o número de parcelas vizinhas, a covariável associada à parcela (i, j) é definida como:

$$Z_{ij} = \sum_{i=1}^n \frac{[Y_{(i)} - (\hat{\mu} + \hat{\tau}_{(i)})]}{n} \quad (5.2)$$

No delineamento completamente casualizado, cada parcela do numerador da equação acima é o erro estimado de uma parcela vizinha. Em outros delineamentos isto não ocorre.

O terceiro passo é conduzir a análise de covariância segundo o modelo:

$$Y_{ij} = \mu + \tau_i + \alpha Z_{ij} + e_{ij} \quad (5.3)$$

Onde α é o coeficiente de regressão associado à covariável Z_{ij} .

Na forma matricial o modelo (5.1) é definido como:

$$Y = D\theta + \varepsilon \quad (5.4)$$

Onde Y é o vetor de observações, D uma matriz de constantes, θ um vetor de parâmetros, incluindo a média geral e os efeitos de tratamentos e ε é o vetor dos erros aleatórios. Seja $\hat{\theta}$ o vetor estimado pelos quadrados mínimos segundo o modelo (5.4), então o estimador de θ , $\hat{\theta}_p$, de Papadakis é o estimador de quadrados mínimos de θ segundo o modelo:

$$Y = D\theta + \alpha N\hat{\varepsilon} + e \quad (5.5)$$

Onde $\hat{\varepsilon}$ é o vetor de erros estimado segundo o modelo (5.4) e N é uma matriz ($n \times n$) de vizinhança, cujos elementos definem quais e quantos são os vizinhos de cada parcela do experimento; o numerador de cada elemento de N diferente de zero é um e o denominador é o número de parcelas vizinhas. Certamente esse número pode variar, pois se a parcela de referência estiver no canto ela poderá ter até três vizinhas, se estiver na borda até cinco.

Se o delineamento é outro, então o modelo final pode ser escrito da forma:

$$Y = X\beta + D\theta + \alpha N(Y - D\hat{\theta}) + e \quad (5.6)$$

Onde X é uma matriz de constantes; β um vetor de parâmetros correspondentes a blocos, linhas ou colunas.

A análise de covariância é feita utilizando o modelo visto anteriormente em (5.6). Bartlett (1978) sugeriu um método iterativo, ou seja, obter $\hat{\theta}_p$ e fazer $\hat{\theta} = \hat{\theta}_p$ no modelo e repetir a análise até que o valor de α se estabilize; Wilkinson (1984), entretanto, mostrou através de simulações que isto leva a um substancial viés positivo no valor do teste F para tratamentos.

Outro detalhe é a criação de mais que uma covariável, por exemplo, no caso discutido anteriormente, as parcelas vizinhas na vertical (2 e 4) podem ter mais influência que as da horizontal (1 e 3) e assim poder-se-ia optar por duas covariáveis. O processo de criação é o mesmo, e o modelo de covariância passa a ter duas covariáveis. Somente um estudo de cada

caso pode orientar a escolha de quantas e quais as covariáveis. Um estudo detalhado do Método de Papadakis também foi feito por Cruz (1998).

5.3 MÉTODO DE WILLIAMS

O método de Papadakis procura envolver a informação da parcela mais próxima na análise do experimento e assim aumentar a precisão; em princípio, ele pode ser utilizado em qualquer experimento de campo. Entretanto, parcelas mais distantes podem também prover informação e, se levadas em conta, podem contribuir para o aumento da precisão. Williams (1986) partiu do seguinte pressuposto: se as parcelas no bloco estão alinhadas em uma só direção, então na análise clássica de experimentos, a variância da diferença entre duas parcelas no mesmo bloco é a mesma independente da distância entre elas. A ideia já tinha sido usada por Wilkinson, Eckert, Hancock e Mayo (1983) que propuseram o método NN (Nearest Neighbour). Segue uma síntese do método de Williams.

Williams (1986) abordou o problema considerando apenas sua utilização na análise de blocos incompletos e em particular para os blocos incompletos resolvíveis. O ponto básico foi envolver a distância entre parcelas dentro do bloco na matriz de variâncias e covariâncias de Y , o vetor de resposta do experimento, já que a variância está relacionada com a distância.

Seguindo a notação de John e Williams (1995) num bloco incompleto resolvível, o modelo adotado é o seguinte:

$$Y = 1\mu + R\pi + X\tau + Z\underline{b} + e \quad (5.7)$$

Onde R é a matriz de incidência das repetições e π é o vetor de efeitos de repetição, considerados como fixos. A matriz de variâncias e covariâncias de Y é dada por:

$$\text{Var}(Y) = ZZ'\sigma_b^2 + I\sigma_e^2 \quad (5.8)$$

Com esse modelo a variância da diferença entre duas parcelas dentro do bloco é $2\sigma_e^2$. No modelo proposto por Williams a matriz $\text{Var}(Y)$ recebe um incremento para que a

variância da diferença entre duas parcelas dentro do bloco leve em conta a distância entre elas. Assim variância de Y passa a ser:

$$\text{Var}(Y) = \sigma_e^2(I + \eta P_B - \phi F) \quad (5.9)$$

Onde,

$$\eta = \frac{k\sigma_b^2}{\sigma_e^2} \quad (5.10)$$

$$P_B = Z(Z'Z)^{-1}Z' \quad (5.11)$$

O parâmetro ϕ está relacionado com a distância entre as parcelas dentro do bloco. A matriz F é um bloco diagonal com componentes $\frac{3L_j}{k^2-1}$ para cada bloco j , considerando que todos os blocos possuem k parcelas; por sua vez, a matriz L_j é uma matriz $(k \times k)$, cujo elemento (i, j) é $|i - j|$, a distância entre as parcelas.

Com esse modelo de variância, a variância da diferença entre duas parcelas dentro do mesmo bloco, separadas por $(d - 1)$ parcelas é $2\sigma_e^2 \left(\frac{1+3\phi d}{k^2-1} \right)$. Assim a variância da diferença entre duas parcelas dentro do bloco está linearmente relacionada com a distância entre elas. Por isso Williams denominou esta técnica como blocos incompletos mais variância linear (BI + VL). Se $\phi = 0$ o modelo volta a ser de blocos incompletos (BI); se $\eta = 0$, mas $\phi > 0$ então o modelo é chamado de variância linear (VL).

Para a estimação dos parâmetros σ_e^2 , η e ϕ , Williams igualou formas quadráticas adequadas aos seus respectivos valores esperados. Elas vão ser usadas nos cálculos, porém os desenvolvimentos não serão feitos aqui, mas estão no trabalho de Williams (1986). Ele definiu a seguinte forma quadrática geral:

$$S(W) = Y'[W - WX(X'WX)^{-1}X'W]Y \quad (5.12)$$

E daqui formulou aquelas relevantes para o trabalho. Basicamente, foi definida uma matriz $\Delta_{(kxk)}$, com segundas diferenças e daí a matriz G , bloco diagonal $\frac{1}{2}(k_j^2 - 1)\Delta$. Com isso ele provou que:

i.

$$\hat{\sigma}_e^2 = \frac{fS(G) - S(I_n - P_B)}{f\text{Traço}(D) - f_b} \quad (5.13)$$

Onde,

$$f = \frac{b - \text{Traço}\{[X'(I_n - P_B)X]^{-1}(X'G^{-1}X)\}}{f_b} \quad (5.14)$$

$$f_b = (n - 1) - (b - 1) - (t - 1) \quad (5.15)$$

$$D = G - GX(X'GX)^{-1}X'G \quad (5.16)$$

ii.

$$\hat{\eta} = \left(\frac{t-1}{s-1}\right) \left(\frac{\hat{\sigma}_r^2 - \hat{\sigma}_e^2}{\hat{\sigma}_e^2}\right) - \frac{1}{f(s-1)} \left(\frac{\hat{\sigma}_b^2 - \hat{\sigma}_e^2}{\hat{\sigma}_e^2}\right) \quad (5.17)$$

Onde,

$$\hat{\sigma}_r^2 = \frac{S(I_n - P_R)}{f_r} \quad (5.18)$$

$$P_R = R'(R'R)^{-1}R \quad (5.19)$$

$$f_r = (r - 1)(t - 1) \quad (5.20)$$

Sendo s o número de blocos incompletos por repetição, considerando um delineamento de blocos incompletos resolvíveis.

iii.

$$\hat{\phi} = \frac{\hat{\sigma}_b^2 - \hat{\sigma}_e^2}{f \hat{\sigma}_e^2} \quad (5.21)$$

Onde,

$$\hat{\sigma}_b^2 = \frac{S(I_n - P_B)}{f_b} \quad (5.22)$$

Estas fórmulas serão implementadas no software SAS (2003) para obtenção das estimativas.

6. APLICAÇÕES E RESULTADOS

Foi visto nos capítulos anteriores que em se tratando de blocos incompletos existem pelo menos três métodos de análise, sendo esses: intrabloco, com recuperação da informação interbloco e análise espacial de Papadakis. Caso seja um delineamento de blocos incompletos resolvíveis, então se tem pelo menos mais duas análises: blocos ao acaso e análise de vizinhança de Williams (embora o modelo de Williams possa ser adaptado para outros delineamentos); o modelo de variância linear (VL) isolado também pode ser considerado como outro método de análise, mas esse não foi exercitado. Neste capítulo são discutidos os resultados da análise de 42 experimentos com variedades de soja, conduzidos pelo então Centro de Pesquisa Agropecuária dos Cerrados/Embrapa (CPAC/Embrapa), hoje, Embrapa Cerrados.

Todos os experimentos foram planejados como blocos incompletos resolvíveis, chamados de latices generalizados de Patterson e Williams (1976), também descrito em John e Williams (1995). Cada experimento é dividido em três grupos com uma repetição, cada grupo de repetição é composto por três blocos com nove parcelas, sendo assim, têm-se 27 tratamentos repetidos três vezes, perfazendo um total de 81 parcelas por experimento. Eles foram realizados no período de 2000/2001. Para cada experimento são feitas cinco análises: blocos ao acaso, intrabloco, com recuperação da informação interbloco (máxima verossimilhança restrita), espacial de Papadakis e vizinhança de Williams. Para todas elas é utilizado o sistema SAS (2003) e a programação segue no Apêndice B.

O objetivo aqui é avaliar o desempenho dos diferentes modelos de análise e para isso é usado o critério chamado de Desvio Padrão Médio (DPM), que consiste no seguinte. Num contexto intrabloco, considere todas as possíveis comparações entre dois tratamentos de um experimento com t tratamentos, isto é, $\binom{t}{2}$ comparações e seja $\text{Var}(\hat{t}) = C^{-1}CC^{-1}\sigma_e^2$, onde \hat{t} é uma solução das equações normais. Então a Variância Média (VM) das $\binom{t}{2}$ comparações é dada por:

$$VM = \frac{\sum_i^t \sum_{j>i}^t \text{Var}(\hat{t}_i - \hat{t}_j)}{\binom{t}{2}} \quad (6.1)$$

O Desvio Padrão Médio (DPM) é dado por \sqrt{VM} .

Este critério foi estudado por alguns autores, como John (1971), e usado por outros como John e Williams (1995) e para facilitar os cálculos, tem-se também o seguinte resultado.

Seja $C\hat{\tau} = Q$ a equação para blocos, então,

$$\sum_i^t \sum_{j>i}^t \text{Var}(\hat{\tau}_i - \hat{\tau}_j) = [t\text{Traço}(C^-) - (\text{soma dos elementos de } C^-)]\sigma_e^2 \quad (6.2)$$

Este resultado foi usado nos programas computacionais.

Assim, para cada análise de um experimento, tem-se um DPM, que é o referencial da eficiência do modelo de análise. No caso do modelo de Williams, as estimativas dos parâmetros η e ϕ para cada experimento também são importantes, pois segundo John e Williams (1995) os seguintes casos podem ocorrer:

- i. Se $\eta > 0$ e $\phi > 0$ então o modelo de Williams pode ser usado (BI + VL);
- ii. Se $\eta > 0$ e $\phi \leq 0$ então o modelo indicado é o blocos incompletos (BI), ou seja, podem ser feitas as análises de blocos ao caso, intrablocos, com recuperação da informação interbloco ou método de Papadakis;
- iii. Se $\eta \leq 0$ e $\phi > 0$ então o modelo é chamado de variância linear (VL), porém nenhum experimento analisado apresenta tal característica, e;
- iv. Se $\eta \leq 0$ e $\phi \leq 0$ então a análise recomendada é a de blocos ao acaso ou método de Papadakis.

Em seguida, na Tabela 11, são apresentados os resultados dos cinco DPM's para cada experimento (Exp.). Para representá-los foram utilizados os seguintes prefixos: DPMBLOC para o DPM obtido na análise de blocos ao acaso, DPMINT para o DPM intrablocos, DPMRECU para o DPM com recuperação da informação interbloco, DPMPAPA para o DPM do método de Papadakis e DPMWILL para o DPM do método de Williams; EFIP e EFIW representam os ganhos em eficiência dos métodos de Papadakis e Williams em relação à análise com recuperação da informação interbloco, respectivamente, e são assim definidas:

$$EFIP = 1 - \frac{DPMPAPA}{DPMRECU} \quad (6.3)$$

$$\text{EFIW} = 1 - \frac{\text{DPMWILA}}{\text{DPMRECU}} \quad (6.4)$$

As estimativas $\hat{\eta}$ e $\hat{\phi}$ são aquelas do modelo de vizinhança de Williams em que $\text{Var}(Y) = \sigma_e^2(I + \eta P_B - \phi F)$

Para melhor visualização dos resultados apresentados a seguir na Tabela 11, considere as seguintes observações: as análises que não são indicadas para um determinado experimento têm seus DPM's na cor vermelha; em relação às demais, é destacado em negrito o menor DPM encontrado e, conseqüentemente, a análise que o possui é a indicada.

Tabela 11 – Resultados de todas as análises

Exp.	DPMBLOC	DPMINT	DPMRECU	DPMPAPA	DPMWILL	EFIP	EPIW	$\hat{\eta}$	$\hat{\phi}$
1	286	312	286	287	259	-	-	-2,033	-1,247
2	254	273	254	257	251	-	0,013	0,252	0,642
3	243	253	243	241	238	0,005	0,021	2,452	1,728
4	296	329	296	302	147	-	0,504	0,283	1,529
5	340	360	340	339	304	0,004	0,107	6,519	5,827
6	276	280	272	273	272	-	-	1,291	0,161
7	349	388	349	337	469	0,034	-	5,540	7,114
8	268	280	267	270	267	-	-	-0,185	-0,397
9	352	391	352	347	1.470	0,014	-	3,346	4,779
10	259	270	258	225	389	0,130	-	-2,278	-2,291
11	208	225	208	208	202	-	0,026	0,056	0,670
12	279	255	252	245	252	0,028	-	3,994	-0,220
13	271	276	267	251	457	0,061	-	-2,019	-2,430
14	276	258	254	254	225	0,001	0,114	17,203	8,596
15	280	262	258	256	254	0,006	0,015	6,053	1,547
16	266	289	266	272	259	-	-	-1,011	-0,300
17	285	293	283	274	283	0,031	-	0,779	0,076
18	241	251	240	241	237	-	0,014	1,843	1,317
19	203	197	193	197	187	-	0,032	5,934	2,616
20	228	222	218	222	215	-	0,011	3,766	1,086
21	343	363	343	349	338	-	0,013	1,223	1,174
22	337	347	334	335	302	-	0,097	7,874	5,917
23	313	307	300	282	300	0,062	-	1,355	-0,462
24	391	375	368	372	355	-	0,035	6,883	2,804
25	366	380	365	370	365	-	-	0,590	0,099
26	414	446	414	409	378	0,013	0,089	3,316	3,586
27	284	274	269	258	253	0,041	0,061	8,584	4,220
28	364	359	351	323	325	0,080	-	0,012	-1,401
29	168	168	163	166	155	-	0,052	6,224	3,588
30	315	331	315	306	261	0,028	0,171	10,889	9,459
31	316	315	307	286	249	0,069	0,190	21,171	14,961
32	341	341	332	294	276	0,115	0,167	17,507	12,179
33	360	379	360	365	351	-	0,024	2,191	1,889
34	355	373	355	323	303	0,090	0,146	9,535	8,280
35	262	283	262	266	246	-	0,064	2,253	2,621
36	315	296	291	280	291	0,039	-	2,696	-0,463
37	302	308	298	300	282	-	0,055	5,407	3,644
38	334	319	313	313	312	0,001	0,003	3,909	0,681
39	365	358	350	339	301	0,033	0,141	15,714	9,941
40	403	379	373	382	368	-	0,013	5,420	1,324
41	317	325	314	320	311	-	-	-0,029	-0,651
42	251	269	251	255	249	-	0,007	0,184	0,490

Citando de maneira geral as análises feitas, tem-se que: em quatro casos a análise indicada é a de blocos ao acaso, a intrablocos não obtém o melhor desempenho em nenhum dos experimentos, a com recuperação da informação interbloco apresenta o menor DPM em duas situações, o método de vizinhança de Papadakis é recomendado em nove casos e, na grande maioria dos experimentos, o método de Williams é o que resulta em um menor DPM, sendo indicado 27 vezes.

Pode-se considerar que as três primeiras análises são as clássicas em experimentos em blocos incompletos resolvíveis: blocos ao acaso, intrablocos e com recuperação da informação interbloco. Comparando estes três tipos de análise pelos seus respectivos DPM's, dezesseis experimentos podem ser analisados como blocos ao acaso.

Confrontando os resultados do método de vizinhança de Williams com os da análise com recuperação da informação interbloco, em quatorze experimentos a eficiência utilizando o método de Williams é maior do que 5% e em um desses casos, chega a 50,4%.

Em quatro casos $\hat{\eta} > 0$ e $\hat{\phi} \leq 0$ e nesses, o método de Papadakis é o que produz o menor DPM. Existem seis casos onde $\hat{\eta} \leq 0$ e $\hat{\phi} \leq 0$, nesses, em quatro ocasiões a análise mais indicada é a de blocos ao acaso e em dois é preferível o uso do método de vizinhança de Papadakis. Não ocorreu nenhum caso onde o modelo de variância linear (VL) é o recomendado, ou seja, $\hat{\eta} \leq 0$ e $\hat{\phi} > 0$. O método de vizinhança de Papadakis não depende de parâmetros como o de Williams e pode ser utilizado em qualquer tipo de experimento de campo; tanto que nas três combinações de $\hat{\eta}$ e $\hat{\phi}$ presentes no estudo, alguns experimentos tiveram ele como o método indicado. A Tabela 12, apresentada abaixo, simplifica as informações apresentadas na Tabela 11.

Tabela 12 – Análises recomendadas conforme estimação de η e ϕ

Análise	$\hat{\eta} > 0$ e $\hat{\phi} > 0$	$\hat{\eta} > 0$ e $\hat{\phi} \leq 0$	$\hat{\eta} \leq 0$ e $\hat{\phi} \leq 0$
Blocos ao acaso	0	0	4
Intrablocos	0	0	-
Com recuperação da informação interbloco	2	0	-
Método de Papadakis	3	4	2
Método de Williams	27	-	-

7. CONCLUSÃO

Neste trabalho foram descritos alguns métodos de análise de experimentos em blocos incompletos: os clássicos e os métodos de análise de vizinhança de Papadakis e de Williams. Os métodos clássicos são bem conhecidos, mas os outros dois não são ou não são utilizados. O método de Williams é mais bem estruturado do que o de Papadakis, tem uma teoria consistente e a partir dos seus resultados pode-se decidir sobre alguma ação, mediante o valor dos parâmetros do modelo. Nos experimentos que foram analisados, cada bloco tinha nove parcelas no sentido longitudinal e neste caso, o ajustamento dos resultados à distância entre parcelas melhorou a eficiência das comparações, o que basicamente propõe o método de Williams. Dos 42 experimentos estudados, 36 apresentaram melhor eficiência com o uso da análise de vizinhança, nove utilizando o método de Papadakis e 27, o método de Williams.

Pelo que foi visto dos resultados sobre eficiência, quatro dos cinco métodos eventualmente vão ser utilizados e isto coloca à disposição dos pesquisadores a opção de selecionar o mais eficiente método na análise dos seus experimentos. Também reforça a recomendação sobre o uso de blocos incompletos e, se possível, blocos incompletos resolvíveis em experimentos de competição de variedades, para grãos principalmente.

A parte computacional, embora não seja complicada, exige algum esforço adicional, tendo em vista que os aplicativos conhecidos não possuem programas específicos para análises de vizinhança. A programação feita no software SAS (2003), no Apêndice B, evidencia este fato.

8. REFERÊNCIAS

- Bailey, R. A. (1991). Strata for randomized experiments. *Journal of the Royal Statistical Society. Series B*, v. 53, n. 1, p. 27-28.
- Bartlett, M. S. (1978). Nearest neighbour models in the analysis of field experiments. *Journal of the Royal Statistical Society. Series B*, vol. 40, n. 2, p. 147-174.
- Bose, R. C., & Nair, K. R. (1939). Partially balanced incomplete block designs. *Sankhya: The Indian Journal of Statistics*, v. 4, n. 3, p. 337-372.
- Cochran, W. G., & Cox, G. M. (1957). *Experimental Designs* (2ª ed.). Oxford: John Wiley & Sons.
- Cox, D. R. (1958). *Planning of Experiments*. New York: John Wiley & Sons.
- Cruz, V. S. (1998). *Estudo de Alguns Modelos Utilizados na Análise Espacial de Experimentos de Campo*. Brasília: Dissertação de Mestrado em Estatística, Universidade de Brasília.
- Federer, W. T. (1963). *Experimental Design: Theory and Application*. Oxford & IBH Publishing Co.
- Fisher, A. R. (1935). *The Design of Experiments* (Ltd. ed.). Edinburgh and London: Oliver and Boyd.
- Hinkelmann, K., & Kempthorne, O. (1994). *Design and Analysis of Experiments: Introduction to Experimental Design* (Vol. 1). New York: John Wiley & Sons.
- IBGE. (2012). *Levantamento Sistemático da Produção Agrícola: Pesquisa mensal de prevenção e acompanhamento das safras agrícolas no ano civil* (Vol. 25). Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística.
- John, J. A., & Williams, E. R. (1995). *Cyclic and Computer Generated Designs* (2ª ed.). Chapman & Hall.
- John, P. W. (1971). *Statistical Design and Analysis of Experiments*. New York: The Macmillan Company.
- Kempthorne, O. (1952). *The Design and Analysis of Experiments*. Oxford: John Wiley & Sons.
- Kuehl, R. O. (2001). *Diseño de Experimentos* (2ª ed.). México: Thomson Learning.

- Nelder, J. A. (1965). The analysis of randomized experiments with orthogonal block structure: I. Block structure and the null analysis of variance, II. Treatment structure and the general analysis of variance. *Proceedings the Royal of Society, A*, v. 283, n. 1393, p. 147-178.
- Papadakis, J. S. (1937). Méthode Statistique pour des expériences sur champ. *Institut d'Amélioration des Plantes à Salonique*, Bulletin, n. 23.
- Patterson, H. D., & Hunter, E. A. (1983). The efficiency of incomplete block designs in National List and Recommended List cereal variety trials. *The Journal of Agricultural Science*, v. 101, n. 2, p. 427-433.
- Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, v. 58, n. 3, p. 545-554.
- Patterson, H. D., & Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika*, v. 63, n. 1, p. 83-92.
- SAS. (2003). *Statistical Analysis System Institute Inc. Cary*. North Carolina: SAS Institute Inc. Version 9.2.
- Searle, S. R., Casella, G., & McCulloch, C. E. (1992). *Variance Components*. New York: John Wiley & Sons.
- Speed, T. P., Williams, E. R., & Patterson, H. D. (1985). A note on the analysis of resolvable block designs. *Journal of the Royal Statistical Society. Series B*, v. 47, n. 2, p. 357-361.
- Wilkinson, G. N. (1984). *Nearest neighbour methodology from designs and analysis of field experiments*. Washington: In International Biometrics Conference.
- Wilkinson, G. N., Eckert, S. R., Hancock, T. W., & Mayo, O. (1983). Nearest Neighbour (NN) analysis of field experiments. *Journal of the Royal Statistical Society, B*, v. 45, n. 2, p. 151-211.
- Williams, E. R. (1986). A neighbour model for field experiments. *Biometrika*, vol. 73, n. 2, p. 279-287.
- Yates, F. (1936). Incomplete randomized blocks. *Annals of Eugenics*, vol. 7, n. 2, p. 121-140.
- Yates, F. (1940). The recovery of inter-block information in balanced incomplete block designs. *Annals of Eugenics*, vol. 10, n. 1, p. 317-325.

APÊNDICE A – Programação em SAS (2003) dos dados de Federer (1963)

```
DATA EXEMPLO;
INPUT BLOCO @;
DO TRAT = 1 TO 10;
INPUT Y @;
IF Y NE . THEN OUTPUT;
END;
DATALINES;
1 9.7 8.7 . 5.4 5.0 . . . . .
2 . 9.6 8.8 . . 5.6 . . . 3.6
3 . 9 . 7.3 . 3.8 4.3 . . .
4 9.3 . 8.7 . 6.8 . 3.8 . . .
5 10 . . 7.5 . . . 4.2 . 2.8
6 . 9.6 . . . . 5.1 4.6 3.6 .
7 . 9.8 . . 7.4 . . 4.4 . 3.8
8 . . . . 9.4 . 6.3 . 5.1 2
9 9.3 9.3 8.2 . . . . . 3.3 .
10 . . . 8.7 9 6 . . 3.3 .
11 9.7 . . . . 6.7 6.6 . . 2.8
12 . . 9.3 8.1 . . . . 3.7 2.6
13 9.8 . . . . 7.3 . 5.4 4 .
14 . . 9 8.3 . . 4.8 3.8 . .
15 . . 9.3 . 8.3 6.3 . 3.8 . .
OPTIONS LS=78 NODATE NONUMBER; RUN;
PROC PRINT DATA=EXEMPLO; RUN;
PROC SORT DATA=EXEMPLO; BY TRAT;
PROC MEANS DATA=EXEMPLO SUM; BY TRAT;
VAR Y;
RUN;
OPTIONS LS=90 NODATE CENTER PS=100 NONUMBER; RUN;
```

```
/******ANÁLISE INTRABLOCO******/
```

```
TITLE ANÁLISE INTRABLOCO;
PROC GLM DATA=EXEMPLO;
CLASS BLOCO TRAT;
MODEL Y = BLOCO TRAT;
LSMEANS TRAT ;
ESTIMATE "1 x 4" TRAT 1 0 0 -1;
RUN;
QUIT;
```

```
/******ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO******/
```

```
TITLE ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO (YATES);
PROC MIXED DATA=EXEMPLO METHOD=TYPE1;
CLASS BLOCO TRAT;
MODEL Y = TRAT;
RANDOM BLOCO;
LSMEANS TRAT ;
```

```

ESTIMATE "1 x 4" TRAT 1 0 0 -1;
RUN;
QUIT;
TITLE ANALISE COM RECUPERAÇÃO INTERBLOCO(NELDER);
PROC MIXED DATA=EXEMPLO2 METHOD=REML;
CLASS BLOCO TRAT;
MODEL Y = TRAT;
RANDOM BLOCO;
LSMEANS TRAT ;
ESTIMATE "1 x 4" TRAT 1 0 0 -1;
RUN;
QUIT;
PROC SORT DATA=EXEMPLO; BY BLOCO TRAT;
PROC PRINT DATA=EXEMPLO; RUN;
OPTIONS LS=78; RUN;

```

/******PREPARAR DATA SET PARA O PROC IML *****/

```

OPTIONS LS=132 PS=62 NODATE NONUMBER; RUN;
PROC IML;
USE EXEMPLO;
READ ALL VAR{TRAT} INTO T;
READ ALL VAR{BLOCO} INTO BL;
READ ALL VAR{Y} INTO Y;
K=4;
V=10;
R=6;
B=15 ;
PRINT T BL Y;
Z=DESIGN(BL);
X=DESIGN(T);
NO=B*K;
UM=J(NO,1,1);
OPTION NOCENTER LS=160;
PRINT UM X Z;
OPTION CENTER LS=90;
N=X*Z;
NN=N*N;
ID1 = I(V);
C = R*ID1 - (1/K)*N*N;
PRINT C;
TT = X*Y;
TB = Z*Y;
PRINT TT TB;
Q =TT-(1/K)*N*TB;
TAU=GINV(C)*Q; /*SOLUÇÃO PARA TAU*/
GERAL=(Y*J(NO,1,1))/NO;
PRINT Q TAU GERAL;
W = UM||X||Z; /*MODELO INTEGRAL*/
XY = W*Y;
PRINT W XY;
GW=GINV(W*W);
BETA=GW*XY; PRINT BETA; /*BETA GERAL*/
ID2=I(NO);
SQERRO=Y*(ID2 - W*GINV(W*W)*W)*Y;
SQTRAT=Q*TAU;

```

```

QMERRO=SQERRO/36;
VAR_GER=GW*W`*W*GW*QMERRO; /*VAR DE BETA GERAL*/
OPTIONS LS=200;
PRINT VAR_GER;
PRINT SQTRAT SQERRO QMERRO;
MEDIA_A=GERAL*J(V,1,1)+TAU;
PRINT TAU MEDIA_A;
VAR=GINV(C)*C*GINV(C); PRINT VAR;
VARTAU=QMERRO*GINV(C)*C*GINV(C);
PRINT VAR;
PRINT VARTAU;

```

/******ANÁLISE INTERBLOCO******/

```

KUM=Z`*UM;
NN=N`;
ZY=TB||KUM||NN;
OPTIONS LS=160 NOCENTER;
CREATE INTER1 FROM ZY;
APPEND FROM ZY;
CLOSE INTER1;
PRINT ZY;
XI=KUM||N`;
GXI=GINV(XI`*XI);
BETA_I=GXI*XI`*TB;
PRINT KUM XI;
ID3=I(B);
SQE_INTE=TB`*(ID3-XI*GINV(XI`*XI)*XI)`*TB; /*SQE*/
PRINT SQE_INTE;
G=Y`*J(NO,1,1);
TAU_I=INV(N*N`)*N*TB-J(V,1,1)*G*(1/(B*K)); /*TAU INTER*/
PRINT TAU_I;
M_AJUS_I=TAU_I+J(10,1,1)*6.528;
PRINT TAU_I M_AJUS_I;
Q_I=N*TB-R*(G/B)*J(V,1,1);
PRINT Q_I;
C_I=N*N`-(R*R/B)*J(V,1,1)*J(V,1,1);
PRINT C_I;
TAU_IC=GINV(C_I)*Q_I;
PRINT TAU TAU_I TAU_IC;
SQTR_I=Q_I*TAU_I; /*SQT*/
PRINT SQTR_I;
M_B=SUM(TB)/B;
SQM_I=B*(M_B**2); PRINT SQM_I;
SQTOT_I=TB`*TB-SQM_I; /*SQTOTAL*/
PRINT SQM_I SQTR_I SQE_INTE SQTOT_I;
QME_I=SQE_INTE/(B-V); /*QME*/
QMTR_I=SQTR_I/(V-1); /*QMT*/
F_I=QMTR_I/QME_I; /* F PARA TRATAMENTO*/
PRINT QMTR_I QME_I F_I;
VAR_I=GINV(C_I)*C_I*GINV(C_I)*QME_I;
PRINT TB;
PRINT SQE_INTE SQTR_I SQTOT_I;
PRINT QMTR_I QME_I F_I;
OPTIONS LS=132;
PRINT VAR_I;

```



```
AZ={1 0 0 -1 0 0 0 0 0};
Z_T1_T4=AZ*TAU_I;
VAR_T1_T4=AZ*VAR_I*AZ`;
EP_Z=SQRT(VAR_T1_T4);
PRINT Z_T1_T4 VAR_T1_T4 EP_Z;
QUIT;
TITLE ANÁLISE INTERBLOCO;
PROC PRINT DATA=INTER1; RUN;
OPTIONS CENTER LS=90;
PROC GLM DATA=INTER1;
MODEL COL1=COL3-COL11/SOLUTION;
RUN; QUIT;
```

APÊNDICE B – Programação em SAS (2003) dos 42 experimentos analisados

```
LIBNAME DADOS "C:/Users/Administrador/Documents/Estatística/8º Semestre/Estágio Supervisionado I e II/Programação"; RUN;  
PROC PRINT DATA=DADOS.TODOS; RUN;  
DATA TODOS; SET DADOS.TODOS; RUN;  
PROC PRINT DATA=TODOS; RUN;
```

```
/******ANÁLISE COMO BLOCOS AO ACASO******/
```

```
PROC SORT DATA=TODOS; BY NUM_EXP; RUN;  
PROC GLM DATA=TODOS OUTSTAT=DPMT NOPRINT; BY NUM_EXP;  
CLASS REP TRAT;  
MODEL REND=REP TRAT;  
RUN;  
QUIT;  
PROC PRINT DATA=DPMT; RUN;  
DATA DPMT1; SET DPMT;  
IF _SOURCE_="ERROR";  
SIGMA=SS/DF;  
DPME=SQRT((2*SS/DF)/3);  
KEEP NUM_EXP DPME SIGMA;  
PROC PRINT DATA=DPMT1; RUN;  
DATA DPMT2; SET DPMT1;  
DPMERAN=DPME;  
SIGMARAN=SIGMA;  
DROP DPME SIGMA;  
TITLE RESULTADOS DA ANÁLISE COMO BLOCOS AO ACASO;  
PROC PRINT DATA=DPMT2; RUN;
```

```
/***RESULTADOS DA ANÁLISE COMO BLOCOS AO ACASO***/
```

```
LIBNAME RESULT "C:/Users/Administrador/Documents/Estatística/8º Semestre/Estágio Supervisionado I e II/Programação"; RUN;  
DATA RESULT.BLOC; SET DPMT2; RUN;  
PROC PRINT DATA=RESULT.BLOC; RUN;  
TITLE TESTE;  
PROC GLM DATA=TODOS; WHERE NUM_EXP=1;  
CLASS REP TRAT;  
MODEL REND=REP TRAT;  
ESTIMATE "1 2" TRAT 1 -1;  
RUN;  
QUIT;
```

```
/******ANÁLISE INTRABLOCO******/
```

```
DATA CRESCE;  
INPUT NUM_EXP COL1 SIGMA;  
DATALINES;  
0 0 0  
PROC PRINT DATA=CRESCE; RUN;
```

```

%MACRO INTRA;
%DO L1 =1 %TO 42;
DATA ANALISE; SET TODOS;
IF NUM_EXP=&L1;
PROC GLM DATA=ANALISE NOPRINT OUTSTAT=DPMTIN;
CLASS REP BLOCO TRAT;
MODEL REND=REP BLOCO(REP) TRAT;
LSMEANS TRAT/COV OUT=COVBIN;
RUN;
QUIT;
DATA COVBIN3; SET COVBIN; KEEP COV1-COV27;
RUN;
PROC IML;
USE COVBIN3;
READ ALL VAR _NUM_ INTO V;
S=SUM(V);
VAR_MEDIA=((27*TRACE(V)-S)/351);
DP_MEDIO=SQRT(VAR_MEDIA);
CREATE CINCO FROM DP_MEDIO;
APPEND FROM DP_MEDIO;
CLOSE CINCO;
QUIT;
DATA DPM; SET CINCO;
NUM_EXP=&L1;
DATA CRESCE; SET CRESCE DPM; RUN;
DATA SIGMAIN; SET DPMTIN;
IF _TYPE_="ERROR";
SIGMA=SS/DF;
NUM_EXP=&L1;
KEEP NUM_EXP SIGMA;
PROC SORT DATA=CRESCE; BY NUM_EXP;
PROC SORT DATA=SIGMAIN; BY NUM_EXP;
DATA CRESCE; MERGE CRESCE SIGMAIN; BY NUM_EXP;
%END;
%MEND INTRA;
RUN;
%INTRA
PROC PRINT DATA=CRESCE; RUN;
DATA CRESCE1; SET CRESCE;
IF NUM_EXP=0 THEN DELETE;
DPMEIN=COL1;
SIGMAIN=SIGMA;
DROP COL1 SIGMA;;
TITLE RESULTADOS DA ANÁLISE INTRABLOCO;
PROC PRINT DATA=CRESCE1; RUN;

/****RESULTADOS DA ANÁLISE INTRABLOCO****/

DATA RESULT.INT; SET CRESCE1; RUN;
PROC PRINT DATA=RESULT.INT; RUN;
TITLE TESTE;
PROC GLM DATA=TODOS OUTSTAT=DPMTIN; WHERE NUM_EXP=2;
CLASS REP BLOCO TRAT;
MODEL REND=REP BLOCO(REP) TRAT;
LSMEANS TRAT/ COV OUT=COVBIN;
ESTIMATE "1 2" TRAT 1 -1;
RUN;
QUIT;

```

```

PROC PRINT DATA=DPMTIN; RUN;
DATA COVBIN3; SET COVBIN; KEEP COV1-COV27;
RUN;
PROC IML;
USE COVBIN3;
READ ALL VAR _NUM_ INTO V;
S=SUM(V); PRINT S;
VAR_MEDIA=((27*TRACE(V)-S)/351);
DP_MEDIO=SQRT(VAR_MEDIA);
CREATE CINCO FROM DP_MEDIO;
APPEND FROM DP_MEDIO;
CLOSE CINCO;
QUIT;
PROC PRINT DATA=CINCO; RUN;
DATA SIGMAIN; SET DPMTIN;
IF _TYPE_="ERROR";
SIGMA=SS/DF;
KEEP SIGMA;
PROC PRINT DATA=SIGMAIN; RUN;
PROC MIXED DATA=TODO; WHERE NUM_EXP=1;
CLASS REP BLOCO TRAT;
MODEL RE=REP BLOCO(REP) TRAT/COVB;
ODS OUTPUT COVB=COVVBM;
ESTIMATE "1 2" TRAT 1 -1;
QUIT;
OPTIONS LS=130;
PROC PRINT DATA=COVVBM; RUN;
DATA COVVBM1; SET COVVBM;
IF EFFECT="Intercept" THEN DELETE;
IF EFFECT="REP" THEN DELETE;
IF EFFECT="BLOCO(REP)" THEN DELETE;
DROP ROW EFFECT REP BLOCO TRAT COL1-COL13;
PROC PRINT DATA=COVVBM1; RUN;
TITLE ANÁLISE INTRABLOCO COM PROC MIXED ;
PROC IML;
USE COVVBM1;
READ ALL VAR _NUM_ INTO V;
S=SUM(V);
VAR_MEDIA=((27*TRACE(V)-S)/351);
PRINT VAR_MEDIA;
DP_MEDIO=SQRT(VAR_MEDIA);
PRINT DP_MEDIO;
QUIT;

```

/******ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO******/

```

DATA CRESCERE;
INPUT NUM_EXP COL1;
DATALINES;
0 0
PROC PRINT DATA=CRESCERE; RUN;
DATA VARIANCIAS;
INPUT NUM_EXP COV Parm $ 4-13 ESTIMATE;
DATALINES;
0 BLOCO(REP) -1
0 RESIDUAL -1

```

```

PROC PRINT DATA=VARIANCIAS; RUN;
%MACRO RECUPE;
%DO L1=1 %TO 42;
PROC PRINTTO NEW UNIT=20; RUN;
DATA ANALISE; SET TODOS;
IF NUM_EXP=&L1 ;
PROC MIXED DATA=ANALISE ;
CLASS REP BLOCO TRAT;
MODEL REND=REP TRAT/COVB ;
RANDOM BLOCO(REP) ;
ODS OUTPUT COVB=COVBMREC;
ODS OUTPUT COVPARMS=VAR_SRE;
RUN;
QUIT;
DATA COVBMREC1; SET COVBMREC;
IF EFFECT="Intercept" THEN DELETE;
IF EFFECT="REP" THEN DELETE;
DROP REP COL1 COL2 COL3 COL4 ROW EFFECT TRAT ;
RUN;
PROC IML;
USE COVBMREC1;
READ ALL VAR _NUM_ INTO V;
S=SUM(V) ;
VAR_MEDIA=((27*TRACE(V)-S)/351);
DP_MEDIO=SQRT(VAR_MEDIA);
CREATE CINCO FROM DP_MEDIO;
APPEND FROM DP_MEDIO;
CLOSE CINCO;
QUIT;
DATA DPM; SET CINCO;
NUM_EXP=&L1;
DATA CRESCERE; SET CRESCERE DPM; RUN;
DATA VARS3; SET VAR_SRE;
NUM_EXP=&L1;
DATA VARIANCIAS; SET VARIANCIAS VARS3; RUN;
%END;
PROC PRINTTO; RUN;
%MEND RECUPE;
RUN;
%RECUPE
PROC PRINT DATA=CRESCERE; RUN;
PROC PRINT DATA=VARIANCIAS; RUN;
DATA CRESCERE1; SET CRESCERE;
IF NUM_EXP=0 THEN DELETE;
DPMERE=COL1;
DROP COL1;
PROC PRINT DATA=CRESCERE1; RUN;

/****RESULTADOS DA ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO I****/

DATA VARI12; SET VARIANCIAS;
SIGMA_BL=ESTIMATE;
IF COVARM="BLOCO(REP)";
IF NUM_EXP=0 THEN DELETE;
DROP ESTIMATE COVARM;
PROC PRINT DATA=VARI12; RUN;

/****RESULTADOS DA ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO II****/

```

```

TITLE RESULTADOS DA ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO;
DATA VARI13; SET VARIANCIAS;
SIGMARE=ESTIMATE;
IF COVPARAM="RESIDUAL";
IF NUM_EXP=0 THEN DELETE;
DROP ESTIMATE COVPARAM;
PROC PRINT DATA=VARI13; RUN;
PROC SORT DATA=CRESCERE1; BY NUM_EXP;
PROC SORT DATA=VARI12; BY NUM_EXP;
PROC SORT DATA=VARI13; BY NUM_EXP;
DATA RECUP; MERGE CRESCERE1 VARI12 VARI13; BY NUM_EXP;
PROC PRINT DATA=RECUP; RUN;

```

```

/****RESULTADOS DA ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO III****/

```

```

DATA JUNTOS; MERGE DPMT2 CRESCE1 RECUP;
PROC PRINT DATA=JUNTOS; RUN;

```

```

/****RESULTADOS DA ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO****/

```

```

DATA RESULT.RECU; SET JUNTOS; RUN;
PROC PRINT DATA=RESULT.RECU; RUN;
TITLE TESTE;
TITLE ANÁLISE COM RECUPERAÇÃO DA INFORMAÇÃO INTERBLOCO;
OPTIONS LS=130 PS=100;
PROC MIXED DATA=TODOS; WHERE NUM_EXP=2;
CLASS REP BLOCO TRAT;
MODEL REND=REP TRAT/COVB;
RANDOM BLOCO(REP);
ODS OUTPUT COVB=COVVBM;
ODS OUTPUT COVPARAMS=VAR_S;
QUIT;
PROC PRINT DATA=COVVBM; RUN;
DATA COVVBM1; SET COVVBM;
IF EFFECT="INTERCEPT" THEN DELETE;
IF EFFECT="REP" THEN DELETE;
DROP REP COL1 COL2 COL3 COL4 ROW EFFECT TRAT;
PROC PRINT DATA=COVVBM1; RUN;
TITLE ANÁLISE COM RECUPERACAO DA INFORMAÇÃO INTERBLOCO;
PROC IML;
USE COVVBM1;
READ ALL VAR _NUM_ INTO V;
A=EIGVAL(V);
S=SUM(V);
VAR_MEDIA=((27*TRACE(V)-S)/351);
PRINT VAR_MEDIA;
DP_MEDIO=SQRT(VAR_MEDIA);
PRINT DP_MEDIO;
QUIT;
PROC PRINT DATA=VAR_S; RUN;

```

```

/*****MÉTODO DE PAPADAKIS*****/

```

```

TITLE MÉTODO DE PAPADAKIS;
DATA CRESCEPA;
INPUT NUM_EXP COL1;

```

```

DATALINES;
0 0
PROC PRINT DATA=CRESCPEPA; RUN;
DATA VARIANCIAP;
INPUT NUM_EXP COVPARM $ 4-13 ESTIMATE;
DATALINES;
0 BLOCO(REP) -1
0 RESIDUAL -1
PROC PRINT DATA=VARIANCIAP; RUN;
%MACRO PAPAD;
%DO L1=1 %TO 42;
PROC PRINTTO NEW UNIT=20; RUN;
DATA ANALISEP; SET TODOS;
IF NUM_EXP=&L1;
C+1;
PARCELA=MOD(C,9);
IF PARCELA=0 THEN PARCELA=9;
PROC PRINT DATA=PAPAD1; RUN;
TITLE ANÁLISE COM RECUPERACAO DA INFORMAÇÃO INTERBLOCO PARA ERROS;
PROC MIXED DATA=ANALISEP;
CLASS REP BLOCO TRAT;
MODEL REND=REP TRAT/OUTPRED=PRE1;
RANDOM BLOCO(REP);
QUIT;
PROC PRINT DATA=PRE1; RUN;
DATA PAPAD2; SET PRE1;
IF REP=1 THEN BLOCO_R=BLOCO;
IF REP=2 THEN BLOCO_R=BLOCO+3;
IF REP=3 THEN BLOCO_R=BLOCO+6;
KEEP LOCAL REP TRAT PARCELA NUM_EXP BLOCO BLOCO_R REND RESID;
PROC PRINT DATA=PAPAD2; RUN;
PROC IML;
USE PAPAD2;
READ ALL VAR{REP} INTO RC;
READ ALL VAR{BLOCO_R} INTO BC;
READ ALL VAR{TRAT} INTO XC;
READ ALL VAR{REND} INTO Y;
READ ALL VAR{RESID} INTO ER;
READ ALL VAR{REP PARCELA BLOCO REND TRAT BLOCO_R RESID} INTO DADOS; PRINT
DADOS;
R=DESIGN(RC);
X=DESIGN(XC);
Z=DESIGN(BC);
N=NROW(R);
NR=NCOL(R);
NT=NCOL(X);
NZ=NCOL(Z);
V=27;
S=3;
K=9;
B=9;
D1=J(K,K,0);
D1[1,2]=1;
D1[2,1]=1;
D1[2,3]=1;
D1[3,2]=1;
D1[3,4]=1;
D1[4,3]=1;

```

```

D1[4,5]=1;
D1[5,4]=1;
D1[5,6]=1;
D1[6,5]=1;
D1[6,7]=1;
D1[7,6]=1;
D1[7,8]=1;
D1[8,7]=1;
D1[8,9]=1;
D1[9,8]=1;
PRINT D1;
SOMA=D1*J(K,1,1);PRINT SOMA;
D2=J(K,K,0);
DO J=1 TO K;
DO L=1 TO K;
D2[L,J]= D1[L,J]/SOMA[L,];
END; END; print D2;
D3=BLOCK(D2,D2,D2);
NN=BLOCK(D3,D3,D3); PRINT NN ER;
XP=NN*ER; PRINT XP;
TUDO=DADOS||ER||XP;
CREATE DADOSP FROM TUDO;
APPEND FROM TUDO;
CLOSE DADOSP;
QUIT;
PROC PRINT DATA=DADOSP; RUN;
DATA DADOSP1; SET DADOSP;
DROP COL8;
REP=COL1; PARCELA=COL2; BLOCO=COL3;
Y=COL4; TRAT=COL5; BLOCO_R=COL6;
ERRO=COL7;
XP=COL9;
DROP COL1-COL7; DROP COL9;
PROC PRINT DATA=DADOSP1; RUN;
TITLE ANÁLISE COM RECUPERACAO DA INFORMAÇÃO INTERBLOCO;
PROC MIXED DATA=DADOSP1;
CLASS REP BLOCO TRAT;
MODEL Y=REP XP TRAT/COVB ;
RANDOM BLOCO(REP);
ODS OUTPUT COVB=COVPAP3;
ODS OUTPUT COVPARMS=VAR_SREP;
QUIT;
DATA COVPAP4; SET COVPAP3;
DROP ROW EFFECT REP TRAT COL1 COL2 COL3 COL4 COL5;
C+1;
IF C < 6 THEN DELETE;
DROP C;
PROC PRINT DATA=COVPAP4; RUN;
PROC PRINTTO; RUN;
TITLE ANÁLISE DE PAPADAKIS COM RECUPERACAO DA INFORMAÇÃO INTERBLOCO;
PROC IML;
USE COVPAP4;
READ ALL VAR _NUM_ INTO V;
A=EIGVAL(V);
S=SUM(V);
VAR_MEDIA=((27*TRACE(V)-S)/351);
PRINT VAR_MEDIA;
DP_MEDIO=SQRT(VAR_MEDIA);

```



```

PRINT DP_MEDIO;
CREATE CINCOP FROM DP_MEDIO;
APPEND FROM DP_MEDIO;
CLOSE CINCOP;
QUIT;
DATA DPMP; SET CINCOP;
NUM_EXP=&L1;
DATA CRESCEPA; SET CRESCEPA DPMP; RUN;
DATA VARS3P; SET VAR_SREP;
NUM_EXP=&L1;
DATA VARIANCIAP; SET VARIANCIAP VARS3P;
RUN;
%END;
%MEND PAPAD;
RUN;
%PAPAD
PROC PRINT DATA=CRESCEPA; RUN;
PROC PRINT DATA=VARIANCIAP; RUN;
DATA DAKS1; SET CRESCEPA;
DPMEDIAPA=COL1;
IF COL1=0 THEN DELETE;
DROP COL1;
PROC PRINT DATA=DAKS1; RUN;
DATA DAKS2; SET VARIANCIAP;
IF NUM_EXP=0 THEN DELETE;
VAR_BLOCO=ESTIMATE;
IF COVPARAM="BLOCO(REP)";
DROP ESTIMATE;
PROC PRINT DATA=DAKS2; RUN;
DATA DAKS3; SET VARIANCIAP;
IF NUM_EXP=0 THEN DELETE;
VAR_RESID= ESTIMATE;
IF COVPARAM="RESIDUAL";
DROP ESTIMATE;
PROC PRINT DATA=DAKS3; RUN;
PROC SORT DATA=DAKS1; BY NUM_EXP;
PROC SORT DATA=DAKS2; BY NUM_EXP;
PROC SORT DATA=DAKS3; BY NUM_EXP;
DATA DAKS4; MERGE DAKS1 DAKS2 DAKS3;
DROP COVPARAM;
PROC PRINT DATA=DAKS4; RUN;

```

/******MÉTODO DE WILLIAMS******/

```

TITLE MÉTODO DE WILLIAMS;
DATA CREWILL;
INPUT COL1 COL2 COL3 COL4 COL5;
DATALINES;
0 0 0 0
PROC PRINT DATA=CREWILL; RUN;
%MACRO LVIB;
%DO L1 =1 %TO 42;
DATA ANALISE; SET TODOS;
IF NUM_EXP=&L1;
BLOCO_R=BLOCO;
IF REP=2 THEN BLOCO_R=BLOCO+3;

```

```

IF REP=3 THEN BLOCO_R=BLOCO+6;
PROC IML;
USE ANALISE;
READ ALL VAR{REP} INTO RC;
READ ALL VAR{BLOCO_R} INTO BC;
READ ALL VAR{TRAT} INTO XC;
READ ALL VAR{REND} INTO Y;
R=DESIGN(RC);
X=DESIGN(XC);
Z=DESIGN(BC);
N=NROW(R);
NR=NCOL(R);
NT=NCOL(X);
NZ=NCOL(Z);
V=27;
S=3;
K=9;
B=9;
UM=J(N,1,1);
IN=I(N);
PG=UM*INV(UM`*UM)*UM`;
PR=R*INV(R`*R)*R`;
PB=(Z*INV(Z`*Z)*Z`);
YC=(IN-PG)*Y;
W=(IN-PB);
FB=N-V-B+1;
S2B=(YC`*(W-W*X*GINV(X`*W*X)*X`*W)*YC)/FB;
W1=(IN-PR);
FR=(NR-1)*(V-1);
S2R=(YC`*(W1-W1*X*GINV(X`*W1*X)*X`*W1)*YC)/FR;
GAMA=((V-1)/(S-1))*((S2R-S2B)/S2B);

/****CÁLCULO DE GAMA DE SPEED, WILLIAMS E PATTERSON****/

L=J(K,K,0);
DO I=1 TO K;
DO J=1 TO K;
L[I,J]=ABS(I-J);
END; END;
DA=(3/(K**2-1))*L;
F=BLOCK(DA,DA,DA,DA,DA,DA,DA,DA,DA);
TESTE=PB*F*PB; PRINT TESTE PB;
DE1=J(8,9,0);
DO D=1 TO 8;
DE1[D,D]=1;
DE1[D,D+1]=-1;
END;
DELTA=(1/2)*DE1`*DE1;
D1=(1/3)*((K**2-1)*DELTA);
G=BLOCK(D1,D1,D1,D1,D1,D1,D1,D1,D1);
D=G-G*X*GINV(X`*G*X)*X`*G;
T1=(GINV(X`*(IN-PB)*X))*(X`*GINV(G)*X);
F1=(B-TRACE(T1))/FB;
SG=(YC`*(G-G*X*GINV(X`*G*X)*X`*G)*YC);
W2=IN-PB;
SW2=(YC`*(W2-W2*X*GINV(X`*W2*X)*X`*W2)*YC);
SIGMA2=(F1*SG -SW2)/(F1*TRACE(D)-FB);
PHI=(S2B - SIGMA2)/(F1*SIGMA2);

```

```
GG2=(1/(F1*(S-1)))*((S2B-SIGMA2)/SIGMA2);
GAMA2=((V-1)/(S-1))*((S2R-SIGMA2)/SIGMA2)-GG2;
PRINT SIGMA2 GAMA2 PHI;
```

```
/**PARÂMETROS DE WILLIAMS (BI + VL)***/
```

```
M_VAR=SIGMA2*(IN+GAMA2*PB-PHI*F);
M_VAR_R=(IN-PR)*M_VAR*(IN-PR);
TAU=(GINV(X`*GINV(M_VAR_R)*X))*(X`*GINV(M_VAR_R)*YC);
CM=GINV(X`*GINV(M_VAR_R)*X);
SOMA=SUM(CM); PRINT SOMA;
VAR_MEDIA=((V*TRACE(CM)-SOMA)/351);
DP_BI_LV=SQRT(VAR_MEDIA);
PRINT VAR_MEDIA DP_BI_LV;
```

```
/**DESVIO PADRÃO MÉDIO DE WILLIAMS (BI + VL)***/
```

```
PRINT SIGMA2 GAMA2 PHI;
NUM_EXP=&L1;
RESUL=NUM_EXP||DP_BI_LV||SIGMA2||GAMA2||PHI;
PRINT RESUL;
CREATE WILLLVIB FROM RESUL;
APPEND FROM RESUL;
CLOSE WILLLVIB;
QUIT;
DATA WILR; SET WILLLVIB ;
DATA CREWILL; SET CREWILL WILR;
%END;
%MEND LVIB;
RUN;
%LVIB
PROC PRINT DATA=WILLLVIB; RUN;
PROC PRINT DATA=CREWILL; RUN;
DATA WLVB1; SET CREWILL;
NUM_EXP=COL1;
DP_BI_LV=COL2;
SIGMA2=COL3;
GAMA2=COL4;
PHI=COL5;
DROP COL1-COL5;
IF NUM_EXP=0 THEN DELETE;
PROC PRINT DATA= WLVB1; RUN;
```

```
/**RESULTADOS DE TODAS AS ANÁLISES***/
```

```
LIBNAME RESULT1 "C:/Users/Administrador/Documents/Estatística/8º Semestre/Estágio Supervisionado I e II/Programação"; RUN;
PROC PRINT DATA= RESULT1.EFICIENCIA2; RUN;
DATA RESULT1.EFICIENCIA3; SET RESULT1.EFICIENCIA2;
DPBLOCO=DPMERAN;
DPINTRA=DPMEIN;
DPPRECUP=DPMERE;
DPPAPAD=DPMEDIAPA;
DPWILLA=DP_BI_LV;
ERROBLO=SIGMARAN;
ERROINT=SIGMAIN;
```

```

ERROREC=SIGMARE;
ERROPAP=VAR_RESID;
ERROWIL=SIGMA2;
BLOCO_REC=SIGMA_BL;
BLOCO_PAD=VAR_BLOCO;
BLOCO_WIL=(SIGMA2*GAMA2)/9;
GAMA=GAMA2;
PHHI=PHI;
DROP DPMERAN SIGMARAN DPMEIN SIGMAIN DPMERE SIGMA_BL
SIGMARE DPMEDIAPA VAR_BLOCO VAR_RESID DP_BI_LV SIGMA2 GAMA2 PHI; RUN;
TITLE RESULTADOS DE TODAS AS ANÁLISES;
PROC PRINT NOOBS DATA=RESULT1.EFICIENCIA3; RUN;
OPTIONS NODATE NONUMBER LS=90;
TITLE RESULTADOS DE TODAS AS ANÁLISES;
PROC PRINT NOOBS DATA=RESULT1.EFICIENCIA3;
VAR NUM_EXP DPBLOCO DPINTRA DPRECUP DPPAPAD DPWILLA GAMA PHHI;
FORMAT DPBLOCO DPINTRA DPRECUP DPPAPAD DPWILLA 4.0;
RUN;

```

ANEXO A – Dados de Federer (1963)

Quadro 12 – Delineamento de blocos incompletos com dez tratamentos em quinze blocos de quatro parcelas

Bloco 1	1	2	4	5
	9,7	8,7	5,4	5
Bloco 2	2	3	6	10
	9,6	8,8	5,6	3,6
Bloco 3	2	4	6	7
	9	7,3	3,8	4,3
Bloco 4	1	3	5	7
	9,3	8,7	6,8	3,8
Bloco 5	1	4	8	10
	10	7,5	4,2	2,8
Bloco 6	2	7	8	9
	9,6	5,1	4,6	3,6
Bloco 7	2	5	8	10
	9,8	7,4	4,4	3,8
Bloco 8	5	7	9	10
	9,4	6,3	5,1	2
Bloco 9	1	2	3	9
	9,3	9,3	8,2	3,3
Bloco 10	4	5	6	9
	8,7	9	6	3,3
Bloco 11	1	6	7	10
	9,7	6,7	6,6	2,8
Bloco 12	3	4	9	10
	9,3	8,1	3,7	2,6
Bloco 13	1	6	8	9
	9,8	7,3	5,4	4
Bloco 14	3	4	7	8
	9	8,3	4,8	3,8
Bloco 15	3	5	6	8
	9,3	8,3	6,3	3,8