

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

Comparação de rols classificatórios de tratamentos e de estimativas
de componentes de variância em grupos de experimentos

Cássio Dessotti

Dissertação apresentada para obtenção do título de Mestre
em Ciências. Área de concentração: Estatística e Experi-
mentação Agronômica

Piracicaba
2009

Cássio Dessotti
Licenciado em Matemática

**Comparação de rols classificatórios de tratamentos e de estimativas
de componentes de variância em grupos de experimentos**

Orientadora:
Prof^ª Dr^ª **SÔNIA MARIA DE STEFANO PIEDADE**

Dissertação apresentada para obtenção do título de Mestre
em Ciências. Área de concentração: Estatística e Experi-
mentação Agrônômica

**Piracicaba
2009**

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

Dessotti, Cássio

Comparação de rols classificatórios de tratamentos e de estimativas de componentes de variância em grupos de experimentos / Cássio Dessotti. - - Piracicaba, 2009.
100 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2009.
Bibliografia.

1. Análise de variância 2. Estatística aplicada 3. Melhoramento genético 4. Modelagem de dados 5. Planejamento e análise de experimentos 6. Softwares I. Título

CDD 519.535
D475c

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

Dedicatória

Primeiramente a **Deus**,
pois sem Ele, jamais teria conseguido
realizar este trabalho!

A meus pais Eliamar e Laerte,
ao meu irmão Fernando,
aos meus avôs Antônio (em memória),
Marcília, Mario e Sinésia,
que eu os **AMO MUITO!**
OFEREÇO.

AGRADECIMENTOS

À minha orientadora professora Dra. Sônia Maria De Stefano Piedade, pelo apoio, compreensão, auxílio e acima de tudo amizade na elaboração deste trabalho.

Agradeço, em especial, o professor Dr. Décio Barbin, pela atenção e incalculável contribuição teórica.

Aos professores doutores Carlos Tadeu dos Santos Dias, Clarice Garcia Borges Demetrio, Gerson Barreto Mourão, Roseli Aparecida Leandro e Silvio Sandoval Zocchi.

Aos professores da UNESP de Ilha Solteira, Dr. Evaristo Bianchini (em memória) e Walter Veriano Valério Filho, pela indicação, motivação e grande amizade.

À pesquisadora Dra. Solidete de Fátima Paziani pelo banco de dados cedido para uso neste trabalho.

As secretárias Solange de Assis Paes Sabadin e Luciane Brajão pela paciência.

A todos meus colegas de pós-graduação, pelas risadas e companheirismo.

Agradeço em especial aos amigos Ana Mayra, Andresa, Douglas, Elizabeth, Elton Araújo, Elton Pereira, Fernanda, Guilherme, Júlio, Leticia, Liliani, Luis Angeloni, Luiz Trench, Mariana, Michele, Paola, Paula Sanches, Raphael, Renata, Simone Sartorio, Thatiane, Thiago, Vanderly e Wilson.

Aos amigos de graduação Darcio, Diógenes, Iguer, Liard, Liliani, Vitor e Wilton.

Aos companheiros de república Celso, Gabriel e Henrique.

A graduanda Patrícia Padilha dos Anjos por se dispor a procurar erros de gramática, e, principalmente, pelos sempre presentes humor e atenção.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq pelo apoio financeiro em forma de bolsa de estudos.

Um agradecimento especial a toda minha família, que com seu apoio incondicional, sempre me deu forças para seguir meu objetivo.

E, finalmente, a todas as pessoas que contribuíram de alguma forma na minha formação e realização deste trabalho.

SUMÁRIO

RESUMO	9
ABSTRACT	11
LISTA DE FIGURAS	13
LISTA DE TABELAS	15
1 INTRODUÇÃO	19
2 REVISÃO DE LITERATURA	21
2.1 Análise de grupos de experimentos	21
2.2 Delineamento em blocos ao acaso	27
2.3 Teste de Tukey	28
2.4 Efeitos fixos e aleatórios de um modelo matemático	29
2.5 Componentes de variância	31
3 METODOLOGIA	35
3.1 Materiais	35
3.1.1 Banco de dados reais	35
3.1.1.1 O milho	35
3.1.1.2 A produção de massa verde - PMV	36
3.1.2 Banco de dados simulados	37
3.2 Métodos	37
3.2.1 Análise de grupos de experimentos reais	37
3.2.1.1 Análise de variância de experimentos em blocos ao acaso - ANOVA individual	39
3.2.1.1.1 Estimadores calculados segundo a análise individual	41
3.2.1.1.2 Análise de variância de grupos de experimentos - ANOVA conjunta	42
3.2.1.1.2.1 Estimadores calculados pós desdobramento da interação significativa $T \times L$	45
3.2.1.1.3 Estimativas de variâncias de estimativas de componentes de variância	47
3.2.2 Análise de grupos de experimentos simulados	49
4 RESULTADOS E DISCUSSÕES	53
4.1 Abordagem do banco de dados na variável produção de massa verde - PMV	53
4.2 Simulação de conjuntos de dados de grupos de experimentos	68
5 CONCLUSÕES	77

REFERÊNCIAS	79
ANEXOS	83

RESUMO

Comparação de rols classificatórios de tratamentos e de estimativas de componentes de variância em grupos de experimentos

As análises de grupos de experimentos, de grande importância em melhoramento genético, são indispensáveis quando se pretende investigar o comportamento de alguns tratamentos em diversos locais de interesse do pesquisador. Nestes casos, parte-se das análises de variância individuais em cada local, para o agrupamento de todos os ensaios em uma única análise. Verifica-se então a veracidade da significância da interação tratamentos versus locais - $T \times L$, sendo esta não-significativa, pode-se obter conclusões generalizadas a respeito do comportamento dos tratamentos. No entanto, o grande interesse está nos casos de interação significativa, em que dois caminhos de destaque surgem para que se conclua a análise, o primeiro, permite que se considerem os resultados e conclusões das análises individuais, com o resíduo específico de cada local, enquanto o segundo aconselha que se desdobrem os graus de liberdade relativos a tratamentos + interação significativa, visando a interpretação dos tratamentos em estudo dentro de cada um dos locais, utilizando o resíduo médio como testador. Partindo do fato de que componentes de variância são variâncias associadas aos efeitos aleatórios de um modelo matemático, que permitem quantificar a variabilidade de tais efeitos, tem-se por objetivo neste trabalho, em grupos de experimentos reais com interação $T \times L$ significativa, comparar os componentes de variância obtidos nas análises individuais utilizando os quadrados médios residuais - QMRes de cada ensaio versus os obtidos pós-desdobramento da interação em questão utilizando o quadrado médio do resíduo médio - QMRM. Tal confronto será fundamentado nas estimativas de variâncias das estimativas destes componentes. Finalmente, em grupos de ensaios reais e simulados, o objetivo será voltado para a comparação de rols classificatórios de tratamentos nas análises individuais versus os rols classificatórios de tratamentos obtidos pós-desdobramento da interação em questão. A montagem destes rols será possível a partir do uso do teste de Tukey, ao nível de 5% de significância, para os cálculos das diferenças mínimas significativas - dms ora com resíduos de análises individuais, ora de conjunta. Todos os cálculos deste trabalho serão realizados no software estatístico R.

Palavras-chave: Melhoramento genético; Análise de grupos de experimentos; Interação significativa; Componentes de variância; Quadrado médio do resíduo; Software R

ABSTRACT

Comparison of treatments classificatory rankings and of variance components estimates in experimental groups

The experimental groups analysis, of great importance in genetic improvement, are essential when intends to investigate the treatments behaviour in many places from researcher interest. In these cases, starts by the individual variance analysis in each place, to the grouping of all experiments in a single analysis. Examine the truth of the significant treatments vs. places interaction - $T \times L$, being this no-significant, is possible to obtain generalized conclusions about the treatments behaviour. However, the interest is in the cases when significant interaction is found, because two eminence ways appear for the analysis conclusions, the first one allow that the individual analysis results and conclusions be considered, with the specific residue from each place, while the second one advise, that the degrees of freedom relative to treatments + significant interaction be unfound, looking at the interpretation of the treatments in study inside each place, using the mean residue how testator. Starting with the fact that variance components are variances associated to the aleatory effects of a mathematical model, that allow the quantifying of such effects, this work objective, in real experimental groups, with significant interaction $T \times L$, is to compare the variance components obtained in individual analysis using the residual mean square - $QMRes$ from each experiment against the obtained after unfolding the interaction in question using the mean residual mean square - $QMRM$. This confrontation will be based in variance estimations of these components estimations. Finally, in real and simulate experimental groups, the objective will be directed to the comparison of treatments classificatory rankings in individual analysis vs. the treatments classificatory rankings obtained after unfolding of the interaction in question. The construction of these rankings will be possible using the Tukey test, with 5% of significance, for the calculation of the significant minimum differences - dms , a time with individual analysis residual, othertime, conjunct. All the calculations from this work will be realized in the R statistical software.

Keywords: Genetic improvement; Experimental groups analysis; Significant interaction; Variance components; Residual mean square; R software

LISTA DE FIGURAS

Figura 1 - (a) Histograma, (b) Q-Q Plot e (c) Box-plot para a variável transformada produção de massa verde - $\ln(\text{PMV})$	54
Figura 2 - Histogramas para a variável transformada produção de massa verde - $\ln(\text{PMV})$ nos municípios de (a) Assis, (b) Mococa, (c) Pindamonhangaba e (d) Votuporanga	54
Figura 3 - Box-plot para a variável transformada produção de massa verde - $\ln(\text{PMV})$ nos municípios de (a) Assis, (b) Mococa, (c) Pindamonhangaba e (d) Votuporanga	55
Figura 4 - Componentes de variância para a variável $\ln(\text{PMV})$ relativos a cultivares - CVC, blocos - CVB e resíduos - CVR, associados a suas respectivas estimativas de variâncias obtidas segundo os métodos de análise individual - AI e pós-desdobramento - APD	65
Figura 5 - Componentes de variância para a variável $\ln(\text{PMV})$ relativos a Assis - CVAS, Mococa - CVMO, Pindamonhangaba - CVPI e Votuporanga - CVVO, associados a suas respectivas estimativas de variâncias obtidas segundo os métodos de análise individual - AI e pós-desdobramento - APD	66

LISTA DE TABELAS

Tabela 1 - Análise de variância de um experimento em blocos ao acaso, com as esperanças dos quadrados médios - E(QM)	40
Tabela 2 - Análise de variância conjunta de um grupo de experimentos em blocos ao acaso, com as esperanças dos quadrados médios - E(QM)	43
Tabela 3 - Análise de variância para o desdobramento da interação significativa T×L de um grupo de experimentos em blocos ao acaso, com as esperanças dos quadrados médios E(QM)	46
Tabela 4 - Testes de normalidade e homogeneidade de variâncias para a variável transformada produção de massa verde - ln(PMV), vista de forma geral, e em cada um dos locais: Assis - <i>AS</i> , Mococa - <i>MO</i> , Pindamonhangaba - <i>PI</i> e Votuporanga - <i>VO</i>	53
Tabela 5 - Análise descritiva da variável resposta original produção de massa verde - PMV, e normalizada - ln(PMV)	55
Tabela 6 - Análises de variância individuais na variável ln(PMV) em cada um dos municípios: Assis - <i>AS</i> , Mococa - <i>MO</i> , Pindamonhangaba - <i>PI</i> e Votuporanga - <i>VO</i>	56
Tabela 7 - Análise de variância conjunta na variável resposta transformada - ln(PMV) nos municípios de Assis, Mococa, Pindamonhangaba e Votuporanga	57
Tabela 8 - Análise de variância sobre a variável ln(PMV), com desdobramento dos graus de liberdade referentes a tratamentos + interação T×L dentro de cada município: Assis - <i>AS</i> , Mococa - <i>MO</i> , Pindamonhangaba - <i>PI</i> e Votuporanga <i>VO</i>	57
Tabela 9 - Logaritmo natural (ln) dos valores médios de produção de massa verde - PMV (kg/ha) avaliados em 16 cultivares nos municípios de Assis - <i>AS</i> , Mococa - <i>MO</i> , Pindamonhangaba - <i>PI</i> e Votuporanga - <i>VO</i> , e de forma geral - <i>GERAL</i>	58
Tabela 10 - Teste de Tukey para comparação dos cultivares segundo o logaritmo natural de suas respectivas médias de produção de massa verde - ln(PMV) nos município de Assis, Mococa, Pindamonhangaba e Votuporanga, a partir do rol classificatório das análises individual - <i>AI</i> versus pós-desdobramento - <i>APD</i>	59

Tabela 11 - Estimativas dos componentes de variância obtidas segundo análises individuais, suas respectivas variâncias, estimativas obtidas segundo a análise pós-desdobramento e suas respectivas variâncias na variável produção de massa verde nos municípios de Assis, Mococa, Pindamonhangaba e Votuporanga . 62

Tabela 12 - Avaliação das metodologias AI e APD na obtenção de menores estimativas de variâncias associadas às estimativas de componentes de variância segundo proporções da metodologia de menor pela de maior variância na variável produção de massa verde nas cidades de Assis, Mococa, Pindamonhangaba e Votuporanga 64

Tabela 13 - Média dos QMRM - \overline{QMRRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 3 ensaios cada, considerando-se $\sigma_{TRA}^2 = 4$, $\sigma_{BLO}^2 = 1$ e $\sigma_{RES}^2 = 1$ para os efeitos dos fatores de variação 69

Tabela 14 - Média dos QMRM - \overline{QMRRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 4 ensaios cada, considerando-se $\sigma_{TRA}^2 = 4$, $\sigma_{BLO}^2 = 1$ e $\sigma_{RES}^2 = 1$ para os efeitos dos fatores de variação 70

Tabela 15 - Média dos QMRM - \overline{QMRRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 3 ensaios cada, considerando-se $\sigma_{TRA}^2 = 64$, $\sigma_{BLO}^2 = 16$ e $\sigma_{RES}^2 = 36$ para os efeitos dos fatores de variação 71

- Tabela 16 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 4 ensaios cada, considerando-se $\sigma_{TRA}^2 = 64$, $\sigma_{BLO}^2 = 16$ e $\sigma_{RES}^2 = 36$ para os efeitos dos fatores de variação 72
- Tabela 17 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 3 ensaios cada, considerando-se $\sigma_{TRA}^2 = 2500$, $\sigma_{BLO}^2 = 225$ e $\sigma_{RES}^2 = 900$ para os efeitos dos fatores de variação 73
- Tabela 18 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 4 ensaios cada, considerando-se $\sigma_{TRA}^2 = 2500$, $\sigma_{BLO}^2 = 225$ e $\sigma_{RES}^2 = 900$ para os efeitos dos fatores de variação 74

1 INTRODUÇÃO

Em melhoramento genético há grande interesse na análise de grupos de experimentos, que se trata de uma avaliação conjunta capaz de permitir tomar decisões abrangentes a toda uma região a partir da instalação de um mesmo experimento em diversos locais de interesse do pesquisador. Para isso, utilizam-se dos mesmos tratamentos e delineamento, de preferência com iguais números de repetições em todos os experimentos, facilitando a análise.

Nesses casos, considera-se desnecessária a utilização de ensaios cuja instalação ou análise estatística seja complexa, podendo-se trabalhar com experimentos inteiramente casualizados ou blocos ao acaso, simplificando os cálculos e reduzindo as despesas, e ainda pode-se repetir o experimento ao longo de alguns anos, aumentando assim a credibilidade dos resultados.

A análise de grupos de experimentos exige ainda a homogeneidade de variâncias, que segundo BOX (1954) será satisfeita se o maior quadrado médio residual - $QMRes$ dentre todos experimentos (locais) vistos separadamente seja no máximo 4 vezes o menor $QMRes$.

A interação tratamentos versus locais é o ponto chave da análise de grupos de experimentos, sendo ela significativa, há indícios de que o comportamento dos tratamentos difere de local para local. Por outro lado, sendo a interação não-significativa, pode-se obter conclusões generalizadas.

Ao se deparar com uma interação significativa, dois caminhos são citados na literatura para que se conclua a análise. O primeiro permite que se considerem os resultados e conclusões das análises individuais, com o resíduo específico de cada local, enquanto o segundo aconselha que se desdobrem os graus de liberdade relativos a tratamentos + interação significativa, visando a interpretação dos tratamentos dentro de cada um dos locais, usando o resíduo médio como testador.

Parte-se do princípio de que componentes de variância são variâncias associadas aos efeitos aleatórios de um modelo matemático, que nos possibilita quantificar a variabilidade de tais efeitos.

Dentro do contexto análise de grupos de experimentos com interação significativa tratamentos versus locais, o presente trabalho tem por objetivos comparar os rols de tratamentos ordenados segundo suas médias, obtidos nas análises individuais - AI versus os provenientes da análise pós-desdobramento - APD, além da meta de comparar as estimativas de componentes de variância obtidas nestas duas análises, em um grupo de experimentos reais.

Têm-se ainda por objetivo, a simulação de conjuntos de dados para que se possa comparar os comportamentos das diferenças mínimas significativas - dms utilizadas no teste de Tukey para comparação de tratamentos, obtidas para análises individuais - AI versus pós-desdobramento - APD da interação, buscando apontar, se possível, dentre essas duas metodologias, a mais sensível em determinada situação.

Tais comparações serão realizadas embasadas em um grupo de experimentos reais de produção de massa verde - PMV de milho, sob o delineamento em blocos ao acaso, e, em grupos de experimentos simulados, no delineamento citado.

2 REVISÃO DE LITERATURA

2.1 Análise de grupos de experimentos

As análises de grupos de experimentos foram estudadas por diversos autores, dentre eles Kempthorne (1952), Cochran & Cox (1957), Campos (1984), Pimentel Gomes (2000) e Barbin (2003) e são de grande interesse, em especial, para os melhoristas genéticos. As estimativas de componentes de variância em experimentos conduzidos em um único ambiente costumam ser superestimadas, pois o fator ambiente costuma influenciar nesses casos. Desta forma, autores como Ramalho (1977) e Regazzi et al. (1999) destacam como fundamental o estudo do componente da interação genótipo versus ambiente.

Fisher (1918) no início do século vinte já trabalhava com o emprego de análise de variância no estudo dos caracteres quantitativos segundo Ramalho et al. (1993).

Bliss (1944) exibiu exemplos interessantes de um conjunto de análises biológicas em um grupo de 16 laboratórios, estimando a potência na preparação de plantas medicinais, já com a idéia de análise de grupos de experimentos.

Pimentel Gomes & Guimarães (1958) destacam-se entre os pioneiros em estudos aprofundados com grupos de experimentos, com o primeiro trabalho referente a análise conjunta de experimentos em blocos completos casualizados com tratamentos comuns, onde foram calculadas estimativas de variâncias de estimativas de contrastes de médias de tratamentos.

Pimentel Gomes (1970) apresenta a análise conjunta de um grupo de experimentos em blocos casualizados com alguns tratamentos comuns, em que o número de repetições para tratamentos varia de um experimento para outro. Mostra, também, o esquema de análise de variância, e que as somas de quadrados são obtidas de modo semelhante ao trabalho desenvolvido por Pimentel Gomes & Guimarães (1958) (ALVES, 1995).

Segundo Barbin (2003), busca-se nesse tipo de estudo, tomar sempre que possível, decisões abrangentes a uma grande área caso se trabalhe com um mesmo experimento em diversas áreas, ou a um grande período de tempo caso se trabalhe com um mesmo experimento ao longo do tempo, ou ambos, apontando mesmo ou similar comportamento por parte dos tratamentos em questão, sempre que a interação se apresentar não-significativa. Caso contrário, esse tipo de conclusão não pode ser tomada.

Regazzi et al. (1999) afirmam que o processo tradicional de investigação das interações genótipo versus ambiente se trata da análise de variância conjunta, e ainda, por meio desta, a magnitude das interações é investigada através da variância dos efeitos de genótipos versus locais, versus anos, versus anos versus locais, e outros segundo o interesse do melhorista.

A interação genótipo x ambiente é definida como sendo o comportamento diferencial de genótipos em diferentes ambientes. Isto significa que os efeitos genéticos e ambientais não são independentes, uma vez que as respostas fenotípicas dos genótipos podem diferir com as variações ambientais. (SOUZA JÚNIOR & VENCOVSKY, 1989).

Segundo Cecon (1992), a influência que o ambiente exerce sobre o genótipo, propiciando a ocorrência de interação genótipo versus ambiente, é um dos maiores problemas que os melhoristas enfrentam na seleção de genótipos, pois a variância genética é afetada pelo ambiente específico em que é realizada a seleção, e o material melhorado geralmente é distribuído para ambientes diversos. Tal interação pode ser detectada estatisticamente a partir da análise conjunta de experimentos realizados em mais de um local, em que os tratamentos sejam constituídos de espécies, cultivares, procedências ou progênies.

Cochran & Cox (1957), Campos (1984) e Pimentel Gomes (2000) propuseram a análise de um conjunto de experimentos em blocos ao acaso com v tratamentos, r repetições conduzido em l locais.

Banzatto & Kronka (1989) alertam que os ensaios individuais não precisam ter obrigatoriamente um número muito grande de parcelas.

Cochran & Cox (1957), ao estudarem as repetições no espaço e no tempo introduziram apenas o caso mais simples de ensaios em cada um dos p locais por y anos, tomando em cada local ensaios independentes (PIEIDADE, 1987).

Ainda segundo Cochran & Cox (1957), a efetividade dos nutrientes nas plantas, de diferentes variedades de uma safra, e diferentes práticas de plantio geralmente variam de campo para campo, e com mais intensidade, de temporada para temporada. Além disso, um ensaio bem conduzido supre informação sobre apenas um local e uma temporada, assim, na análise de grupos de experimentos, os ensaios são instalados em uma série de locais distintos, em áreas que se deseja obter informações, e são repetidos por uma determinada quantidade de temporadas.

Pode-se também, segundo esses autores, fazer uso da análise de grupos de experimentos quando não se tem interesse em fazer inferência sobre alguma população específica, e sim, em estudar a influência de fatores externos sobre as respostas de algum experimento, ou seja, confiando no fato de que os tratamentos devem apresentar mesmo comportamento ao longo de toda a área de estudo, faz-se então, tal tipo de análise, para investigar a possível interferência do calor ou da umidade por exemplo nesses tratamentos. Neste caso, repetições de tais experimentos em diferentes locais são necessárias, para que haja variações nos fatores externos de interesse.

Banzatto & Kronka (1989) ressaltam que as análises conjuntas podem ser realizadas com os dados originais, os totais de tratamentos ou ainda as médias de cada tratamento em cada experimento.

Kempthorne (1952) exhibe um método para análise de um grupo de experimentos instalados em diversos locais, levando em conta a informação de cada parcela dos ensaios ao invés do total ou das médias de tratamentos em cada local, esta escolha é a mais comum entre os pesquisadores.

Segundo Yates & Cochran (1938), a análise estatística apropriada para os dados de uma série de experimentos depende do objetivo da pesquisa.

Campos (1984) destaca que não é raro a ocorrência de instalação de grupos de experimentos, utilizando para tal propósito, locais distintos, visando maior abrangência das conclusões obtidas.

Esse autor afirma ainda, que alguns critérios devem ser obedecidos no agrupamento dos ensaios, dentre os quais:

- por setores geográficos
- por ano agrícola
- por afinidade quanto a alguma característica de interesse
- pela grandeza dos quadrados médios dos resíduos das análises individuais.

Pimentel Gomes (2000) sugere que os experimentos sejam preferivelmente de delineamento simples, numerosos e baratos, para que o custo não seja excessivo.

Barbin (2003) destaca ainda que se deve, sempre que praticável, obedecer o mesmo delineamento em todos os locais, incluir os mesmos tratamentos e, se possível, com iguais números de repetições, o que facilitará a análise de variância conjunta.

Segundo Vencovsky & Barriga (1992), um fator que costuma complicar esse tipo de análise, é o caso de se dispor de ensaios com número diferente de blocos ou repetições, para contornar tal problema, deve-se procurar os processos biométricos adequados expostos no trabalho de Cochran & Cox (1957).

Cecon (1992) alerta que em grupos de experimentos, inicia-se fazendo a análise e interpretação dos dados de cada ensaio, para depois verificar se as diferenças entre tratamentos são consistentes em todos os locais, isto é, se eles variam da mesma forma de local para local, para finalmente realizar a estimação e comparação dos efeitos médios de tratamentos sobre o conjunto de experimentos.

Diversos trabalhos dentre os quais Box (1954), Cochran & Cox (1957), Pimentel Gomes (2000) e Barbin (2003) destacam que a homogeneidade de variâncias deve ser satisfeita a fim de que os experimentos possam ser reunidos em uma única análise, porém não existe um consenso a respeito de quando aceitar tal homogeneidade.

Alguns autores como Cochran & Cox (1957), costumavam testar a homogeneidade de variâncias através do teste de Bartlett (1937), porém, Pimentel Gomes (2000) apontou este teste como muito sensível à falta de normalidade dos dados, devendo então ser preterido.

Segundo Cochran (1954), pode-se utilizar todos os locais em uma única análise, independente de haver ou não homogeneidade de variâncias, adotando n_r graus de liberdade para o resíduo médio da análise conjunta. Em busca de contornar a falta de homogeneidade de variâncias, o autor propôs um ajuste no número de graus de liberdade do resíduo, segundo a equação 1 porém esse método não se revelou promissor para Pimentel Gomes (2000) nos casos de grandes discrepâncias entre os quadrados médios residuais.

$$n_r = \frac{[QMRes(L_1) + \dots + QMRes(L_k)]^2}{\frac{[QMRes(L_1)]^2}{n_1} + \dots + \frac{[QMRes(L_k)]^2}{n_k}}. \quad (1)$$

Sendo L_k o local k , para $k = 1, \dots, K$, e n_k , o número de graus de liberdade do resíduo associado a L_k .

Partindo do princípio de que em todos os experimentos os tratamentos possuem o mesmo número de parcelas (repetições ou blocos), estudos realizados por Box (1954) revelam que, dessa forma, se o maior QMRes dentre todos os ensaios analisados, individualmente, for no máximo 4 vezes o menor QMRes, não há problemas em se realizar a análise conjunta, agrupando os locais em um único estudo.

Pimentel Gomes (2000) por sua vez, afirma que se o quociente entre o maior e o menor QMRes for menor que 7, quase sempre a análise conjunta poderá ser realizada sem grandes problemas.

Pimentel Gomes (2000) salientou ainda, que quando esta relação entre quadrados médios residuais das análises individuais for além do estipulado, convém considerar separadamente subgrupos de experimentos com quadrados médios residuais não muito heterogêneos para construir as análises conjuntas (nesse caso, cada subgrupo geraria uma análise).

Barbin (2003) reforça tal ideia, afirmando que nos casos em que esta homogeneidade não é satisfeita, pode-se eliminar o experimento cujo QMRes seja discrepante, ou ainda, agrupar os experimentos que possuam QMRes similares, levando em consideração os critérios enumerados por Campos (1984) citados nesse trabalho.

A detecção da significância para a interação não esclarece, contudo, as implicações que esta possa ter sobre o melhoramento, de forma que estudos de detalhamento deste componente de variação devem também ser realizados (VENCOVSKY & GERALDI, 1977).

Eberhart & Russel (1966) destacam, que em grupos de experimentos, as interações estão normalmente presentes para qualquer material genético com o qual o melhorista possa vir a trabalhar, sugerindo então, que um programa de melhoramento poderia ser implantado para o desenvolvimento de variedades restritas a determinados ambientes, porém, alertam que isto levaria a grandes despesas financeiras e de material genético.

Kempthorne (1952) considera que na análise conjunta, existem duas possíveis dificuldades para a interpretação dos resultados na análise de variância. A primeira dificuldade é que a variância residual (σ^2) não é constante entre os experimentos, e a segunda, é que o componente de variância referente à interação de tratamentos com locais (σ_{ti}^2) depende da combinação de tratamentos e locais (OLIVEIRA & NOGUEIRA, 2007).

Sendo a interação tratamentos versus locais significativa, não se pode tirar conclusões gerais a toda área em estudo, o comportamento dos tratamentos difere de local para local, pode-se então, simplesmente, aceitar os resultados das análises individuais, porém, segundo Barbin (2003), vale observar, se é também significativo o efeito de tratamentos na análise conjunta, pois essa combinação de significâncias pode indicar um destaque de um ou mais tratamentos em relação aos demais em estudo.

Outro caminho ao deparar-se com a interação em questão significativa, é, segundo Barbin (2003), desdobrar os graus de liberdade de tal interação, estudando-se os tratamentos dentro de locais, com a vantagem, sobre as análises individuais, além de se trabalhar com um único QMRes, pode-se contar com maior número de graus de liberdade para o resíduo, sendo uma análise mais sensível, que a grosso modo possibilita detectar diferenças significativas que com um número de graus de liberdade residual menor não se obteria.

Para se verificar uma possível significância desejada, admite-se:

- H_0 : hipóteses nula - não existência de diferença significativa entre tratamentos, não existência de interação tratamentos x locais,...

Sendo F_{calc} o valor de F calculado de acordo com o quadro da análise de variância na causa de variação desejada (tratamentos, tratamentos versus locais, ...), e F_{tab} o valor observado em uma tabela estatística F com n_1 graus de liberdade referentes a causa de variação em questão, e n_2 graus de liberdade referentes ao denominador utilizado na obtenção de F_{calc} (QM Res, QMT×L, ...), tem-se, segundo Banzatto & Kronka (1989):

- Se $F_{calc} > F_{tab}$, rejeita-se H_0 ;

- Se $F_{calc} \leq F_{tab}$, aceita-se H_0 .

Assim, a consulta da tabela estatística F para tratamentos, no caso de desdobramento, fornecerá um valor menor em relação ao das análises individuais, valor este a ser comparado ao valor de F calculado segundo as expressões do quadro de análise de variância, para verificação de existência ou não de significância no teste.

Costuma-se tomar $\alpha = 0,05$ como nível de significância para H_0 , sendo que o valor da probabilidade do teste F menor ou igual a α , admite-se diferenças significativas entre os fatores em estudo, logo, rejeita-se H_0 . A interpretação da interação deve ser anterior a dos fatores vistos de forma individual, pois convida o pesquisador a fazer um desdobramento.

Tanto nas análises individuais, quanto na pós-desdobramento, grande interesse existe em relação à comparação de tratamentos, quando este efeito se apresenta significativo segundo o teste F, e um critério de comparação é o teste de Tukey, citado por, entre outros, Banzatto & Kronka (1989), que testa todos os possíveis contrastes entre duas médias.

2.2 Delineamento em blocos ao acaso

Segundo Vieira (2006), os experimentos em blocos ao acaso surgiram na área agrícola, com o termo bloco criado para designar uma faixa de terra de fertilidade pouco variável, de características similares, que recebia um “bloco” de unidades onde se manejava tratamentos diferentes. Diversos trabalhos utilizam tal delineamento até os dias de hoje, já que é de instalação e análise estatística não muito complexas.

Vieira (2006) ressalta ainda, que em ensaios em blocos ao acaso, a casualização não é completa já que primeiramente é preciso organizar os blocos, o que é restrito, para depois sortear-se os tratamentos dentro de cada bloco.

Banzatto & Kronka (1989) destacam que nesse tipo de delineamento, os três princípios básicos da experimentação são satisfeitos: o “controle local”, pois os blocos funcionam com tal finalidade, a “repetição”, sendo que o número de parcelas por bloco deve ser um múltiplo do número de tratamentos (geralmente igual ao número de tratamentos, sendo nesse caso bloco sinônimo de repetição), e a “casualização”, já que os tratamentos são distribuídos entre as parcelas dentro do bloco de forma aleatória.

Os mesmos autores frisam ainda, dentre as principais vantagens desse tipo de delineamento: os blocos podem controlar diferenças ambientais; o delineamento permite uma boa variação sobre o número de tratamentos e blocos; melhora a estimativa da variância residual pois a variação ambiental entre blocos é isolada; análise de variância simples. Como principal desvantagem em relação a outros delineamentos, se tem a redução do número de graus de liberdade do resíduo, já que parte desse número é destinada à blocos.

A análise de variância de experimentos em blocos ao acaso sem repetições (número de parcelas por bloco igual ao número de tratamentos) é feita da mesma forma (seção 3.2.1.1), descrita no trabalho de Barbin (2003), quer os efeitos sejam fixos, quer sejam aleatórios. As hipóteses em teste são, porém, diferentes.

A princípio, tomam-se I tratamentos, sendo o efeito de tratamentos fixos, a hipótese a ser testada é a de que as médias destes são iguais:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_I. \quad (2)$$

De outro lado, sendo o efeito de tratamentos aleatório, a hipótese a ser testada é a de que a variabilidade dos tratamentos é a mesma:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2. \quad (3)$$

Rejeitando-se a hipótese H_0 da equação 2 segundo a consulta de uma estatística F, o teste de Tukey é uma boa alternativa para verificar quais tratamentos se destacam em relação aos demais.

2.3 Teste de Tukey

O teste de Tukey é um dos principais no quesito comparação de médias de tratamentos, sendo utilizado em diversas pesquisas no ramo experimental, descrito em diversos trabalhos e implementado em softwares estatísticos, como o R por exemplo.

Face a um teste F significativo para tratamentos, é inegável a existência de pelo menos um contraste entre duas médias diferindo de zero, já que tal situação acarreta na rejeição de $H_0 : t_1 = t_2 = \dots = t_I$, considerando neste caso I tratamentos.

De modo geral, este teste consiste no cálculo de uma diferença mínima significativa - dms, diferença essa utilizada na comparação de cada um dos contrastes de médias de tratamentos visando ordenar os tratamentos segundo suas médias.

$$dms = q\sqrt{\frac{QMRes}{J}}. \quad (4)$$

De posse de $I = 2$ tratamentos, pode-se montar 1 contraste para comparar suas médias, ou seja, $c_2 = 1$. Assim, tomados I tratamentos, o número de contrastes possíveis - c_I para os comparar dois a dois é dado pela fórmula $c_I = c_{I-1} + (I - 1)$, para $I \geq 3$.

Cada contraste entre duas médias, deve ser calculado em módulo, por exemplo, $|m_r - m_s|$, para r e s variando de 1 a I , com $r \neq s$. Em seguida, compara-se cada um desses módulos a dms, sendo a dms calculada segundo a equação 4, única para todos os contrastes.

Na equação acima, q é a amplitude total estudentizada obtida em tabelas específicas, geralmente aos níveis de 5 e 1% de probabilidade, com n = número de tratamentos e n' = número de graus de liberdade do resíduo. E ainda, J é o número de blocos.

Vale lembrar que a equação 4 não é apropriada no caso de delineamentos desbalanceados, outra fórmula seria utilizada, sendo o teste de Tukey aproximado em tal caso.

2.4 Efeitos fixos e aleatórios de um modelo matemático

Segundo Vencovsky & Barriga (1992), um modelo matemático é dito fixo caso todas as causas de variação em questão, excedendo o erro (que é sempre tomado aleatório), sejam fixos. Um modelo é considerado aleatório, caso todos os efeitos em estudo, excedendo a média que é sempre fixa, sejam aleatórios. Enfim, um modelo é misto, caso além da média que é fixa, e do erro que é aleatório, existam ao menos mais um efeito efeito fixo e mais um aleatório. Alguns cuidados devem ser tomados na decisão entre um efeito ser fixo ou aleatório.

Barbin (1993) considera efeitos fixos, aqueles que dentro do modelo representam apenas os fatores em estudo. Os resultados obtidos não podem ser estendidos para outros fatores semelhantes. Já os efeitos aleatórios são aqueles que entram no modelo através de uma amostra retirada de uma população, podendo então representar essa população.

Se um nutricionista quiser comparar a qualidade do pão fabricado por cinco padarias que escolheu, estará fazendo um ensaio com tratamentos de efeitos fixos. No entanto, se sortear cinco padarias de uma lista telefônica para comparar a variabilidade de qualidade do pão fabricado por elas, estará fazendo um ensaio em que os efeitos de tratamentos são aleatórios (VIEIRA, 2006).

A autora destaca ainda, que segundo a literatura, são mais encontrados ensaios com tratamentos de efeitos fixos, o que é adotado por muitos estatísticos. As diferenças nos resultados das análises são pequenas, desde que o número de tratamentos que podem ser submetidos a sorteio seja pequeno. Finaliza ressaltando a importância em optar por tratamentos de efeitos “fixos” ou “aleatórios” em função da forma como o experimento foi planejado.

Pimentel Gomes (1968) sobre o fato de tomarmos o efeito de tratamentos como fixo, exemplifica com o caso de os tratamentos serem três variedades de milho em estudo, cuja produção queremos avaliar. Por outro lado, em se tratando de três linhagens ou híbridos, escolhidos entre muitos outros criados ou que se possam criar, e dos quais representam uma amostra, então o efeito de tratamentos nesse caso é considerado aleatório com média 0 e variância σ_t^2 .

De acordo com Vencovsky & Barriga (1992), se tivermos em análise conjunta, ensaios em diferentes anos, o efeito de anos será considerado fixo se estes forem atípicos, perante as condições prevaescentes. Em caso contrário, poderemos considerar como aleatórios esses efeitos de anos.

Barbin (1993) dentre outros autores costumam tomar efeitos de blocos e locais como aleatórios, já que blocos tem por característica a representação de todo um local, enquanto que locais em análise de grupos de experimentos tem por característica a representação de toda uma região.

O trabalho de Pimentel Gomes (1968) destaca que, nos casos mais comuns da experimentação agrícola, com adubos, variedades, inseticidas, fungicidas, etc, o modelo usado e apropriado é o modelo fixo, porém em alguns casos especiais como na análise de grupos de experimentos e nos ensaios relativos a melhoramentos de plantas e de animais.

Vencovsky & Barriga (1992) afirmam que a decisão a respeito de um efeito ser fixo ou aleatório não é fruto da escolha do estatístico, e sim consequência da própria natureza da pesquisa. Deve-se levar em conta, por exemplo, se um material genético a ser estudado permite que se tire conclusões abrangentes ou específicas.

Segundo esses autores, quando o efeito de genótipos (tratamentos) é fixo, geralmente compara-se as médias, por meio de testes apropriados, como Tukey por exemplo. Diante de um modelo aleatório, é interessante estimar os componentes de variância de genótipos, de locais e da interação genótipos \times locais, sendo que a magnitude de cada um deles é extremamente importante para o melhorista, pois, o componente de genótipos, por exemplo, serve para obtenção do coeficiente de herdabilidade e parâmetros afim.

2.5 Componentes de variância

Segundo Barbin (1993), componentes de variância são variâncias associadas aos efeitos aleatórios de um modelo matemático.

Rao (1997) salienta, que os modelos de componentes de variância que possuem efeitos genéticos e ambientais são utilizados para a análise de experimentos conduzidos em laboratórios e casas de vegetação visando estudar mudanças evolucionárias de diferentes tratamentos de uma geração para outra. Além disso, na área agrícola, onde experimentos são conduzidos em busca de melhorias nas variedades dos alimentos, fertilizantes e métodos de cultivo, os componentes de variância são utilizados para esse tipo de análise, além de permitirem a generalização dos resultados.

O autor vai mais longe, e afirma que componentes de variância são úteis além do campo genético, na análise experimental e observacional de dados educacionais, de saúde, psicológicos, industriais, além de ser vastamente aplicada em econometria na análise de dados provenientes de séries temporais e outros tipos de pesquisas.

Para Barbin (1993), a análise de variância tem como um dos principais objetivos, a obtenção da estatística F, porém, a partir dessa análise, obtêm-se também as estimativas dos componentes de variância, que é de grande interesse no melhoramento genético, vegetal ou animal. Ressalta ainda, que a análise de variância é sempre feita considerando-se o modelo como fixo.

O método da análise de variância ou dos momentos, merece destaque devido ao grande uso na estimação de componentes de variância, principalmente em modelos aleatórios balanceados. Tal método, descrito por Valério Filho (1991), consiste, a partir do modelo matemático escolhido, em se obter as esperanças matemáticas dos quadrados médios dos fatores de variação de uma análise de variância, cada uma dessas esperanças matemáticas (dada por uma expressão envolvendo componentes de variância) é igualada ao respectivo quadrado médio, possibilitando o cálculo de cada um dos estimadores dos componentes de variância do modelo.

O autor ressalta ainda, que em caso de delineamento balanceado, aquele em que todos os tratamentos possuem o mesmo número de repetições, o método da análise de variância produz estimadores com propriedades estatísticas desejáveis.

Barbin (2003) define a expressão da variância - s^2 de uma população na qual não se conhece a média verdadeira, como segue na equação 5, sendo n o número de observações, x_i o valor da i -ésima observação e \bar{x} a estimativa da média.

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}. \quad (5)$$

Espera-se que uma estimativa de componente de variância seja sempre um valor maior ou igual a zero, pois, segundo a equação 5, o cálculo de uma variância vem de um somatório de quadrados, ou seja, fatores não-negativos, porém, na prática isso pode acabar não acontecendo, um motivo é quando se depara com um valor muito alto para o QMRes, já que este, pelo método da análise de variância, aparece na maioria das expressões dos estimadores de componentes de variância com sinal negativo.

Segundo Vencovsky & Barriga (1992), diante de um modelo aleatório, é interessante estimar os componentes de variância de genótipos, residual e das interações genótipos versus locais, genótipos versus anos e genótipos versus locais versus anos. A magnitude referente a cada uma dessas estimativas é informação crucial para o melhorista em relação à questão de estabilidade fenotípica dos cultivares, pois uma simples análise de variância conjunta não permite a estimação de parâmetros de estabilidade sem o auxílio dessas estimativas.

Regazzi et al. (1999) destacam que sendo o efeito de tratamentos fixo, costuma-se realizar testes de hipóteses utilizando combinações lineares dos mesmos, sendo este um problema relativamente simples. Porém, sendo o efeito de tratamentos aleatório, o interesse está na estimação dos componentes de variância que são de grande interesse no melhoramento genético vegetal, pois o método de melhoramento e a população a serem utilizados dependem do conhecimento de certos parâmetros genéticos cujas estimativas podem ser obtidas através dos componentes de variância, nesse caso o problema é mais complexo, em especial quando envolve análise conjunta de experimentos.

Vieira (2006) alerta que quando se trabalha com um modelo fixo, o objetivo é estimar e realizar testes de hipóteses sobre os efeitos fixos envolvidos. No caso de um modelo aleatório, objetiva-se a predição dos efeitos aleatórios e a estimação dos componentes de variância. Assim, em um modelo misto, os dois objetivos recém citados são relevantes.

Marcelino & Iemma (2000) afirmam que a experiência do pesquisador e a estrutura dos dados deve influenciar na escolha da estrutura mais apropriada nos sistemas computacionais estatísticos, já que métodos diferentes de estimação de componentes de variância costumam proporcionar estimativas diferentes.

Yates & Cochran (1938) salientam que estimativas de componentes de variância só podem ser associadas se provenientes de experimentos semelhantes.

Cox & Solomon (2003), ressaltam que a corrente de softwares para estimação de componentes de variância abrangendo S-PLUS, R, GenStat, SAS, SPSS e Stata, vem tendendo em direção à estimação de parâmetros de efeitos aleatórios no campo clássico de componentes de variância e covariância.

Existem diversos métodos de estimação de componentes de variância, dentre eles, o método da máxima verossimilhança - ML, o de Fisher (1918) e o de Hartley & Rao (1967), que por sua vez é bastante usado em modelos mistos apesar de ser um pouco tendencioso. Destaque também para os métodos I, II e III de Henderson (1953), e alguns provenientes de funções quadráticas, como pode-se citar os métodos do estimador quadrático não-viesado de norma mínima - MINQUE (Rao, 1971a), e o de variância mínima - MIVQUE (Rao, 1971b). Algumas dessas metodologias foram comparadas por Marcelino & Iemma (2000), no trabalho desenvolvido por estes, com modelos mistos desbalanceados.

Esses autores ressaltam a peculiaridade do método da máxima verossimilhança que é a vantagem de nunca fornecer estimativas negativas, porém, exige normalidade dos dados, o que algumas vezes não acontece em se tratando de dados reais.

Duarte (2000) acrescenta que em se tratando de modelos mistos, é fundamental a estimação dos componentes de variância, e afirma que avaliados os principais métodos de estimação - ANOVA, MIVQUE, ML e REML, através de simulação computacional, notou-se conflitos de resultados, principalmente em caso de delineamentos desbalanceados, devido a escassez de informações específicas.

Rao (1997), expõe os métodos de estimação ML, REML, MINQUE e MIVQUE, descrevendo, a partir de formas quadráticas e derivadas, as deduções realizadas para obtenção dos estimadores de componentes de variância desejados.

Merece destaque o método prático de Hicks (1973), aplicado em modelos balanceados para obtenção das esperanças dos quadrados médios dos fatores de variação do modelo em estudo, fornece os mesmos resultados do método da análise de variância com mais dinamismo, sem exigir grande quantidade de cálculos, e a partir desse, foi criado o método prático de Hicks modificado, citado por Barbin (1993).

Pelo método prático de Hicks (1973), as expressões de estimadores de componentes de variância apontam, que o teste F para tratamentos deve ser feito com o QMRes no denominador, quando o modelo é misto considerando o efeito de tratamentos fixo e o de locais aleatório. Do mesmo modo, o teste F para locais é praticado com o QMRes no denominador, se o modelo for misto com o efeito de tratamentos aleatório e o de locais fixo. Devido a essa falta de critério, modificou-se o método prático de Hicks (1973) conforme alerta Barbin (1993).

Assim, segundo o método prático de Hicks modificado, citado por Barbin (1993), quando se trabalha com grupos de experimentos, na estatística experimental, os testes F para tratamentos e locais são realizados com o quadrado médio da interação tratamentos versus locais - $QMT \times L$, conforme afirmam outros autores, dentre eles Pimentel Gomes (2000).

Segundo Banzatto & Kronka (1989), se o quadrado médio da interação em estudo não for muito discrepante do quadrado médio do resíduo, pode-se testar tanto tratamentos quanto locais versus o quadrado médio residual, proporcionando então uma maior precisão em virtude do maior número de graus de liberdade associado ao efeito residual.

Optar por modelo fixo, aleatório ou misto não influencia o cálculo das somas de quadrados da análise de variância em nenhum dos casos (simples, conjunta e desdobramento), por outro lado, altera a forma como são calculadas as estimativas dos componentes de variância pré e pós desdobramento da interação significativa, já que o método prático de Hicks modificado altera os resultados conforme é alterado um efeito de fixo para aleatório ou vice-versa.

Marcelino & Iemma (2000) afirmam, que quando se trabalha com conjuntos de dados reais, pode-se comparar os métodos de estimativas de componentes de variância de interesse, porém não se pode apontar a melhor ou pior metodologia devido ao desconhecimento dos reais valores de tais componentes. Tal confirmação pode ser obtida caso se simule um conjunto de dados de uma determinada distribuição para a qual os componentes de variância fossem conhecidos à priori.

3 METODOLOGIA

3.1 Materiais

3.1.1 Banco de dados reais

A partir de um conjunto de dados cedidos pela pesquisadora zootecnista Solidete de Fátima Paziani da Agência Paulista de Tecnologia dos Agronegócios - APTA, em que um grupo de experimentos de milho foi estudado na safra 2004/05, consideremos como variável de interesse neste trabalho, a produção de massa verde - PMV coletada no ponto de ensilagem da planta.

A configuração dos dados caracteriza um grupo de 4 ensaios de blocos ao acaso, em que foi cultivada a planta em questão nos municípios: Assis, Mococa, Pindamonhangaba e Votuporanga, todos localizados no estado de São Paulo.

Em cada local, se fazem presentes os mesmos 16 cultivares (tratamentos), que não serão descritos ao longo deste trabalho, pois o foco está na metodologia que entorna a análise de grupos de experimentos quanto a dois caminhos de comparação de médias de tratamentos e na obtenção das estimativas de componentes de variância.

3.1.1.1 O milho

O milho é uma importante fonte energética para o homem. A sua casca é rica em fibras, fundamental para a eliminação das toxinas do organismo humano. Seu grão é constituído de carboidratos, proteínas e vitaminas do complexo B. Este cereal possui bom potencial calórico, sendo constituído de grandes quantidades de açúcares e gorduras. O milho contém vários sais minerais como ferro, fósforo, potássio e zinco, no entanto é rico em ácido fítico, que dificulta a absorção destes sais.

A cultura do milho é uma das mais antigas do mundo, sua importância econômica é caracterizada pelas diversas formas de sua utilização, que vai desde a alimentação animal até a indústria de alta tecnologia. Apesar de apenas cerca de 30% do milho em grão mundial ser destinado à alimentação humana, contra 70% destinado à alimentação animal, em se tratando de derivados de milho, constitui fator importante de uso desse cereal em regiões com baixa renda.

Dentro da evolução mundial de produção de milho, o Brasil é destaque como terceiro maior produtor, perdendo apenas para Estados Unidos e China. A produção mundial ficou em torno de 590 milhões de toneladas no ano 2000, sendo grande parte desse total mérito dos Estados Unidos, China e Brasil que produziram aproximadamente 253, 105 e 32,3 milhões de toneladas respectivamente.

Em 2001, o Brasil apresentou a safra recorde de 41,5 milhões de toneladas, porém, apesar de estar entre os três maiores produtores deste grão, o Brasil não se destaca entre os países com maior nível de produtividade (kg/ha), e tem nos últimos anos ficado abaixo da produtividade média mundial.

No Brasil, o estado do Paraná liderou a safra agrícola de 2008, em se tratando de produção de milho. Cultivado em todo o país, este cereal é utilizado tanto como alimento, quanto para usos alternativos. A maior parte de sua produção é utilizada como ração de bovinos, suínos, aves e peixes.

Melo et al (1999) afirmam que embora os cultivares de milho possuam atributos capazes de proporcionar altas produtividades e qualidade, a produção final não depende unicamente dos cultivares, mas também, da interação entre genótipo e ambiente.

3.1.1.2 A produção de massa verde - PMV

A produção de massa verde - PMV do milho, geralmente medida em quilogramas por hectare (*kg/ha*), é uma importante característica representativa da qualidade de uma safra de tal cereal.

O milho possui um papel de destaque entre as plantas forrageiras por apresentar alto rendimento de massa verde por hectare, além de boas qualidades nutricionais para o gado, possibilitando boas produções e alto valor nutritivo de silagem (MELO et al., 1999).

Segundo Paziani & Duarte (2006), em se tratando de produção de volumosos para os animais, principalmente para a época seca do ano, a produtividade de massa verde é importante. Uma elevada produtividade é desejável pois reduz os custos de produção da silagem, sendo que a de melhor valor nutritivo, permitirá melhor desempenho animal, seja na produção de carne ou leite.

3.1.2 Banco de dados simulados

Serão realizadas simulações de conjuntos de dados experimentais, especificamente, de grupos de experimentos, em que cada grupo simulado será constituído por predeterminados número de locais - L e número de tratamentos - I por ensaio (os mesmos tratamentos em todos os locais). Além disso, cada experimento será sempre caracterizado por um delineamento em blocos ao acaso com um número de blocos - J também predeterminado.

3.2 Métodos

3.2.1 Análise de grupos de experimentos reais

Considerou-se a princípio, o banco de dados descrito em 3.1.1, em que uma série de passos serão praticados, visando atingir os objetivos do presente trabalho.

Alguns procedimentos serão adotados neste estudo, entre eles, a realização das análises individuais nas cidades de Assis, Mococa, Pindamonhangaba e Votuporanga, e de uma análise conjunta reunindo todos estes municípios em um único estudo. Será verificada a significância da interação tratamentos versus locais, significância essa que possibilitará o desdobramento dos graus de liberdade relativos a tratamentos + tratamentos versus locais trabalhando-se então com o resíduo médio no cálculo das estimativas dos componentes de variância e no teste de Tukey após o desdobramento.

Em se tratando de modelos aleatórios, para o efeito de tratamentos, o teste de hipóteses aplicado a partir de uma estatística F , permite verificar a veracidade da afirmação $\sigma_t^2 = 0$ (equação 3). Em contrapartida, sendo o modelo matemático fixo, a hipótese a ser testada é a de que $t_1 = \dots = t_I$, para todo $i = 1, \dots, I$ (equação 2).

Desta forma, vale ressaltar que tanto o modelo aleatório quanto o fixo serão abordados neste trabalho. O modelo aleatório permite a obtenção de componentes de variância para todos efeitos em questão, enquanto o modelo fixo, possibilita a aplicação do teste de Tukey para comparação de tratamentos pré e pós-desdobramento.

Ora o objetivo deste estudo estará voltado para a comparação de rols de tratamentos entre as análises individuais - AI e pós-desdobramento - APD, admitindo o efeito de tratamentos como fixo, outrora, voltado para a comparação de estimativas de componentes de variância segundo suas estimativas de variâncias, nestas mesmas duas metodologias.

Baseados nas análises de variância individuais, no método prático de Hicks modificado e admitindo o modelo matemático como aleatório, pode-se calcular as estimativas dos componentes de variância relativas a tratamentos, blocos e resíduo, utilizando os QMRs dos respectivos locais. Obtêm-se, então, em cada local: $\hat{\sigma}_t^2$, $\hat{\sigma}_b^2$ e $\hat{\sigma}_e^2$. Parte-se para a análise conjunta e desdobramento da interação significativa, alcançando-se a partir do QMRM: $\hat{\sigma}_{t(L_k)}^2$, para $k = 1, \dots, 4$ locais, $\hat{\sigma}_{b(l)}^2$ e $\hat{\sigma}_{e^*}^2$ únicos, independente de local. Estas últimas 3 estimativas correspondem aos mesmos fatores de variação das expressões referentes às análises individuais.

Assim, pode-se pensar em comparar, em cada um dos locais em estudo, os componentes de variância: $\hat{\sigma}_t^2$ versus $\hat{\sigma}_{t(L_k)}^2$, $\hat{\sigma}_b^2$ versus $\hat{\sigma}_{b(l)}^2$ e $\hat{\sigma}_e^2$ versus $\hat{\sigma}_{e^*}^2$ para $k = 1, \dots, 4$ através de suas variâncias, priorizando sempre os componentes que possuam menores variâncias associadas segundo a equação 16.

Serão comparados também, os resultados obtidos pelo teste de Tukey antes e após o desdobramento da interação significativa, buscando, caso existam, diferenças entre as listas de classificação dos tratamentos segundo suas médias, nas análises individuais versus pós-desdobramento. Vale lembrar que nas análises individuais trabalha-se com os QMRs de cada local nos testes F e de Tukey, por outro lado, na análise pós-desdobramento, utiliza-se o QMRM para tais testes.

Todas estas etapas serão cumpridas com o uso de uma ferramenta extremamente prática e eficiente: o software *R*.

A partir do método da análise de variância, obtêm-se as expressões dos estimadores de componentes de variância nas análises individuais, conjunta e na obtida após o desdobramento da interação significativa da análise conjunta. Para tal propósito, como o próprio nome da metodologia sugere, é necessária a construção dos quadros de análise de variância - ANOVA em cada caso, como será detalhado nas próximas subseções. Vale ressaltar que o método prático de Hicks modificado, descrito por Barbin (1993), permite a obtenção dos estimadores que serão estudados nesse trabalho, com a vantagem de não ser necessário um número excessivo de cálculos envolvendo esperanças matemáticas que o método da análise de variância exige, porém, através do método da análise de variância pode-se visualizar passo a passo a construção e justificativa de cada expressão.

3.2.1.1 Análise de variância de experimentos em blocos ao acaso - ANOVA individual

Banzatto & Kronka (1989), Pimentel Gomes (2000), Barbin (2003), entre outros, consideram o modelo matemático para um experimento em blocos ao acaso como:

$$y_{ij} = m + t_i + b_j + e_{ij}, \text{ para } i = 1, 2, \dots, I; j = 1, 2, \dots, J.$$

Sendo:

I =número de tratamentos;

J =número de blocos;

y_{ij} =valor observado na parcela que recebeu o tratamento i no bloco j ;

Visando a comparação de estimativas de variâncias de estimativas de componentes de variância, utilizou-se o modelo aleatório, admitindo-se as seguintes recomendações para uso nos cálculos das expressões das E(QM) conforme o método da análise de variância:

m é uma constante inerente a todas as observações, sendo que $E(m) = m$ e $E(m^2) = m^2$;

t_i com $i = 1, \dots, I$ é o efeito de tratamentos, aleatório, normalmente e independentemente distribuído com média 0 e variância σ_t^2 . Logo, $E(t_i) = 0$ e $E(t_i^2) = \sigma_t^2$;

b_j com $j = 1, \dots, J$ é o efeito de blocos, aleatório, normalmente e independentemente distribuído com média 0 e variância σ_b^2 . Logo, $E(b_j) = 0$ e $E(b_j^2) = \sigma_b^2$;

e_{ij} com $i = 1, \dots, I$ e $j = 1, \dots, J$ é o efeito residual, normalmente e independentemente distribuído com média 0 e variância σ_e^2 . Logo, $E(e_{ij}) = 0$ e $E(e_{ij}^2) = \sigma_e^2$.

Com base nessas informações, pode-se fazer uso do método da análise de variância, a partir da construção do quadro de análise de variância com uma coluna extra para as esperanças dos quadrados médios - E(QM), que deverá ser obtida para todos os fatores de variação. De acordo com esse método, para cada lacuna desta coluna, uma série de cálculos é necessária, por exemplo, para a esperança do quadrado médio de tratamentos no caso de um ensaio em blocos ao acaso (tabela 1), segundo a equação 6, utilizando as informações citadas no início desta seção e substituindo y_{ij} por $m + t_i + b_j + e_{ij}$, chega-se, após alguns cálculos, a $E(\text{QM Trat}) = \sigma_e^2 + J\sigma_t^2$. O mesmo raciocínio é utilizado para obtenção das demais expressões das E(QM).

$$E\left(\frac{SQTrat}{GLTrat}\right) = \frac{1}{I-1} E\left\{\frac{1}{J} \sum_{i=1}^I \left[\left(\sum_{j=1}^J y_{ij}\right)^2\right] - \frac{1}{IJ} \left(\sum_{i,j=1}^{I,J} y_{ij}\right)^2\right\}. \quad (6)$$

As mesmas expressões das E(QM) da tabela 1, seriam obtidas baseando-se no método prático de Hicks modificado, citado por Barbin (1993), através de uma série de procedimentos baseados em algumas informações do modelo matemático.

Tabela 1 - Análise de variância de um experimento em blocos ao acaso, com as esperanças dos quadrados médios - E(QM)

Causa de Variação	GL	SQ	QM	E(QM)	F_{calc}
Tratamentos	I - 1	S_1	$Q_1 = \frac{S_1}{I-1}$	$\sigma_e^2 + J\sigma_t^2$	$\frac{Q_1}{Q_3}$
Blocos	J - 1	S_2	$Q_2 = \frac{S_2}{J-1}$	$\sigma_e^2 + I\sigma_b^2$	
Resíduo	(I - 1)(J - 1)	S_3	$Q_3 = \frac{S_3}{(I-1)(J-1)}$	σ_e^2	
Total	IJ - 1	S_4			

Considerando GL o número de graus de liberdade, SQ a soma de quadrados e QM o quadrado médio relativos a cada fator de variação.

As expressões utilizadas para os cálculos das somas de quadrados - SQ da tabela 1, seguem abaixo:

$$S_1 = SQTrat = \frac{1}{J} \sum_{i=1}^I \left[\left(\sum_{j=1}^J y_{ij}\right)^2\right] - \frac{1}{IJ} \left(\sum_{i,j=1}^{I,J} y_{ij}\right)^2;$$

$$S_2 = SQBlocos = \frac{1}{I} \sum_{j=1}^J \left[\left(\sum_{i=1}^I y_{ij}\right)^2\right] - \frac{1}{IJ} \left(\sum_{i,j=1}^{I,J} y_{ij}\right)^2;$$

$$S_4 = SQTotal = \sum_{i,j=1}^{I,J} y_{ij}^2 - \frac{1}{IJ} \left(\sum_{i,j=1}^{I,J} y_{ij}\right)^2;$$

$$S_3 = SQRes = S_4 - S_1 - S_2 = \sum_{i,j=1}^{I,J} y_{ij}^2 - \frac{1}{J} \sum_{i=1}^I \left[\left(\sum_{j=1}^J y_{ij}\right)^2\right] +$$

$$- \frac{1}{I} \sum_{j=1}^J \left[\left(\sum_{i=1}^I y_{ij}\right)^2\right] + \frac{1}{IJ} \left(\sum_{i,j=1}^{I,J} y_{ij}\right)^2.$$

O teste F será realizado apenas para tratamentos. Tem-se como propósito uma comparação entre os rols de tratamentos segundo suas médias obtidos nas análises individual versus pós-desdobramento.

Além desse fator, vale destacar que os testes de Tukey para comparação de tratamentos segundo suas médias deve ser realizado sobre o modelo fixo, já que têm-se nesse caso, para o delineamento em blocos ao acaso, $E(\text{QM Trat}) = \sigma_e^2 + J\Phi_t$, sendo Φ_t dado pela equação 7, e $E(\text{QM Res}) = \sigma_e^2$, ou seja, o quociente $F = \frac{QM\text{Trat}}{QM\text{Res}}$ determina que o que está sendo testado pela estatística F é a veracidade da afirmação $\Phi_t = 0$ (hipótese nula H_0). Aceitar a hipótese nula nesse caso a um certo nível de significância de interesse, implica em admitir que todos os tratamentos não diferem entre si estatisticamente, ou ainda, todos os efeitos de tratamentos são nulos.

$$\Phi_t = \frac{\sum_{i=1}^I t_i^2}{I - 1}. \quad (7)$$

3.2.1.1.1 Estimadores calculados segundo a análise individual

De acordo com as expressões das esperanças de quadrados médios - E(QM) abordadas na tabela 1, pode-se obter as expressões dos estimadores dos componentes de variância $\hat{\sigma}_t^2$, $\hat{\sigma}_b^2$ e $\hat{\sigma}_e^2$ relativos a tratamentos, blocos e resíduo respectivamente, para um ensaio em blocos ao acaso de um modelo aleatório. Assim, os estimadores dos componentes de variância que serão comparados aos da análise conjunta têm as seguintes expressões:

$$\hat{\sigma}_t^2 = \frac{Q_1 - Q_3}{J}; \quad (8)$$

$$\hat{\sigma}_b^2 = \frac{Q_2 - Q_3}{I}; \quad (9)$$

$$\hat{\sigma}_e^2 = Q_3. \quad (10)$$

3.2.1.2 Análise de variância de grupos de experimentos - ANOVA conjunta

A análise conjunta de um grupo de experimentos possibilita uma visão ampla de mais de um ensaio em uma única avaliação para que se possa tirar conclusões abrangentes. Tal agrupamento é possível, sempre que se satisfaçam algumas condições já citadas nesse trabalho.

Barbin (2003) entre outros, adotam o seguinte modelo matemático para se analisar um grupo de experimentos em blocos ao acaso:

$$y_{ikj} = m + t_i + l_k + b_{j(k)} + tl_{ik} + e_{ikj}, \text{ para } i = 1, 2, \dots, I; k = 1, 2, \dots, K \text{ e } j = 1, 2, \dots, J.$$

Sendo:

I =número de tratamentos;

K =número de locais onde se repetiu o mesmo experimento;

J =número de blocos por local;

y_{ikj} =valor observado na parcela que recebeu o tratamento i , local k e bloco j ;

Visando a comparação de estimativas de variâncias de estimativas de componentes de variância, utilizou-se o modelo aleatório, admitindo-se as seguintes recomendações para uso nos cálculos das expressões das E(QM) conforme o método da análise de variância:

m é uma constante inerente a todas as observações. $E(m) = m$ e $E(m^2) = m^2$;

t_i com $i = 1, \dots, I$ é o efeito de tratamentos, aleatório, normalmente e independentemente distribuído com média 0 e variância $\sigma_{t^*}^2$. Logo, $E(t_i) = 0$ e $E(t_i^2) = \sigma_{t^*}^2$;

l_k com $k = 1, \dots, K$ é o efeito de locais, aleatório, normalmente e independentemente distribuído com média 0 e variância σ_l^2 . Logo, $E(l_k) = 0$ e $E(l_k^2) = \sigma_l^2$;

$b_{j(k)}$ com $k = 1, \dots, K$ e $j = 1, \dots, J$ é o efeito de blocos dentro de locais, aleatório, normalmente e independentemente distribuído com média 0 e variância $\sigma_{b(l)}^2$. Logo, $E(b_{j(k)}) = 0$ e $E(b_{j(k)}^2) = \sigma_{b(l)}^2$;

tl_{ik} com $i = 1, \dots, I$ e $k = 1, \dots, K$ é o efeito da interação tratamentos x locais, aleatório proveniente de formadores são aleatórios, normalmente e independentemente distribuído com média 0 e variância σ_{tl}^2 . Logo, $E(tl_{ik}) = 0$ e $E(tl_{ik}^2) = \sigma_{tl}^2$;

e_{ikj} com $i = 1, \dots, I$; $j = 1, \dots, J$ e $k = 1, \dots, K$ é o efeito residual, aleatório, normalmente e independentemente distribuído com média 0 e variância $\sigma_{e^*}^2$. Logo, $E(e_{ikj}) = 0$ e $E(e_{ikj}^2) = \sigma_{e^*}^2$.

Observação: Consideremos $\sigma_{t^*}^2$ e $\sigma_{e^*}^2$ como representantes das estimativas de componentes de variância para tratamentos e residual respectivamente no caso de análise conjunta, e para evitar equívocos, σ_t^2 e σ_e^2 representam tais estimativas porém provenientes de análises individuais.

Finalmente, pode-se utilizar o método da análise de variância para obter as expressões das E(QM) dos fatores de variação, e incluí-las no quadro de análise de variância conjunta (tabela 2).

Tabela 2 - Análise de variância conjunta de um grupo de experimentos em blocos ao acaso, com as esperanças dos quadrados médios - E(QM)

Causa de Variação	GL	SQ	QM	E(QM)	F_{calc}
Tratamentos (T)	I - 1	S_5	$Q_5 = \frac{S_5}{I-1}$	$\sigma_{e^*}^2 + J\sigma_{tl}^2 + KJ\sigma_{t^*}^2$	$\frac{Q_5}{Q_8}$
Blocos d. Locais	K(J - 1)	S_6	$Q_6 = \frac{S_6}{K(J-1)}$	$\sigma_{e^*}^2 + I\sigma_{b(l)}^2$	$\frac{Q_6}{Q_8}$
Locais (L)	K - 1	S_7	$Q_7 = \frac{S_7}{K-1}$	$\sigma_{e^*}^2 + J\sigma_{tl}^2 + I\sigma_{b(l)}^2 + IJ\sigma_l^2$	$\frac{Q_7}{Q_8}$
Interação T \times L	(I - 1)(K - 1)	S_8	$Q_8 = \frac{S_8}{(I-1)(K-1)}$	$\sigma_{e^*}^2 + J\sigma_{tl}^2$	$\frac{Q_8}{QMRM}$
Resíduo Médio	K(I - 1)(J - 1)	S_9	$QMRM$	$\sigma_{e^*}^2$	
Total	IKJ - 1	S_{10}			

Neste caso, para a obtenção das somas de quadrados - SQ da tabela 2, são utilizadas as seguintes expressões:

$$S_5 = SQT_{trat} = \frac{1}{KJ} \sum_{i=1}^I \left[\left(\sum_{k,j=1}^{K,J} y_{ikj} \right)^2 \right] - \frac{1}{IKJ} \left(\sum_{i,k,j=1}^{I,K,J} y_{ikj} \right)^2;$$

$$S_6 = SQ_{Blocos d. Locais} = \frac{1}{I} \sum_{k,j=1}^{K,J} \left[\left(\sum_{i=1}^I y_{ikj} \right)^2 \right] - \frac{1}{IKJ} \left(\sum_{i,k,j=1}^{I,K,J} y_{ikj} \right)^2;$$

$$S_7 = SQ_{Locais} = \frac{1}{IJ} \sum_{k=1}^K \left[\left(\sum_{i,j=1}^{I,J} y_{ikj} \right)^2 \right] - \frac{1}{IKJ} \left(\sum_{i,k,j=1}^{I,K,J} y_{ikj} \right)^2;$$

$$S_8 = SQT_{xL} = \frac{1}{KJ} \sum_{i=1}^I \left[\left(\sum_{k,j=1}^{K,J} y_{ikj} \right)^2 \right] + \frac{1}{IJ} \sum_{k=1}^K \left[\left(\sum_{i,j=1}^{I,J} y_{ikj} \right)^2 \right] +$$

$$+ \frac{1}{J} \sum_{i,k=1}^{I,K} \left[\left(\sum_{j=1}^J y_{ikj} \right)^2 \right] - \frac{1}{IKJ} \left(\sum_{i,k,j=1}^{I,K,J} y_{ikj} \right)^2;$$

$$\begin{aligned}
S_{10} &= SQTotal = \sum_{i,k,j=1}^{I,K,J} y_{ikj}^2 - \frac{1}{IKJ} \left(\sum_{i,k,j=1}^{I,K,J} y_{ikj} \right)^2; \\
S_9 &= SQResMedio = S_{10} - S_5 - S_6 - S_7 - S_8 = \\
&= \sum_{i,k,j=1}^{I,K,J} y_{ikj}^2 - \frac{1}{I} \sum_{k,j=1}^{K,J} \left[\left(\sum_{i=1}^I y_{ikj} \right)^2 \right] - \frac{1}{J} \sum_{i,k=1}^{I,K} \left[\left(\sum_{j=1}^J y_{ikj} \right)^2 \right] + \\
&+ \frac{3}{IKJ} \left(\sum_{i,k,j=1}^{I,K,J} y_{ikj} \right)^2 - \frac{2}{IJ} \sum_{k=1}^K \left[\left(\sum_{i,j=1}^{I,J} y_{ikj} \right)^2 \right] - \frac{2}{KJ} \sum_{i=1}^I \left[\left(\sum_{k,j=1}^{K,J} y_{ikj} \right)^2 \right].
\end{aligned}$$

Para o delineamento em blocos ao acaso, o número de graus de liberdade relativo ao resíduo médio é igual a soma dos graus de liberdade dos resíduos de cada análise individual. Além disso, havendo a homogeneidade de variâncias, o QMRes da análise conjunta é dado pela média aritmética dos QMRes das análises individuais (equação 11), denominado quadrado médio do resíduo médio - QMRM, o que torna desnecessário o cálculo da expressão S_9 .

$$QMRes = \frac{QMResL_1 + \dots + QMResL_K}{K}. \quad (11)$$

A partir das expressões das esperanças de quadrados médios - E(QM) abordadas na tabela 2, pode-se obter as expressões dos cinco estimadores de componente de variância em questão - $\hat{\sigma}_{t^*}^2$, $\hat{\sigma}_l^2$ e $\hat{\sigma}_{tl}^2$, $\hat{\sigma}_{b(l)}^2$ e $\hat{\sigma}_{e^*}^2$ relativos respectivamente a tratamentos proveniente de análise conjunta, locais, interação T×L, blocos dentro de locais e resíduo médio, para um conjunto de ensaios em blocos ao acaso, como segue abaixo:

$$\begin{aligned}
\hat{\sigma}_{t^*}^2 &= \frac{Q_5 - Q_8}{KJ}; \\
\hat{\sigma}_l^2 &= \frac{Q_7 - Q_6 - Q_8 + QMRM}{IJ}; \\
\hat{\sigma}_{tl}^2 &= \frac{Q_8 - QMRM}{J}; \\
\hat{\sigma}_{b(l)}^2 &= \frac{Q_6 - QMRM}{I}; \\
\hat{\sigma}_{e^*}^2 &= QMRM.
\end{aligned}$$

Vale salientar ainda, que os testes F utilizados para testar a significância dos efeitos de tratamentos, blocos ou locais costumam ser realizados com o $QMT \times L$, enquanto, por outro lado, o teste F para verificar a significância da interação $T \times L$ é efetuado com o QMRM no denominador.

O efeito de blocos dentro de locais geralmente não apresenta importância prática, segundo Barbin (2003), podendo então ser desprezado do quadro de análise de variância. Porém, nesse caso, deseja-se obter também o componente de variância relativo a blocos dentro de locais, optou-se então, por manter tal efeito no quadro.

Verificada a significância da interação tratamentos versus locais, faz-se necessário um desdobramento de tal, para obtenção das expressões dos estimadores dos componentes de variância relativos a tratamentos dentro de cada um dos locais ($\hat{\sigma}_{t(L_k)}^2$, para $k = 1, \dots, K$), a serem comparados aos provenientes das análises individuais. Tal desdobramento faz-se útil também, a fim de verificar o comportamento das médias dos tratamentos dentro de cada local visando comparação com tais comportamentos nas análises individuais.

3.2.1.2.1 Estimadores calculados pós desdobramento da interação significativa $T \times L$

Considera-se, o modelo aleatório $y_{ikj} = m + t_{i(1)} + t_{i(2)} + \dots + t_{i(K)} + l_k + b_{j(k)} + e_{ikj}$ correspondente ao desdobramento dos graus de liberdade de t_i (tratamentos) + tl_{ik} (interação tratamentos versus locais) do modelo referente à análise conjunta de grupos de experimentos em blocos ao acaso.

Segundo Barbin (2003), após desdobrada a interação significativa tratamentos versus locais, a soma $t_i + tl_{ik}$ é substituída por $t_{i(1)} + t_{i(2)} + \dots + t_{i(K)}$.

Tomando as mesmas informações utilizadas na análise conjunta, pode-se fazer uso tanto do método da análise de variância, quanto do método prático de Hicks modificado para obter as expressões das E(QM) dos fatores de variação, e incluí-las no quadro de análise de variância para o desdobramento (tabela 3).

Tabela 3 - Análise de variância para o desdobramento da interação significativa $T \times L$ de um grupo de experimentos em blocos ao acaso, com as esperanças dos quadrados médios $E(QM)$

Causa de Variação	GL	QM	$E(QM)$
Tratamentos d. L_1	$I - 1$	$Q_{1(1)}$	$\sigma_{e^*}^2 + J\sigma_{t(L_1)}^2$
...
Tratamentos d. L_K	$I - 1$	$Q_{1(K)}$	$\sigma_{e^*}^2 + J\sigma_{t(L_K)}^2$
Blocos d. Locais	$K(J - 1)$	Q_6	$\sigma_{e^*}^2 + I\sigma_{b(l)}^2$
Locais (L)	$K - 1$	Q_7	$\sigma_{e^*}^2 + J\sigma_{il}^2 + I\sigma_{b(l)}^2 + IJ\sigma_l^2$
Resíduo Médio	$K(I - 1)(J - 1)$	QM_{RM}	$\sigma_{e^*}^2$
Total	$IKJ - 1$		

Observações:

1) Nota-se que $Q_{1(k)}$, para $k = 1, \dots, K$ é o quadrado médio de tratamentos em cada um dos K locais em estudo, calculados segundo a fórmula de Q_1 (tabela 1);

2) De forma similar, $\sigma_{t(L_k)}^2$, para $k = 1, \dots, K$ é o componente de variância referente a tratamento dentro de cada um dos K locais em estudo.

Analogamente ao caso das análises individuais, vale ressaltar que os testes de Tukey para comparação de tratamentos segundo suas médias deve ser realizado sobre o modelo fixo devido a fatores já tratados, sendo $E(QM \text{ Trat d. } L_k) = \sigma_{e^*}^2 + J\Phi_t$, em que Φ_t é dado pela equação 7, e $E(QM \text{ Res}) = \sigma_{e^*}^2$. A hipótese nula a ser testada é a de que $\Phi_t = 0$, ou seja, inexistência de diferença significativa entre os tratamentos em estudo.

Segundo as expressões das esperanças de quadrados médios - $E(QM)$, que constam nas tabelas 2 e 3, é possível obter as expressões dos estimadores dos componentes de variância $\hat{\sigma}_{t(L_k)}^2$, para $k = 1, \dots, K$, $\hat{\sigma}_{b(l)}^2$ e $\hat{\sigma}_{e^*}^2$ correspondentes respectivamente a tratamentos dentro de cada local, blocos dentro de locais e resíduo médio, obtidos após uma análise conjunta de um grupo de experimentos em blocos ao acaso, com desdobramento da interação significativa tratamentos versus locais.

Assim, os estimadores dos componentes de variância que serão comparados aos das análises individuais têm as seguintes expressões:

$$\hat{\sigma}_{t(L_k)}^2 = \frac{Q_{1(k)} - QMRM}{J}, \text{ para } k = 1, \dots, K; \quad (12)$$

$$\hat{\sigma}_{b(l)}^2 = \frac{Q_6 - QMRM}{I}; \quad (13)$$

$$\hat{\sigma}_{e^*}^2 = QMRM. \quad (14)$$

3.2.1.3 Estimativas de variâncias de estimativas de componentes de variância

De acordo com Barbin (1993), em modelos balanceados, a estimativa de variância da estimativa de um componente de variância para um fator x, representada por $\hat{V}(\hat{\sigma}_x^2)$, é dada pela equação 15, sendo considerados os fatores de variação A e B, e seus respectivos quadrados médios - QMA e QMB e graus de liberdade - GLA e GLB, além do número natural C.

$$\hat{\sigma}_x^2 = \frac{QMA - QMB}{C} \Rightarrow \hat{V}(\hat{\sigma}_x^2) = \frac{2}{C^2} \left[\frac{(QMA)^2}{GLA + 2} + \frac{(QMB)^2}{GLB + 2} \right]. \quad (15)$$

O autor alerta ainda, que qualquer quadrado médio - QM de uma análise de variância associado a um determinado fator de variação x, multiplicado pelos seus respectivos graus de liberdade - GL e dividido pela respectiva esperança do quadrado médio - E(QM) possui distribuição qui-quadrado - χ^2 , então:

$$\frac{fQM}{E(QM)} \sim \chi_f^2$$

Consideremos o QMRes, associado a f GL, assim:

$$E(QMRes) = \sigma_e^2;$$

obtendo-se então a seguinte estimativa do componente de variância referente ao resíduo:

$$\hat{\sigma}_e^2 = QMR_{es}.$$

Na distribuição χ^2 , a média é igual ao número de GL e a variância é igual a 2 vezes o número de GL, tem-se:

$$V\left(\frac{f\hat{\sigma}_e^2}{\sigma_e^2}\right) = 2f;$$

e verifica-se que:

$$\hat{V}(\hat{\sigma}_e^2) = \frac{2(\sigma_e^2)^2}{f}.$$

A estimativa não-tendenciosa nesse caso, tal que $E[\hat{V}(\hat{\sigma}_e^2)] = V(\hat{\sigma}_e^2)$, é:

$$\hat{V}(\hat{\sigma}_e^2) = \frac{2(\hat{\sigma}_e^2)^2}{f+2}. \quad (16)$$

Desta forma, sendo o componente de variância relativo a um determinado fator x , da forma $\sigma_x^2 = QMA$, sua estimativa de variância é dada segundo a equação 16.

Teremos desta forma, as estimativas de variância de estimativas de componentes de variância para as expressões descritas nas equações 8, 9, 10, 12, 13 e 14, dadas respectivamente por:

$$\hat{V}(\hat{\sigma}_t^2) = \frac{2}{J^2} \left[\frac{(Q_1)^2}{I+1} + \frac{(Q_3)^2}{(I-1)(J-1)+2} \right]; \quad (17)$$

$$\hat{V}(\hat{\sigma}_b^2) = \frac{2}{I^2} \left[\frac{(Q_2)^2}{J+1} + \frac{(Q_3)^2}{(I-1)(J-1)+2} \right]; \quad (18)$$

$$\hat{V}(\hat{\sigma}_e^2) = \frac{2(Q_3)^2}{(I-1)(J-1)+2}; \quad (19)$$

$$\hat{V}(\hat{\sigma}_{t(L_k)}^2) = \frac{2}{J^2} \left[\frac{(Q_{1(k)})^2}{I+1} + \frac{(QMRM)^2}{K(I-1)(J-1)+2} \right]; \quad (20)$$

$$\hat{V}(\hat{\sigma}_{b(l)}^2) = \frac{2}{I^2} \left[\frac{(Q_6)^2}{K(J-1)+2} + \frac{(QMRM)^2}{K(I-1)(J-1)+2} \right]; \quad (21)$$

$$\hat{V}(\hat{\sigma}_{e^*}^2) = \frac{2(QMRM)^2}{K(I-1)(J-1)+2}. \quad (22)$$

A partir do conjunto de dados de produção de massa verde - PMV de milho, pode-se comparar as estimativas de variância de estimativas de componentes de variância obtidas via análises individuais - AI nos 4 municípios versus via análise pós desdobramento - APD da interação significativa TxL.

Desta forma, o interesse está nas comparações das estimativas de variância de estimativas de componentes de variância para tratamentos em cada local (equação 17) versus as obtidas para tratamentos desdobrados dentro de locais (equação 20), das estimativas para blocos dentro de cada local (equação 18) versus a estimativa única de blocos dentro de locais via análise conjunta (equação 21), e finalmente, das estimativas residuais nos 4 cidades (equação 19) versus a estimativa única residual alcançada pós-desdobramento de TxL (equação 22).

3.2.2 Análise de grupos de experimentos simulados

Serão considerados, na programação no software *R* (anexo B):

- i) O modelo matemático: $y_{ij} = m + t_i + b_j + e_{ij}$, com $i = 1, \dots, I$ tratamentos e $j = 1, \dots, J$ blocos para cada local, além do fato de se simular 1000 grupos de experimentos com L locais cada ($L = 3, 4$), para cada caso desejado;
- ii) Admitiu-se como valor inerente a todas as observações $m = 200$;
- iii) O número de tratamentos (I) por experimento variará entre 4 e 20, sendo que os mesmos tratamentos serão estudados em todos os locais. O número de blocos por ensaio variará entre 3 e 6. O número de locais tratados em cada grupo de experimentos será 3 ou 4. Algumas combinações destas possibilidades serão investigadas neste estudo de simulação;

- iv) Os valores correspondentes aos efeitos de tratamentos, blocos e resíduo serão obtidos de forma aleatória, todos eles serão gerados segundo a distribuição normal com média 0 e variância σ^2 . Serão fixados para a simulação destes efeitos, três trios de variâncias do tipo $(\sigma_{TRA}^2, \sigma_{BLO}^2, \sigma_{RES}^2)$, sendo eles, $S1 = (4, 1, 1)$, $S2 = (64, 16, 36)$ e $S3 = (2500, 225, 900)$;
- v) A soma dos efeitos de tratamentos deve ser nula, ou seja, $\sum_{i=1}^I t_i = 0$;
- vi) A soma dos efeitos de blocos deve ser nula, ou seja, $\sum_{j=1}^J b_j = 0$;
- vii) A soma dos efeitos residuais deve ser nula, ou seja, $\sum_{j=1}^J \sum_{i=1}^I e_{ij} = 0$;
- viii) Calculados todos os efeitos dos fatores em questão, é possível se obter o conjunto de dados completos a partir da equação do item (i), e assim, construir em cada local, a análise de variância individual;
- ix) A reunião de todos os locais em uma única análise possibilita o cálculo da estatística F, para que se possa visualizar a existência ou não de interação significativa T×L. Neste caso, tomaremos como nível de significância $\alpha = 0,05$;
- x) A interação tratamentos versus locais, ponto fundamental deste estudo deve ser verificada, um contador de grupos de experimentos com interação significativa foi implantado na programação visando calcular o número de grupos com interação significativa dentre cada 1000 simulados;
- xi) O foco deste estudo de simulação é a comparação entre as diferenças mínimas significativas - dms calculadas segundo as análises individuais com seus respectivos quadrados médios residuais e valores de amplitude estudentizada “q” versus as dms alcançadas segundo a análise pós-desdobramento do grupo de experimentos, com o quadrado médio do resíduo médio - QMRM no denominador e um valor “q” menor. Desta forma, serão verificados o número de vezes que a dms de análise individual se apresenta menor que o proveniente da análise conjunta, através de um indicador implantado na programação;

- xii) Serão calculadas também a média de QMRM de cada 1000 grupos simulados, para investigar se estes valores podem influenciar na comparação de dms;
- xiii) A semente “196”, foi escolhida de forma aleatória na programação em R, para possibilitar a reprodução das simulações e obtenção dos mesmos resultados que serão alcançados neste trabalho, a partir das linhas de comando descritas no anexo B.

Há grande diferença entre os trios de variâncias utilizados na distribuição normal para gerar os efeitos dos fatores de variação citados no item (iv), tal discrepância visa trabalhar-se com distintos quadrados médios de resíduo ao se utilizar um trio em relação a outro.

Devido ao fato de as somas exibidas nos itens (v), (vi) e (vii) deverem ser nulas, ao se desejar simular n efeitos para um determinado fator, serão simulados $n - 1$, e o n -ésimo será aquele que tornará esta soma nula.

Pelas combinações baseadas no item (iii), é possível afirmar que mudanças nos números de tratamentos, blocos e locais modificará também os graus de liberdade dos resíduos das análises individuais e conjunta. Este fato alterará também os valores da amplitude total estudentizada “q” para uso no teste de Tukey, já que tais valores dependem diretamente do quadrado médio residual. Em tal tabela de amplitude estudentizada, fixado um número de tratamentos qualquer, conforme se aumenta o número de graus de liberdade do resíduo, se diminui o valor de “q”, e vice-versa.

Vale salientar que obter uma dms em análise individual menor que a dms calculada na análise conjunta, implica em uma maior sensibilidade do teste de Tukey proveniente de análise individual. A comparação de dms obtidos pelos dois caminhos em questão, terá possivelmente como fatores decisivos o número de graus de liberdade residuais das análises individuais e conjunta.

4 RESULTADOS E DISCUSSÕES

4.1 Abordagem do banco de dados na variável produção de massa verde - PMV

O modelo matemático de delineamento em blocos ao acaso foi utilizado nos 4 ensaios, com 16 tratamentos e 4 blocos cada. Consideraram-se para este caso de produção de massa verde - PMV, modelos aleatórios em que apenas a média - m em todos os locais é fixa, enquanto os efeitos de tratamentos - t_i , de blocos - b_j e residual - e_{ij} são aleatórios para posteriores obtenções de estimativas de componentes de variância.

Analisando o conjunto de dados reunindo os 4 locais, obteve-se pelo teste de Shapiro-Wilk, p-valor igual a 0,0028 para a hipótese nula H_{0_1} : dados dispostos segundo uma distribuição normal, apontando a significância de tal teste, sendo cabível então, uma transformação nos dados para prosseguir o estudo.

Box & Cox (1964) exibem uma família de transformações visando reverter esse tipo de situação. Embora neste caso, para o teste de Bartlett (1950), com 15 GL associados, o p-valor para a hipótese nula H_{0_2} : existência de homogeneidade de variâncias, foi igual a 0,0244, indicando a não rejeição de H_{0_2} , optou-se pela transformação dos dados devido à não normalidade destes, visando o cumprimento das exigências do modelo de análise de variância.

O software *R*, baseado na família Box & Cox (1964), apontou $\lambda = 0$, ou seja, indicou a transformação logarítmica como uma boa alternativa para a normalização dos dados. Optou-se pela transformação $\ln(\text{PMV})$, e, a partir desta alteração no conjunto completo original, obteve-se normalidade para os dados segundo Shapiro-Wilk, e, além disso, manteve-se a homogeneidade de variâncias de tratamentos, antes existente, segundo Bartlett (1950).

Tabela 4 - Testes de normalidade e homogeneidade de variâncias para a variável transformada produção de massa verde - $\ln(\text{PMV})$, vista de forma geral, e em cada um dos locais: Assis - *AS*, Mococa - *MO*, Pindamonhangaba - *PI* e Votuporanga - *VO*

Testes pré-ANOVA	LOCALS				
	<i>AS</i>	<i>MO</i>	<i>PI</i>	<i>VO</i>	<i>GERAL</i>
Shapiro-Wilk	0,0709	0,0439	0,7429	0,9322	0,6133
Bartlett	0,2787	0,2071	0,0697	0,9969	0,1002

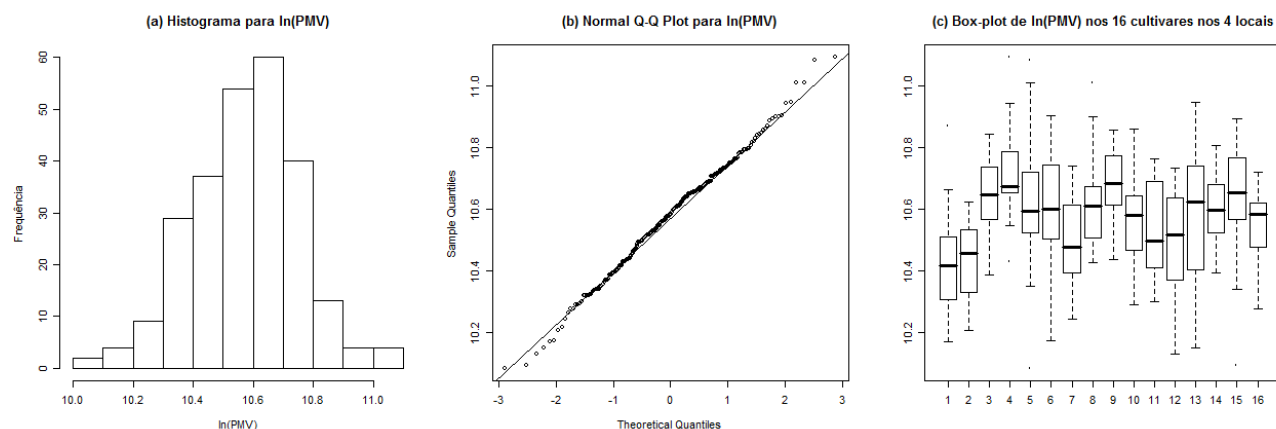


Figura 1 - (a) Histograma, (b) Q-Q Plot e (c) Box-plot para a variável transformada produção de massa verde - $\ln(\text{PMV})$

Pode-se visualizar pela figura 1, o que os testes de Shapiro-Wilk e Bartlett (1950) já afirmaram segundo a tabela 4: a existência de normalidade nas figuras 1-a e 1-b, e homogeneidade de variâncias na figura 1-c para a variável transformada produção de massa verde - $\ln(\text{PMV})$ abrangendo os municípios em estudo.

A seguir, estudaram-se os 4 locais separadamente segundo os pré-requisitos da análise de variância - ANOVA acima citados, obtendo-se então resultados descritos na tabela 4. Tal situação pode ser visualizada também nas figuras 2 e 3.

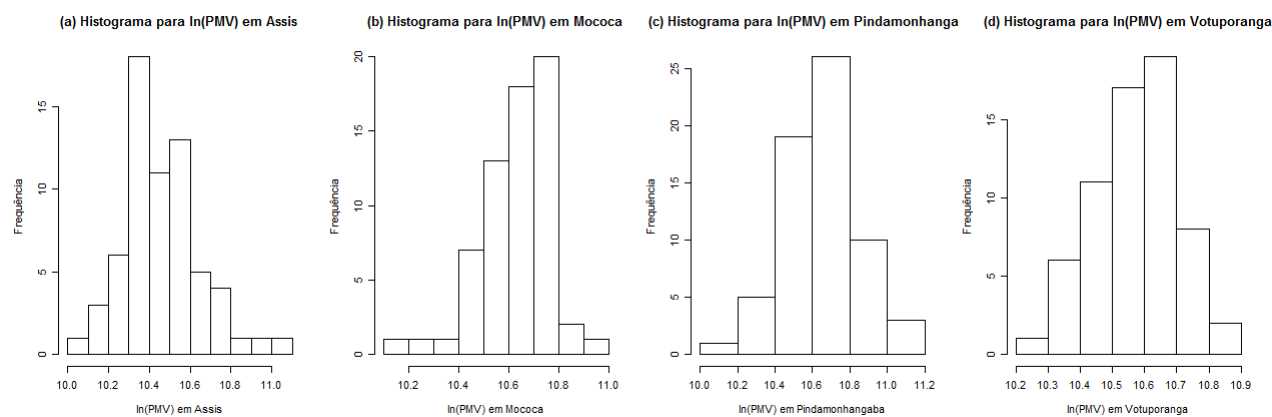


Figura 2 - Histogramas para a variável transformada produção de massa verde - $\ln(\text{PMV})$ nos municípios de (a) Assis, (b) Mococa, (c) Pindamonhangaba e (d) Votuporanga

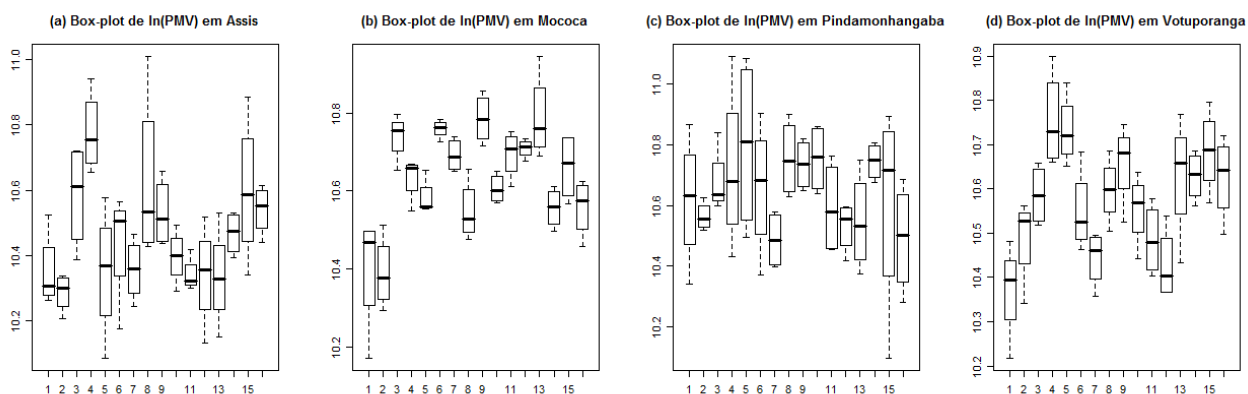


Figura 3 - Box-plot para a variável transformada produção de massa verde - $\ln(\text{PMV})$ nos municípios de (a) Assis, (b) Mococa, (c) Pindamonhangaba e (d) Votuporanga

Ainda segundo os resultados apresentados na tabela 4 e ilustrados nas figuras 2 e 3, a transformação \ln foi suficiente para a normalização do conjunto de dados de PMV em todos municípios, uma vez que, em Assis, Pindamonhangaba e Votuporanga, o teste de Shapiro-Wilk foi não significativo ao nível de 5% de probabilidade, contra 1% de probabilidade na cidade de Mococa.

Além disso, na mesma tabela, nota-se, nestas cidades, a não significância do teste de Bartlett (1950) ao nível de 5% de probabilidade, ou seja, a hipótese nula foi aceita e a homogeneidade de variância está presente em todos ensaios, mesmo não sendo este último fato tão claro pela figura 3.

A tabela 5 exhibe uma análise descritiva, com o intuito de se conhecer melhor o comportamento do conjunto de dados pré e pós-transformação \ln .

Tabela 5 - Análise descritiva da variável resposta original produção de massa verde - PMV, e normalizada - $\ln(\text{PMV})$

	Medidas de Posição				Medidas de Dispersão	
	Média	Mediana	Mínimo	Máximo	Variância	Desvio Padrão
PMV	39840,80	39470,25	24000,00	65620,00	50623954,00	7115,05
$\ln(\text{PMV})$	10,5770	10,5833	10,0858	11,0917	0,0317	0,1781

Satisfeitas as exigências, construiu-se computacionalmente (software *R*) o quadro com as análises de variância individuais, como pode-se visualizar na tabela 6.

Tabela 6 - Análises de variância individuais na variável $\ln(\text{PMV})$ em cada um dos municípios:
Assis - *AS*, Mococa - *MO*, Pindamonhangaba - *PI* e Votuporanga - *VO*

Causa de Variação	GL	QM <i>AS</i>	QM <i>MO</i>	QM <i>PI</i>	QM <i>VO</i>
Tratamentos	15	0,0763**	0,0596**	0,0417 ^{NS}	0,0485**
Blocos	3	0,0482	0,0150	0,0662	0,0631
Resíduo	45	0,0198	0,0049	0,0315	0,0047

Nota: ** Teste F significativo a 1% de probabilidade;

^{NS} Teste F não significativo.

Por conseguinte, vale ressaltar que $\ln(\text{PMV})$ apresentou distribuição normal e homogeneidade de variâncias conforme desejado, o que possibilitou a construção dos quadros de análise de variância onde obtiveram-se testes significativos para os tratamentos em Assis, Mococa e Votuporanga segundo a estatística F na tabela 6, implicando que nesses 3 municípios há diferenças relevantes entre os tratamentos, desigualdades essas detectadas pelo teste de Tukey e exibidas na tabela 10. Nota-se que em Pindamonhangaba os tratamentos possuem efeitos estatisticamente similares.

Segundo a tabela 6, tem-se que o quociente entre o maior (Pindamonhangaba) e o menor (Votuporanga) QMRes é igual a $6,7021 < 7$, o que segundo Pimentel Gomes (2000) permite a construção da análise de variância conjunta (tabela 7), em que Assis, Mococa, Pindamonhangaba e Votuporanga entram em uma única análise visando a observação da interação de interesse tratamentos versus locais. Caso se trabalhasse com o conjunto de dados original (sem transformação), obteria-se como quociente de QMRes , $7,4480 > 7$, o que não possibilitaria o agrupamento dos 4 locais em uma mesma análise segundo o mesmo autor.

As análises descritas na tabela 6 permitiram concluir sobre a possibilidade de reunir os 4 ensaios em um único estudo. Possibilitaram também a elaboração de testes de Tukey para comparação dos tratamentos segundo suas médias, vide tabela 10, e construção do rol classificatório dos tratamentos em cada análise individual a serem comparados, em breve, a cada rol classificatório de tratamentos na análise pós desdobramento.

Os valores obtidos nas tabelas 6, 7 e 8 foram calculados segundo os procedimentos de construção das tabelas 1, 2 e 3 respectivamente.

Tabela 7 - Análise de variância conjunta na variável resposta transformada - $\ln(\text{PMV})$ nos municípios de Assis, Mococa, Pindamonhangaba e Votuporanga

Causa de Variação	GL	QM	F
Tratamentos	15	0,1079	2,7386**
Blocos d. Locais	12	0,0481	1,2208 ^{NS}
Locais	3	0,4615	11,7132**
$T \times L$	45	0,0394	2,5921**
Resíduo Médio	180	0,0152	

Nota: ** Teste F significativo a 1% de probabilidade;

^{NS} Teste F não significativo.

A tabela 7 exhibe um estudo conjunto dos locais em questão visando obter e desdobrar a interação significativa tratamentos versus locais - $T \times L$, para que se possa verificar o rol dos tratamentos segundo suas médias dentro de cada local, além das estimativas dos componentes de variância a serem comparadas as das análises individuais.

O teste F foi significativo para a interação $T \times L$ ao nível de 1% de probabilidade conforme a tabela 7, pois $F_{calc} = 2,59 > 1,69 = F_{tab}(45, 180)$. Assim, pode-se afirmar com 99% de confiança que os tratamentos estão sendo influenciados pelos locais, ou seja, se comportam de maneiras distintas de um município para outro.

Tabela 8 - Análise de variância sobre a variável $\ln(\text{PMV})$, com desdobramento dos graus de liberdade referentes a tratamentos + interação $T \times L$ dentro de cada município: Assis - AS, Mococa - MO, Pindamonhangaba - PI e Votuporanga VO

Causa de Variação	GL	QM	F
Tratamentos dentro de AS	15	0,0763	5,0197**
Tratamentos dentro de MO	15	0,0596	3,9210**
Tratamentos dentro de PI	15	0,0417	2,7434**
Tratamentos dentro de VO	15	0,0485	3,1908**
Resíduo Médio	180	0,0152	

Nota: ** Teste F significativo a 1% de probabilidade.

Mesmo sendo o teste F para tratamentos significativo ao nível de 1% de probabilidade, não se pode tomar decisões abrangentes a toda a área em estudo devido a significância da interação. É cabível em tal situação, um desdobramento das cultivares dentro de cada uma das cidades para um estudo mais detalhado, conforme a tabela 8. Esse tipo de desdobramento é muito utilizado na literatura. Barbin (2003) entre outros, descrevem tal investigação.

A significância do teste F para locais ao nível de 1% de probabilidade, por sua vez, é de pouco interesse e neste caso será desprezada.

Tabela 9 - Logaritmo natural (ln) dos valores médios de produção de massa verde - PMV (kg/ha) avaliados em 16 cultivares nos municípios de Assis - AS, Mococa - MO, Pindamonhangaba - PI e Votuporanga - VO, e de forma geral - GERAL

Cultivares	<i>LOCAIS</i>				
	<i>AS</i>	<i>MO</i>	<i>PI</i>	<i>VO</i>	<i>GERAL</i>
1	10,3508	10,4005	10,6191	10,3719	10,4356
2	10,2866	10,3899	10,5641	10,4894	10,4325
3	10,5823	10,7397	10,6787	10,5862	10,6467
4	10,7756	10,6328	10,7212	10,7547	10,7211
5	10,3506	10,5817	10,8003	10,7327	10,6163
6	10,4376	10,7594	10,6595	10,5493	10,6014
7	10,3573	10,6911	10,4861	10,4436	10,4946
8	10,6251	10,5473	10,7562	10,5969	10,6314
9	10,5303	10,7848	10,7358	10,6577	10,6772
10	10,3960	10,6057	10,7549	10,5543	10,5777
11	10,3417	10,6942	10,5935	10,4853	10,5287
12	10,3400	10,7085	10,5312	10,4279	10,5019
13	10,3342	10,7884	10,5476	10,6300	10,5750
14	10,4685	10,5559	10,7451	10,6285	10,5995
15	10,6004	10,6608	10,6061	10,6850	10,6381
16	10,5407	10,5577	10,4922	10,6256	10,5541

Curiosamente, os testes F para tratamentos dentro de cada um dos locais foram significativos a 1% de probabilidade segundo a tabela 8, fato este que ocorreu também nas análises individuais, exceto na cidade de Pindamonhangaba. Assim, o teste de Tukey é importante na construção do rol classificatório nos 4 locais em questão enquanto foi responsável pela montagem das listas classificatórias dos 3 ensaios significativos das análises individuais.

Tabela 10 - Teste de Tukey para comparação dos cultivares segundo o logaritmo natural de suas respectivas médias de produção de massa verde - $\ln(\text{PMV})$ nos município de Assis, Mococa, Pindamonhangaba e Votuporanga, a partir do rol classificatório das análises individual - AI versus pós-desdobramento - APD

Assis		Mococa		Pindamonhangaba		Votuporanga	
AI	APD	AI	APD	AI	APD	AI	APD
$m_4(a)^*$	$m_4(c)$	$m_{13}(f)$	$m_{13}(k)$	$m_5(n)$	$m_5(o)$	$m_4(q)$	$m_4(v)$
$m_8(ab)$	$m_8(cd)$	$m_9(fg)$	$m_9(k)$	$m_8(n)$	$m_8(op)$	$m_5(q)$	$m_5(vx)$
$m_{15}(ab)$	$m_{15}(cd)$	$m_6(fgh)$	$m_6(k)$	$m_{10}(n)$	$m_{10}(op)$	$m_{15}(qr)$	$m_{15}(vxy)$
$m_3(ab)$	$m_3(cde)$	$m_3(fgh)$	$m_3(k)$	$m_{14}(n)$	$m_{14}(op)$	$m_9(qrs)$	$m_9(vxyw)$
$m_{16}(ab)$	$m_{16}(cde)$	$m_{12}(fghi)$	$m_{12}(k)$	$m_9(n)$	$m_9(op)$	$m_{13}(qrs)$	$m_{13}(vxyw)$
$m_9(ab)$	$m_9(cde)$	$m_{11}(fghi)$	$m_{11}(kl)$	$m_4(n)$	$m_4(op)$	$m_{14}(qrs)$	$m_{14}(vxyw)$
$m_{14}(ab)$	$m_{14}(de)$	$m_7(fghi)$	$m_7(klm)$	$m_3(n)$	$m_3(op)$	$m_{16}(qrs)$	$m_{16}(vxyw)$
$m_6(ab)$	$m_6(de)$	$m_{15}(fghi)$	$m_{15}(klm)$	$m_6(n)$	$m_6(op)$	$m_8qrst)$	$m_8(vxyw)$
$m_{10}(b)$	$m_{10}(de)$	$m_4(fghi)$	$m_4(klm)$	$m_1(n)$	$m_1(op)$	$m_3qrst)$	$m_3(vxyw)$
$m_7(b)$	$m_7(de)$	$m_{10}(ghi)$	$m_{10}(klm)$	$m_{15}(n)$	$m_{15}(op)$	$m_{10}(rst)$	$m_{10}(vxyw)$
$m_1(b)$	$m_1(de)$	$m_5(hi)$	$m_5(klm)$	$m_{11}(n)$	$m_{11}(op)$	$m_6(rst)$	$m_6(vxyw)$
$m_5(b)$	$m_5(de)$	$m_{16}(ij)$	$m_{16}(klm)$	$m_2(n)$	$m_2(op)$	$m_2(stu)$	$m_2(vxyw)$
$m_{11}(b)$	$m_{11}(de)$	$m_{14}(ij)$	$m_{14}(klm)$	$m_{13}(n)$	$m_{13}(op)$	$m_{11}(stu)$	$m_{11}(vxyw)$
$m_{12}(b)$	$m_{12}(de)$	$m_8(ij)$	$m_8(klm)$	$m_{12}(n)$	$m_{12}(op)$	$m_7(tu)$	$m_7(xyw)$
$m_{13}(b)$	$m_{13}(de)$	$m_1(j)$	$m_1(lm)$	$m_{16}(n)$	$m_{16}(p)$	$m_{12}(tu)$	$m_{12}(yw)$
$m_2(b)$	$m_2(d)$	$m_2(j)$	$m_2(m)$	$m_7(n)$	$m_7(p)$	$m_1(u)$	$m_1(w)$

Nota: * Cada $m_i()$, para $i = 1, \dots, 16$ representa a média do cultivar i segundo a Tabela 9. Entre parênteses, letras iguais correspondem a efeitos semelhantes para os cultivares em questão segundo o teste de Tukey ao nível de 5% de probabilidade. Cada coluna representa o rol classificatório de cultivares em ordem decrescente segundo suas médias.

Visando comparar as duas metodologias de construção de rol classificatório de tratamentos segundo suas médias, foram calculadas as médias gerais de todos os cultivares e dentro dos 4 municípios em estudo de acordo com a tabela 9.

Tem-se segundo Tukey, que as diferenças mínimas significativas - dms calculadas segundo a equação 4, e utilizadas na verificação da significância de contrastes envolvendo duas médias de tratamentos são dadas por 0,3604 (análise individual em Assis), 0,1800 (análise individual em Mococa), 0,4548 (análise individual em Pindamonhangaba), 0,1753 (análise individual em Votuporanga) e, finalmente, 0,3031 (dms relativa à análise pós-desdobramento comum aos 4 municípios utilizando-se o mesmo QMRM).

Cabe então neste momento, um confronto entre os resultados obtidos pelos testes de Tukey realizados sobre as análises individuais versus os obtidos sobre a análise pós-desdobramento. A tabela 10 ilustra tal comparação de listas classificatórias de tratamentos de acordo com suas respectivas médias descritas na tabela 9 e agrupadas segundo o teste de Tukey ao nível de 5% de probabilidade, levando-se em consideração as dms acima citadas.

A tabela 10 possibilita observar diferenças existentes entre um rol classificatório proveniente de uma análise individual - AI e um outro proveniente de uma análise pós-desdobramento - APD da interação significativa tratamentos versus locais. Os municípios apresentaram diferenças de uma lista classificatória para outra, tais como, alguns tratamentos que por AI não diferiram significativamente uns dos outros, por APD foram apontados como de efeitos distintos, e vice-versa.

Vale salientar que na cidade de Assis, por exemplo, no rol resultante de uma AI, o cultivar 4 que possui média $m_4 = 10,7756$ não difere estatisticamente dos cultivares 14 e 6, que possuem médias $m_{14} = 10,4685$ e $m_6 = 10,4376$ segundo Tukey, ao nível de 5% de probabilidade. Porém, no rol resultante de uma APD, o tratamento 4 difere do 14, e difere do 6, ou seja, o valor absoluto da diferença de suas médias tomadas duas a duas é maior que a diferença mínima significativa - dms.

Diga-se de passagem que no município de Pindamonhangaba, a análise de variância individual não se apresentou significativa para tratamentos segundo as tabelas 6 e 10 (5ª coluna). Tal fato revelou que os 16 tratamentos neste local e metodologia não diferem entre si segundo a estatística de Tukey ao nível de 5% de probabilidade.

Em contrapartida, neste mesmo município, porém, no rol resultante de uma APD, nota-se diferença entre os tratamentos 5 e 16 conforme as letras entre parênteses da tabela 10 indicam, mesmo se tratando de dois tratamentos de valores médios aparentemente pouco discrepantes ($m_5 = 10,8003$ e $m_{16} = 10,4922$).

Destaque para o fato de que as dms nas cidades de Assis e Pindamonhangaba foram superiores à dms pós-desdobramento, implicando uma menor sensibilidade para apontar diferenças significativas entre tratamentos segundo Tukey nessas duas AI em relação a APD. Porém, as cidades de Mococa e Votuporanga apresentaram dms inferiores à dms pós-desdobramento, e, conseqüentemente, maior sensibilidade para tal teste.

Este caso, em particular, exhibe as diferenças que ocasionalmente podem aparecer entre rols classificatórios provenientes de análises individuais versus pós-desdobramento da interação significativa tratamentos versus locais. É importante lembrar que este resultado baseou-se em um conjunto de locais que apresentou homogeneidade de variâncias.

Vários outros fatos curiosos desta mesma espécie podem ser encontrados na tabela 10. Em se tratando de rol classificatório de tratamentos, pode-se colocar a “culpa” da diferença de resultados da análise individual - AI para a pós-desdobramento - APD no quadrado médio do resíduo - QMRes. Nota-se que sendo I , J e K o número de tratamentos, blocos e locais respectivamente, trabalha-se na obtenção do QMRes na AI, com $(I - 1)(J - 1)$ graus de liberdade residuais, por outro lado, na APD, utilizam-se $K(I - 1)(J - 1)$ GL residuais, sendo $I, J, K \geq 2$, ou seja, $K(I - 1)(J - 1) > (I - 1)(J - 1)$.

Devido ao uso de um número de graus de liberdade residual maior na análise pós-desdobramento - APD em relação à análise individual - AI, neste caso, 180×45 não implica necessariamente, que o QMRes na APD de um determinado local será menor comparado ao da AI, mesmo sendo o valor q da estatística de Tukey na APD sempre menor que na AI.

Nota-se neste caso, que as diferenças mínimas significativas - dms nas cidades de Mococa e Votuporanga apresentaram-se inferiores na AI em relação à APD, enquanto nas cidades de Assis e Pindamonhangaba apresentaram-se inferiores na APD em relação à AI. Em suma, a AI foi mais sensível ao teste de Tukey para comparação de tratamentos segundo suas médias, nos municípios de Mococa e Votuporanga onde obtêm-se mais diferenças significativas entre tratamentos. Diversamente, a APD foi mais sensível em Assis e Pindamonhangaba.

Com a finalidade de enriquecer o presente trabalho, simulações de conjuntos de dados deste mesmo delineamento, com efeito significativo da interação tratamentos versus locais e homogeneidade de variâncias foram realizadas sob diferentes graus de liberdade para o resíduo, visando obter maior abrangência nos resultados (seção 4.2).

Partiu-se então para o cálculo das estimativas de componentes de variância a partir do método da análise de variância, baseando-se nos quadros de análise de variância individual (tabela 6), conjunta (tabela 7) e pós-desdobramento (tabela 8), para que se possa em seguida, realizar uma comparação destas estimativas.

Tabela 11 - Estimativas dos componentes de variância obtidas segundo análises individuais, suas respectivas variâncias, estimativas obtidas segundo a análise pós-desdobramento e suas respectivas variâncias na variável produção de massa verde nos municípios de Assis, Mococa, Pindamonhangaba e Votuporanga

Componentes de variância	Estimativas de componentes de variância e respectivas variâncias	
	Análises individuais	Análise pós-desdobramento
$\hat{\sigma}_{t(AS)}^2$	0,0141(4, 3896 $\times 10^{-5}$)*	0,0153(4, 3016 $\times 10^{-5}$)
$\hat{\sigma}_{t(MO)}^2$	0,0137 (2, 6213 $\times 10^{-5}$)	0,0111 (2, 6308 $\times 10^{-5}$)
$\hat{\sigma}_{t(PI)}^2$	0,0025 (1, 5413 $\times 10^{-5}$)	0,0066 (1, 2936 $\times 10^{-5}$)
$\hat{\sigma}_{t(VO)}^2$	0,0109 (1, 7327 $\times 10^{-5}$)	0,0083 (1, 7427 $\times 10^{-5}$)
$\hat{\sigma}_{b(AS)}^2$	0,0018 (0, 3693 $\times 10^{-5}$)	0,0021 (0, 1302 $\times 10^{-5}$)
$\hat{\sigma}_{b(MO)}^2$	0,0006 (0, 0355 $\times 10^{-5}$)	0,0021 (0, 1302 $\times 10^{-5}$)
$\hat{\sigma}_{b(PI)}^2$	0,0022 (0, 7014 $\times 10^{-5}$)	0,0021 (0, 1302 $\times 10^{-5}$)
$\hat{\sigma}_{b(VO)}^2$	0,0036 (0, 6229 $\times 10^{-5}$)	0,0021 (0, 1302 $\times 10^{-5}$)
$\hat{\sigma}_{e(AS)}^2$	0,0198 (1, 6620 $\times 10^{-5}$)	0,0152 (0, 2544 $\times 10^{-5}$)
$\hat{\sigma}_{e(MO)}^2$	0,0049 (0, 1035 $\times 10^{-5}$)	0,0152 (0, 2544 $\times 10^{-5}$)
$\hat{\sigma}_{e(PI)}^2$	0,0315 (4, 2175 $\times 10^{-5}$)	0,0152 (0, 2544 $\times 10^{-5}$)
$\hat{\sigma}_{e(VO)}^2$	0,0047 (0, 0932 $\times 10^{-5}$)	0,0152 (0, 2544 $\times 10^{-5}$)

Nota: * y(z): y indica a estimativa de componente de variância de acordo com a linha em que se encontra, enquanto z representa a estimativa da variância associada a tal estimativa de componente de variância calculada segundo a equação 16.

Construídos os quadros de ANOVA e conseqüentemente calculados os quadrados médios dos fatores de variação e suas respectivas esperanças, foi possível obter os estimadores e calcular as estimativas de componentes de variância dentro de cada município segundo as equações 8, 9 e 10, contra as estimativas pós-desdobramento da interação significativa pelas equações 12, 13 e 14. A partir destas estimativas, pôde-se avaliar as estimativas de variâncias associadas à estas, dispostas na tabela 11, e alcançadas segundo a equação 16.

Face ao exposto na tabela 11, verificam-se certas diferenças, mesmo que pequenas, dignas de discussão, entre as estimativas de componentes de variância e, inclusive, entre as estimativas de variâncias de estimativas de componentes de variância alcançadas via análises individuais - AI e via análises pós-desdobramento - APD.

Vale frisar que a estimativa $\hat{\sigma}_{b(l)}^2$ para $l = AS, MO, PI$ e VO é igual a 0,0021, e é obtida ao se calcular a média aritmética das estimativas $\hat{\sigma}_b^2$ nos 4 locais. Não o bastante, a estimativa $\hat{\sigma}_{e^*}^2 = 0,0152$, também é a média aritmética dos valores estimados de $\hat{\sigma}_e^2$ nos 4 municípios.

Analisando a tabela 11 mais a fundo, nota-se uma certa vantagem em se optar pelas estimativas de variâncias provenientes da APD, pois, além do desdobramento proporcionar estimativas de variâncias de estimativas de componentes de variância menores que as derivadas de análises individuais em sua maioria (7×5), trabalha-se com um único valor para blocos dentro de locais, e um único valor para a estimativa residual.

Apesar de exigir uma maior quantidade de cálculos, uma única estimativa representando vários locais, pode ser útil caso se queira trabalhar com esses mesmos cultivares em uma outra cidade com características próximas às destas.

Ao se observarem dados com uma variância muito baixa, espera-se obter quadrados médios das análises de variância pequenos, neste caso por exemplo, todos foram menores que 1, acarretando em estimativas de componentes de variância e estimativas de variância de estimativas de componentes de variância muito baixas, o que ainda mantém a comparação entre as duas metodologias cabível, pois busca-se a menor variância sempre que possível, e caso se dispusesse de um banco de dados com maior variância, poderiam talvez, ser verificadas desigualdades maiores.

Tabela 12 - Avaliação das metodologias AI e APD na obtenção de menores estimativas de variâncias associadas às estimativas de componentes de variância segundo proporções da metodologia de menor pela de maior variância na variável produção de massa verde nas cidades de Assis, Mococa, Pindamonhangaba e Votuporanga

Componentes de variância - CV	Metodologia de menor variância associada aos CV's	Metodologia de menor variância em relação à de maior (%)
$\hat{\sigma}_{t(AS)}^2$	APD	97,9953
$\hat{\sigma}_{t(MO)}^2$	AI	99,6389
$\hat{\sigma}_{t(PI)}^2$	APD	83,9291
$\hat{\sigma}_{t(VO)}^2$	AI	99,4262
$\hat{\sigma}_{b(AS)}^2$	APD	35,2559
$\hat{\sigma}_{b(MO)}^2$	AI	27,2657
$\hat{\sigma}_{b(PI)}^2$	APD	18,5629
$\hat{\sigma}_{b(VO)}^2$	APD	20,9022
$\hat{\sigma}_{e(AS)}^2$	APD	15,3069
$\hat{\sigma}_{e(MO)}^2$	AI	32,2368
$\hat{\sigma}_{e(PI)}^2$	APD	6,0320
$\hat{\sigma}_{e(VO)}^2$	AI	30,9211

Na tabela 12, tem-se que a metodologia de menor estimativa de variância de estimativa de componente de variância na respectiva linha correspondente a x% da estimativa obtida pela outra metodologia. Nota-se mais claramente a vantagem em se optar pelas estimativas calculadas segundo a análise pós desdobramento. Tomando as 4 menores proporções nesta tabela, tem-se $x < 25$, ou seja, a metodologia em questão foi superior por apresentar uma variância inferior a 1/4 da outra nestes casos. Estas 4 maiores diferenças de estimativas de variâncias, apontam superioridade para a análise pós-desdobramento - APD.

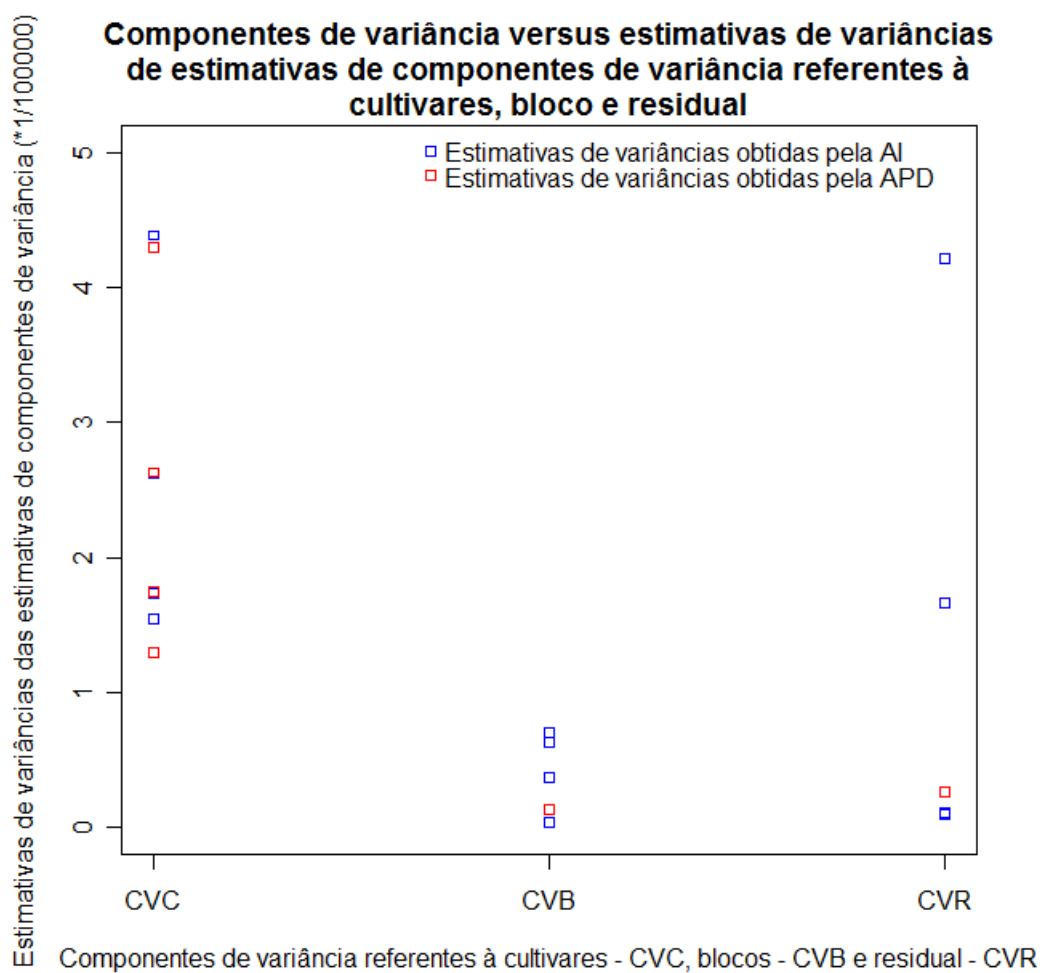


Figura 4 - Componentes de variância para a variável $\ln(\text{PMV})$ relativos a cultivares - CVC, blocos - CVB e resíduos - CVR, associados a suas respectivas estimativas de variâncias obtidas segundo os métodos de análise individual - AI e pós-desdobramento - APD

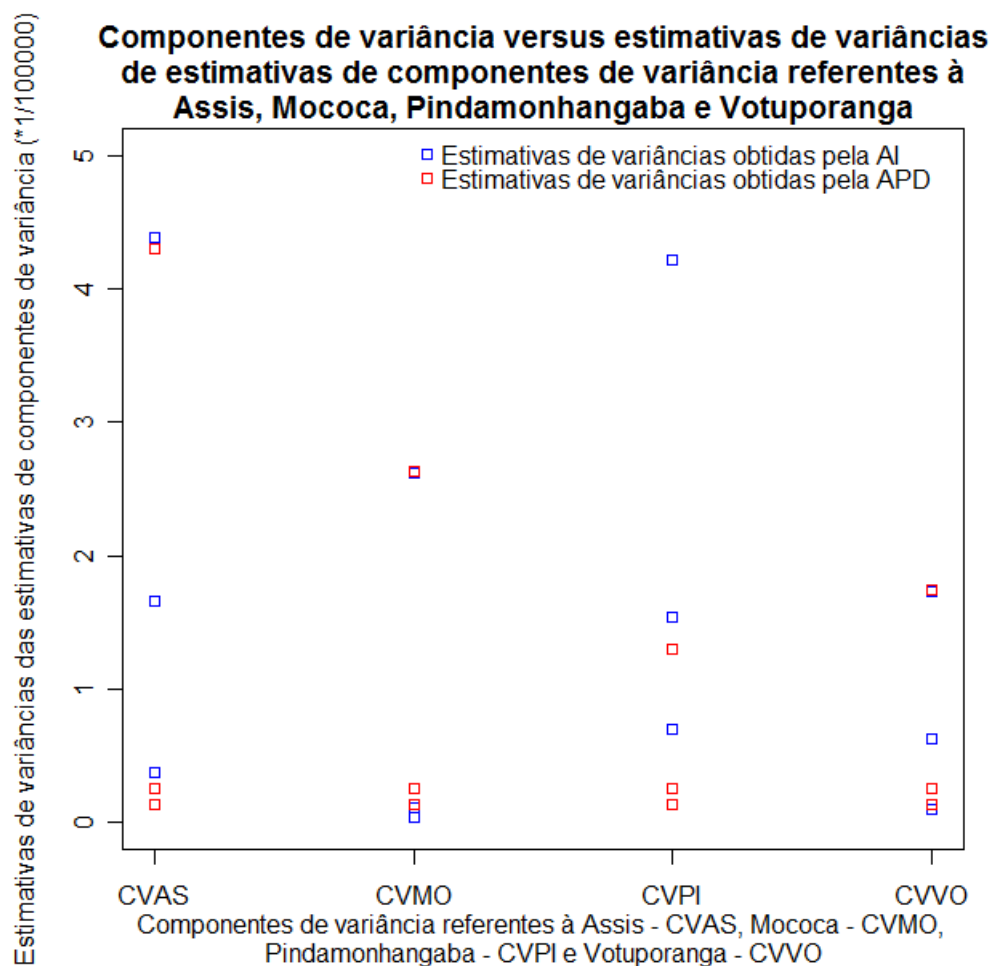


Figura 5 - Componentes de variância para a variável $\ln(\text{PMV})$ relativos a Assis - CVAS, Mococa - CVMO, Pindamonhangaba - CVPI e Votuporanga - CVVO, associados a suas respectivas estimativas de variâncias obtidas segundo os métodos de análise individual - AI e pós-desdobramento - APD

Outro fator que aponta tal superioridade segundo a tabela 12, é que as AI apresentaram em média 57,90% da estimativa de variância das APD quando foram melhores, em contrapartida, as APD tiveram, em média, 39,71% da variância das AI quando possuíram menores estimativas de variâncias de estimativas de componentes de variância.

A figura 4 exibe a comparação entre as estimativas de variâncias de estimativas de componentes de variância dentro de cada fator de variação presente tanto nos modelos aleatórios de análise individual - AI, quanto nos de análise pós-desdobramento - APD, sendo eles tratamentos, blocos e resíduo.

Em se tratando de cultivares ou tratamentos, nota-se que a maior variância de estimativa de componente de variância foi obtida a partir de uma AI considerando-se os 4 locais. Para blocos, os valores estão bem próximos, porém a variância obtida por AI também foi a maior. Finalmente para o resíduo, a estimativa de variância obtida por APD ficou acima de duas obtidas por AI e abaixo de outras duas, no entanto, bem próxima das baixas estimativas e distante das altas, conforme a figura 4.

Segundo a figura 5, é possível visualizar um confronto entre as estimativas de variâncias de estimativas de componentes de variância segundo as metodologias de AI e APD, dentro de cada município em estudo: Assis - AS, Mococa - MO, Pindamonhangaba - PI e Votuporanga - VO.

Admitindo-se em cada coluna da figura 5 todos os fatores de variação e buscando-se comparar apenas a variabilidade do referido local, é possível afirmar que em Assis, a maior variância de estimativa de componente de variância encontrada é referente a uma AI, enquanto as duas menores são de APD. Em Mococa a situação está mais equilibrada, entretanto, pela AI obteve-se a menor estimativa de variância. No município de Pindamonhangaba, um valor muito alto obtido pela AI destaca-se entre os demais. Por fim, na cidade de Votuporanga, o equilíbrio entre as estimativas de variâncias de estimativas de componentes de variância é muito grande, AI e APD apresentaram valores muito próximos nesse município.

4.2 Simulação de conjuntos de dados de grupos de experimentos

Realizou-se então uma comparação entre as diferenças mínimas significativas obtidas nas análises individuais - dmsI versus as dms alcançadas pós-desdobramento - dmsC de grupos de ora 3, ora 4 ensaios. Para isso, foram realizadas diversas simulações de 1000 conjuntos de dados cada, variando o número de tratamentos e blocos a fim de atingir diferentes graus de liberdade residuais que afetam por sua vez os valores da amplitude total estudentizada “q”, utilizada no cálculo das dms.

A partir das distribuições normais $N(0, \sigma_{TRA}^2)$, $N(0, \sigma_{BLO}^2)$ e $N(0, \sigma_{RES}^2)$, gerou-se os efeitos de tratamentos, blocos e residual, respectivamente, para cada local dentro de cada ensaio. Nas tabelas 13 e 14, optou-se pelo trio $(S1 = (\sigma_{TRA}^2, \sigma_{BLO}^2, \sigma_{RES}^2) = (4, 1, 1))$, nas tabelas 15 e 16, utilizou-se o trio $S2 = (64, 16, 36)$, enquanto nas tabelas 17 e 18 o trio adotado foi $S3 = (2500, 225, 900)$.

As tabelas 13, 15 e 17 apresentam simulações de grupos de 3 ensaios, enquanto as tabelas 14, 16 e 18 exibem simulações de grupos de 4 experimentos.

Verifica-se nas 6 tabelas acima citadas, um número muito elevado de interações tratamentos versus locais significativos - $n(T \times L^*)$. A escolha de uma variância maior para simular os efeitos de tratamentos em relação aos demais pode ter sido importante nesse sentido.

Nas tabelas 13 e 14, a média dos quadrados médios do resíduo médio - \overline{QMRM} obtida está sempre em torno de 2, o número de locais por ensaio não afetou alteração de uma tabela pra outra. Situações similares podem ser notadas entre as tabelas 15 e 16 e entre as tabelas 17 e 18.

Além disso, em se tratando de diferenças mínimas significativas - dms, nota-se pelas tabelas 13, 15 e 17, relativas a grupos de 3 ensaios, que estão sendo efetuadas 3000 comparações entre dms provenientes de análises individuais - dmsI versus pós-desdobramento - dmsC. Em cada grupo, compara-se a dms do local 1 versus a de tratamentos desdobrados dentro do local 1, e de maneira análoga para os locais 2 e 3. De forma análoga, nas tabelas 14, 16 e 18, referentes a grupos de 4 experimentos, o processo recém-citado é efetuado, porém neste caso, são investigadas 4000 comparações entre dms.

Tabela 13 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 3 ensaios cada, considerando-se $\sigma_{TRA}^2 = 4$, $\sigma_{BLO}^2 = 1$ e $\sigma_{RES}^2 = 1$ para os efeitos dos fatores de variação

GLResI	GLResC	I	J	n(T×L*)	\overline{QMRM}	n(dmsI<dmsC)
6	18	4	3	940	2,00	1027
9	27	4	4	988	1,98	1140
12	36	5	4	992	2,00	1215
15	45	6	4	997	2,01	1290
16	48	5	5	994	1,99	1339
21	63	8	4	1000	2,00	1346
24	72	7	5	1000	2,01	1421
25	75	6	6	998	2,00	1407
27	81	10	4	1000	1,98	1417
28	84	8	5	1000	2,02	1425
30	90	11	4	1000	2,00	1455
33	99	12	4	1000	2,00	1439
36	108	10	5	999	2,02	1481
40	120	11	5	1000	2,00	1496
44	132	12	5	1000	2,02	1491
45	135	16	4	1000	2,03	1501
45	135	10	6	1000	2,06	1536
52	156	14	5	1000	2,04	1517
75	225	16	6	1000	2,01	1592
95	285	20	6	1000	2,02	1617

Tabela 14 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 4 ensaios cada, considerando-se $\sigma_{TRA}^2 = 4$, $\sigma_{BLO}^2 = 1$ e $\sigma_{RES}^2 = 1$ para os efeitos dos fatores de variação

GLResI	GLResC	TRAT	BLOC	n(T×L*)	\overline{QMRM}	n(dmsI<dmsC)
6	24	4	3	988	1,97	1384
9	36	4	4	997	1,99	1618
12	48	5	4	999	2,01	1722
15	60	6	4	1000	2,01	1750
16	64	5	5	1000	2,00	1863
21	84	8	4	1000	2,00	1923
24	96	7	5	1000	2,00	2004
25	100	6	6	1000	2,02	1977
27	108	10	4	1000	1,99	2030
28	112	8	5	1000	2,02	2051
30	120	11	4	1000	2,00	2029
33	132	12	4	1000	2,02	2065
36	144	10	5	1000	2,03	2067
40	160	11	5	1000	1,98	2102
44	176	12	5	1000	2,01	2107
45	180	16	4	1000	2,00	2082
45	180	10	6	1000	2,05	2163
52	208	14	5	1000	2,02	2184
75	300	16	6	1000	2,03	2254
95	380	20	6	1000	2,03	2278

Tabela 15 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 3 ensaios cada, considerando-se $\sigma_{TRA}^2 = 64$, $\sigma_{BLO}^2 = 16$ e $\sigma_{RES}^2 = 36$ para os efeitos dos fatores de variação

GLResI	GLResC	TRAT	BLOC	n(T×L*)	\overline{QMRM}	n(dmsI<dmsC)
6	18	4	3	803	71,88	1029
9	27	4	4	904	71,42	1140
12	36	5	4	940	72,18	1214
15	45	6	4	966	72,23	1293
16	48	5	5	956	71,53	1339
21	63	8	4	986	72,02	1346
24	72	7	5	990	72,22	1421
25	75	6	6	985	72,04	1408
27	81	10	4	989	71,17	1418
28	84	8	5	991	72,69	1426
30	90	11	4	996	72,06	1456
33	99	12	4	995	72,02	1439
36	108	10	5	997	72,59	1483
40	120	11	5	995	71,93	1494
44	132	12	5	1000	72,85	1492
45	135	16	4	997	72,97	1501
45	135	10	6	998	74,01	1537
52	156	14	5	995	73,35	1519
75	225	16	6	1000	72,40	1591
95	285	20	6	1000	72,77	1617

Tabela 16 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 4 ensaios cada, considerando-se $\sigma_{TRA}^2 = 64$, $\sigma_{BLO}^2 = 16$ e $\sigma_{RES}^2 = 36$ para os efeitos dos fatores de variação

GLResI	GLResC	TRAT	BLOC	n(T×L*)	\overline{QMRM}	n(dmsI<dmsC)
6	24	4	3	906	70,78	1384
9	36	4	4	965	71,66	1618
12	48	5	4	986	72,27	1719
15	60	6	4	992	72,32	1751
16	64	5	5	991	71,96	1863
21	84	8	4	994	71,95	1926
24	96	7	5	998	71,95	2004
25	100	6	6	997	72,80	1978
27	108	10	4	1000	71,64	2030
28	112	8	5	998	72,91	2052
30	120	11	4	999	72,00	2030
33	132	12	4	998	72,75	2064
36	144	10	5	1000	73,02	2069
40	160	11	5	998	71,43	2101
44	176	12	5	1000	72,20	2104
45	180	16	4	1000	71,93	2083
45	180	10	6	1000	73,83	2163
52	208	14	5	1000	72,81	2185
75	300	16	6	1000	72,99	2254
95	380	20	6	1000	73,18	2277

Tabela 17 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 3 ensaios cada, considerando-se $\sigma_{TRA}^2 = 2500$, $\sigma_{BLO}^2 = 225$ e $\sigma_{RES}^2 = 900$ para os efeitos dos fatores de variação

GLResI	GLResC	TRAT	BLOC	n(T×L*)	\overline{QMRM}	n(dmsI<dmsC)
6	18	4	3	893	1797,02	1029
9	27	4	4	971	1785,44	1140
12	36	5	4	982	1804,58	1213
15	45	6	4	987	1805,86	1293
16	48	5	5	989	1788,30	1340
21	63	8	4	997	1800,57	1346
24	72	7	5	998	1805,55	1420
25	75	6	6	996	1801,12	1409
27	81	10	4	999	1779,25	1419
28	84	8	5	1000	1817,24	1426
30	90	11	4	999	1801,44	1456
33	99	12	4	1000	1800,51	1439
36	108	10	5	999	1814,81	1483
40	120	11	5	999	1798,34	1494
44	132	12	5	1000	1821,36	1492
45	135	16	4	1000	1824,17	1500
45	135	10	6	1000	1850,38	1538
52	156	14	5	1000	1833,80	1518
75	225	16	6	1000	1810,01	1592
95	285	20	6	1000	1819,34	1617

Tabela 18 - Média dos QMRM - \overline{QMRM} , número de interações T×L significativas - n(T×L*) e de dms obtidas via AI menor que via APD - n(dmsI<dmsC) em predeterminados número de tratamentos - I, blocos - J, GL residual individual - GLResI e conjunto - GLResC, para simulados 1000 grupos de 4 ensaios cada, considerando-se $\sigma_{TRA}^2 = 2500$, $\sigma_{BLO}^2 = 225$ e $\sigma_{RES}^2 = 900$ para os efeitos dos fatores de variação

GLResI	GLResC	TRAT	BLOC	n(T×L*)	\overline{QMRM}	n(dmsI<dmsC)
6	24	4	3	963	1769,45	1383
9	36	4	4	985	1791,51	1618
12	48	5	4	997	1806,72	1719
15	60	6	4	998	1808,10	1750
16	64	5	5	999	1799,08	1863
21	84	8	4	1000	1798,65	1926
24	96	7	5	1000	1798,81	2005
25	100	6	6	999	1820,09	1978
27	108	10	4	1000	1791,10	2030
28	112	8	5	1000	1822,80	2051
30	120	11	4	1000	1800,05	2030
33	132	12	4	1000	1818,81	2064
36	144	10	5	1000	1825,56	2070
40	160	11	5	1000	1785,84	2102
44	176	12	5	1000	1805,13	2105
45	180	16	4	1000	1798,18	2083
45	180	10	6	1000	1845,72	2163
52	208	14	5	1000	1820,19	2185
75	300	16	6	1000	1824,81	2255
95	380	20	6	1000	1829,44	2277

Segundo as tabelas 13, 15 e 17, um ensaio com 16 tratamentos, 4 blocos e 45 graus de liberdade do resíduo na análise individual - AI, parece ser o ponto de equilíbrio, já que este tipo de experimento apresentou, respectivamente, 1501, 1501 e 1500 como número de dms obtidas via AI menor que as alcançadas via APD - $n(\text{dmsI} < \text{dmsC})$, verificadas 1000 simulações de grupos de 3 experimentos (3000 comparações entre os 2 tipos de dms). Um experimento com 10 tratamentos e 6 blocos também possui 45 graus de liberdade mas não foi tão equilibrado na comparação de interesse.

De acordo com as tabelas 14, 16 e 18, um experimento com 7 tratamentos, 5 blocos e 24 graus de liberdade residuais na AI, pode ser considerado ponto de equilíbrio, pois este tipo de experimento apresentou, respectivamente, 2004, 2004 e 2005 como número de dms obtidas via AI menor que as calculadas via APD - $n(\text{dmsI} < \text{dmsC})$, verificadas 1000 simulações de grupos de 4 experimentos (4000 comparações).

Consideraram-se então, intervalos de simulações em que o $n(\text{dmsI} < \text{dmsC})$ está entre a metade da quantidade simulada (1500 ou 2000) $\pm 5\%$, nestes casos, não se pode apontar a metodologia mais sensível para o teste de Tukey. Assim, nos casos de grupos de 3 ensaios, aceitou-se as simulações em que $1425 \leq n(\text{dmsI} < \text{dmsC}) \leq 1575$. De outra parte, nos grupos de 4 experimentos, aceitou-se as simulações em que $1900 \leq n(\text{dmsI} < \text{dmsC}) \leq 2100$.

Nas circunstâncias predefinidas nas tabelas 13, 15 e 17, verifica-se que ensaios compreendidos no intervalo de 28 a 52 graus de liberdade residuais em AI, apresentaram grande equilíbrio em relação ao número de dms obtidas via AI menor que as alcançadas via APD - $n(\text{dmsI} < \text{dmsC})$, tirando a possibilidade de se escolher nestes casos, a metodologia mais sensível para o teste de Tukey para comparação de tratamentos e construção de rols classificatórios.

Sob as circunstâncias predeterminadas nas tabelas 14, 16 e 18, nota-se que experimentos compreendidos no intervalo entre 21 e 36 graus de liberdade do resíduo em AI apresentaram um grande equilíbrio em se tratando do número de dms obtidas via AI menor que as calculadas via APD - $n(\text{dmsI} < \text{dmsC})$.

Curiosamente, foram observados nas 6 últimas tabelas, valores discrepantes entre as duas simulações envolvendo 45 graus de liberdade residuais para AI, sendo que $n(\text{dmsI} < \text{dmsC})$ para 16 tratamentos foi sempre menor que $n(\text{dmsI} < \text{dmsC})$ para 10 tratamentos neste caso.

No quesito comparação de rols de tratamentos, vale destacar, que as simulações realizadas com números de graus de liberdade residuais em AI inferiores às representantes dos intervalos de equilíbrio acima citados, apresentam mais sensibilidade, nas diferenças mínimas significativas via análises pós-desdobramento - dmsC utilizando o QMRM. Isto se aplica a $n(\text{dmsI} < \text{dmsC}) < 1425$ para grupos de 3 experimentos e $n(\text{dmsI} < \text{dmsC}) < 1900$ para grupos de 4 ensaios.

Por outro lado, simulações com números de graus de liberdade residuais superiores aos dos intervalos de equilíbrio acima descritos, apresentam mais sensibilidade nas dmsI, utilizando o QMRes específico de cada local. Tal fato ocorre para $n(\text{dmsI} < \text{dmsC}) > 1575$ em grupos de 3 ensaios e $n(\text{dmsI} < \text{dmsC}) > 2100$ em grupos de 4 experimentos.

Em situações fora da região de equilíbrio, por exemplo, nos casos de grupos de 3 experimentos, 3000 comparações são realizadas, e caso obtido $n(\text{dmsI} < \text{dmsC}) > 1500$, afirma-se que, obtiveram-se, na maioria dos casos, $\text{dmsI} < \text{dmsC}$, em outras palavras, a metodologia dmsI é mais sensível para o teste de Tukey para comparações de médias de tratamentos neste caso, possibilitando mais diferenças significativas de contrastes de duas médias que pela metodologia dmsC.

Além dos intervalos de equilíbrio obtidos, sob as mesmas condições trabalhadas no conjunto de dados reais, ou seja, com 16 tratamentos e 4 blocos em 4 locais, de acordo com as tabelas 14, 16 e 18, chega-se a números de dms obtidas via AI menores que via APD - $n(\text{dmsI} < \text{dmsC})$ próximos à metade das comparações simuladas, fato este, que impossibilita a escolha da metodologia mais sensível para a comparação de tratamentos. Vale lembrar que no conjunto de dados reais, foram obtidas 2 dmsI menores que dmsC, dentre as 4 possíveis.

5 CONCLUSÕES

Com base nos resultados do banco de dados reais, na variável transformada $\ln(\text{PMV})$, sob o grupo de 4 experimentos em questão, e sendo a interação $T \times L$ significativa, pode-se concluir:

- Os rols classificatórios de tratamentos organizados a partir das dms calculadas segundo o teste de Tukey ao nível de 5% para comparação de médias de tratamentos, foram diferentes nas AI em relação às APD. Nos municípios de Mococa e Votuporanga, obtiveram-se dmsI inferiores a dmsC, implicando em uma vantagem em se trabalhar com as AI, por outro lado, nas cidades de Assis e Pindamonhangaba, os valores de dmsI foram superiores aos de dmsC, resultando em ser mais adequado o uso de APD nestes locais;
- Os componentes de variância obtidos para os fatores de variação do modelo em estudo (tratamentos, blocos e resíduo), foram mais satisfatórios quando calculados via APD utilizando o QMRM.

Embasados nas análises de grupos de experimentos simulados, sendo 1000 simulações para cada situação predefinida, conclui-se:

- Tanto para grupos de 3, quanto de 4 ensaios, observaram-se intervalos de equilíbrio, cujo números de dms relativas a AI inferiores às relativas a APD - $n(\text{dmsI} < \text{dmsC})$, era ao redor da metade do número de simulações. Para estas ocasiões, pode-se trabalhar tanto com os rols classificatórios de tratamentos via AI utilizando os resíduos específicos, quanto via APD com o QMRM como testador;
- Ao deparar-se com um $n(\text{dmsI} < \text{dmsC})$ inferior à metade possível, pode-se dizer que foram encontradas mais dmsI maiores que dmsC. Julga-se cabível nestes casos, o uso do teste de Tukey para construção de rols classificatórios de tratamentos, sobre as APD, visando assim, maior quantidade de diferenças entre tratamentos;
- Por outro lado, ao obter-se um $n(\text{dmsI} < \text{dmsC})$ superior à metade possível, afirma-se que foram obtidas mais dmsI menores que dmsC. Assim, aponta-se como mais apropriado, o uso do teste de Tukey para construção de rols classificatórios de tratamentos, sobre as AI nestes casos.

REFERÊNCIAS

ALVES, G.F. **Estimadores de variâncias de estimadores de contrastes de médias de tratamentos em grupos de experimentos em parcelas subdivididas com alguns tratamentos comuns nas parcelas**. 67 p. (Dissertação de Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1995.

APTA (Agência Paulista de Tecnologia e Agronegócios), URL. Disponível em: <<http://www.aptaregional.sp.gov.br/>>. Acessado em: 29 jun. 2009.

BANZATTO, D.A.; KRONKA, S.N. **Experimentação agrícola**. Jaboticabal: FUNEP, 1989. 247p.

BARBIN, D. **Componentes de variância: teoria e aplicações**. 2 ed. rev. e amp. Piracicaba: FEALQ, 1993. 108p.

BARBIN, D. **Planejamento e análise estatística de experimentos agrônômicos**. Araçatuba: Midas, 2003. 208p.

BARTLETT, M.S. Tests of significance in factor analysis. **The British Journal of Mathematical and Statistical Psychology**. Leicester, v.3, p.77-85, 1950.

BLISS, C.I. The U.S.P. collaborative cat assay for digitalis. **Journal of the American Pharmaceutical Association**. Washington, v.33, p.225-245, 1944.

BOX, G.E.P. Some theorems on quadratic forms applied in the study of analysis of variance problems. **The Annals of Mathematical Statistics**. Baltimore, v.25, p.290-302, 1954.

BOX, G.E.; COX, D.R. An analysis of transformations. **Journal of the Royal Statistical Society: B**. London, v.26, p.211-252, 1964.

CAMPOS, H. **Estatística aplicada à experimentação com cana-de-açúcar**. Piracicaba: FEALQ, 1984. 292p.

CECON, P.R. **Alternativas de análise de experimentos em látex e aplicações no melhoramento vegetal**. 109p. (Dissertação de Doutorado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1992.

COCHRAN, W.G. The combination of estimates from different experiments. **Biometrics**. Arlington, v.10, p.101-129, 1954.

COCHRAN, W.G.; COX, G.M. **Experimental designs**. 2 ed. New York: John Wiley, 1957. 611p.

COX, D.R.; SOLOMON P.J. **Components of variance**. London: Chapman & Hall/CRC, 2003. 169p.

DUARTE, J.B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal**. 293p. (Dissertação de Doutorado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2000.

EBERHART, S.A.; RUSSEL. W.A. Stability parameters for comparing varieties. **Crop Science**. Madison, v.6, p.36-40, 1966.

FISHER, R.A. The correlation between relatives on the supposition of Medelian inheritance. **Transactions of the Royal Society of Edinburgh**. Edinburgh, v.52, p.399-433, 1918.

HARTLEY, H.O.; RAO, J.N.K. Maximum-likelihood estimation for the mixed analysis of variance models. **Biometrika**. London, v.54, p.93-108, 1967.

HENDERSON, C.R. Estimation of variance and variance components. **Biometrics**. New York, v.9, p.226-252, 1953.

HICKS, C.R. **Fundamental concepts in the design of experiments**. 2 ed., New York: Holt, Rinehart and Winston, 1973. 349p.

KEMPTHORNE, O. **The design and analysis of experiments**. New York: John Wiley, 1952. 631p.

MARCELINO, S.D. do R.; IEMMA, A.F. Métodos de estimação de componentes de variância em modelos mistos desbalanceados. **Scientia Agricola**. Piracicaba, v.57, n.4, p.643-652, out./dez. 2000.

MELO, W.M.C.; PINHO, R.G.V.; CARVALHO, M.L.M.; PINHO, E.V.R.V. Avaliação de cultivares de milho para produção de silagem na região de Lavras - MG. **Ciência e Agrotecnologia**. Lavras, v.23, n.1, p.31-39, jan/mar. 1999.

OLIVEIRA, W.A.; NOGUEIRA, M.C.S. Aplicação do resíduo específico na análise de grupos de experimentos. **Bragantia**. Campinas, v.66, n.4, p.737-744, 2007.

PAZIANI, S.F.; DUARTE, A.P. Avaliação de cultivares de milho e sorgo para silagem. **Relatório Científico da Agência Paulista de Tecnologia dos Agronegócios - APTA**. Campinas, 2006.

PIEIDADE, S.M.S. **Análise de grupos de experimentos em faixas**. 148p. (Dissertação de Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1987.

PIMENTEL GOMES, F. **Componentes de variância**. Piracicaba, 1968. 30p.

PIMENTEL GOMES, F. An extension of the method of joint analysis of experiments in complete randomized blocks. **Biometrics**. Tallahassee, v.26, p.332-336, 1970.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 14 ed. Piracicaba: Degaspari, 2000. 477p.

PIMENTEL GOMES, F.; GUIMARÃES, R.F. Joint analysis of experiments in complete blocks designs with some common treatments. **Biometrics**. Tallahassee, v.14, p.521-526, 1958.

R. Development Core Team. R Foundation for Statistical Computing. **R: A language and environment for statistical computing**, Vienna, Austria. ISBN 3-900051-07-0, URL. Disponível em: <<http://R-project.org>>. Acessado em: 15 mar. 2008.

RAMALHO, M.A.P. **Eficiência relativa de alguns processos de seleção intrapopulacional no milho baseados em famílias não endógamas**. 121p. (Dissertação de Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1987.

RAMALHO, M.A.P.; SANTOS, J.B. dos; ZIMMERMANN, M.J.O. **Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro**. Goiânia: UFG, 1993. 271p.

RAO, C.R. Estimation of variance and covariance components - MINQUE Theory. **Journal of Multivariate Analysis**. New York, v.1, p.257-275, 1971a.

RAO, C.R. Minimum variance quadratic unbiased estimation of variance components. **Journal of Multivariate Analysis**. New York, v.1, p.445-456, 1971b.

RAO, P.S.R.S. **Variance components estimation: mixed models, methodologies and applications**. London: Chapman & Hall, 1997. 204p.

REGAZZI, A.J.; SILVA, H.D.; VIANA, J.M.S.; CRUZ, C.D. Análise de experimentos em látice quadrado com ênfase em componentes de variância. **Pesquisa agropecuária brasileira**. Brasília, v.34, n.11, p.1987-1997, 1999.

SOUZA JÚNIOR, C.L.; VENCOVSKY, R. Covariância entre parentes na presença da interação genótipo-ambiente. **In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA**. 3., Lavras, 1989. Resumos... Lavras: ESAL, 1989. p.50-51

VALÉRIO FILHO, W.V. **Comparação de métodos para estimação de componentes da variância através de simulação de dados**. 160p. (Dissertação de Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1991.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão preto: SBG, 1992. 496p. Publicado na Revista Brasileira de Genética, Ribeirão Preto, 1992.

VENCOVSKY, R.; GERALDI, I.O. Um modelo multiplicativo aplicado à análise de produção de grãos. **Relatório Científico do Departamento de Genética**. Piracicaba, v.11, p.157-165, 1977.

VIEIRA, S. **Análise de variância: (ANOVA)**. São Paulo: Atlas, 2006. 204p.

YATES, F.; COCHRAN, W.Z. The analysis of groups of experiments. **Journal of Agricultural Science**. London, v.28, p.556-580, 1938.

ANEXOS

ANEXO A - Programa (R) para construção dos rols classificatórios de tratamentos, e cálculo das estimativas de componentes de variância e suas respectivas variâncias no grupo de experimentos em estudo, segundo as análises individuais e pós desdobramento da interação significativa tratamentos versus locais

```
#####
##### Construção dos rols classificatórios de tratamentos segundo #####
##### suas médias no grupo de experimentos em estudo, nas análises #####
##### individuais e pós-desdobramento da interação significativa #####
##### tratamentos versus locais #####
#####

##### Leitura dos dados de PMV #####

dados <- read.table("pmv.txt", head=T)
verde <- dados$pmv

##### Separação dos grupos de experimentos (4 locais) #####

verde_AS <- verde[seq(1,64)]
verde_M0 <- verde[seq(65,128)]
verde_PI <- verde[seq(129,192)]
verde_V0 <- verde[seq(193,256)]

##### Fatores de variação da análise conjunta #####

trat <- as.factor(dados$cult)
bloc <- as.factor(dados$bloco)
loc <- as.factor(dados$local)

##### Fatores de variação das análises individuais #####

trat_i <- trat[seq(1,64)]
bloc_i <- bloc[seq(1,64)]

i <- 16 # Número de tratamentos (cultivares)
j <- 4 # Número de blocos
l <- 4 # Número de locais (municípios)

##### Histograma e teste de normalidade (Shapiro-Wilk) para a #####
##### variável resposta verde (PMV) #####

hist(verde)
shapiro.test(verde)
```

```
##### Boxplot e teste de Bartlett para homogeneidade de variâncias #####
x11()
with(dados, boxplot(verde~trat))
title("Homogeneidade Verde")
bartlett.test(verde ~ trat)

##### Busca de "lambda" segundo a família Box-Cox de transformações #####
require(MASS)
x11()
boxcox(verde ~ trat + bloc + loc, data=dados, lambda = seq(-2,2))
title("Transf. Verde")

##### Transformação ln da variável resposta verde #####
verde_t <- log(verde)
dados_t <- data.frame(loc, trat, bloc, verde_t)

##### Histograma e teste de normalidade para ln(verde) = verde_t #####
x11()
hist(verde_t)
shapiro.test(verde_t)

##### Teste de Bartlett para homogeneidade de variâncias de verde_t #####
x11()
with(dados, boxplot(verde_t~trat))
title("Homogeneidade Verde_t")
bartlett.test(verde_t ~ trat)

##### variável verde transformada no local 1 - Assis #####
verde_AS_t <- verde_t[seq(1,64)]
##### variável verde transformada no local 2 - Mococa #####
verde_MO_t <- verde_t[seq(65,128)]
##### variável verde transformada no local 3 - Pindamonhangaba #####
verde_PI_t <- verde_t[seq(129,192)]
##### variável verde transformada no local 4 - Votuporanga #####
verde_VO_t <- verde_t[seq(193,256)]

##### Teste de Shapiro-Wilk de normalidade para cada município #####
shapiro.test(verde_AS_t)
shapiro.test(verde_MO_t)
shapiro.test(verde_PI_t)
shapiro.test(verde_VO_t)

##### Testes de homogeneidade de variâncias para cada município #####
bartlett.test(verde_AS_t ~ trat_i)
bartlett.test(verde_MO_t ~ trat_i)
bartlett.test(verde_PI_t ~ trat_i)
bartlett.test(verde_VO_t ~ trat_i)
```

```
##### Histograma, Q-Q Plot e boxplot para verde_t (banco completo) #####
```

```
x11()
par(mfrow=c(1,3))
hist(verde_t, xlab="ln(PMV)", ylab="Frequência",
main="(a) Histograma para ln(PMV)")
qqnorm(verde_t, main="(b) Normal Q-Q Plot para ln(PMV)")
qqline(verde_t, col=1)
with(dados_t, boxplot(verde_t~trat), xlab="cultivares")
title("(c) Box-plot de ln(PMV) nos 16 cultivares nos 4 locais")
```

```
##### Histogramas de PMV nos 4 municípios #####
```

```
x11()
par(mfrow=c(1,4))
hist(verde_AS_t,xlab="ln(PMV) em Assis", ylab="Frequência",
main="(a) Histograma para ln(PMV) em Assis")
hist(verde_MO_t,xlab="ln(PMV) em Mococa", ylab="Frequência",
main="(b) Histograma para ln(PMV) em Mococa")
hist(verde_PI_t,xlab="ln(PMV) em Pindamonhangaba",ylab="Frequência",
main="(c) Histograma para ln(PMV) em Pindamonhangaba")
hist(verde_VO_t,xlab="ln(PMV) em Votuporanga", ylab="Frequência",
main="(d) Histograma para ln(PMV) em Votuporanga")
```

```
##### Boxplot para visualizar a homogeneidade de variâncias nos 4 locais #####
```

```
x11()
par(mfrow=c(1,4))
with(dados_t, boxplot(verde_AS_t~trat_i))
title("(a) Box-plot de ln(PMV) em Assis")
with(dados_t, boxplot(verde_MO_t~trat_i))
title("(b) Box-plot de ln(PMV) em Mococa")
with(dados_t, boxplot(verde_PI_t~trat_i))
title("(c) Box-plot de ln(PMV) em Pindamonhangaba")
with(dados_t, boxplot(verde_VO_t~trat_i))
title("(d) Box-plot de ln(PMV) em Votuporanga")
```



```
##### Análises de variância individuais #####
```

```
anava_AS <- aov(verde_t[seq(1,64)] ~ trat_i + bloc_i, data=dados_t)
summary(anava_AS)
anava_MO <- aov(verde_t[seq(65,128)] ~ trat_i + bloc_i, data=dados_t)
summary(anava_MO)
anava_PI <- aov(verde_t[seq(129,192)] ~ trat_i + bloc_i, data=dados_t)
summary(anava_PI)
anava_VO <- aov(verde_t[seq(193,256)] ~ trat_i + bloc_i, data=dados_t)
summary(anava_VO)
```

```
##### Condição para agrupamento dos 4 locais em uma única análise #####
```

```
qm_res <- c(anova(anava_AS)$Mean[3], anova(anava_MO)$Mean[3],
anova(anava_VO)$Mean[3], anova(anava_PI)$Mean[3])
homog <- max(qm_res)/min(qm_res)
##### homog deve ser menor que 7 para se aceitar a homog. de variâncias #####
homog < 7
```

```
##### Análise de variância conjunta #####
```

```
anava_cjt <- aov(verde_t ~ loc/bloc + trat + loc + trat:loc, data=dados_t)
summary(anava_cjt)
```

```
##### Desdobramento dos tratamentos dentro de locais (municípios) #####
```

```
dad <- transform(dados_t, lt = loc:trat)
desd <- aov(verde_t ~ loc/bloc + loc + lt, data = dad)
summary(desd, split =
list(lt = list(TdAS = 1:15, TdMO = 16:30, TdPI = 31:45, TdVO = 46:60)))
```

```
##### Médias dos 16 tratamentos nos 4 municípios #####
```

```
m_AS <- with(dados_t,tapply(verde_AS_t, trat_i,mean))
m_MO <- with(dados_t,tapply(verde_MO_t, trat_i,mean))
m_PI <- with(dados_t,tapply(verde_PI_t, trat_i,mean))
m_VO <- with(dados_t,tapply(verde_VO_t, trat_i,mean))
```

```
##### Cálculo dos "q" para o teste de Tukey #####
```

```
##### Para as análise individuais com 16 e 45 graus de liberdade #####
```

```
qtab_pre <- qtukey(0.95, i, (i-1)*(j-1))
```

```
##### Para a análise pós-desdobramento com 16 e 180 graus de liberdade #####
```

```
qtab_pos <- qtukey(0.95, i, (i-1)*(j-1)*1)
```

```

##### Cálculo das "dms" para as comparações dos tratamentos pré e pós- #####
##### desdobramento da interação significativa tratamentos versus locais #####

dms_pre_AS <- qtab_pre*sqrt(anova(anava_AS)$Mean[3]/j)
dms_pre_MO <- qtab_pre*sqrt(anova(anava_MO)$Mean[3]/j)
dms_pre_PI <- qtab_pre*sqrt(anova(anava_PI)$Mean[3]/j)
dms_pre_VO <- qtab_pre*sqrt(anova(anava_VO)$Mean[3]/j)

dms_pos <- qtab_pos*sqrt(anova(anava_cjt)$Mean[5]/j)

##### Contrastes de médias de tratamentos (duas a duas) #####

##### Assis #####
dif_AS <- matrix(nrow=16,ncol=16)
for(a1 in 1:16){
for(b1 in 1:16){
dif_AS[a1,b1] = abs(m_AS[a1] - m_AS[b1])}}
dif_AS
##### Mococa #####
dif_MO <- matrix(nrow=16,ncol=16)
for(a2 in 1:16){
for(b2 in 1:16){
dif_MO[a2,b2] = abs(m_MO[a2] - m_MO[b2])}}
dif_MO
##### Pindamonhangaba #####
dif_PI <- matrix(nrow=16,ncol=16)
for(a3 in 1:16){
for(b3 in 1:16){
dif_PI[a3,b3] = abs(m_PI[a3] - m_PI[b3])}}
dif_PI
##### Votuporanga #####
dif_VO <- matrix(nrow=16,ncol=16)
for(a4 in 1:16){
for(b4 in 1:16){
dif_VO[a4,b4] = abs(m_VO[a4] - m_VO[b4])}}
dif_VO

```

```
##### Comparação entre os tratamentos (cultivares) nas análises #####
##### individuais (dms_pre) e pós-desdobramento (dms_pos) #####
##### segundo Tukey ao nível de 5% de probabilidade #####
```

```
##### Assis pré e pós-desdobramento #####
dif_AS > dms_pre_AS
dif_AS > dms_pos
##### Mococa pré e pós-desdobramento #####
dif_MO > dms_pre_MO
dif_MO > dms_pos
##### Pindamonhangaba pré e pós-desdobramento #####
dif_PI > dms_pre_PI
dif_PI > dms_pos
##### Votuporanga pré e pós-desdobramento #####
dif_VO > dms_pre_VO
dif_VO > dms_pos
```

```
#####
##### Cálculo das estimativas de componentes de variância no grupo de #####
##### experimentos em estudo, obtidas segundo análises individuais e pós- #####
##### desdobramento da interação significativa tratamentos versus locais #####
#####
```

```
##### Componentes de variância (tratamentos, blocos e residual) #####
##### nas análises individuais nos 4 municípios #####
```

```
##### Assis #####
sig2t_AS <- (anova(anava_AS)$Mean[1]-anova(anava_AS)$Mean[3])/j
sig2b_AS <- (anova(anava_AS)$Mean[2]-anova(anava_AS)$Mean[3])/i
sig2e_AS <- anova(anava_AS)$Mean[3]
```

```
##### Mococa #####
sig2t_MO <- (anova(anava_MO)$Mean[1]-anova(anava_MO)$Mean[3])/j
sig2b_MO <- (anova(anava_MO)$Mean[2]-anova(anava_MO)$Mean[3])/i
sig2e_MO <- anova(anava_MO)$Mean[3]
```

```
##### Pindamonhangaba #####
sig2t_PI <- (anova(anava_PI)$Mean[1]-anova(anava_PI)$Mean[3])/j
sig2b_PI <- (anova(anava_PI)$Mean[2]-anova(anava_PI)$Mean[3])/i
sig2e_PI <- anova(anava_PI)$Mean[3]
```

```
##### Votuporanga #####
sig2t_VO <- (anova(anava_VO)$Mean[1]-anova(anava_VO)$Mean[3])/j
sig2b_VO <- (anova(anava_VO)$Mean[2]-anova(anava_VO)$Mean[3])/i
sig2e_VO <- anova(anava_VO)$Mean[3]
```

```

##### Estimativas de componentes de variância #####
##### (tratamentos dentro de locais, blocos dentro de locais e residual) #####
##### para a análise de variância pós-desdobramento #####

sig2tdAS <- (anova(anava_AS)$Mean[1] - anova(anava_cjt)$Mean[5])/j
sig2tdMO <- (anova(anava_MO)$Mean[1] - anova(anava_cjt)$Mean[5])/j
sig2tdPI <- (anova(anava_PI)$Mean[1] - anova(anava_cjt)$Mean[5])/j
sig2tdVO <- (anova(anava_VO)$Mean[1] - anova(anava_cjt)$Mean[5])/j
sig2b_cjt <- (anova(anava_cjt)$Mean[3] - anova(anava_cjt)$Mean[5])/i
sig2e_cjt <- anova(anava_cjt)$Mean[5]

##### Graus de liberdade para tratamentos, blocos e resíduo a serem #####
##### utilizados no cálculo das estimativas de variância das estimativas #####
##### de componentes de variância pré e pós-desdobramento da #####
##### interação significativa tratamentos versus locais #####

gl_t_ind <- anova(anava_AS)$Df[1]
gl_t_cjt <- anova(anava_cjt)$Df[2]
gl_b_ind <- anova(anava_AS)$Df[2]
gl_b_cjt <- anova(anava_cjt)$Df[3]
gl_e_ind <- anova(anava_AS)$Df[3]
gl_e_cjt <- anova(anava_cjt)$Df[5]

##### Estimativas de variâncias de estimativas de componentes de variância #####
##### provenientes de análises individuais #####
##### Componente de variância relativo à tratamentos nos 4 locais #####

v_sig2t_AS <- (2/(j^2))*(((anova(anava_AS)$Mean[1])^2)/(gl_t_ind + 2) +
((anova(anava_AS)$Mean[3])^2)/(gl_e_ind + 2))
v_sig2t_MO <- (2/(j^2))*(((anova(anava_MO)$Mean[1])^2)/(gl_t_ind + 2) +
((anova(anava_MO)$Mean[3])^2)/(gl_e_ind + 2))
v_sig2t_PI <- (2/(j^2))*(((anova(anava_PI)$Mean[1])^2)/(gl_t_ind + 2) +
((anova(anava_PI)$Mean[3])^2)/(gl_e_ind + 2))
v_sig2t_VO <- (2/(j^2))*(((anova(anava_VO)$Mean[1])^2)/(gl_t_ind + 2) +
((anova(anava_VO)$Mean[3])^2)/(gl_e_ind + 2))

##### Componentes de variância relativos à blocos nos 4 locais #####

v_sig2b_AS <- (2/(i^2))*(((anova(anava_AS)$Mean[2])^2)/(gl_b_ind + 2) +
((anova(anava_AS)$Mean[3])^2)/(gl_e_ind + 2))
v_sig2b_MO <- (2/(i^2))*(((anova(anava_MO)$Mean[2])^2)/(gl_b_ind + 2) +
((anova(anava_MO)$Mean[3])^2)/(gl_e_ind + 2))
v_sig2b_PI <- (2/(i^2))*(((anova(anava_PI)$Mean[2])^2)/(gl_b_ind + 2) +
((anova(anava_PI)$Mean[3])^2)/(gl_e_ind + 2))
v_sig2b_VO <- (2/(i^2))*(((anova(anava_VO)$Mean[2])^2)/(gl_b_ind + 2) +
((anova(anava_VO)$Mean[3])^2)/(gl_e_ind + 2))

```

```

##### Componentes de variância relativos ao resíduo nos 4 locais #####

v_sig2e_AS <- (2*(sig2e_AS^2))/(gl_e_ind + 2)
v_sig2e_MO <- (2*(sig2e_MO^2))/(gl_e_ind + 2)
v_sig2e_PI <- (2*(sig2e_PI^2))/(gl_e_ind + 2)
v_sig2e_VO <- (2*(sig2e_VO^2))/(gl_e_ind + 2)

### Estimativas de variâncias de estimativas de componentes de variância ###
### provenientes de análise pós-desdobramento ###
### Componentes de variância relativos à tratamentos dentro de cada local ###

v_sig2tdAS <- (2/(j^2))*(((anova(anava_AS)$Mean[1])^2)/(gl_t_cjt+2)+
((anova(anava_cjt)$Mean[5])^2)/(gl_e_cjt+2))
v_sig2tdMO <- (2/(j^2))*(((anova(anava_MO)$Mean[1])^2)/(gl_t_cjt+2)+
((anova(anava_cjt)$Mean[5])^2)/(gl_e_cjt+2))
v_sig2tdPI <- (2/(j^2))*(((anova(anava_PI)$Mean[1])^2)/(gl_t_cjt+2)+
((anova(anava_cjt)$Mean[5])^2)/(gl_e_cjt+2))
v_sig2tdVO <- (2/(j^2))*(((anova(anava_VO)$Mean[1])^2)/(gl_t_cjt+2)+
((anova(anava_cjt)$Mean[5])^2)/(gl_e_cjt+2))

##### Componente de variância relativo à blocos dentro de locais #####

v_sig2b_cjt <- (2/(i^2))*(((anova(anava_cjt)$Mean[3])^2)/(gl_b_cjt+2)+
((anova(anava_cjt)$Mean[5])^2)/(gl_e_cjt+2))

##### Componente de variância residual pós-desdobramento #####

v_sig2e_cjt <- (2*(sig2e_cjt^2))/(gl_e_cjt + 2)

#### Vetores de estimativas de componentes de variância e de estimativas ####
#### de variâncias de estimativas de componentes de variância ####
#### Vetores provenientes de análises individuais ####

sig2_ind <- c(sig2t_AS, sig2t_MO, sig2t_PI, sig2t_VO, sig2b_AS, sig2b_MO,
sig2b_PI, sig2b_VO, sig2e_AS, sig2e_MO, sig2e_PI, sig2e_VO)
v_sig2_ind <- c(v_sig2t_AS, v_sig2t_MO, v_sig2t_PI, v_sig2t_VO,
v_sig2b_AS, v_sig2b_MO, v_sig2b_PI, v_sig2b_VO,
v_sig2e_AS, v_sig2e_MO, v_sig2e_PI, v_sig2e_VO)

##### Vetores provenientes de análise conjunta #####

sig2_cjt <- c(sig2tdAS, sig2tdMO, sig2tdPI, sig2tdVO, sig2b_cjt, sig2b_cjt,
sig2b_cjt, sig2e_cjt, sig2e_cjt, sig2e_cjt, sig2e_cjt)
v_sig2_cjt <- c(v_sig2tdAS, v_sig2tdMO, v_sig2tdPI, v_sig2tdVO,
v_sig2b_cjt, v_sig2b_cjt, v_sig2b_cjt, v_sig2b_cjt,
v_sig2e_cjt, v_sig2e_cjt, v_sig2e_cjt, v_sig2e_cjt)

```

```
#### Tabela de estimativas de componentes de variância e de estimativas ####
#### de variâncias associadas a tais estimativas ####
```

```
saida <- data.frame(sig2_ind,v_sig2_ind,sig2_cjt,v_sig2_cjt)
```

```
##### Vetores de variância das estimativas de componentes de variância #####
##### pré e pós desdobramento, dividido por 105 a fim de facilitar #####
##### a visualização dos gráficos #####
```

```
resp_ind <- (saida$v_sig2_ind)/(10-5)
resp_cjt <- (saida$v_sig2_cjt)/(10-5)
```

```
##### Eixos x dos gráficos que vem a seguir representando os fatores #####
##### (tratamentos, blocos e resíduo) aleatórios envolvidos na #####
##### comparação pré versus pós-desdobramento #####
```

```
fatores <- c(rep(1,4),rep(2,4),rep(3,4))
munic <- rep(c(1:4),3)
```

```
##### Gráfico de componentes de variância versus estimativas de #####
##### variâncias de estimativas de componentes de variância #####
##### referentes à cultivares, bloco e residual #####
```

```
x11()
plot(c(1,3) ,c(0,5) , xaxt="n", col="white",
xlab= "Componentes de variância referentes à
cultivares - CVC, blocos - CVB e resíduos - CVR",
ylab= "Estimativas de variâncias das estimativas de
componentes de variância (*1/100000)",
main= "Componentes de variância versus estimativas de variâncias
de estimativas de componentes de variância referentes à
cultivares, blocos e resíduos")
axis(1, at=1:3, labels=c("CVC" ,"CVB", "CVR"))
points(fatores, resp_ind, pch=15, col="blue")
points(fatores, resp_cjt, pch=15, col="red")
legend("top", bty = "n", legend = c("Estimativas de variâncias
obtidas pela AI", "Estimativas de variâncias obtidas pela APD"),
col = c("blue", "red"), pch=15)
```

```
##### Gráfico de componentes de variância versus estimativas de #####  
##### variâncias de estimativas de componentes de variância #####  
##### referentes à Assis, Mococa, Pindamonhangaba e Votuporanga #####  
  
x11()  
plot(c(1,4) ,c(0,5) , xaxt="n", col="white",  
xlab= "Componentes de variância referentes à Assis - CVAS, Mococa - CVMO,  
Pindamonhangaba - CVPI e Votuporanga - CVVO",  
ylab= "Estimativas de variâncias das estimativas de  
componentes de variância (*1/100000)",  
main= "Componentes de variância versus estimativas de variâncias  
de estimativas de componentes de variância referentes à  
Assis, Mococa, Pindamonhangaba e Votuporanga")  
axis(1, at=1:4, labels=c("CVAS" ,"CVMO", "CVPI", "CVVO"))  
points(munic, resp_ind, pch=15, col="blue")  
points(munic, resp_cjt, pch=15, col="red")  
legend("top", bty = "n", legend = c("Estimativas de variâncias  
obtidas pela AI", "Estimativas de variâncias obtidas pela APD"),  
col = c("blue", "red"), pch=15)  
#####
```

ANEXO B - Programa (R) para simulação de 1000 grupos de experimentos contendo 4 locais cada. Cada local constitui um delineamento em blocos ao acaso com 16 tratamentos (os mesmos 16 nos 4 locais), e 4 blocos. Desta forma, uma comparação entre as diferenças mínimas significativas - dms calculadas segundo Tukey, antes e após o desdobramento da interação significativa tratamentos versus locais é efetuada

```
#### Escolha de "semente" para possibilitar a reprodução dos resultados ####
set.seed(196)

##### Número de tratamentos i = 16, número de blocos j = 4, #####
##### média geral m = 20, número de locais L = 4 #####

i <- 16
j <- 4
m <- 200
L <- 4

#### Desvios-padrão para sorteio dos efeitos de tratamentos, #####
#### blocos e resíduo, respectivamente #####

dp_t <- 2
dp_b <- 1
dp_r <- 1

##### Criando coluna com os índices de tratamentos #####

a=NULL
a.p=NULL
for(k in 1:i){
a=rep(k,j)
a.p=c(a.p,a)}
trat <- factor(a.p)

##### Criando coluna com os índices de blocos #####

g <- seq(1,j,1)
bloco <- factor(rep(g,i))
```



```
##### Definindo as variáveis de interesse que guardarão valores a cada #####
##### novo conjunto de dados simulado: "int_sig" guarda o número de #####
##### interações TxL significativas, "dmsL1" guarda o número de #####
##### vezes que dms das análises individuais é menor que o das #####
##### análises pós-desdobramento no local 1, assim como nos outros #####
##### 3 locais, "qmrn" guarda o quadrado médio do resíduo médio de #####
##### cada grupo de experimentos gerado #####
```

```
int_sig=NULL
dmsL1=NULL
dmsL2=NULL
dmsL3=NULL
dmsL4=NULL
qmrn=NULL
```

```
##### F tabelado para TxL #####
```

```
F_tab <- qf(0.95, (i-1)*(L-1), (i-1)*(j-1)*L)
```

```
##### Cálculo das dms para o Teste de Tukey #####
```

```
qtab_pre <- qtukey(0.95, i, (i-1)*(j-1))
qtab_pos <- qtukey(0.95, i, (i-1)*(j-1)*L)
```

```
##### Início da função responsável pela geração e manutenção de alguns #####
##### resultados de interesse #####
```

```
for(w in 1:1000){
```

```
##### LOCAL 1 #####
```

```
##### Efeitos dos tratamentos (soma dos efeitos de trat. deve ser nula) #####
```

```
ef_t_1 <- rnorm(i-1,0,dp_t)
s0_t_L1 <- -sum(ef_t_1)
z1 <- rep(c(ef_t_1,s0_t_L1),j)
x1=NULL
x1.y1=NULL
for(s1 in 1:i){
x1=rep(z1[s1],j)
x1.y1=c(x1.y1,x1)}
ef_t_L1 <- (x1.y1)
```

```
##### Efeitos dos blocos (soma dos efeitos de blocos deve ser nula) #####
```

```
ef_b_1 <- rnorm(j-1,0,dp_b)
s0_b_L1 <- -sum(ef_b_1)
ef_b_L1 <- rep(c(ef_b_1,s0_b_L1),i)
```

```
##### Efeitos residuais (soma dos efeitos residuais deve ser nula) #####
```

```
erro_1 <- rnorm(i*j-1,0,dp_r)
s0_e_L1 <- -sum(erro_1)
erro_L1 <- c(erro_1,s0_e_L1)
```

```

##### Dados simulados para o local 1 #####

media <- rep(m,i*j)
y_L1 <- round(media + ef_t_L1 + ef_b_L1 + erro_L1, 2)
dados_L1 <- data.frame(trat, bloco, y_L1)

##### Análise de variância individual (local 1) #####

an_L1 <- aov(y_L1 ~ bloco + trat, data = dados_L1)
summary(an_L1)

##### LOCAL 2 #####
##### Efeitos dos tratamentos (soma dos efeitos de trat. deve ser nula) #####

ef_t_2 <- rnorm(i-1,0,dp_t)
s0_t_L2 <- -sum(ef_t_2)
z2 <- rep(c(ef_t_2,s0_t_L2),j)
x2=NULL
x2.y2=NULL
for(s2 in 1:i){
x2=rep(z2[s2],j)
x2.y2=c(x2.y2,x2)}
ef_t_L2 <- (x2.y2)

##### Efeitos dos blocos (soma dos efeitos de blocos deve ser nula) #####

ef_b_2 <- rnorm(j-1,0,dp_b)
s0_b_L2 <- -sum(ef_b_2)
ef_b_L2 <- rep(c(ef_b_2,s0_b_L2),i)

##### Efeitos residuais (soma dos efeitos residuais deve ser nula) #####

erro_2 <- rnorm(i*j-1,0,dp_r)
s0_e_L2 <- -sum(erro_2)
erro_L2 <- c(erro_2,s0_e_L2)

##### Dados simulados para o local 2 #####

media <- rep(m,i*j)
y_L2 <- round(media + ef_t_L2 + ef_b_L2 + erro_L2, 2)
dados_L2 <- data.frame(trat, bloco, y_L2)

##### Análise de variância individual (local 2) #####

an_L2 <- aov(y_L2 ~ bloco + trat, data = dados_L2)
summary(an_L2)

```

```
##### LOCAL 3 #####
##### Efeitos dos tratamentos (soma dos efeitos de trat. deve ser nula) #####

ef_t_3 <- rnorm(i-1,0,dp_t)
s0_t_L3 <- -sum(ef_t_3)
z3 <- rep(c(ef_t_3,s0_t_L3),j)
x3=NULL
x3.y3=NULL
for(s3 in 1:i){
x3=rep(z3[s3],j)
x3.y3=c(x3.y3,x3)}
ef_t_L3 <- (x3.y3)

##### Efeitos dos blocos (soma dos efeitos de blocos deve ser nula) #####

ef_b_3 <- rnorm(j-1,0,dp_b)
s0_b_L3 <- -sum(ef_b_3)
ef_b_L3 <- rep(c(ef_b_3,s0_b_L3),i)

##### Efeitos residuais (soma dos efeitos residuais deve ser nula) #####

erro_3 <- rnorm(i*j-1,0,dp_r)
s0_e_L3 <- -sum(erro_3)
erro_L3 <- c(erro_3,s0_e_L3)

##### Dados simulados para o local 3 #####

media <- rep(m,i*j)
y_L3 <- round(media + ef_t_L3 + ef_b_L3 + erro_L3, 2)
dados_L3 <- data.frame(trat, bloco, y_L3)

##### Análise de variância individual (local 3) #####

an_L3 <- aov(y_L3 ~ bloco + trat, data = dados_L3)
summary(an_L3)

##### LOCAL 4 #####
##### Efeitos dos tratamentos (soma dos efeitos de trat. deve ser nula) #####

ef_t_4 <- rnorm(i-1,0,dp_t)
s0_t_L4 <- -sum(ef_t_4)
z4 <- rep(c(ef_t_4,s0_t_L4),j)
x4=NULL
x4.y4=NULL
for(s4 in 1:i){
x4=rep(z4[s4],j)
x4.y4=c(x4.y4,x4)}
ef_t_L4 <- (x4.y4)

##### Efeitos dos blocos (soma dos efeitos de blocos deve ser nula) #####

ef_b_4 <- rnorm(j-1,0,dp_b)
s0_b_L4 <- -sum(ef_b_4)
ef_b_L4 <- rep(c(ef_b_4,s0_b_L4),i)
```

```

##### Efeitos residuais (soma dos efeitos residuais deve ser nula) #####

erro_4 <- rnorm(i*j-1,0,dp_r)
s0_e_L4 <- -sum(erro_4)
erro_L4 <- c(erro_4,s0_e_L4)

##### Dados simulados para o local 4 #####

media <- rep(m,i*j)
y_L4 <- round(media + ef_t_L4 + ef_b_L4 + erro_L4, 2)
dados_L4 <- data.frame(trat, bloco, y_L4)

##### Análise de variância individual (local 4) #####

an_L4 <- aov(y_L4 ~ bloco + trat, data = dados_L4)
summary(an_L4)

##### Análise Conjunta: #####

loc_cjt <- factor(c(rep(1,i*j),rep(2,i*j),rep(3,i*j),rep(4,i*j)))
trat_cjt <- rep(trat, L)
bloco_cjt <- rep(bloco, L)

y <- c(y_L1,y_L2,y_L3,y_L4)
dados <- data.frame(loc_cjt, trat_cjt, bloco_cjt, y)

an_cjt <- aov(y ~ loc_cjt/bloco_cjt + trat_cjt + loc_cjt + trat_cjt:loc_cjt,
data=dados)
summary(an_cjt)

##### Quadrado médio do resíduo médio #####

qrmr[w] <- anova(an_cjt)$Mean[5]
qrmr.draw=qrmr

##### Diferenças mínimas significativas obtidas via análises individuais #####

delta_pre_L1 <- qtab_pre*sqrt(anova(an_L1)$Mean[3]/j)
delta_pre_L2 <- qtab_pre*sqrt(anova(an_L2)$Mean[3]/j)
delta_pre_L3 <- qtab_pre*sqrt(anova(an_L3)$Mean[3]/j)
delta_pre_L4 <- qtab_pre*sqrt(anova(an_L4)$Mean[3]/j)

##### Diferença mínima significativa obtida via análise conjunta #####
##### utilizando o quadrado médio do resíduo médio - QMRM #####

delta_pos <- qtab_pos*sqrt(anova(an_cjt)$Mean[5]/j)

##### Verificando se TxL é significativo (1) ou não (0) #####

int_sig[w] <- ifelse(F_tab < anova(an_cjt)$F[4], 1, 0)
int_sig.draw=int_sig[w]

```

```
##### Verificando se a dms alcançada pré-desdobramento é menor (1) #####  
##### ou maior (0) que a obtida via análise pós-desdobramento #####  
##### em cada um dos 4 locais #####
```

```
dmsL1[w] <- ifelse(delta_pre_L1 < delta_pos, 1, 0)  
dmsL1.draw=dmsL1[w]  
dmsL2[w] <- ifelse(delta_pre_L2 < delta_pos, 1, 0)  
dmsL2.draw=dmsL2[w]  
dmsL3[w] <- ifelse(delta_pre_L3 < delta_pos, 1, 0)  
dmsL3.draw=dmsL3[w]  
dmsL4[w] <- ifelse(delta_pre_L4 < delta_pos, 1, 0)  
dmsL4.draw=dmsL4[w]
```

```
}
```

```
##### A seguir, têm-se respectivamente, a média dos QMRM dos 1000 #####  
##### grupos de experimentos, a quantidade total de interações T x L #####  
##### significativas, e, a quantidade total de valores de dms obtidos #####  
##### via análises individuais menores que os alcançados via análise #####  
##### conjunta, dentre os 4000 possíveis #####
```

```
med_qrm <- mean(qrm)  
n_int_sig <- sum(int_sig)  
n_dms_pre_menor_pos <- sum(dmsL1,dmsL2,dmsL3,dmsL4)
```

```
med_qrm  
n_int_sig  
n_dms_pre_menor_pos
```