

## Gráficos *biplot* e *joint plot* para o estudo da interação tripla

### Biplot and joint plot graphics to study the triple interaction

Lúcio Borges de Araújo<sup>I</sup> Mario Varela Nualles<sup>II</sup>  
Mirian Fernandes Carvalho Araújo<sup>III</sup>  
Carlos Tadeu dos Santos Dias<sup>IV</sup>

#### RESUMO

*Este trabalho teve os seguintes objetivos: propor uma sistemática para o estudo e a interpretação da estabilidade e adaptabilidade fenotípica, por meio do modelo de Tucker3; avaliar o uso do gráfico joint plot; e comparar os resultados obtidos pela metodologia de três entradas com os resultados obtidos pelos modelos AMMI e pela análise biplot. Os dados utilizados são relativos a experimentos com 13 genótipos de feijão conduzidos em nove experimentos. Os resultados indicaram que o gráfico joint plot facilita o entendimento da interação tripla e traz ao pesquisador informações mais reais sobre a interação tripla do que a modelagem AMMI de duas entradas; o genótipo 6 foi o que menos contribuiu para a interação, e os genótipos 12, 9 e 5 foram os que mais contribuíram para a interação.*

**Palavras-chave:** interação genótipos  $\times$  ambientes  $\times$  anos, modelo Tucker3, estabilidade.

#### ABSTRACT

*The present research had the following objectives: to propose a systematics for the study and the interpretation of the phenotypic stability and adaptability, through Tucker3 models; assess the use of the joint plot; comparing the results obtained by Tucker3 model with the results obtained by biplot analysis and AMMI models. The data used are related to the experiments with 13 genotypes of beans that had been lead in 9 distinct experiments. The results indicated that the joint plot graph facilitates the understanding of triple interaction and bring to the researcher more real information about triple interaction, than AMMI modeling of two entries. The genotype 6 was the one which least contributed for the triple interaction*

*and genotypes 12, 9 and 5 were those e which had contributed more for the interaction.*

**Key words:** genotypes  $\times$  locations  $\times$  years interaction, Tucker3 model, stability.

#### INTRODUÇÃO

Os experimentos multiambientais (MET) são conduzidos através de vários anos para os principais produtos agrícolas no mundo, constituindo um passo caro, mas essencial para a liberação de um novo genótipo de um produto agrícola e, conseqüentemente, a recomendação de cultivar. Por isso, melhoria nos métodos usados para análise de dados deve ser de interesse da comunidade de melhoristas. O objetivo primário de um MET é identificar cultivares superiores. A prática mais comum usada para esse fim é comparar o rendimento de um genótipo em vários ambientes de teste (normalmente combinações de locais e anos). A segunda utilidade de análise de dados multiambientais, antes de fazer a avaliação de cultivares, deveria ser investigar as relações entre os ambientes de teste e a possibilidade de diferenciação do mega-ambiente (YAN & HUNT, 2002).

Para a descrição da resposta média de genótipos em ambientes e para o estudo e interpretação

<sup>I</sup>Departamento de Ciências Exatas (LCE), Universidade de São Paulo (USP), Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Piracicaba, SP, Brasil. Faculdade de Matemática (FAMAT), Universidade Federal de Uberlândia (UFU), Uberlândia, MG, Brasil.

<sup>II</sup>Departamento de Matemática (DEMAT), Instituto Nacional de Ciências Agrícolas (INCA), Havana, Cuba.

<sup>III</sup>FAMAT, UFU, Uberlândia, MG, Brasil.

<sup>IV</sup>LCE, USP, ESALQ. Av. Pádua Dias, 11, 13418-900, Piracicaba, SP, Brasil. E-mail: ctsdias@esalq.usp.br. Autor para correspondência

interação genótipo x ambiente (GE), em METs de experimentos agrícolas, duas classes de modelos são comumente utilizadas: modelos lineares e modelos lineares-bilineares. A princípio, as abordagens para a análise da interação GE incluem a apresentação dos dados em tabela de duas entradas (matriz), e cada casela dessa tabela contém a resposta média de cada genótipo em cada ambiente.

Considerando o caso em que os METs são avaliados através de vários anos (ou seja, genótipos x locais x anos) (GLA), os dados podem ser organizados em arranjo de três entradas que, neste caso, cada entrada se refere a genótipos, locais e anos.

Em alguns casos, o investigador pode estar interessado em saber se existe uma estrutura comum encoberta pelos locais com relação aos anos e como os vários genótipos respondem através da estrutura formada por ambientes e anos. Alguns genótipos podem responder com altas respostas em alguns locais, mas não em outros, e alguns locais podem estar mais associados com alguns genótipos do que a outros por alguns anos. Um procedimento para ganhar uma compreensão clara em arranjo GLA de três entradas é determinar uma estrutura dimensional menor, expressada em componentes principais, para a interação genótipos x locais x anos e então estudar as relações entre esses componentes. Essa aproximação é mais útil que combinar dois dos três fatores de maneira que os dados formem um arranjo de duas entradas. Outro procedimento menos útil é excluir um fator diretamente (por exemplo anos) e analisar um arranjo de duas entradas dos genótipos x locais em cada ano e, nesse caso, o problema está em encontrar uma interpretação global para os anos.

Para os dados organizados em arranjo de três entradas, existem alguns modelos para analisá-los, como, por exemplo, os modelos propostos por TUCKER (1966), que fornecem uma decomposição trilinear dos dados organizados no arranjo.

Na maioria dos estudos, em razão da falta de uma ferramenta adequada para estudar a interação entre genótipos, locais e anos, os pesquisadores combinam os fatores locais e anos (criando outro fator, denominado ambiente) e fazem o estudo da interação genótipos x ambientes utilizando uma decomposição em valores singulares, chamada de modelos AMMI (Additive main effects and multiplicative interaction) por GAUCH & ZOBEL (1988). No entanto, de acordo com VARELA et. al. (2006), essa combinação leva a uma perda de informação quando se ajusta um modelo de duas entradas e quando se faz a estimação dos efeitos da interação. Então, faz-se necessário o uso ferramentas que permitam o desdobramento e a interpretação da interação tripla.

Assim, o trabalho tem os objetivos de propor uma sistemática para o estudo e a interpretação da estabilidade e adaptabilidade fenotípica, por meio do modelo de Tucker3, e avaliar o uso do gráfico joint plot (KROONENBERG 1983).

## MATERIAL E MÉTODOS

Os dados utilizados são relativos a experimentos com 13 genótipos de feijão conduzidos em nove experimentos distintos e constituídos pelos anos agrícolas de 2000/2001, 2001/2002 e 2005/2006, nos Municípios de Dourados e Aquidauana. Esses experimentos foram instalados na época das águas (Dourados) e também na época da seca (Dourados e Aquidauana). Cada local é constituído de um município e uma época de instalação. Além disso, em cada experimento, foi utilizado um delineamento aleatorizado em blocos, com três blocos em cada experimento. Neste trabalho, foi considerada a variável produtividade de grãos, em t ha<sup>-1</sup>.

O modelo AMMI pressupõe componentes aditivos para os efeitos principais de genótipos e ambientes e componentes multiplicativos para o efeito de interação. Então, a resposta média sobre  $b$  blocos do  $i$ -ésimo genótipo no  $j$ -ésimo ambiente é representada

$$\text{por: } Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^q \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij} + \varepsilon_{ij}, \text{ em}$$

que:  $Y_{ij}$ : é a resposta média do  $i$ -ésimo genótipo no  $j$ -ésimo ambiente, com  $i = 1, 2, \dots, g$  e  $j = 1, 2, \dots, e$ ;  $\mu$  é uma constante, geralmente a média;  $g_i$ : é o efeito do  $i$ -ésimo genótipo;  $e_j$ : é o efeito do  $j$ -ésimo ambiente;  $\lambda_k$ : é a raiz quadrada do  $k$ -ésimo autovalor da matriz  $(\mathbf{GE})(\mathbf{GE})^t$  (ou  $(\mathbf{GE})^t(\mathbf{GE})$ ), com  $k = 1, 2, \dots, q$  e em que  $q < p$  determina uma aproximação de mínimos quadrados para a matriz  $\mathbf{GE}$  pelos  $q$  primeiros termos da DVS e  $p = \min\{g-1, e-1\}$ ;  $\alpha_{ik}$ : é o  $i$ -ésimo elemento do vetor coluna  $\alpha_k$  associado a  $\lambda_k$ ;  $\gamma_{jk}$ : é o  $j$ -ésimo elemento do vetor linha  $\gamma_k$  associado a  $\lambda_k$ ;  $\rho_{ij}$ : é o resíduo adicional;  $\varepsilon_{ij}$ : é erro experimental associado ao  $i$ -ésimo genótipo no  $j$ -ésimo ambiente, sendo assumido como independente e

$$\varepsilon_{ij} \sim N(0, \frac{\sigma^2}{b}); (ge)_{ij} = \sum_{k=1}^q \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij}.$$

A matriz  $\mathbf{GE}^{k \times j}$  é a interação entre genótipos x ambientes (matriz de resíduos), em que cada elemento  $(ge)_{ij}$  é encontrado pela seguinte relação:  $(ge)_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$ , em que:  $Y_{ij}$ : é a média das repetições do genótipo  $i$  no ambiente  $j$ , com  $i = 1, 2, \dots, g$  e  $j = 1, 2, \dots, e$ ;  $\bar{Y}_{i.}$ : é a média do genótipo  $i$ ;  $\bar{Y}_{.j}$ : é a média do ambiente  $j$ ;  $\bar{Y}_{..}$ : é a média geral do experimento.

Uma possível generalização do modelo de componentes principais para dados de duas entradas

é para um arranjo de três entradas  $\underline{\mathbf{X}}$ , com elementos

$x_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk}$ , em que  $e_{ijk}$  é um elemento do arranjo  $\underline{\mathbf{E}}$  ( $I \times J \times K$ );  $a_{ip}$ ,  $b_{jq}$  e  $c_{kr}$  são elementos típicos das matrizes de cargas  $\mathbf{A}$  ( $I \times P$ ),  $\mathbf{B}$  ( $J \times Q$ ) e  $\mathbf{C}$  ( $K \times R$ ); e  $g_{pqr}$  é um elemento típico do arranjo núcleo  $\underline{\mathbf{G}}$  ( $P \times Q \times R$ ). Esse é o modelo Tucker3 de  $\mathbf{X}$  ( $P, Q, R$ ), em que  $P, Q, R$  indicam o número de componentes em cada entrada.

Uma vez fixado o número de componentes nas matrizes  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$ , a estimação dos parâmetros  $a_{ip}$ ,  $b_{jq}$  e  $c_{kr}$  do modelo de Tucker3 é realizada por meio do método iterativo de mínimos quadrados alternados, em que cada conjunto de parâmetros é estimado condicionalmente aos demais parâmetros. Assim, a estimação é repetida iterativamente até não serem observadas mudanças significativas nos valores dos parâmetros, e como solução inicial utilizaram-se os valores propostos por TUCKER (1966). Para determinar a melhor solução do modelo de Tucker3, utilizou-se o método proposto por TIMMERMAN & KIERS (2000).

O modelo AMMI para a interação tripla pressupõe componentes aditivos para os efeitos principais de genótipos, locais e anos e componentes multiplicativos para os efeitos de interações duplas e triplas. Então, a resposta média sobre  $b$  repetições ou blocos do  $i$ -ésimo genótipo no  $j$ -ésimo ambiente é representada por:

$$Y_{ijk} = \mu + g_i + l_j + a_k + (gl)_{ij} + (ga)_{ik} + (la)_{jk} + \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr} g_{pqr} + \rho_{ijk} + \varepsilon_{ijk}$$

em que:  $Y_{ijk}$ : é a resposta média do  $i$ -ésimo genótipo no  $j$ -ésimo local, no  $k$ -ésimo ano, com  $i = 1, 2, \dots, g$ ,  $j = 1, 2, \dots, l$  e  $k = 1, 2, \dots, a$ ;  $\mu$ : é uma constante, geralmente a média;  $g_i$ : é o efeito do  $i$ -ésimo genótipo;  $l_j$ : é o efeito do  $j$ -ésimo local;  $a_k$ : é o efeito do  $k$ -ésimo ano;  $(gl)_{ij}$ : é o efeito da interação do  $i$ -ésimo genótipo com o  $j$ -ésimo local;  $(ga)_{ik}$ : é o efeito da interação do  $i$ -ésimo genótipo com o  $k$ -ésimo ano;  $(la)_{jk}$ : é o efeito da interação do  $j$ -ésimo local com o  $k$ -ésimo ano;  $a_{ip}$ : é o  $i$ -ésimo elemento da  $p$ -ésima coluna da matriz de componentes  $\mathbf{A}$ ;  $b_{jq}$ : é o  $j$ -ésimo elemento da  $q$ -ésima coluna da matriz de componentes  $\mathbf{B}$ ;  $c_{kr}$ : é o  $k$ -ésimo elemento da  $r$ -ésima coluna da matriz de componentes  $\mathbf{C}$ ;  $\rho_{ijk}$ : é o resíduo adicional da interação tripla que não é explicado pelo modelo Tucker3 ( $P, Q, R$ );  $(gla)_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr} g_{pqr} + \rho_{ijk}$ ;  $\varepsilon_{ijk}$ : é erro experimental associado ao  $i$ -ésimo genótipo no  $j$ -ésimo local, no  $k$ -ésimo ano, assumindo-se ser  $\varepsilon_{ijk} \sim N(0, \frac{\sigma^2}{b})$  e todos os  $\varepsilon_{ijk}$  independentes;

Seja  $\underline{\mathbf{Z}}$  um arranjo com as interações entre genótipos x locais x anos (arranjo de resíduos), de modo

que cada elemento  $(gla)_{ijk}$  do arranjo de três entradas  $\underline{\mathbf{Z}}$  é estimado como:

$(gla)_{ijk} = Y_{ijk} - \bar{Y}_{ij.} - \bar{Y}_{i.k} - \bar{Y}_{.jk} + \bar{Y}_{i..} + \bar{Y}_{.j.} + \bar{Y}_{..k} - \bar{Y}_{...}$ , em que:  $(gla)_{ijk}$ : é o efeito da interação tripla estimada para o genótipo  $i$  no local  $j$  e no ano  $k$ ;  $Y_{ijk}$ : é a média das  $b$  repetições do genótipo  $i$  no local  $j$  e no ano  $k$ ;  $\bar{Y}_{ij.}$ : é a média dos elementos da  $i$ -ésima linha com a  $j$ -ésima coluna do arranjo de interação;  $\bar{Y}_{i.k}$ : é a média dos elementos da  $i$ -ésima linha com o  $k$ -ésimo tubo do arranjo de interação;  $\bar{Y}_{.jk}$ : é a média dos elementos da  $j$ -ésima coluna com o  $k$ -ésimo tubo do arranjo de interação;  $\bar{Y}_{i..}$ : é a média dos elementos da  $i$ -ésima fatia horizontal do arranjo de interação;  $\bar{Y}_{.j.}$ : é a média dos elementos da  $j$ -ésima fatia vertical do arranjo de interação;  $\bar{Y}_{..k}$ : é a média dos elementos da  $k$ -ésima fatia frontal do arranjo de interação;  $\bar{Y}_{...}$ : é a média geral do experimento.

*Biplot* é uma representação gráfica em que as linhas e as colunas são apresentadas em um gráfico com duas ou três dimensões, e a construção do *biplot* é baseada na decomposição em valores singulares (DVS) da matriz de dados (GABRIEL, 1971). Assim, para o caso da interação entre genótipos e ambientes, é possível representar num único gráfico esses efeitos, fazendo a DVS da matriz residual  $\mathbf{GE}$ .

Um *joint biplot* (KROONENBERG, 1983) é utilizado para representar graficamente os modelos de Tucker3, sendo semelhante a um *biplot* padrão e todos os princípios de interpretação do *biplot* padrão podem ser utilizados. A diferença nessa construção é que o *joint biplot* é construído como um *biplot* para dois fatores dada a matriz de componente do modelo Tucker3 referente ao terceiro fator (terceira entrada), ou seja, cada *joint biplot* é construído usando diferentes fatias do arranjo núcleo. Para construir um *joint biplot* após ajustar um modelo de Tucker3, é necessário obter uma matriz  $\mathbf{D}_r = \mathbf{A} \mathbf{G}_r \mathbf{B}' = \mathbf{A}_r^* \mathbf{B}_r^{**}$  de dimensão  $I \times J$ , com  $r = 1, 2, \dots, R$  e então, por meio da DVS, representa-se  $\Delta_r$  por meio de um *biplot*, obtendo-se então o *joint plot* (KROONENBERG, 1994). Para cada fatia do núcleo,  $\mathbf{G}_r$  é necessário construir um *joint biplot* para a matriz de componentes  $\mathbf{A}^*$  ( $J \times P$ ) e  $\mathbf{B}^*$  ( $J \times Q$ ).

Quanto à interpretação de um *joint plot* (VARELA, et al., 2006), suponha um gráfico que é projetado sobre o  $r$ -ésimo componente principal da terceira entrada, tal que no *joint plot* aparecem todos os níveis das duas primeiras entradas. Em seguida, selecione, a partir de matrizes  $\mathbf{C}$  (matriz das componentes principais da terceira entrada), os níveis desse fator com maior peso no  $r$ -ésimo componente (positivos ou negativos). Suponha que a matriz  $\mathbf{C}$  tem um valor positivo e elevado associado ao  $k$ -ésimo nível da terceira entrada, então proximidades entre os níveis

da primeira e da segunda entrada indicam que a interação tripla entre o  $i$ -ésimo nível da primeira entrada,  $j$ -ésimo nível da segunda entrada e  $k$ -ésimo nível da terceira entrada é positiva. Em contrapartida, se o  $i$ -ésimo nível do primeiro fator está muito longe do  $j$ -ésimo nível do segundo fator, isso indica que a interação tripla associada a esses três fatores é negativa. Caso a matriz **C** tenha um valor alto negativo associado ao  $k$ -ésimo nível do terceiro fator, as interações triplas serão o oposto de quando o valor é positivo. Em geral, os níveis de uma entrada localizada no centro do *joint plot* são considerados um conjunto que tem um desempenho médio em todos os outros modos.

## RESULTADOS E DISCUSSÃO

A análise de variância conjunta considerando 13 genótipos e nove ambientes (combinação de três locais e três anos) mostra que a interação genótipos  $\times$  ambientes é significativa ( $SQ_{G \times E} = 58,64$ , 96 graus de liberdade) e representa 28 % da variabilidade total dos dados. A soma de quadrados do resíduo é de 8,51, com 216 graus de liberdade.

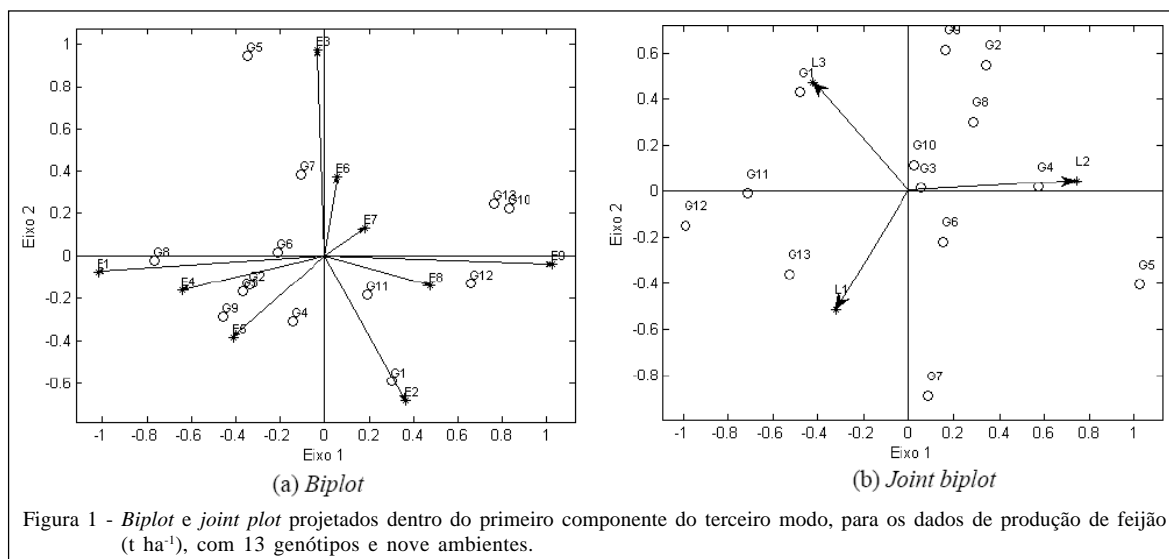
A matriz de interações **GE** é apresentada na tabela 1 e a próxima etapa da análise correspondem ao ajuste da interação pela decomposição em valor singular

aplicada à matriz **GE**. A matriz **GE** terá posto  $p = \min(12,8) = 8$ , conseqüentemente a  $SQ_{G \times E}$  pode ser decomposta em até oito componentes. Assim, um modelo AMMI com dois componentes explica 64,5% da soma de quadrados da interação entre genótipos e ambientes, como resposta-padrão, e 35,5% dessa soma de quadrados é de ruídos presentes nos dados. A última etapa da análise AMMI consiste na representação gráfica dos genótipos e ambientes em um gráfico denominado de *biplot*. Para o conjunto de dados em questão (que é composto por 13 genótipos e nove ambientes), um *biplot* de duas dimensões representará 64,5% da  $SQ_{G \times E}$ .

A figura 1a ilustra o biplot resultante e a partir dele são feitas as devidas interpretações a fim de identificar genótipos e ambientes que menos contribuem para a interação entre genótipos  $\times$  ambientes. Logo, por esse gráfico, nota-se que os genótipos que menos contribuíram para a interação (pontos próximos da origem, que indicam quais são genótipos estáveis) foram os genótipos  $G_2, G_3, G_4, G_6, G_7$  e  $G_{11}$ ; mas, para fins de recomendação de cultivares, deseja-se um alto desempenho na produtividade, que pode ser avaliada pelas médias (DUARTE; VENKOVSKY, 1999). Assim, entre esses genótipos destacam-se os genótipos  $G_2$  e  $G_3$ , que tiveram a primeira e a quarta maior média, respectivamente, em

Tabela 1 - Valores estimados da interação dupla de 13 genótipos e nove ambientes (combinação de três locais e três anos) e da interação tripla dos 13 genótipos, três locais e três anos, para a produção, em ton ha<sup>-1</sup>.

Ambiente/ Genótipo	-----Interação dupla-----									-----Interação tripla-----								
	$E_1$	$E_2$	$E_3$	$E_4$	$E_5$	$E_6$	$E_7$	$E_8$	$E_9$	$A_1 L_1$	$A_1 L_2$	$A_1 L_3$	$A_2 L_1$	$A_2 L_2$	$A_2 L_3$	$A_3 L_1$	$A_3 L_2$	$A_3 L_3$
$G_1$	-0,41	0,47	-0,53	-0,03	0,27	-0,09	-0,31	0,33	0,30	0,00	0,27	-0,27	0,17	-0,14	-0,04	-0,17	-0,14	0,30
$G_2$	0,66	-0,21	-0,21	0,17	0,03	-0,25	-0,04	0,00	-0,15	0,31	-0,23	-0,08	-0,08	0,11	-0,03	-0,24	0,12	0,11
$G_3$	0,29	0,07	-0,04	0,13	0,53	-0,07	-0,12	-0,47	-0,31	0,09	-0,08	-0,01	-0,17	0,29	-0,12	0,08	-0,21	0,13
$G_4$	-0,06	-0,25	-0,28	0,14	0,59	-0,33	-0,08	0,58	-0,32	0,14	-0,36	0,23	0,00	0,15	-0,15	-0,14	0,21	-0,08
$G_5$	0,20	-0,80	0,94	-0,05	0,05	0,62	-0,31	-0,22	-0,42	0,14	-0,59	0,45	-0,20	0,17	0,04	0,06	0,42	-0,48
$G_6$	-0,01	-0,17	-0,01	0,02	0,51	0,16	0,04	-0,32	-0,23	0,04	-0,12	0,08	-0,23	0,27	-0,05	0,19	-0,16	-0,03
$G_7$	0,13	-0,27	0,35	0,22	-0,54	-0,38	0,85	-0,13	-0,23	-0,34	-0,03	0,37	0,05	0,01	-0,06	0,29	0,02	-0,31
$G_8$	0,77	-0,03	0,26	0,76	0,20	-0,23	-0,56	-0,61	-0,56	0,11	-0,21	0,10	0,20	0,10	-0,30	-0,31	0,11	0,19
$G_9$	0,79	-0,07	-0,66	0,30	-0,30	0,14	0,25	0,19	-0,64	0,33	-0,03	-0,30	-0,19	-0,29	0,48	-0,13	0,32	-0,19
$G_{10}$	-0,73	-0,30	-0,08	-0,63	-0,40	0,24	0,19	0,85	0,87	0,03	0,02	-0,05	0,02	-0,19	0,16	-0,06	0,17	-0,11
$G_{11}$	-0,23	0,68	-0,13	-0,30	-0,07	0,41	-0,10	-0,28	0,01	-0,13	0,46	-0,33	-0,10	-0,20	0,30	0,23	-0,27	0,04
$G_{12}$	-0,56	0,54	0,04	-0,31	-0,33	-0,37	0,23	-0,23	1,00	-0,35	0,54	-0,19	0,24	0,01	-0,25	0,11	-0,56	0,44
$G_{13}$	-0,85	0,33	0,35	-0,43	-0,54	0,14	-0,03	0,32	0,70	-0,36	0,35	0,01	0,28	-0,30	0,02	0,08	-0,05	-0,03



termos da produtividade, enquanto que os genótipos  $G_4$ ,  $G_6$ ,  $G_7$  e  $G_{11}$ , que também são estáveis, tiveram algumas das piores médias. Os demais genótipos tiveram adaptações específicas a determinados ambientes (coordenadas dos genótipos estão próximas a coordenadas dos ambientes), ou seja,  $G_1$  adaptou-se especificamente ao ambiente  $E_2$ ,  $G_5$  adaptou-se bem ao ambiente  $E_3$ ,  $G_8$  adaptou-se aos ambientes  $E_1$  e  $E_4$ ,  $G_9$  adaptou-se aos ambientes  $E_4$  e  $E_5$ ,  $G_{10}$  adaptou-se ao ambiente  $E_9$ ,  $G_{12}$  adaptou-se aos ambientes  $E_8$  e  $E_9$  e  $G_{13}$  adaptou-se ao ambiente  $E_9$ .

Por outro lado, o ordenamento dos genótipos num ambiente estável será de maior confiança para o melhorista. Assim, entre os ambientes destacam-se  $E_6$  (combinação do local Aquidauana na época das “secas” com o ano de 2001/2002),  $E_7$  (combinação do local Dourados na época das “águas” com o ano de 2005/2006) e  $E_8$  (combinação do local Dourados na época das “secas” com o ano de 2005/2006) como ambientes estáveis, enquanto que os demais locais são não estáveis.

A análise de variância conjunta mostra que o efeito da interação tripla (genótipos x locais x anos) é significativa ( $SQ_{G \times L \times A} = 18,27$  com 48 graus de liberdade e a soma de quadrados do resíduo é de 8,51, com 216 graus de liberdade).

A grande vantagem desse modelo, com relação aos outros modelos multiplicativos de duas entradas, é a possibilidade do estudo simultâneo de diversos fatores. Isso significa, por exemplo, que é possível fazer uma decomposição da interação tripla entre genótipos x locais x anos, tornando as conclusões mais precisas e reais do que aquelas obtidas com modelos multiplicativos para duas entradas. Dessa

forma, é possível construir um arranjo cúbico de dimensão (13x3x3) com os efeitos das interações triplas entre genótipos x locais x anos, de modo que, nas linhas, estão os genótipos, nas colunas estão os locais e, nos tubos, estão os anos. Esses efeitos das interações triplas são apresentados na tabela 1.

Na comparação dos resultados estimados da interação dupla com os efeitos da interação tripla, percebe-se uma das principais diferenças entre os modelos com três fatores e os modelos com dois fatores. Os resultados da análise realizada a partir de um modelo com dois fatores, sendo que um destes fatores foi obtido pela combinação de outros dois fatores (locais e anos), parecem que não são apropriados, pois os efeitos das interações mostraram-se diferentes. As principais diferenças foram encontradas no genótipo 10, no local 3 e no ano 3, em que a estimação da interação dupla foi de 0,87, e a interação tripla foi de -0,11, e essa diferença não foi refletida no *biplot* (Figura 1a). Resultados similares foram observados para outras combinações, por exemplo: genótipo 8, no local 3 e no ano 3; genótipo 10, no local 1 e no ano 1. De uma maneira geral, na comparação dessas duas tabelas com os efeitos de interação, pode-se observar que há vários valores estimados que são altos para a interação dupla e que não aparecem (estão próximos de zero) para a interação tripla; por outro lado, existem vários valores estimados que são baixos na interação dupla (próximos de zero) e na interação tripla têm valores altos. Assim, a ideia de fazer a combinação de dois fatores traz prejuízos para a análise da interação entre os fatores em questão (genótipo, local e ano), superestimando ou subestimando a interação. Portanto, é preciso utilizar uma metodologia adequada para interpretar a interação tripla.

Para selecionar o melhor modelo de Tucker3, foi utilizado o procedimento de Timmerman-Kiers, já para o conjunto de dados em questão o procedimento sugere selecionar o modelo de Tucker3 (3,2,2). Esse modelo explica 90,38% da soma de quadrados da interação tripla, e as três componentes,  $p_1$ ,  $p_2$  e  $p_3$ , da matriz **A**, explicam 52,05, 21,34 e 17,00%, respectivamente. As duas componentes,  $q_1$  e  $q_2$ , da matriz **B**, explicam 64,48 e 25,90%, respectivamente, e na matriz **C** as duas componentes  $r_1$  e  $r_2$  explicam 67,12 e 23,26%, respectivamente.

$$\begin{aligned}
 \mathbf{A} &= \begin{pmatrix} p_1 & p_2 & p_3 \\ G_1 & 0,2633 & 0,0854 & -0,2875 \\ G_2 & -0,1587 & 0,0489 & -0,4110 \\ G_3 & -0,0006 & -0,2277 & -0,1433 \\ G_4 & -0,2806 & -0,1755 & -0,1231 \\ G_5 & -0,5519 & -0,0694 & 0,2732 \\ G_6 & -0,0696 & -0,1914 & 0,0626 \\ G_7 & -0,0646 & -0,2284 & 0,5738 \\ G_8 & -0,1035 & -0,2491 & -0,3853 \\ G_9 & -0,1383 & 0,6940 & -0,0839 \\ G_{10} & -0,0376 & 0,2526 & 0,0580 \\ G_{11} & 0,3361 & 0,3268 & 0,2070 \\ G_{12} & 0,5517 & -0,3125 & -0,0566 \\ G_{13} & 0,2543 & 0,0463 & 0,3161 \end{pmatrix}; \quad \mathbf{B} = \begin{pmatrix} q_1 & q_2 \\ L_1 & \begin{pmatrix} -0,1498 & 0,8026 \\ 0,7700 & -0,2716 \\ -0,6202 & -0,5311 \end{pmatrix} \\ L_2 & \\ L_3 & \end{pmatrix}; \\
 \mathbf{C} &= \begin{pmatrix} r_1 & r_2 \\ A_1 & \begin{pmatrix} -0,7903 & 0,2051 \\ 0,2175 & -0,7870 \\ 0,5728 & 0,5819 \end{pmatrix} \\ A_2 & \\ A_3 & \end{pmatrix}; \quad \mathbf{G} = \begin{pmatrix} p_1 & p_2 & p_3 \\ G_1 & \begin{pmatrix} -1,6769 & 0,4374 & -0,1383 & -0,3829 \\ -0,2772 & -0,4625 & 0,9434 & 0,3448 \\ 0,1206 & 0,8824 & 0,3372 & 0,3579 \end{pmatrix} \\ G_2 & \\ G_3 & \end{pmatrix}
 \end{aligned}$$

O arranjo núcleo **G** apresenta as relações entre as componentes e entre esses valores a relação mais importante é entre os primeiros componentes de cada fator,  $g_{111} = -1,6769$ , sendo que essa quantia indica que a combinação do primeiro componente dos genótipos com o primeiro componente dos locais com o primeiro componente dos anos explicam juntos 46,17% da  $SQ_{G \times L \times A}$ , e a relação menos importante é a relação entre o terceiro componente dos genótipos com o primeiro componente dos locais com o primeiro componente dos anos ( $g_{311} = 0,1206$ ), que explicam juntos 0,23% da  $SQ_{G \times L \times A}$ .

Percebe-se que o primeiro componente da matriz **C** é caracterizado por um contraste entre o ano 1 (-0,7903) e o ano 3 (0,5728), e o segundo componente é caracterizado por um contraste entre o ano 2 (-0,7870) e o ano 3 (0,5819). Assim, ao ser construído um *joint plot*, que projeta os genótipos e locais dentro do primeiro componente dos anos, as conclusões serão restritas somente ao ano 1 e ao ano 3 (Figura 1b), mas, ao serem projetados os genótipos e locais dentro do segundo componente dos anos, as conclusões serão válidas para o ano 2 e ano 3.

O *joint plot* (Figura 1b) corresponde ao *biplot* da matriz  $\mathbf{A}_1 = \mathbf{A} \mathbf{G}_1 \mathbf{B}'$ , em que,

$$\mathbf{G}_1 = \begin{pmatrix} -1,6769 & 0,4374 \\ -0,2772 & -0,4625 \\ 0,1206 & 0,8824 \end{pmatrix} \quad \text{é a primeira fatia frontal do}$$

arranjo núcleo **G**, obtido ao ajustar modelo de Tucker3 (3,2,2). Esse *joint plot* é projetado dentro do primeiro componente do fator ano ( $r_1$ ) e esse componente explica 67,12 % da  $SQ_{G \times L \times E}$ , e o primeiro componente desse gráfico corresponde a 49,88, e o segundo componente explica 17,23 % da soma de quadrados da interação tripla. O gráfico representa a interação entre genótipo x ambiente no ano 1 (2000/2001) e no ano 3 (2005/2006). Assim, em relação ao ano 1 ( $c_{11}$  é negativo), há as seguintes relações:  $G_1$  teve uma interação negativa com  $L_3$  (Aquidauana na época das secas), positiva com  $L_2$  (Dourados na época da secas) e  $G_1$  não interage com o local  $L_1$  (Dourados na época das águas); em relação aos genótipos 11, 12 e 13, observa-se uma interação positiva com  $L_2$  e negativa com  $L_1$  e  $L_3$ ; os genótipos 4 e 5 tiveram uma interação positiva com  $L_1$  e  $L_3$  e negativa com  $L_2$ ; para os genótipos 2, 8 e 9, nota-se que a interação é negativa com  $L_2$  e  $L_3$  e positiva com  $L_1$ ; e o genótipo 7 teve interação negativa com  $L_1$ , positiva com  $L_3$  e não interagiu com o local  $L_2$ ; ainda nesse gráfico, com relação ao ano 3 ( $c_{31}$  é positivo), as relações são opostas das observadas no ano 1.

Assim como VARELA et al. (2006) encontraram problemas de superestimação e subestimação da interação para o modelo AMMI de duas entradas, quando se faz a combinação de dois fatores (locais e anos), neste trabalho, também foram encontrados os mesmos problemas na estimação dos efeitos da interação. Portanto, a utilização dessa técnica oferece desvantagens, ou seja, essa metodologia pode fornecer resultados e conduzir a conclusões que não são observáveis no conjunto de dados.

Diante desse problema, faz-se necessário aplicar uma metodologia adequada e conseqüentemente a extensão dos modelos AMMI de duas entradas para o modelo AMMI de três entradas, oferece uma aproximação natural para avaliar a resposta de genótipos em diferentes locais e diferentes anos.

Com relação ao outro modelo AMMI de três entradas que utilizou o modelo de Tucker3 para encontrar as matrizes **A**, **B** e **C** e depois construir o *joint plot*, pode-se relatar a seguinte vantagem: o fato de o modelo de Tucker3 ter recuperado uma alta quantidade da soma de quadrados da interação genótipos x locais x anos. Outra vantagem dessa metodologia é que os *joint plot* são fáceis de interpretar. Em relação às desvantagens, cita-se o fato de que o número de *joint plot* a ser construído é igual ao número de componentes que têm o fator que receberá a projeção e, portanto, o fator que receberá a projeção será aquele que tem o menor número de componentes, assim, à medida que aumentar o número de *joint plot* ficará mais difícil agrupar as conclusões para o conjunto de dados.

Além disso, os resultados indicam que o genótipo 6 é o mais estável (o genótipo que menos contribuiu para interação), e os genótipos 12, 9 e 5 são os que mais contribuíram para a interação.

## CONCLUSÕES

A sistemática de análise proposta com o uso dos modelos de Tucker3 mostrou-se eficiente e adequada para separar o padrão de resposta técnica e o ruído contidos no arranjo de interação tripla, bem como a interpretação da estabilidade e adaptabilidade fenotípica. O gráfico *joint plot* facilita o entendimento da interação tripla, além de trazer informações adicionais sobre estas, comparativamente aos gráficos *biplot* obtidos pela modelagem AMMI de duas entradas, permitindo identificar quais combinações de genótipos, locais e anos contribuem ou não para essa interação.

## AGRADECIMENTO

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro, aos revisores e editores por várias sugestões que melhoraram o artigo.

## REFERÊNCIAS

- DUARTE, J.B.; VENCOVSKY, R. **Interação genótipo x ambiente: uma introdução à análise "AMMI"**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 60p. (Série Monografias, 9).
- GABRIEL, K.R. The biplot graphic display of matrices with applications to principal components analysis. **Biometrika**, Cambridge, v.58, p.453-467, 1971. Disponível em: <<http://biomet.oxfordjournals.org/cgi/content/abstract/58/3/453>>. Acesso em: 01 jun 2009. doi:10.2307/2334381.
- GAUCH, H.G.; ZOBEL, R.W. Predictive and postdictive success of statistical analysis of yield trials. **Theoretical and Applied Genetics**, New York, v.76, p.1-10, 1988. Disponível em: <<http://www.springerlink.com/content/qq03t432n0315424/>>. Acesso em: 01 jun 2009. doi: 10.1007/BF00288824.
- KROONENBERG, P.M. **Three-mode principal component analysis: theory and applications**. Leiden: DSWO, 1983. 398p.
- KROONENBERG, P.M. The TUCKALS line: a suite programs for three-way data analysis. **Computational Statistics and Data Analysis**, Amsterdam, v.18,p.73-96, 1994. Disponível em: <[http://www.sciencedirect.com/science?\\_ob=ArticleURL&\\_udi=B6V8V-45GN701-6&\\_user=10&\\_coverDate=08%2F31%2F1994&\\_rdoc=1&\\_fmt=high&\\_orig=search&\\_sort=d&\\_docanchor=&view=c&\\_acct=C000050221&\\_version=1&\\_urlVersion=0&\\_useid=10&md5=8282db7b7791bfe534e2aa91e6440150](http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V8V-45GN701-6&_user=10&_coverDate=08%2F31%2F1994&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000050221&_version=1&_urlVersion=0&_useid=10&md5=8282db7b7791bfe534e2aa91e6440150)>. Acesso em: 01 jun 2009. doi:10.1016/0167-9473(94)90133-3.
- TIMMERMAN, M.E.; KIERS, H.A.L. Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima. **British Journal of Mathematical and Statistical Psychology**, London, v.53, n.1, p.1-16, 2000. Disponível em: <<http://www.ingentaconnect.com/content/bpsoc/bjmsp/2000/00000053/00000001/art00001?token=006519d826b2d783b72752d455e4e2663433b393f6a3340257d7241255e437a63736a6f2c47464c7a663f70532379ec942dc0>>. Acesso em: 01 jun 2009. doi:10.1348/000711000159132.
- TUCKER, L. Some mathematical notes on three-mode factor analysis. **Psychometrika**. New York, v.31, p.279-311, 1966. Disponível em: <<http://www.springerlink.com/content/87x451g73474n136/>>. Acesso em: 01 jun 2009. doi:10.1007/BF02289464.
- VARELA, M. et. al. Analysis of a three-way interaction including multi-attributes. **Australian Journal of Agricultural Research**, Sydney, v.57, p.1185-1193, 2006. Disponível em: <<http://www.publish.csiro.au/?paper=AR06081>>. Acesso em: 01 jun 2009. doi:10.1071/AR06081.
- YAN, W.; HUNT L.A. Biplot analysis of multi-environment trial data. In: KANG, M.S. **Quantitative genetics, genomics and plant breeding**. New York: CAB, 2002. p.289-303.