



ROGÉRIO LUNEZZO DE OLIVEIRA

**ÍNDICE DE SELEÇÃO NO ESTUDO DA
ADAPTABILIDADE E ESTABILIDADE EM
MILHO**

**LAVRAS-MG
2013**

ROGÉRIO LUNEZZO DE OLIVEIRA

**ÍNDICE DE SELEÇÃO NO ESTUDO DA ADAPTABILIDADE E
ESTABILIDADE EM MILHO**

Tese apresentada à Universidade Federal de Lavras
como parte das exigências do Programa de Pós-
Graduação em Genética e Melhoramento de Plantas,
área de concentração em Genética e Melhoramento de
Plantas, para a obtenção do título de Doutor.

Orientador

Dr. Renzo Garcia Von Pinho

Coorientador

Dr. Daniel Furtado Ferreira

**LAVRAS – MG
2013**

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e
Serviços da Biblioteca Universitária da UFLA**

Oliveira, Rogério Lunezzo de.

Índice de seleção no estudo da adaptabilidade e estabilidade em
milho / Rogério Lunezzo de Oliveira. – Lavras : UFLA, 2013.
66 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2013.
Orientador: Renzo Garcia Von Pinho.
Bibliografia.

1. *Zea mays*. 2. Ideótipo. 3. Análise multivariada. 4. Interação
GxA. 5. GGEbiplot. I. Universidade Federal de Lavras. II. Título.

CDD – 631.53

ROGÉRIO LUNEZZO DE OLIVEIRA

**ÍNDICE DE SELEÇÃO NO ESTUDO DA ADAPTABILIDADE E
ESTABILIDADE EM MILHO**

Tese apresentada à Universidade Federal de Lavras
como parte das exigências do Programa de Pós-
Graduação em Genética e Melhoramento de Plantas,
área de concentração em Genética e Melhoramento de
Plantas, para a obtenção do título de Doutor.

APROVADA em 26 de Julho de 2013.

Dr. Antônio Carlos de Oliveira	EMBRAPA MILHO E SORGO
Dr. José Airton Rodrigues Nunes	UFLA
Dr. Adriano Teodoro Bruzi	UFLA
Dr. Samuel Pereira de Carvalho	UFLA

Dr. Renzo Garcia Von Pinho
Orientador

**LAVRAS – MG
2013**

A Deus, pois sem Ele não teria chegado onde estou.

OFEREÇO

*A minha esposa por ser minha companheira de todas as horas e grande
incentivadora desta minha trajetória
Ao meu pai Evandro (in memorian), pelos exemplos de perseverança, coragem e
honestidade e a minha mãe, por toda luta que passou ao lado dele, sem recuar,
para sustentar e educar a mim e meu irmão.
Ao meu irmão pela grande amizade e apoio dado nos momentos que precis*

DEDICO

AGRADECIMENTOS

A Deus por ser o fiel cumpridor de suas promessas, sendo esta mais uma de suas promessas cumpridas em minha vida.

À Universidade Federal de Lavras e ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos.

À Empresa Advanta, na pessoa do Dr. Wagner Mateus Costa Melo, por disponibilizar os dados utilizados neste trabalho.

A todos os professores do Programa de Pós-Graduação em Genética e Melhoramento que, sem dúvida, são os principais contribuidores para o conhecimento em Melhoramento de Plantas e exemplos de profissionalismo e dedicação que adquiri.

Ao Prof. Renzo por todos os anos de orientação, amizade, ensinamentos, conselhos e confiança em mim depositados.

Ao Prof. Daniel, por ser o criador da proposta deste trabalho e pelos valiosos conhecimentos estatísticos transmitidos durante toda nossa convivência. Aos Professores José Airton Rodrigues Nunes, Adriano Teodoro Bruzi e Samuel Pereira de Carvalho e ao Dr. Antônio Carlos de Oliveira por aceitarem participar da banca da minha Defesa de Tese e pelas valiosas contribuições dadas a esta.

À Lilian, Secretária do Programa de Pós-Graduação pelo grande auxílio dado e pela atenção. Muito Obrigado!

Ao Prof. Márcio Balestre pela amizade, parceria e apoio durante toda minha Pós-Graduação;

A todos os amigos que fiz na UFLA durante os cursos do mestrado e doutorado, pela boa convivência e amizade compartilhadas. A todos que de alguma maneira contribuíram para este trabalho o meu Muito Obrigado!

RESUMO

Neste trabalho é proposto um método alternativo para avaliar a estabilidade e adaptabilidade de híbridos de milho, utilizando um índice de seleção de distância para o ideótipo (ISDI). Foram utilizados dados de sete variáveis (produtividade de grão - t/ha % de grãos ardidos, % de acamamento, % de quebramento, % de tombamento, nota de Ferrugem Comum e nota de Cercospora), obtidos por meio da avaliação de 25 híbridos de milho em seis locais (ambientes) da Região Sul do Brasil (Vacaria-RS, Abelardo Luz-SC, Candió-PR, Canoinhas-SC, Castro-PR e Ponta Grossa-PR). Foi realizada a análise de variância multivariada (MANOVA) para se obter a matriz de variâncias e covariâncias residuais entre as variáveis avaliadas. Foi estabelecido também o valor do ideótipo para cada variável e, em seguida, foi estimado o ISDI por meio da Distância Generalizada de Mahalanobis, para cada parcela dos ensaios. Realizou-se a análise de variância conjunta para o ISDI e para a produtividade a fim de verificar a significância da interação genótipos por ambientes (GxA), para cada uma dessas variáveis. Uma vez constatada a significância da interação (GxA) para ambas as variáveis, procedeu-se a análise GGE biplot para cada uma das variáveis, a fim de comparar a acurácia preditiva dos modelos GGE, o agrupamento de ambientes e selecionar os cinco melhores híbridos, avaliando-se no biplot a distância em relação ao “genótipo ideal”. A interação (GxA) foi altamente significativa para as duas variáveis analisadas. O modelo GGE2 (com dois componentes principais) obteve uma acurácia preditiva (PRECORR) de 0,8913 para o ISDI e 0,8709 para a produtividade (t/ha). Os dois primeiros componentes principais do GGEbiplot somados, captaram 92,14% e 87,24% para o ISDI e produtividade. Analisando-se o ISDI foram obtidos dois grupos de ambientes, enquanto que analisando-se a produtividade todos os ambientes ficaram no mesmo grupo. Foram selecionados, por meio da análise GGEbiplot, os híbridos 12, 21, 19, 8 e 22 avaliando o ISDI. Por outro lado a mesma análise, avaliando somente a produtividade, levou à seleção dos híbridos 21, 23, 24, 22 e 25. A avaliação da adaptabilidade e estabilidade do ISDI proporcionou a seleção de híbridos que aliam bons desempenhos na maioria das variáveis avaliadas, sendo mais recomendável utilizá-lo do que analisar cada variável isoladamente. A avaliação da adaptabilidade e estabilidade do ISDI proporcionou a seleção de híbridos que aliam adaptabilidade e estabilidade na maioria das variáveis avaliadas, sendo mais recomendável utilizá-lo do que analisar cada variável isoladamente. A produtividade de grãos não deve ser considerada isoladamente como um índice de seleção, pois nem todos os híbridos mais produtivos foram os melhores nas outras variáveis analisadas.

Palavras-chave: Índice de seleção. Ideótipo. Milho. Estabilidade. GGEbiplot.

ABSTRACT

This study proposes an alternative method to evaluate the stability and adaptability of corn hybrids, using a distance selection index for the ideotype (DSII). It was used data from seven variables (grain productivity - t/ha % of damaged kernels, % of mulching, % of breaking, % of tipping, Common Rust note and Cercospora note), obtained through the evaluation of 25 corn hybrids in six locations (environments) of Southern Brazil (Vacaria-RS, Abelardo Luz-SC, Candi-PR, Canoinhas-SC, Castro-PR and Ponta Grossa-PR). Multivariate analysis of variance (MANOVA) was taken to obtain the matrix of variances and residual covariances between the assessed variables. I was also established the ideotype value for each variable, and then estimated the DSII through Mahalanobis Generalized Distance for each part of the tests. It was conducted the joint variance analysis for the DSII and the productivity in order to verify the interaction genotypes per environments significance (GxA), for each of these variables. Once confirmed the interaction significance (GxA) for both variables, GGE biplot analysis was carried for each of the variables, in order to compare the GEE models predictive accuracy, the environments grouping and select the five best hybrids, evaluating in the biplot the distance compared to the “ideal genotype”. The interaction (GxA) was highly significant for both analyzed variables. The GGE2 model (with two main components) obtained a predictive accuracy (PRECORR) of 0.8913 for DSII and 0.8709 for the productivity (t/ha). The two first main GGEbiplot components added, captured 92.14 and 87.24 for the DSII and the productivity. Analyzing the DSII it was obtained two environment groups, while analyzing the productivity all environments remained in the same group. It was selected, through the GGEbiplot analysis, the hybrids, 12, 21, 19, 8 and 22 evaluating the DSII. On the other hand, the same analysis, evaluating only the productivity, led to the selection of the following hybrids, that is, 21, 23, 24, 22 and 25. The stability and adaptability evaluation of the DSII allowed the selection of hybrids that combine good performances in most evaluated variables, being more recommended to use it than to analyze each variable separately. The adaptability and stability evaluation of the DSII allowed the selection of hybrids that combine adaptability and stability in most evaluated variables, being more recommended to use it than to analyze each variable separately. Grain productivity may not be considered separately as a selection index, once not all of the most productive hybrids were the best ones in the other analyzed variables.

Keywords: Selection index. Ideotype. Corn. Stability. GGEbiplot.

LISTA DE FIGURAS

- Figura 1 Distância entre os genótipos G_1 e G_2 , com base em dois caracteres correlacionados, X_1 e X_2 17
- Figura 2 Gráfico GGEbiplot, avaliando o ISDI com o agrupamento dos seis locais (Vacaria-RS, Abelardo Luz-SC, Candoi-PR, Canoinhas-SC, Castro-PR e Ponta Grossa-PR) identificados, respectivamente, de E1 a E6, sendo que os grupos de ambientes foram identificados pelos algarismos romanos I e II 46
- Figura 3 Gráfico GGEbiplot, com a distância dos 25 híbridos (identificados de G1 a G25) em relação ao “genótipo ideal”, quanto ao ISDI..... 47
- Figura 4 Gráfico GGEbiplot, avaliando a produtividade de grãos (t/ha), com o agrupamento dos seis locais (Vacaria-RS, Abelardo Luz-SC, Candoi-PR, Canoinhas-SC, Castro-PR e Ponta Grossa-PR) identificados, respectivamente, de E1 a E6..... 48
- Figura 5 Gráfico GGEbiplot, com a distância dos 25 híbridos (numerados de G1 a G25 em relação ao “genótipo ideal”, quanto à produtividade de grãos (t/ha)..... 49

LISTA DE TABELAS

Tabela 1	Dados de Latitude, Longitude e Altitude (m) dos seis locais onde os ensaios foram conduzidos	31
Tabela 2	Matriz de variâncias e covariância residuais dos sete caracteres avaliados nos seis locais	41
Tabela 3	Resumo da análise de variância conjunta dos dados do índice de seleção de distância para o ideótipo (ISDI), para os 25 híbridos avaliados e os seis locais de avaliação	42
Tabela 4	Resumo da análise de variância conjunta dos dados de produtividade de grãos (kg/ha), para os 25 híbridos avaliados e os seis locais de avaliação	43
Tabela 5	Classificação decrescente dos híbridos quanto à sua produtividade média de grãos (kg/ha) e suas respectivas médias do ISDI, considerando os seis locais de avaliação	43
Tabela 6	Estimativas obtidas para $PRESS_m$ (soma de quadrados do erro das predições) e $PRECORR_{(m)}$ (acurácia preditiva) na validação cruzada dos modelos GGE1 e GGE2 (Biplot com 1 e 2 componentes principais, respectivamente) para o ISDI e para a produtividade (ton/ha)	45
Tabela 7	Rankeamento considerando a média dos seis locais de avaliação, para cada um dos híbridos envolvidos nas seleções, por meio da análise GGEbiplot do ISDI e da produtividade (t/ha) para cada uma das sete variáveis avaliadas	51

SUMÁRIO

1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	14
2.1	Índices de seleção	14
2.2	Interação genótipos x ambientes	21
2.3	Meios de atenuar o efeito da interação genótipos por ambientes	23
2.3.1	Identificação de cultivares específicas para cada ambiente	23
2.3.2	Identificação de cultivares com maior estabilidade fenotípica	24
3	MATERIAL E MÉTODOS	31
3.1	Caracterização dos ensaios e locais	31
3.2	Determinação do ideótipo para cada característica	32
3.3	Obtenção da matriz de variâncias e covariâncias	33
3.4	Obtenção do Índice de Seleção de Distância para o Ideótipo	34
3.5	Análises de variância e análises por meio do modelo SREG (GGE)	35
3.6	Validação cruzada do modelo GGE2	37
3.7	Análise do GGE biplot e comparação dos híbridos selecionados	38
4	RESULTADOS E DISCUSSÃO	40
5	CONCLUSÕES	54
	REFERÊNCIAS	55
	ANEXOS	60

1 INTRODUÇÃO

Em programas de melhoramento genético, muitas vezes torna-se necessário obter medidas de vários caracteres para selecioná-los simultaneamente. Tal necessidade levou ao desenvolvimento dos índices de seleção que normalmente são combinações lineares das medidas fenotípicas, obtidas dos diversos caracteres, o que permite realizar a seleção utilizando um único valor. A proposta de aplicação de índices de seleção foi apresentada inicialmente por Smith (1936), com base em trabalho de Fisher (1936). Conhecido também como índice otimizado, foi modificado posteriormente, sendo que a maioria dessas modificações também se baseava em combinações lineares dos valores fenotípicos observados. Os mais comuns são os de Brim, Johnson e Cockerham (1959), Kempthorne e Nordskog (1959), Pesek e Baker (1969), Smith, Hallauer e Russell (1981) e Tai (1977).

Outros autores apresentaram índices não lineares, também conhecidos como índices não-paramétricos, cujo objetivo da aplicação é diferente do que é almejado na aplicação dos índices lineares ou paramétricos, os quais visam a melhoria do valor genotípico. No caso dos índices não-paramétricos, o objetivo é classificar os genótipos. Elston (1963) propôs um índice multiplicativo que considera todos os caracteres com o mesmo peso econômico. Wricke e Weber (1986) sugeriram o uso das distâncias Euclidiana e a de Mahalanobis para classificar os genótipos para várias características, simultaneamente, baseando-se na distância desses em relação a um genótipo ideal (ideótipo), definido pelo pesquisador. Mulamba e Mock (1978) desenvolveram um índice bastante simples, que utiliza a soma do número de ordem que o genótipo apresenta para cada caráter; quanto menor o valor obtido nessa soma, melhor a classificação. Em todos os índices não-paramétricos citados, não são necessárias estimativas

de parâmetros genéticos, como para os índices lineares, tornando-os mais práticos.

Nos programas de melhoramento de milho, os genótipos são avaliados em diferentes ambientes, por isso deve-se considerar a contribuição da interação genótipos por ambientes (GxA) para os fenótipos observados. Para amenizar a interação GxA pode-se recomendar cultivares específicas para cada ambiente ou usar cultivares com ampla adaptabilidade e boa estabilidade ou, ainda, estratificar a região considerada em sub-regiões com características ambientais semelhantes, fazendo com que a interação seja não significativa dentro dessas sub-regiões (TERASAWA JÚNIOR; VENCOVSKY; KOEHLER, 2008).

Existem várias metodologias para avaliação da interação GxE. No entanto, os métodos de maior uso são os baseados em regressão simples e múltipla. Apesar do emprego generalizado, os métodos baseados em regressão possuem limitações que têm sido frequentemente relatadas na literatura. Crossa (1990) sugere que a aplicação de métodos multivariados pode ser útil para melhor explorar as informações contidas nos dados. Recomenda, então, técnicas como a análise de componentes principais (ACP), a análise de agrupamentos e o procedimento AMMI (Additive Main Effects and Multiplicative Interaction Analysis), que estão ganhando grande aplicabilidade nos últimos anos.

Recentemente, uma modificação da análise AMMI convencional, proposta por Yan et al. (2000), denominada de GGE biplot, tem sido utilizada para estudo da interação GxE. A análise GGE agrupa o efeito de genótipo, que é um efeito aditivo na análise AMMI, com a interação GxE efeito multiplicativo e submete esses efeitos a um modelo multiplicativo de regressão para locais (SREG - Sites Regression). A principal vantagem dessa técnica em relação à análise AMMI, reside no fato de que o método GGE biplot explica sempre uma parcela intermediária da soma de quadrados de genótipos (G) + interação (GxE)

em relação aos modelos AMMI1 (com um componente principal) e AMMI2 (com dois componentes principais) (YAN et al., 2007).

Alguns trabalhos como o de Farshadfar (2008) e Garcia e Souza Júnior (1999) incorporaram nos índices de seleção, parâmetros de estabilidade e adaptabilidade para selecionar os indivíduos mais estáveis e de melhor desempenho. No entanto, esses trabalhos que aplicaram índices de seleção não-paramétricos com parâmetros de estabilidade incorporados, o fizeram com as médias das cultivares em todos os ambientes. Sendo assim, a estabilidade e adaptabilidade do índice de seleção não foi avaliada nesses trabalhos.

Não foram encontrados, na literatura, trabalhos que tenham analisado esses índices de seleção ao nível de parcela e feitas as análises de adaptabilidade e estabilidade com esses índices, utilizando-se métodos multivariados como o GGE biplot. Dessa maneira, a aplicabilidade de um índice de seleção não-paramétrico poderia ser melhor avaliada e explorada. Por esse motivo, o objetivo principal deste trabalho foi avaliar a adaptabilidade e estabilidade do índice de seleção de distância para o ideótipo. O segundo objetivo foi selecionar os melhores híbridos, usando-se essa alternativa e comparando-a à seleção, avaliando-se a adaptabilidade e estabilidade por meio da produtividade, para verificar se se trata, isoladamente, de um bom índice de seleção.

2 REFERENCIAL TEÓRICO

2.1 Índices de seleção

A seleção direcionada a um ou poucos caracteres pode promover alterações desfavoráveis em outros, devido a existência de correlações genéticas negativas entre esses caracteres. Com o objetivo de amenizar esse problema, os melhoristas têm empregado os índices de seleção, os quais permitem a seleção com base em um conjunto de variáveis de importância agrônômica (CRUZ; REGAZZI, 2001).

Na literatura existem muitos índices de seleção descritos, os quais são compostos pelas médias fenotípicas. Dessa forma o índice de seleção é a combinação linear desses valores fenotípicos, resultando num único valor que resume os pontos positivos e negativos de cada genótipo para vários caracteres (GARCIA; SOUZA JÚNIOR, 1999).

Devido à dificuldade de se estabelecer pesos econômicos aos vários caracteres, a utilização dos índices de seleção paramétricos tem sido limitada. Além disso, a eficiência do índice dependerá de estimativas precisas de variâncias e covariâncias genéticas e fenotípicas (CRUZ; REGAZZI, 2001; GARCIA, 1998).

Como alternativa, podem ser utilizados os índices não-paramétricos, dentre os quais destacam-se os baseados em distâncias em relação a um ideótipo, como a euclidiana e a generalizada de Mahalanobis, recomendados por Farias (2005) e Santos (2005). A primeira alternativa é preferível quando não houver correlações entre os caracteres que são utilizados para compor o índice e por ser de mais fácil interpretação dos valores obtidos. Wricke e Weber (1986) salientaram que o índice de seleção, baseado na distância média euclidiana a um ideótipo, tem por pressuposto que os valores fenotípicos sejam altamente

correlacionados com os genotípicos e obtêm-se, a partir das médias fenotípicas, distâncias de cada genótipo a um ideótipo, o qual é fixado pelo melhorista para todos os caracteres considerados. A distância euclidiana possui os seguintes pontos negativos: é alterada com a mudança de escala de medição dos caracteres, com o número de caracteres e não considera a existência de correlações entre os caracteres (CRUZ; REGAZZI, 2001).

Segundo Santos (2005), o problema da escala pode ser resolvido com a padronização dos dados, ou seja, dividir cada observação pelo seu respectivo desvio padrão conforme (1)

$$Z_{ij} = \frac{X_{ij}}{s_j} \quad (1)$$

em que:

Z_{ij} : observação fenotípica estandardizada do caráter j, medido no genótipo i;

X_{ij} : fenótipo do caráter j no genótipo i;

s_j : desvio padrão do caráter j.

Essa padronização evita que caracteres medidos numa escala maior exerçam maior peso no valor de distância e permite, também, incluir num mesmo valor de distância caracteres medidos em escalas diferentes.

O problema do número de caracteres pode ocorrer quando um ou mais caracteres não são medidos num dado genótipo. Esse problema pode ser contornado dividindo-se a distância euclidiana pelo número de caracteres e, dessa forma, a distância euclidiana média padronizada entre o genótipo i e o ideótipo ℓ é definida pela expressão (2)

$$dm_{Gi\ell} = \sqrt{\frac{1}{n} \sum_{j=1}^n (Z_{ij} - Z_{\ell j})^2} \quad (2)$$

em que

$dm_{Gi\ell}$: distância euclidiana entre o genótipo G_i e o ideótipo ℓ ;

Z_{ij} : observação fenotípica padronizada do caráter j , medido no genótipo i ; $j = 1, 2, \dots, n$;

$Z_{\ell j}$: observação fenotípica padronizada do caráter j , medido no ideótipo ℓ , $j = 1, 2, \dots, n$.

Dessa forma, segundo Santos (2005), pode ser realizada a comparação de valores de distâncias obtidos a partir de números diferentes de observações. No entanto, conforme o mesmo autor, a distância euclidiana tem outra limitação que só pode ser contornada usando-se outra medida de distância. É que ela pressupõe que os caracteres são independentes entre si, ou seja, não correlacionados. Essa pressuposição geralmente não pode ser, principalmente quando se trabalha com vários caracteres, como é o caso dos índices de seleção.

Dessa forma, quando se têm à disposição dados de vários caracteres com repetições, a aplicação da Distância Generalizada de Mahalanobis é mais adequada, pois essa considera as correlações existentes entre os caracteres.

Quando dois caracteres são independentes um do outro, a Distância Euclidiana estimada entre dois genótipos, com base neles, é uma medida válida. No entanto, quando muitos caracteres são avaliados, dificilmente todas as correlações serão nulas. Considerando dois caracteres correlacionados, seja devido à pleiotropia ou à ligação gênica, a relação entre eles é a mostrada na Figura 1. Nessa Figura, X_{11} e X_{21} são as médias do caráter 1, observadas nos

genótipos 1 (G_1) e 2 (G_2), respectivamente, X_{12} e X_{22} representam as médias do caráter 2, observadas nos genótipos 1 e 2, respectivamente, e as coordenadas Y_{11} , Y_{21} , Y_{12} e Y_{22} correspondem, nessa ordem, a X_{11} , X_{21} , X_{12} e X_{22} , considerando a correlação existente entre os caracteres X_1 e X_2 .

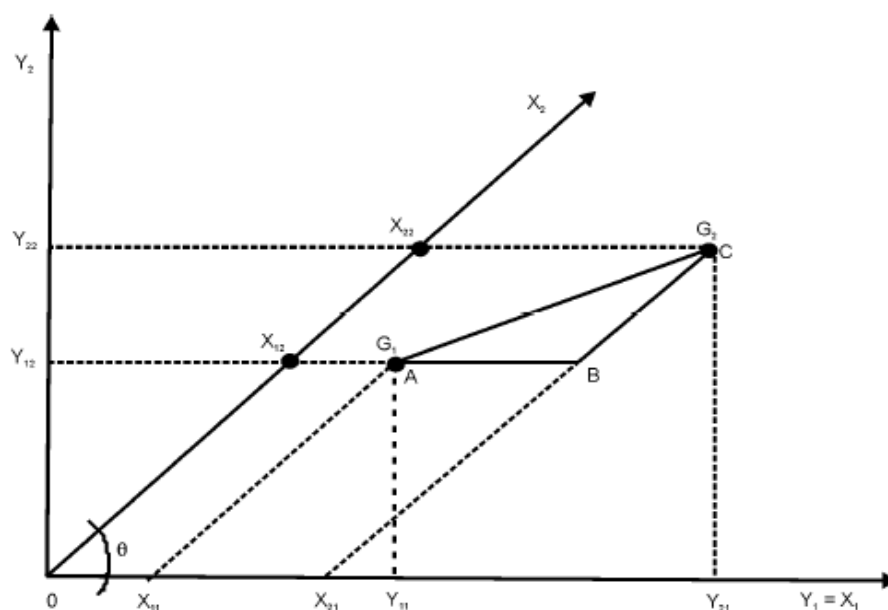


Figura 1 Distância entre os genótipos G_1 e G_2 , com base em dois caracteres correlacionados, X_1 e X_2

Fonte: Cruz e Regazzi (2001)

Pela Figura 1 observa-se que devido à correlação existente entre os caracteres X_1 e X_2 , o ângulo formado pelas observações desses caracteres é $\theta < 90^\circ$. O cosseno do ângulo θ corresponde à correlação entre X_1 e X_2 e, dessa forma, não é válida a aplicação do teorema de Pitágoras para a estimação da distância entre dois genótipos, baseando-se em observações desses dois caracteres. No entanto, a partir de X_1 e X_2 pode-se criar dois outros “caracteres” Y_1 e Y_2 , independentes entre si.

Observando-se a Figura 1, tem-se que $Y_{11} = X_{11} + (Y_{11} - X_{11})$. Além disso, no triângulo formado pelos vértices X_{11} , Y_{11} e A , o segmento $(Y_{11} - X_{11})$ é o cateto adjacente ao ângulo θ . Sendo assim, por meio da função cosseno, deduz-se que $(Y_{11} - X_{11}) = X_{12} \cos \theta$ e, portanto, $Y_{11} = X_{11} + X_{12} \cos \theta$.

De acordo com a Figura 1, o segmento $Y_{11} - A$ tem o mesmo comprimento do segmento $Y_{12} - O$, da mesma forma que $X_{12} - O$ tem o mesmo comprimento do segmento $X_{11} - A$. Sendo assim, o segmento $Y_{11} - A$ corresponde ao cateto oposto a θ assim, pela função seno, têm-se que $Y_{12} = X_{12} \sin \theta$.

Seguindo o mesmo raciocínio usado anteriormente na dedução dos “caracteres” Y , correspondentes ao genótipo G_2 por meio do triângulo de vértices X_{21} , Y_{21} e C , chega-se às expressões (3) e (4)

$$Y_{21} = X_{21} + X_{22} \cos \theta, \text{ e} \quad (3)$$

$$Y_{22} = X_{22} \sin \theta \quad (4)$$

Assim, pode-se representar os genótipos G_1 e G_2 por meio das seguintes coordenadas

$$G_1: (Y_{11}, Y_{12}) : (X_{11} + X_{12} \cos \theta, X_{12} \sin \theta);$$

$$G_2: (Y_{21}, Y_{22}) : (X_{21} + X_{22} \cos \theta, X_{22} \sin \theta).$$

Após a aplicação do teorema de Pitágoras aos “caracteres” não correlacionados Y_1 e Y_2 , chega-se à Distância Generalizada de Mahalanobis (MAHALANOBIS, 1936) entre os genótipos G_1 e G_2 , com base nos caracteres X_1 e X_2 , conforme a expressão (5)

$$D_{G_1G_2}^2 = (Y_{21} - Y_{11})^2 + (Y_{22} - Y_{12})^2 \quad (5)$$

$$\begin{aligned}
&= [(X_{21} + X_{22} \cos \theta) - (X_{11} + X_{12} \cos \theta)]^2 + [(X_{22} \sin \theta) - (X_{12} \sin \theta)]^2 \\
&= (X_{21} - X_{11})^2 + (X_{22} - X_{12})^2 + 2(X_{21} - X_{11})(X_{22} - X_{12}) \cos \theta.
\end{aligned}$$

Para simplificar a expressão acima, substitui-se $(X_{21} - X_{11})$ por d_1 e $(X_{22} - X_{12})$ por d_2 , e assim chega-se à expressão (6)

$$D_{G1G2}^2 = d_1^2 + d_2^2 - 2d_1d_2 \cos \theta \quad (6)$$

De acordo com a expressão anterior, quando comparada à Distância Euclidiana, a Distância Generalizada de Mahalanobis apresenta um termo que inclui o $\cos \theta$, que nada mais é que a correlação entre os caracteres X_1 e X_2 . Se esses caracteres forem completamente independentes entre si, $\theta = 90^\circ$, e como $\cos 90^\circ = 0$, dessa forma a distância de Mahalanobis torna-se igual à distância euclidiana.

A expressão (5) pode ser escrita na forma matricial, da seguinte maneira

$$\begin{aligned}
D_{G1G2}^2 &= [(Y_{21} - Y_{11})(Y_{22} - Y_{12})][(Y_{21} - Y_{11})(Y_{22} - Y_{12})]' \\
&= [(X_{21} - X_{11})(X_{22} - X_{12})]S^{-1}[(X_{21} - X_{11})(X_{22} - X_{12})]'
\end{aligned}$$

em que S^{-1} é a inversa da matriz de variâncias e covariâncias residuais ou, conforme a convenção feita anteriormente, chega-se à expressão (7)

$$D_{G1G2}^2 = [d_1 d_2]S^{-1}[d_1 d_2]' = [d_1 d_2] \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} [d_1 d_2]'. \quad (7)$$

Resolvendo esse sistema matricial chega-se à seguinte expressão

$$D_{G1G2}^2 = \frac{d_1^2 \sigma_2^2 + d_2^2 \sigma_1^2 - 2d_1 d_2 \sigma_{12}}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}$$

que, simplificando torna-se (8)

$$D_{G1G2}^2 = \frac{1}{\text{sen}^2 \theta} \left(\frac{d_1^2}{\sigma_1^2} + \frac{d_2^2}{\sigma_2^2} - \frac{2d_1 d_2 \cos \theta}{\sigma_1 \sigma_2} \right) \quad (8)$$

conforme apresentado por Arunachalam (1981).

Como apresentado anteriormente, $d_1 = (X_{21} - X_{11})$ e $d_2 = (X_{22} - X_{12})$, desse modo a expressão anterior pode ser reescrita da seguinte forma

$$D_{G1G2}^2 = \frac{1}{\text{sen}^2 \theta} \left(\frac{(X_{21} - X_{11})^2}{\sigma_1^2} + \frac{(X_{22} - X_{12})^2}{\sigma_2^2} - \frac{2(X_{21} - X_{11})(X_{22} - X_{12}) \cos \theta}{\sigma_1 \sigma_2} \right) \quad (9)$$

Analisando a expressão (9), observa-se que quando dois caracteres possuem escalas semelhantes e supondo-se que os dados tenham sido obtidos com a mesma precisão, o caráter que possuir menor variância do resíduo, ou seja, maior herdabilidade, exercerá maior peso sobre o valor da distância. Pode-se observar também, que como o termo que contém o $\cos \theta$ é subtraído do restante da expressão, dois caracteres correlacionados entre si terão um valor de distância menor do que teriam dois caracteres independentes entre si (JOHNSON; WICHERN, 1998).

Generalizando para um número qualquer de caracteres, a representação matricial da Distância Generalizada de Mahalanobis fica conforme (10) (CRUZ; REGAZZI, 2001).

$$D^2 = \delta \psi^{-1} \delta', \quad (10)$$

em que

δ : vetor-linha, composto pelos desvios (vetor d , na notação desse trabalho), de dimensões $1 \times n$, sendo n o número de caracteres;

ψ : matriz de variâncias e covariâncias dos resíduos, de dimensões $n \times n$ (matriz S , na notação adotada anteriormente).

2.2 Interação genótipos x ambientes

Considerando um caráter como por exemplo a produtividade de grãos, o valor observado para esse caráter, denominado de fenótipo (F), é função do genótipo (G), do ambiente (A) e da interação genótipos por ambientes ($G \times A$) (RAMALHO et al., 2012). A ocorrência de interação entre genótipos e ambientes é consequência do comportamento não coincidente dos genótipos nos diferentes ambientes testadores (CRUZ; CARNEIRO, 2003). Caracteres muito afetados pelo ambiente, como é o caso dos caracteres quantitativos, frequentemente apresentam interação ($G \times A$) significativa (BERNARDO, 2002).

De acordo com Allard e Bradshaw (1964), as variações ambientais que contribuem para a interação com os genótipos podem ser classificadas em dois tipos: previsíveis e não previsíveis. O primeiro tipo abrange aquelas características do clima cuja variação é previsível, tais como local, solo, comprimento do dia, insolação e também as variações ambientais determinadas pela ação do homem, como, época e densidade de semeadura, níveis de adubação e outros tratamentos culturais. As variações imprevisíveis são as oscilações no clima que não podem ser previstas, tais como volume e distribuição das chuvas, variações da temperatura e outros.

A avaliação das variações previsíveis pode ser feita individualmente ou de forma conjunta em relação à sua interação com os genótipos. Desta forma, interações individuais, tais como genótipos \times densidade de semeadura e

genótipos x época de semeadura ou estudos envolvendo todos esses fatores podem ser feitos generalizando a interação para genótipos x ambientes. No entanto, as interações genótipos x anos, genótipos x locais e as interações mais complexas como a interação tripla genótipos x locais x anos são compostas, em sua maior parte, por variações imprevisíveis (FEHR, 1987).

Para a detecção da interação (GxA) é necessário que genótipos diferentes sejam avaliados em dois ou mais ambientes contrastantes, pois a avaliação em apenas um ambiente não permite isolar o componente da interação, fazendo com que a variância genética seja superestimada. Consequentemente, a herdabilidade também fica superestimada e também o ganho esperado com a seleção, o qual é diretamente proporcional à herdabilidade (TERASAWA JÚNIOR; VENCOVSKY; KOEHLER, 2008).

Segundo Robertson (1959), quanto à sua natureza, a interação (GxA) pode ser de dois tipos: simples ou complexa. Nas interações simples a classificação dos genótipos quanto ao seu desempenho nos diferentes ambientes de avaliação é mantida. Essas interações são ocasionadas pela diferença de variabilidade dos genótipos nos ambientes de avaliação, as quais não representam problemas ao melhoramento (CRUZ; CARNEIRO, 2003).

Por outro lado, as interações complexas ocorrem quando não há correlação entre o desempenho dos genótipos nos diferentes ambientes (VENCOVSKY; BARRIGA, 1992). Neste caso, a classificação dos genótipos quanto ao seu desempenho não é coincidente nos diferentes ambientes de avaliação. Interações dessa natureza são muito importantes no melhoramento, afetando a eficiência da seleção e a precisão das recomendações varietais (BAKER, 1988; CROSSA; CORNELIUS, 1997).

A ocorrência de interação (GxA) de natureza complexa é esperada em programas comerciais voltados à obtenção de novas cultivares. Por este motivo são demandados grandes investimentos na avaliação dos genótipos em muitos

ambientes contrastantes, por sucessivos anos (BERNARDO, 2002). O mesmo autor relata que os híbridos comerciais de milho nos EUA são testados em cerca de 120 a 2100 locais diferentes antes de sua comercialização. No Brasil, a ocorrência de interações complexas em milho já é esperada entre e dentro das regiões de cultivo. Este fato pode ser verificado analisando-se os resultados de ensaios de híbridos comerciais, conduzidos sob coordenação nacional (EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA, 2012).

2.3 Meios de atenuar o efeito da interação genótipos por ambientes

Para atenuar a interação genótipos x ambientes podem ser tomadas as seguintes medidas: a) identificar uma cultivar específica para cada ambiente; b) identificar cultivares com maior estabilidade fenotípica, e; c) realizar o zoneamento ecológico ou estratificação ambiental (RAMALHO et al., 2012).

2.3.1 Identificação de cultivares específicas para cada ambiente

Mesmo essa opção sendo possível na teoria, sua aplicação é praticamente inviável. Neste caso, os genótipos são avaliados em vários ambientes e por meio da análise dos dados são identificadas as melhores cultivares para cada ambiente. Pelo ambiente ser muito restrito, qualquer variação imprevisível, como um ano com menor pluviosidade por exemplo, pode fazer com que o genótipo não apresente um bom desempenho novamente (RAMALHO et al., 2012).

2.3.2 Identificação de cultivares com maior estabilidade fenotípica

Antes de apresentar os métodos que podem ser aplicados para avaliação da estabilidade fenotípica é necessário conceituá-la (RAMALHO et al., 2012). Segundo Lin, Binns e Lefkovitch (1986), existem basicamente três conceitos de estabilidade:

- a) Tipo 1 - A cultivar será considerada estável se sua variância entre ambientes é pequena.
- b) Tipo 2 - A cultivar será considerada estável se sua resposta ao ambiente for paralela ao desempenho médio de todos os genótipos avaliados.
- c) Tipo 3 - A cultivar é estável se o quadrado médio dos desvios da regressão, que avalia a estabilidade, é pequeno.

A estabilidade do tipo 1, denominada por Becker (1981) “estabilidade no sentido biológico”, corresponde à cultivar cujo desempenho é constante com a variação do ambiente. Esse não é o tipo de estabilidade buscada pelos melhoristas, porque o genótipo não responde à melhoria do ambiente. Além disso, essa estabilidade normalmente está associada a uma menor produtividade média (RAMALHO et al., 2012).

O segundo tipo de estabilidade, denominada por Becker (1981) “estabilidade no sentido agrônômico”, é aquela onde a cultivar acompanha o desempenho médio obtido nos ambientes. Essa é a estabilidade desejada pelos melhoristas, pois permite a identificação de cultivares estáveis e que podem manter-se entre as melhores em todos os ambientes de avaliação (RAMALHO et al., 2012). Segundo Lin, Binns e Lefkovitch (1986) esse tipo de estabilidade é restrita aos genótipos avaliados, sendo assim não pode ser generalizada.

Existem vários métodos que podem ser utilizados para avaliação da estabilidade tipo 2. Um dos métodos mais utilizados é o proposto por Wricke e Weber (1986), cujo parâmetro é denominado “ecoalência” (W_i), o qual mede o quanto cada cultivar contribui para a interação. Dessa forma, quanto mais próximo de zero for a estimativa de W_i menor é a contribuição dessa cultivar para a interação sendo, portanto, mais estável.

Outras metodologias baseadas em regressão foram propostas, dentre elas destaca-se a proposta por Finlay e Wilkinson (1963), por ter introduzido o conceito de índice ambiental, que é função da média de todas as cultivares em cada ambiente. O método nada mais é que uma análise de regressão linear, onde o índice ambiental é a variável independente e a produtividade média de cada cultivar, em cada ambiente, é a variável dependente (RAMALHO et al., 2012). Finlay e Wilkinson (1963) propuseram ainda a transformação logarítmica dos dados para aumentar a linearidade da regressão e a homogeneidade das variâncias residuais. Essa metodologia permite ainda que a interação genótipos x ambientes seja desdobrada em duas partes: a) devido à regressão linear e b) devido aos desvios da regressão. Sendo que, por meio do coeficiente de regressão e da produtividade média da cultivar, a sua estabilidade e adaptabilidade são obtidas (MARANHA, 2005).

Na estabilidade do tipo 3, quando o quadrado médio do desvio da regressão do genótipo é pequeno, esse é considerado estável. A metodologia proposta por Eberhart e Russel (1966), pode ser utilizada para avaliar esse tipo de estabilidade. Essa metodologia é semelhante à de Finlay e Wilkinson (1963), tendo como diferenças o fato dos dados serem analisados sem transformação e também pela estimativa do desvio da regressão (MARANHA, 2005).

De acordo com Crossa (1990), a principal limitação de uso do coeficiente de regressão linear, como metodologia para estimar a estabilidade fenotípica de cultivares, está na dependência que existe entre a variável que

mede o índice ambiental e a produtividade média da cultivar. Esse fato viola um princípio da análise de regressão que é a independência entre variáveis dependentes e independentes. Segundo o mesmo autor, quando locais com desempenhos contrastantes são incluídos na análise, a adaptação de um determinado genótipo pode se dar devido ao seu desempenho nos ambientes extremos, mascarando os resultados.

Para contornar os problemas relacionados às estimativas do índice ambiental, Toler (1990) propôs uma nova metodologia, a qual utiliza modelos de regressão não lineares nos parâmetros. Além disso, esse método apresenta testes de hipóteses mais rigorosos para os padrões de respostas, o que possibilita a classificação dos genótipos em vários grupos. O método de Toler (1990) é, do ponto de vista estatístico, um aprimoramento das análises de estabilidade que utilizam regressão.

Compartilhando a mesma opinião de Crossa (1990) e Duarte e Vencovsky (1999) afirmaram que os métodos baseados em regressão tendem a explicar a variação devido à interação numa única dimensão, simplificando os modelos de resposta, quando, na verdade, a interação pode ser bastante complexa. Além disso, esses métodos, de forma geral, não permitem identificar se interações específicas de genótipos com ambientes são positivas ou negativas, dificultando a capitalização dos efeitos positivos da interação. Crossa (1990) propõe que a aplicação de métodos multivariados pode ser útil para explorar melhor as informações contidas nos dados. O autor recomenda, principalmente, o método AMMI o qual vem sendo muito aplicado nos últimos anos.

Na análise AMMI, os efeitos principais (genótipos e ambientes) são associados a componentes aditivos, enquanto que os efeitos da interação são associados a componentes multiplicativos (DUARTE; VENCOVSKY, 1999). O método AMMI, em primeiro lugar estima os efeitos principais através da análise de variância. Os resíduos provenientes da aditividade desses efeitos são

particionados em um termo multiplicativo denominado padrão, que resulta da análise de componentes principais (ACP), e um desvio do modelo, também chamado de ruído (ANICCHIARICO, 2002). Segundo Zobel, Wright e Gauch (1988), o método AMMI possibilita uma análise mais detalhada da interação (GxA), garante a seleção de genótipos mais produtivos, fornece estimativas mais precisas das respostas genotípicas e facilita a interpretação gráfica dos resultados da análise. No que se refere ao detalhamento da interação (GxA), essa análise se faz via decomposição da soma de quadrados original (SQGxA), em uma porção denominada padrão e em outra chamada ruído. De acordo com Oliveira, Duarte e Pinheiro (2003), o padrão permite identificar os fatores ambientais e genotípicos mais diretamente relacionados à interação obtida após o descarte de ruídos adicionais ao erro experimental.

Recentemente, Yan et al. (2000) propôs uma modificação da análise AMMI convencional, denominada GGE biplot (Genotype and Genotypes by Environments Interaction) a qual tem sido amplamente utilizada para estudo da interação (GxA). A análise GGE agrupa o efeito de genótipo, que é um efeito aditivo na análise AMMI com a interação (GxA) efeito multiplicativo e submete esses efeitos a uma análise que segue o modelo SREG (Sites Regression). O modelo SREG é um modelo multiplicativo de regressão para locais. Seu biplot é chamado de GGE biplot, conforme descrito por Yan et al. (2000). Esse método tem como principal vantagem, comparado à análise AMMI, o fato de que o método GGE biplot explica sempre uma porção intermediária da soma de quadrados de genótipos (G) + interação (GxA) em relação aos modelos AMMI1 e AMMI2 (modelos AMMI com um e dois componentes principais, respectivamente).

Utilizando o modelo AMMI2, para tornar possível a análise de adaptabilidade e estabilidade no mesmo biplot, como no método GGE biplot, Gauch (1992) desenvolveu um gráfico 3D onde o efeito de genótipo é a terceira

dimensão perpendicular ao plano IPCA1 (primeiro componente principal da interação) vs. IPCA2 (segundo componente principal da interação). Mesmo esse modelo AMMI2, captando uma porção maior da interação, comparado ao modelo AMMI1, a interpretação desse gráfico 3D é complexa e pouco prática (OLIVEIRA et al., 2010).

Outra vantagem do GGE biplot, quando comparado ao gráfico AMMI1 é que, no AMMI1 os eixos estão em escalas diferentes (abscissa – produtividade média e ordenada – IPCA1). Consequentemente, o desempenho de um genótipo em um dado ambiente não pode ser visualizada com precisão, fato observado também no gráfico AMMI2 (3D). Por outro lado, no GGE biplot ambos os eixos estão na mesma escala. Portanto, o gráfico do GGE biplot mostra não só o desempenho médio e a estabilidade de cada genótipo avaliado mas também o desempenho relativo de cada genótipo, em um dado ambiente (OLIVEIRA et al., 2010).

Com relação à escolha do número ideal de componentes a serem adicionados, Gauch e Zobel (1997) comentam que a avaliação de modelos por meio de testes de F, denominadas de “pós-ativas”, buscam um modelo para explicar a maior parte da variação observada nos dados. Sendo assim, os autores argumentam que esses métodos não promovem a seleção de modelos parcimoniosos (com menor número de componentes principais) sendo, portanto, mais suscetíveis a incluir o ruído. Por outro lado, métodos de avaliação “preditiva” mensuram a capacidade de um modelo obter predições, excluindo dados na análise, simulando respostas futuras, ou seja, que ainda não foram medidas. Sendo assim, a escolha do modelo por meio de uma avaliação “preditiva” é mais vantajosa.

Geralmente, para fazer predições é necessário usar métodos estatísticos computacionalmente intensivos. Nesse caso os resultados precisam ter maior generalidade possível, sendo assim, a escolha do modelo ou avaliação do

desempenho de um preditor não deve basear-se em pressuposições de distribuição. Assim, métodos que são essencialmente baseados em dados e livres de distribuições teóricas terão a maior generalidade. Esses métodos utilizam reamostragem em determinado conjunto de dados, usando técnicas como o jackknife, o bootstrap e a validação cruzada (DIAS; KRZANOWSKI, 2003).

Gabriel (2002) propôs um algoritmo para a validação cruzada utilizando-se de uma partição na matriz de dados, via decomposição por valores singulares (DVS). O algoritmo desenvolve-se a partir da submatriz $X_{\setminus 11}$ dada por (11)

$$X_{G \times A} = \begin{bmatrix} x_{11} & X_1' \\ X_1 & X_{\setminus 11} \end{bmatrix} \quad (11)$$

em que: $X_{\setminus 11} = \sum_{k=1}^r u_{(k)} d_{(k)} v_{(k)}' = UDV'$, sendo r o número de componentes multiplicativos em análise; u_k , d_k e v_k correspondentes, respectivamente aos elementos λ_k , γ_{ik} e α_{jk} ($k = 1, 2, \dots, r$) oriundos da DVS; $U = [u_1, u_2, \dots, u_r]$, $V = [v_1, v_2, \dots, v_r]$ e $D = [d_1, d_2, \dots, d_r]$. O valor predito para x_{11} é dado por: $\hat{x}_{11} = x_1' V D^{-1} U' x_1$. O resíduo da validação cruzada é obtido por $e_{11} = x_{11} - \hat{x}_{11}$.

Da mesma forma calculam-se todos os valores \hat{x}_{ij} ajustados pela validação cruzada e resíduos $e_{ij} = x_{ij} - \hat{x}_{ij}$ para todos os outros elementos x_{ij} , ($i = 1, 2, \dots, g$) onde g é o número de genótipos; ($j = 1, 2, \dots, a$) onde a é o número de ambientes. Os resíduos e os valores ajustados são resumidos por:

$$PRESS_m = \frac{1}{ga} \sum_{i=1}^g \sum_{j=1}^a e_{ij}^2 \text{ e } PRECORR_{(m)} = Corr(x_{ij}, \hat{x}_{ij} | \forall i, j), \text{ respectivamente.}$$

Sendo assim, o número de componentes principais corresponde ao r que proporcionar o menor $PRESS_m$, conseqüentemente, o maior valor de $PRECORR_{(m)}$ (GABRIEL, 2002).

Dessa maneira torna-se possível não só a avaliação, a estabilidade e adaptabilidade do índice de seleção de distância para o ideótipo, mas também verificar a acurácia preditiva do método GGE biplot, analisando esse índice. Podendo-se compará-la à acurácia preditiva obtida na análise feita rotineiramente pelos melhoristas por meio da produtividade de grãos.

3 MATERIAL E MÉTODOS

3.1 Caracterização dos ensaios e locais

Foram utilizados dados da avaliação de sete características, de 25 híbridos de milho submetidos a ensaios de avaliação de desempenho, em seis locais da Região Sul do Brasil (Vacaria-RS, Abelardo Luz-SC, Candió-PR, Canoinhas-SC, Castro-PR e Ponta Grossa-PR) durante a safra 2010/2011. Os detalhes dos locais onde os ensaios foram conduzidos estão na Tabela 1.

Tabela 1 Dados de Latitude, Longitude e Altitude (m) dos seis locais onde os ensaios foram conduzidos

Local	Latitude/Longitude	Altitude (m)
Vacaria - RS	S 28.22'19,0"/W 51.01'02,6"	934
Abelardo Luz - SC	S 26.35'16,9"/W 52.23'50,1"	775
Candió - PR	S 25.37'18,0"/W 52.01'36"	920
Canoinhas - SC	S 26.20'26,0"/W 50.26'15,0"	803
Castro - PR	S 24.47'26,2"/W 50.00'41,3"	1004
Ponta Grossa - PR	S 25.04'69"/W 50.13'16"	851

Os ensaios foram elaborados no delineamento de blocos completos casualizados, com duas repetições. As parcelas foram compostas por quatro linhas de 5m, com espaçamento de 70cm entre linhas. As adubações realizadas seguiram as recomendações feitas para cada local por meio da análise de solo. Foram realizados os tratos culturais necessários para o controle da lagarta-do-cartucho (*Spodoptera frugiperda*), da lagarta-da-espiga (*Helicoverpa zea*) e das plantas invasoras.

As características avaliadas foram: produtividade de grãos – kg/ha, a percentagem de grãos ardidos, a percentagem de acamamento, a percentagem de quebramento e a percentagem de tombamento, nota de Ferrugem Comum (*Puccinia sorghi*) e nota de Cercosporiose (*Cercospora zae-maydis*).

A percentagem de grãos ardidos foi determinada através da proporção de grãos com sintoma de podridão numa amostra de 1000 grãos. A percentagem de acamamento foi calculada através da proporção de plantas com inclinação maior que 30°, devido ao enfraquecimento das raízes em relação ao total de plantas da parcela. Já a percentagem de tombamento correspondeu à proporção de plantas com inclinação maior que 30°, devido ao enfraquecimento da base do colmo em relação ao total de plantas da parcela. Sendo assim, a diferença entre acamamento e tombamento é que, no primeiro caso a planta inteira fica prostrada, enquanto que no segundo caso o colmo enfraquece na base e se dobra, causando o tombamento da planta.

A nota de severidade das doenças baseou-se na escala diagramática proposta pela Agrocerec (1996), que varia de 1 a 9, sendo que na nota 1 a folha apresenta-se sem sintomas e a nota 9 corresponde ao material com mais de 80% da área foliar afetada.

3.2 Determinação do ideótipo para cada característica

O ideótipo para produtividade (kg/ha) foi determinado buscando-se o valor da parcela mais produtiva de todos os ensaios e utilizando-se do próximo milho acima desse valor. Sendo assim, como na parcela mais produtiva obteve-se 16959 kg/ha, o ideótipo foi de 17000 kg/ha. Para os caracteres percentagem de grãos ardidos, percentagem de acamamento, percentagem de tombamento e percentagem de quebramento o ideótipo foi 0%. Para as notas de Ferrugem Comum e Cercosporiose o ideótipo obteve a nota 1.

3.3 Obtenção da matriz de variâncias e covariâncias

Para a obtenção da matriz de variâncias e covariâncias entre os caracteres avaliados, foi realizada a análise de variância multivariada (MANOVA), utilizando-se os dados de todos os híbridos, caracteres e locais. Para isso utilizou-se o modelo, conforme Ferreira (2011), expresso na forma vetorial na expressão (13)

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \epsilon_{ijk}, \quad (13)$$

em que o vetor $Y_{ijk} = [Y_{ijk1}, Y_{ijk2}, \dots, Y_{ijkp}]$ é referente às observações multivariadas associada ao i -ésimo híbrido ($i = 1, 2, \dots, 25$), ao j -ésimo local ($j = 1, 2, \dots, 6$) na k -ésima repetição dessa combinação dos níveis i e j dos dois fatores; μ é o vetor de constantes do modelo linear multivariado dado por $\mu_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{ip}]$, α_i é o vetor de efeitos do i -ésimo híbrido dado por $\alpha_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ip}]$, β_j é o vetor de efeitos do j -ésimo local dado por $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jp}]$, δ_{ij} é o vetor dos efeitos da interação entre o i -ésimo híbrido e o j -ésimo local dado por $\delta_{ij} = [\delta_{ij1}, \delta_{ij2}, \dots, \delta_{ijp}]$ e ϵ_{ijk} é o vetor de efeitos do erro experimental não observável, correspondente à observação Y_{ijk} .

Para a MANOVA foi utilizado o pacote manova do software R v 3.0, cujo detalhamento está no Anexo A (R DEVELOPMENT CORE TEAM, 2011).

3.4 Obtenção do Índice de Seleção de Distância para o Ideótipo

Foi obtido o índice de seleção de distância, em relação ao ideótipo (ISDI), utilizando-se os dados à nível de parcela das sete características avaliadas nos 25 híbridos. Sendo assim, para cada parcela de avaliação foi obtido um índice baseado na Distância Generalizada de Mahalanobis. Para se obter a Distância Generalizada de Mahalanobis, a nível de parcela, foi utilizado o modelo matricial conforme a expressão (12)

$$D_{G\ell}^2 = \text{diag}(IS^{-1}I'), \quad (12)$$

onde o termo $IS^{-1}I'$ é escrito da seguinte forma:

$$\begin{pmatrix} (Y_{1111} - Y_{1\ell}) & \cdots & (Y_{n111} - Y_{n\ell}) \\ \vdots & \ddots & \vdots \\ (Y_{1her} - Y_{1\ell}) & \cdots & (Y_{nher} - Y_{n\ell}) \end{pmatrix}_{(n \times h)} \begin{pmatrix} V_1 & \cdots & COV_{1n} \\ \vdots & \ddots & \vdots \\ COV_{n1} & \cdots & V_n \end{pmatrix}_{(n \times n)}^{-1} \begin{pmatrix} (Y_{1111} - Y_{1\ell}) & \cdots & (Y_{1her} - Y_{n\ell}) \\ \vdots & \ddots & \vdots \\ (Y_{n111} - Y_{n\ell}) & \cdots & (Y_{nher} - Y_{n\ell}) \end{pmatrix}_{(n \times h)}$$

em que

$D_{G\ell}^2$ é a matriz das distâncias dos h ($h = 1, 2, \dots, 25$) híbridos em relação ao ideótipo ℓ

Y_{ijkl} observação do caráter i ($i = 1, 2, \dots, n$) onde $n = 7$, do híbrido j ($j = 1, 2, \dots, h$), no local k ($k = 1, 2, \dots, e$) onde $e = 6$, na repetição l ($l = 1, 2$) V_1, V_2, \dots, V_n são as variâncias dos resíduos (QMerro da análise de variância multivariada) dos n caracteres avaliados

$COV_{12}, COV_{2n}, \dots, COV_{(n-1)n}$ são as covariâncias entre os resíduos desses caracteres.

Conforme representado acima, a matriz I (matriz dos desvios de cada observação em relação ao ideótipo) tem dimensão (N x n) onde N é o total de observações para cada caractere, ou seja, $h * e * r = 300$, a matriz S (matriz de variâncias e covariâncias) tem dimensão (n x n). Realizando todas as operações matriciais o termo detalhado acima resulta numa matriz de distâncias de dimensão (NxN), cuja diagonal principal corresponde às distâncias de cada parcela.

3.5 Análises de variância e análises por meio do modelo SREG (GGE)

Após a obtenção do ISDI para cada parcela, foi realizado o teste de normalidade de Shapiro-Wilk para esta variável e também para a produtividade de grãos. Em seguida foi realizada a análise de variância conjunta para verificar a presença de interação (GxA) para o ISDI e para a produtividade de grãos. Uma vez constatada a presença de interação GxA (teste F significativo), foi realizada a análise de adaptabilidade e estabilidade para as duas variáveis, que permitiu mensurar a adaptação e a estabilidade de cada híbrido sob teste, para o ISDI e para a produtividade. Tal avaliação foi feita utilizando o método GGE biplot. Como a interpretação do ISDI é o inverso da produtividade, ou seja, quanto menor o valor do índice melhor é o desempenho, o valor de cada parcela foi subtraído de 2000. Esse valor foi determinado usando-se o mesmo princípio utilizado na determinação de ideótipo de produtividade, ou seja, utilizando o próximo milhar acima do maior valor do índice, considerando todas as parcelas de todos os ensaios. Desta forma, a interpretação gráfica pôde ser realizada da mesma forma que para a produtividade. A análise GGE biplot foi realizada por meio do pacote GGEBiplotGUI do software R v 3.0, conforme detalhado no Anexo A (R DEVELOPMENT CORE TEAM, 2011).

A análise por meio do método GGE biplot foi realizada conforme apresentado por Oliveira et al. (2010), considerando o modelo simplificado para dois componentes principais

$$\bar{Y}_{ij} - \mu_j = \lambda_1 \gamma_{i1} \alpha_{j1} + \lambda_2 \gamma_{i2} \alpha_{j2} + \rho_{ij} + \bar{\varepsilon}_{ij},$$

em que

\bar{Y}_{ij} é a média do híbrido i no local j , para o ISDI ou produtividade;

μ_j é a média do local j ;

$\lambda_1 \gamma_{i1} \alpha_{j1}$ é o primeiro componente principal (PCA1), do efeito de genótipos (G) + interação (GxA);

$\lambda_2 \gamma_{i2} \alpha_{j2}$ é o segundo componente principal (PCA2), do efeito de genótipos (G) + interação (GxA);

λ_1 e λ_2 são os autovalores associados ao PCA1 e ao PCA2;

γ_{i1} e γ_{i2} são os escores do PCA1 e do PCA2, respectivamente, para genótipos;

α_{j1} e α_{j2} são os escores do PCA1 e do PCA2, respectivamente, para ambientes;

ρ_{ij} é o resíduo da interação genótipo x ambiente, também conhecido por 'ruído', correspondente aos componentes principais não retidos no modelo;

$\bar{\varepsilon}_{ij}$ é o residual do modelo com distribuição normal, com média zero e variância σ^2/r (onde σ^2 é a variância do erro entre parcelas para cada ambiente e r é o número de repetições).

3.6 Validação cruzada do modelo GGE2

A acurácia gráfica dos métodos de identificação de mega-ambientes e genótipos vencedores foi testada pelo procedimento de validação cruzada, proposto por Gabriel (2002). Com esse objetivo, as estatísticas PRESSm e PRESScorr foram usadas para medir a discrepância entre os valores observados e preditos e a correlação preditiva (DIAS; KRZANOWSKI, 2003).

A validação cruzada foi feita utilizando-se de uma partição na matriz de dados via decomposição por valores singulares (DVS). O algoritmo desenvolve-se a partir da submatriz $X_{\setminus 11}$ dada por

$$X_{G \times A} = \begin{bmatrix} x_{11} & X_1' \\ X_1 & X_{\setminus 11} \end{bmatrix} \quad (14)$$

em que $X_{\setminus 11} = \sum_{k=1}^r u_{(k)} d_{(k)} v_{(k)}' = UDV'$, sendo r o número de componentes multiplicativos em análise; u_k , d_k e v_k correspondentes, respectivamente aos elementos λ_k , γ_{ik} e α_{jk} ($k = 1, 2, \dots, r$) oriundos da DVS; $U = [u_1, u_2, \dots, u_r]$, $V = [v_1, v_2, \dots, v_r]$ e $D = [d_1, d_2, \dots, d_r]$. O valor predito para x_{11} é dado por $\hat{x}_{11} = x_1' V D^{-1} U' x_1$. O resíduo da validação cruzada é obtido por $e_{11} = x_{11} - \hat{x}_{11}$.

De forma similar, calculam-se todos os valores \hat{x}_{ij} ajustados pela validação cruzada e resíduos $e_{ij} = x_{ij} - \hat{x}_{ij}$ para todos os outros elementos x_{ij} , ($i = 1, 2, \dots, g$) onde g é o número de genótipos e ($j = 1, 2, \dots, a$) onde a é o número de ambientes. Os resíduos e os valores ajustados são resumidos por

$$PRESS_m = \frac{1}{ga} \sum_{i=1}^g \sum_{j=1}^a e_{ij}^2 \text{ e } PRECORR_{(m)} = Corr(x_{ij}, \hat{x}_{ij} | \forall i, j) \text{ respectivamente.}$$

O número de fatores multiplicativos correspondem ao r que proporcionou o menor $PRESS_m$ e conseqüentemente o maior $PRECORR_{(m)}$. Essa análise de validação cruzada foi realizada por meio do PROC IML do pacote estatístico do software SAS v 9.0, cuja rotina foi detalhada no Anexo B (SAS INSTITUTE, 2000).

3.7 Análise do GGE biplot e comparação dos híbridos selecionados

O gráfico GGE biplot foi construído a partir dos dados da decomposição das médias empíricas, apresentando graficamente qual é o híbrido com melhor desenvolvimento (OLIVEIRA et al., 2010). Para analisar as respostas, é feito um polígono unindo os pontos dos genótipos mais afastados da origem em cada quadrante, sendo que todos os outros híbridos estarão incluídos dentro do polígono formado. Posteriormente, é traçada uma perpendicular para cada lado do polígono com relação à origem e, assim, os ambientes e híbridos são separados em setores. Cada setor representa um estrato ambiental, no qual o genótipo localizado no vértice do polígono representa o genótipo campeão para esse setor ou estrato ambiental, por esse motivo essa abordagem gráfica foi denominada por Yan et al. (2000) como “Which Won Where”, que traduzido significa “Qual Venceu Onde”. Os genótipos em cada setor são os que apresentam o melhor desenvolvimento nos ambientes incluídos naquele setor.

Os híbridos com maior adaptabilidade e estabilidade foram identificados por meio da distância em relação ao “genótipo ideal” que foi plotado no centro de um “alvo” no biplot. Dessa forma, os híbridos plotados mais próximo ao

centro desse “alvo” são os que possuem a melhor combinação entre adaptabilidade e estabilidade, conforme apresentado por Balestre et al. (2010).

O gráfico biplot também pode indicar a existência de um ambiente ideal, dado por aquele que tem, no primeiro componente principal (PCA1), um escore alto, apresentando alta capacidade para discriminar os genótipos. Além disso, se o segundo componente principal (PCA2) for próximo de zero, isso indica que o ambiente proporciona a maior produtividade média.

Foram selecionados os cinco melhores híbridos avaliando-se o ISDI e a produtividade por meio da distância em relação ao “genótipo ideal” no biplot. Em seguida foram comparadas as classificações, obtidas nas duas análises, observando-se os dados de todas as variáveis originais, a fim de avaliar a eficiência do uso do ISDI.

Para essa comparação foram obtidas as médias de cada híbrido, considerando os seis locais, para cada variável. Essas médias foram ordenadas em ordem decrescente para a produtividade de grãos e crescente para as demais variáveis. Em seguida, cada híbrido recebeu um valor de rank que variou de 1 a 25 para cada variável, conforme a posição deste após a ordenação das médias. Foram então somados os valores de rank de cada um dos melhores híbridos selecionados nas análises GGEbiplot, para as sete variáveis, obtendo então o índice de soma dos ranks de Mulamba e Mock (1978).

Dessa forma, os híbridos com os menores valores do índice de Mulamba e Mock (1978) são os de melhor desempenho, considerando todas características avaliadas. Além disso foi comparado também o agrupamento dos ambientes nas duas situações.

4 RESULTADOS E DISCUSSÃO

A partir da análise de variância multivariada (MANOVA) foi obtida a matriz de variâncias e covariâncias residuais, envolvendo os seis locais de avaliação (Tabela 2).

Tabela 2 Matriz de variâncias e covariância residuais dos sete caracteres avaliados nos seis locais

	Produtividade (kg/ha)	Grãos Ardidos (%)	Acamamento (%)	Quebramento (%)	Tombamento (%)	Doença 1¹	Doença 2²
Produtividade	145685.7415	-364.0067	-210.5553	34.6681	-176.3815	-5.2515	-2.1502
Grãos Ardidos	-364.0067	23.3332	-1.6230	2.7520	1.1249	0.0969	0.1097
Acamamento	-210.5553	-1.6230	13.9921	-1.9613	12.0125	-0.0823	0.0203
Quebramento	34.6681	2.7520	-1.9613	27.0457	25.1025	0.1250	0.1542
Tombamento	-176.3815	1.1249	12.0125	25.1024	37.1152	0.0427	0.1742
Doença 1	-5.251514	0.0969	-0.0823	0.1250	0.0427	0.0864	0.0120
Doença 2	-2.1502	0.1097	0.0203	0.1542	0.1742	0.0120	0.0850

¹ Nota de Ferrugem Comum; ² Nota de Cercosporiose

Em seguida foi obtida a matriz dos desvios dos dados observados em relação ao ideótipo já estabelecido. De posse dessas matrizes foram obtidos os ISDI para cada parcela de todos os locais, os quais, antes da análise de variância foram submetidos ao Teste de Normalidade proposto por Shapiro-Wilk. O valor da estatística W foi de 0,9562 com um p-valor de $7,936 \times 10^{-8}$ sendo, portanto, rejeitada a hipótese de normalidade dos dados dos índices.

Como o objetivo dessa análise de variância foi somente o de verificar a significância do teste de F de Snedecor para a interação (GxA), foi realizada a análise de variância conjunta convencional, pois o teste de F é muito robusto não alterando portanto a significância do valor de F calculado. Mas recomenda-se que num próximo trabalho seja realizada a análise de variância via Bootstrap não-paramétrico, conforme descrita por Edington e Onghena (2007).

Desta forma foi realizada a análise de variância conjunta com os índices de seleção (Tabela 3).

Tabela 3 Resumo da análise de variância conjunta dos dados do índice de seleção de distância para o ideótipo (ISDI), para os 25 híbridos avaliados e os seis locais de avaliação

Fontes de Variação	GL	SQ	QM	F	p-valor
Genótipos (G)	24	17830100	742921	340,285	0,000 ***
Ambientes (A)	5	2378491	475698	217,887	0,000 ***
Bloco/Ambientes	6	15055	2509	1,1449	0,337 NS [†]
Interação (GxA)	120	11285092	94042	43,075	0,000 ***
Resíduo	144	314385	2183		
TOTAL	299				

[†] Não significativo; *** Significativo a 0,1% ; * Significativo a 5% de probabilidade.

A interação (GxA) foi significativa a 0,1% de probabilidade, demonstrando que o desempenho dos híbridos não foi coincidente nos diferentes ambientes de avaliação quanto ao ISDI (Tabela 3). Os resultados da análise de

variância para a produtividade de grãos foram semelhantes e estão apresentados na Tabela 4.

Tabela 4 Resumo da análise de variância conjunta dos dados de produtividade de grãos (kg/ha), para os 25 híbridos avaliados e os seis locais de avaliação

Fontes de Variação	GL	SQ	QM	F	p-valor
Genótipos (G)	24	535817675	22325736	148,887	0,000 ***
Ambientes (A)	5	472429471	94485894	630,110	0,000 ***
Bloco/Ambientes	6	114406	19068	0,127	0,993 NS ¹
Interação (GxA)	120	254371726	2119764	14,136	0,000 ***
Resíduo	144	21592990	149951		
TOTAL	299				
CV²(%)	3,3				

¹ Não significativo; *** Significativo a 0,1% ; * Significativo a 5% de probabilidade.

² Coeficiente de variação experimental (%)

As fontes de variação Genótipos, Ambientes e sua interação foram significativos a 0,1% de probabilidade. O coeficiente de variação foi muito baixo (3,3%) o que demonstra a alta precisão experimental dos ensaios.

A média dos híbridos, envolvendo os seis locais para a produtividade de grãos (kg/ha) e para o ISDI (já subtraído de 2000, para tornar a interpretação semelhante à da produtividade) foram apresentados na Tabela 5.

Tabela 5 Classificação decrescente dos híbridos quanto à sua produtividade média de grãos (kg/ha) e suas respectivas médias do ISDI, considerando os seis locais de avaliação

Híbridos	Produtividade (kg/ha)	ISDI
21	14713,84	1784,55
23	13923,10	1530,36
24	13903,24	1576,26
22	13844,71	1702,09
12	13166,87	1826,48
25	13146,49	1605,71
18	12589,40	1706,14
7	12448,18	1331,71

“continua”

Tabela 5 “conclusão”

Híbridos	Produtividade (kg/ha)	ISDI
19	12347,06	1733,43
3	11637,13	1067,36
20	11612,62	1534,22
8	11210,67	1654,61
4	11013,32	997,00
1	10999,73	1159,46
11	10984,10	1632,94
16	10956,94	1528,95
13	10870,05	1630,95
5	10850,68	1049,96
15	10636,55	1621,16
2	10603,71	1118,82
9	10563,95	1519,58
6	10558,86	1192,99
17	10334,76	1532,99
10	10280,67	1510,38
14	10059,97	1258,74

Como podemos observar na Tabela 5, os híbridos mais produtivos não são, necessariamente, os que possuem os maiores valores de ISDI. Isto se deve ao fato de que os híbridos mais produtivos nem sempre são os melhores nas outras características avaliadas e contempladas no ISDI.

Em seguida foi realizada a análise GGEbiplot a fim de se avaliar a adaptabilidade e a estabilidade dos híbridos avaliados quanto ao ISDI e a produtividade de grãos.

Primeiramente foi realizada a validação cruzada do modelo GGEbiplot para o ISDI e para a produtividade de grãos para verificar a acurácia preditiva desse modelo para as duas variáveis analisadas (Tabela 6). A produtividade foi transformada em t/ha antes da realização da análise.

Tabela 6 Estimativas obtidas para $PRESS_m$ (soma de quadrados do erro das predições) e $PRECORR_{(m)}$ (acurácia preditiva) na validação cruzada dos modelos GGE1 e GGE2 (Biplot com 1 e 2 componentes principais, respectivamente) para o ISDI e para a produtividade (ton/ha)

Modelo	ISDI		Produtividade	
	$PRESS_m$	$PRECORR_{(m)}$	$PRESS_m$	$PRECORR_{(m)}$
GGE1	36324,678	0,8086993	1,0547498	0,8656868
GGE2	21597,326	0,8913184	1,0179156	0,8708956

Neste trabalho foram apresentados somente os resultados da validação cruzada para os dois primeiros componentes principais, já que foi escolhido o modelo reduzido devido a dificuldade de se avaliar o biplot em mais de duas dimensões.

Nota-se, que para o modelo GGE2 a acurácia preditiva ($PRECORR_{(m)}$) foi alta para as duas variáveis analisadas, além disso os valores foram próximos quando se compara as duas variáveis (Tabela 4). Oliveira et al. (2010) e Balestre et al. (2010) avaliando estabilidade e adaptabilidade da produtividade de grãos em milho e arroz, respectivamente, obtiveram um $PRECORR$ de 0,8617 e 0,90, respectivamente. Valores estes muito próximos aos obtidos neste trabalho. Sendo assim, o modelo GGE2 foi acurado para as duas variáveis.

Em seguida, foi obtido o biplot e realizada a abordagem “Which Won Where” para identificar os grupos de ambientes ou mega-ambientes, avaliando-se o ISDI (Figura 2). Nessa figura os híbridos foram identificados de G1 a G25.

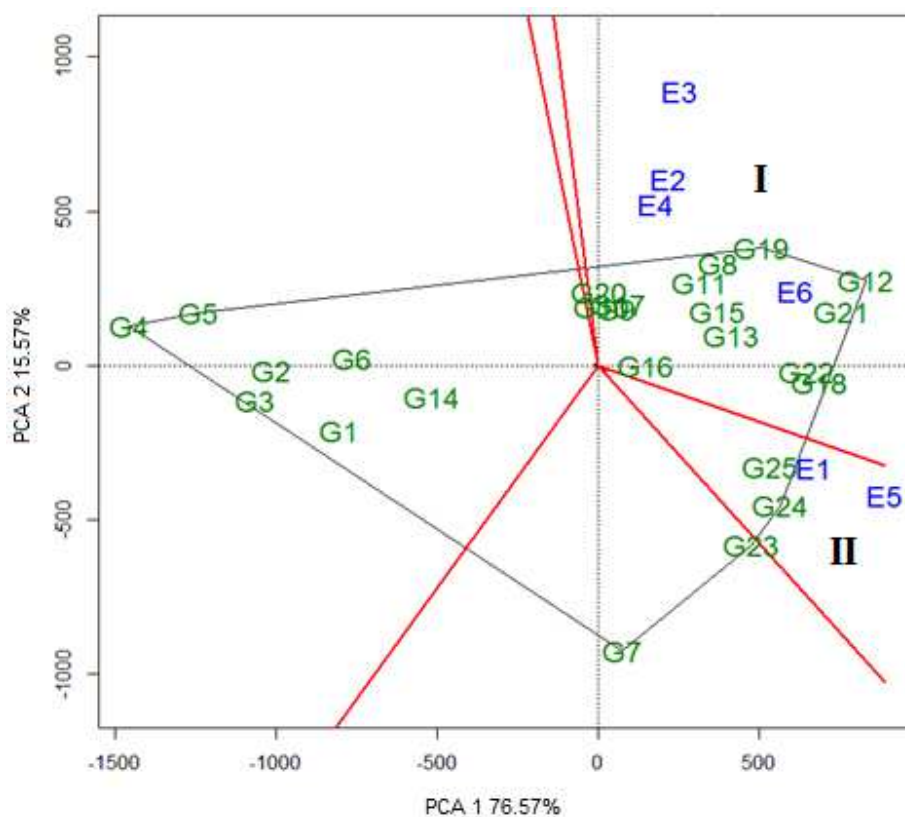


Figura 2 Gráfico GGEbiplot, avaliando o ISDI com o agrupamento dos seis locais (Vacaria-RS, Abelardo Luz-SC, Candió-PR, Canoinhas-SC, Castro-PR e Ponta Grossa-PR) identificados, respectivamente, de E1 a E6, sendo que os grupos de ambientes foram identificados pelos algarismos romanos I e II

O primeiro resultado a ser ressaltado é a alta explicação da Soma de Quadrados de G + GxA apresentada pelos dois componentes principais do biplot (PCA 1 e PCA 2) (Figura 1). Os dois componentes somados captaram 92,14%, mostrando que a análise foi muito eficiente. Quanto ao agrupamento, foram obtidos dois grupos, identificados no gráfico pelos numerais romanos I e II. O grupo I foi composto apenas pelos locais 2, 3, 4 e 6 (Abelardo Luz-SC, Candió-PR, Canoinhas-SC e Ponta Grossa-PR, respectivamente) no qual o híbrido 12 foi

o vencedor. O grupo II foi composto pelos locais 1 e 5 (Vacaria-RS e Castro-PR, respectivamente) no qual o híbrido 24 foi o vencedor. Nessa análise podemos observar também que o local 6 (Ponta Grossa-PR) foi o que apresentou maior capacidade de discriminar genótipos, devido ao seu alto score no primeiro componente principal (Axis 1).

O gráfico obtido na Figura 1 foi utilizado também para a avaliação da adaptabilidade e estabilidade dos híbridos, por meio da abordagem de distância em relação ao “genótipo ideal”, apresentado na (Figura 3).

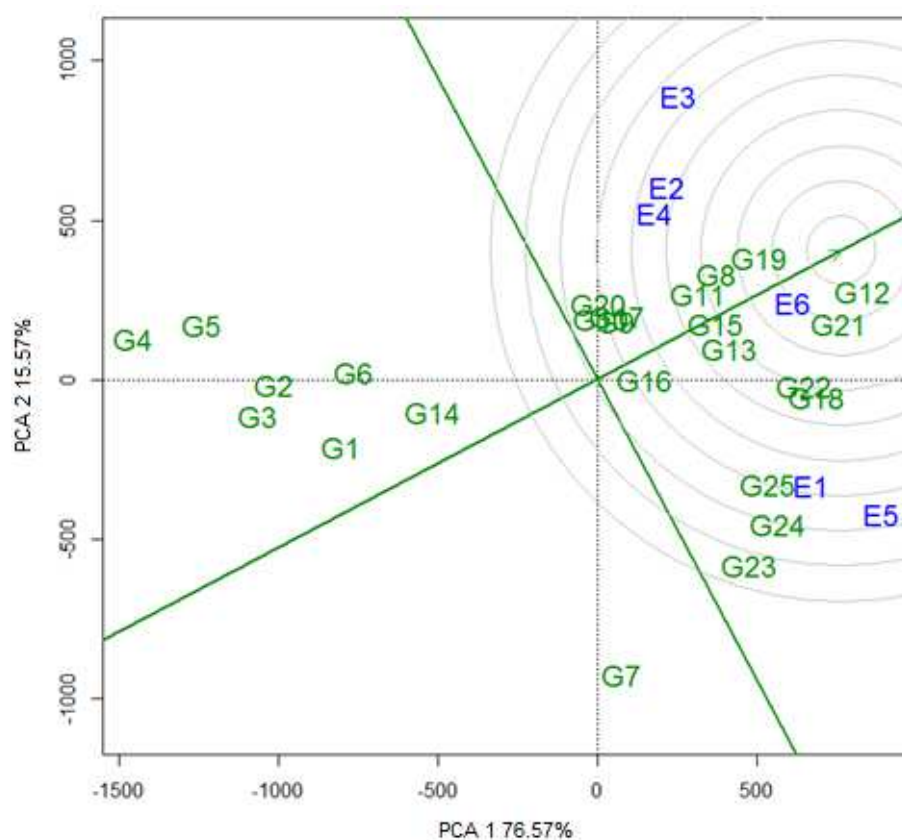


Figura 3 Gráfico GGEbiplot, com a distância dos 25 híbridos (identificados de G1 a G25) em relação ao “genótipo ideal”, quanto ao ISDI

Observa-se que os híbridos 12, 21, 19, 8 e 22 são, nesta ordem, os mais próximos ao “genótipo ideal”, portanto são os cinco melhores híbridos combinando adaptabilidade e estabilidade (Figura 3).

Foi realizada a análise GGE biplot também para os dados de produtividade, transformados em ton/ha, para melhorar a escala dos componentes principais do biplot. Com relação à porção da soma de quadrados de $G + (G \times A)$ explicada pelos dois componentes principais (Figura 3), essa também foi alta (87,24%), porém um pouco menor que a encontrada na análise utilizando o ISDI (92,14%). Da mesma forma foi realizada a abordagem “Which Won Where”, por meio da qual constatou-se que todos os locais ficaram no mesmo grupo (Figura 4).

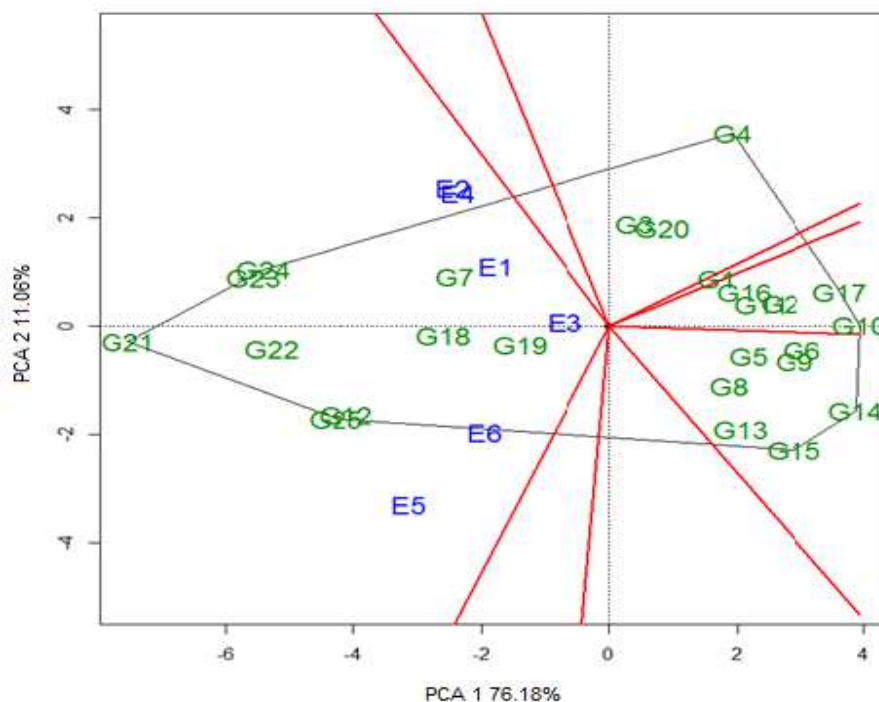


Figura 4 Gráfico GGEbiplot, avaliando a produtividade de grãos (t/ha), com o agrupamento dos seis locais (Vacaria-RS, Abelardo Luz-SC, Candió-PR, Canoinhas-SC, Castro-PR e Ponta Grossa-PR) identificados, respectivamente, de E1 a E6

No entanto, na análise com o ISDI foram identificados dois grupos. Essa diferença se deve ao fato do ISDI agregar outras variáveis, não somente a produtividade de grãos, o que torna a interação (GxA) mais complexa quando comparada à análise somente com a produtividade. Deste modo, isso torna vantajosa a utilização do ISDI no agrupamento de ambientes.

Em seguida foi feita a avaliação da adaptabilidade e estabilidade dos híbridos para a produtividade de grãos, por meio da distância desses em relação ao “genótipo ideal” (Figura 5).

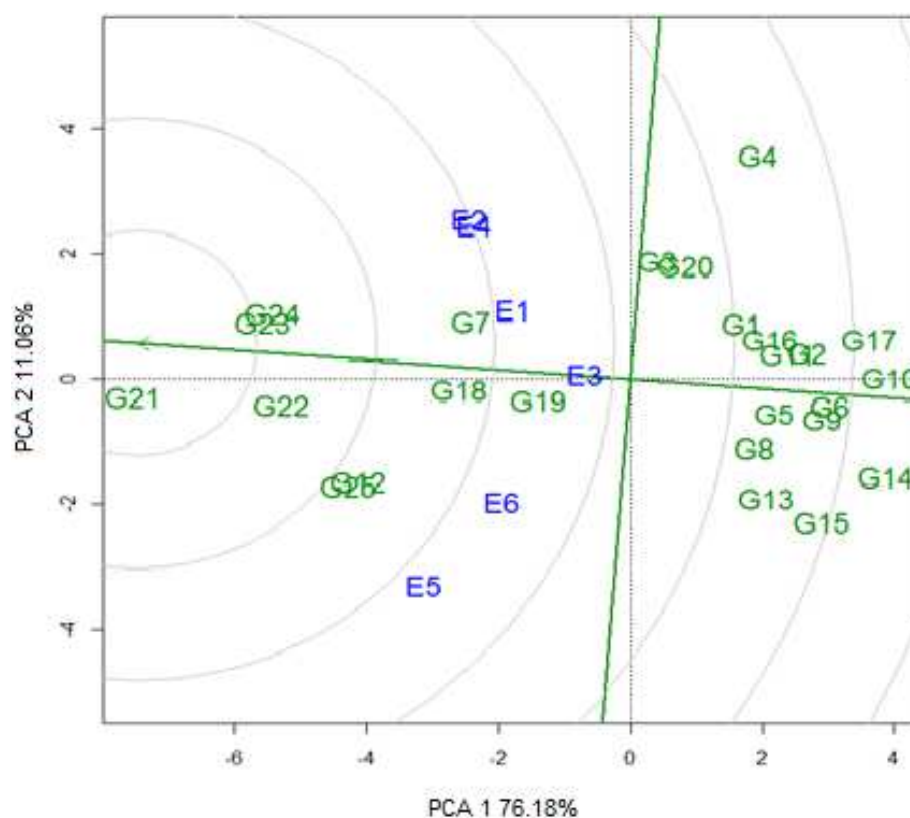


Figura 5 Gráfico GGEbiplot, com a distância dos 25 híbridos (numerados de G1 a G25) em relação ao “genótipo ideal”, quanto à produtividade de grãos (t/ha)

Os cinco híbridos mais próximos do “genótipo ideal” foram, nesta ordem, os híbridos 21, 23, 24, 22 e 25. Comparando essa seleção com a obtida pela análise com o ISDI (híbridos 12, 21, 19, 8 e 22), constata-se que somente os híbridos 21 e 22 foram comuns às duas seleções. Houve também uma mudança na posição dos híbridos 21 e 22 comparando as duas seleções. Para diagnosticar a razão dessa diferença foram obtidas as médias de cada híbrido, considerando os seis locais para cada uma das sete variáveis avaliadas. Em seguida, os híbridos envolvidos nas duas seleções foram ranqueados para cada variável. Para facilitar o diagnóstico os ranks desses híbridos foram somados, conforme o índice de Mulamba e Mock (1978). Esses resultados estão apresentados na Tabela 7.

Tabela 7 Ranqueamento considerando a média dos seis locais de avaliação, para cada um dos híbridos envolvidos nas seleções, por meio da análise GGEbiplot do ISDI e da produtividade (t/ha) para cada uma das sete variáveis avaliadas

Híbridos que apareceram na seleção por meio da produtividade (t/ha)									
Híbridos	Rank 1¹	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Soma²	Rank da Soma
21	1	5	6	10	10	9	12	53	2
22	4	4	21	20	20	10	15	94	7
23	2	11	19	2	8	6	22	70	6
24	3	1	9	12	11	2	19	57	4
25	6	12	5	14	12	4	17	70	6
Híbridos que apareceram na seleção por meio do ISDI									
8	12	22	11	3	5	12	4	69	5
12	5	9	10	6	6	3	1	40	1
19	9	16	4	8	4	8	6	55	3
21	1	5	6	10	10	9	12	53	2
22	4	4	21	20	20	10	15	94	7

¹ – Os Ranks de 1 a 7 correspondem, respectivamente, aos ranks das variáveis produtividade (ton/ha), % de grãos ardidos, % de acamamento, % de quebramento, % de tombamento, nota de Ferrugem Comum e nota de Cercosporiose.

² – Soma dos ranks de 1 a 7

Constata-se, que considerando os ranks das sete variáveis avaliadas para cada híbrido, os híbridos selecionados somente na avaliação por meio da produtividade (23, 24 e 25) apresentaram altos valores de rank para pelo menos duas variáveis (valores em negrito na Tabela 7). Sendo, que para a variável Doença 2 esses apresentaram os maiores valores de rank dentre todos os híbridos selecionados.

Por outro lado, os híbridos selecionados somente pelo ISDI (8, 12 e 19) apresentaram baixos valores de rank para a maioria das variáveis avaliadas. Sendo, que para as variáveis % de Quebramento, % de Tombamento e nota de Cercosporiose, esses híbridos apresentaram os menores valores de rank dentre todos os híbridos selecionados.

Garcia e Souza Júnior (1999) também utilizaram o ISDI para a seleção de híbridos de milho e incorporaram a esse os parâmetros de adaptabilidade e estabilidade b e R^2 , respectivamente, obtidos pelo método de Eberhart e Russel. Esse procedimento levou à seleção de híbridos estáveis e adaptados somente para a produtividade de espigas (kg/ha), diferente da análise realizada neste trabalho, a qual avaliou a adaptabilidade e estabilidade do ISDI estimado a nível de parcela, permitindo a seleção de híbridos adaptados e estáveis para todas as variáveis analisadas. Além disso, esses autores estimaram o ISDI por meio da Distância Euclidiana, o que interferiu negativamente na seleção obtida por eles, já que essa distância só pode ser aplicada a caracteres independentes entre si, o que não foi o caso dos caracteres utilizados por eles.

Por este motivo esses autores recomendaram que a correlação entre as variáveis avaliadas deveria ser considerada no índice de seleção. Essa correção foi realizada neste trabalho, pois o ISDI foi estimado por meio da distância de mahalanobis, a qual é ponderada pela matriz de variâncias e covariâncias.

De modo geral, a alternativa de análise da estabilidade e adaptabilidade do índice de seleção de distância para o ideótipo se mostrou mais eficiente do que

a análise tradicional, ou seja, por meio da adaptabilidade e estabilidade somente da produtividade de grãos e das médias das outras variáveis para cada híbrido.

No entanto, recomenda-se para trabalhos futuros que seja realizada a análise de variância do ISDI, por meio do método bootstrap não-paramétrico para a obtenção das médias ajustadas de cada híbrido para o ISDI.

Outra alternativa seria a aplicação do método REML/BLUP para a obtenção dos valores genéticos de cada característica, considerando-se as covariâncias entre as características para depois estimar o ISDI, o qual passaria a ser um índice de seleção genotípico. Dessa forma, poderia ser comparado o ganho genético da seleção por meio do ISDI genotípico com o ganho genético por meio da seleção tradicional.

Para tornar a seleção por meio do ISDI ainda mais acurada e confiável pode-se simular diferentes pesos econômicos para as características consideradas mais importantes pelo melhorista, identificando, assim, a combinação de pesos que torna a seleção por meio do índice mais eficiente.

5 CONCLUSÕES

A avaliação da adaptabilidade e estabilidade do ISDI proporcionou a seleção de híbridos que aliam adaptabilidade e estabilidade na maioria das variáveis avaliadas, sendo mais prático utilizá-lo do que analisar cada variável isoladamente.

A produtividade de grãos por si só não deve ser considerada como um índice de seleção, pois nem todos os híbridos mais produtivos foram os melhores nas outras variáveis analisadas.

REFERÊNCIAS

- AGROCERES. **Guia agrocere de sanidade**. São Paulo: Sementes Agrocere, 1996. 72 p.
- ALLARD, R. W.; BRADSHAW, A. D. Implications of genotype-environmental interactions in applied plant breeding. **Crop Science**, Madison, v. 4, n. 5, p. 503-508, Sept. 1964.
- ANICCHIARICO, P. **Genotype x environment interaction**: challenges and opportunities for plant breeding and cultivar recommendations. Rome: FAO, 2002. 115 p.
- ARUNACHALAM, V. Genetic distance in plant breeding. **Indian Journal of Genetics and Plant Breeding**, New Delhi, v. 41, n. 2, p. 226-236, 1981.
- BAKER, R. J. Tests for crossover genotype-environmental interactions. **Canadian Journal of Plant Science**, Ottawa, v. 45, n. 2, p. 405-410, Apr. 1988.
- BALESTRE, M. et al. Stability and adaptability of upland rice genotypes. **Crop Breeding and Applied Biotechnology**, Londrina, v. 10, p. 357-363, 2010.
- BECKER, H. C. Correlations among some statistical measures of phenotypic stability. **Euphytica**, Wageningen, v. 30, n. 3, p. 835-840, Dec. 1981.
- BERNARDO, R. **Breeding for quantitative traits in plants**. Minnesota: Stemma, 2002. 369 p.
- BRIM, C. A.; JOHNSON, H. W.; COCKERHAM, C. C. Multiple selection criteria in soybeans. **Agronomy Journal**, Madison, v. 51, p. 42-46, 1959.
- CROSSA, J.; CORNELIUS, P. L. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of errors variances. **Crop Science**, Madison, v. 37, n. 2, p. 406-415, Mar. 1997.
- CROSSA, J. Statistical analysis of multilocations trials. **Advances in Agronomy**, Newark, v. 44, n. 1, p. 55-85, Nov. 1990.
- CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2003. 585 p.

CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2001. 390 p.

DIAS, C. T. S.; KRZANOWSKI, W. J. Model selection and cross validation in additive main effect and multiplicative interaction models. **Crop Science**, Madison, v. 43, p. 865-873, 2003.

DUARTE, J. B.; VENCOSKY, R. **Interação genótipos x ambientes: uma introdução à análise AMMI**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. (Série Monografias, 9).

EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, n. 1, p. 36-40, Jan. 1966.

EDINGTON, E. S.; ONGHENA, P. **Randomization tests**. 4th ed. London: Chapman and Hall, 2007. 343 p.

ELSTON, R. C. A weight free index for the purpose of ranking of selection with respect to several traits at a time. **Biometrics**, Alexandria, v. 19, p. 85-97, 1963.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Resultados dos ensaios nacionais**. 2011. Disponível em: <<http://www.cnpms.embrapa.br/ensaio/index.php>>. Acesso em: 22 jan. 2012.

FARIAS, F. J. C. **Índice de seleção em cultivares de algodoeiro**. 2005. 121 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2005.

FARSHADFAR, E. Incorporation of AMMI stability value and grain yield in a single non-parametric index (GSI) in Bread Wheat. **Pakistan Journal of Biological Sciences**, Pakistan, v. 11, n. 14, p. 1791-1796, 2008.

FEHR, W. R. **Principles of cultivar development**. New York: Macmillan, 1987. 487 p.

FERREIRA, D. F. Análise de variância multivariada. In: _____. **Estatística multivariada**. 2. ed. rev. e ampl. Lavras: UFLA, 2011. p. 329-331.

FINLAY, K. W.; WILKINSON, G. N. The analysis of adaptation in a plant-breeding programme. **Australian Journal Agricultural Research**, Collingwood, v. 14, n. 6, p. 742-754, Nov. 1963.

FISHER, R. A. The use of multiple measurements in taxonomic problems. **Annals Eugenics**, San Francisco, v. 7, p. 179-189, 1936.

GABRIEL, K. R. Le biplot–outil d’exploration de données multidimensionnelles. **Journal de la Société Française de Statistique**, Paris, v. 143, p. 5-55, 2002.

GARCIA, A. A. F. **Índice para a seleção de cultivares**. 1998. 112 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 1998.

GARCIA, A. A. F.; SOUZA JÚNIOR, C. L. Comparação de índices de seleção não-paramétricos para a seleção de cultivares. **Bragantia**, Campinas, v. 58, n. 2, p. 253-267, 1999.

GAUCH, H. G. **Statistical analysis of regional yield trials: AMMI analysis of factorial designs**. New York: Elsevier, 1992. 278 p.

GAUCH, H. G.; ZOBEL, R. W. Identifying mega-environments and targeting genotypes. **Crop Science**, Madison, v. 37, p. 311-326, Mar. 1997

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. New Jersey: Prentice-Hall, 1998. 816 p.

KEMPTHORNE, O.; NORDSKOG, A. W. Restricted selection indices. **Biometrics**, Alexandria, v. 15, p. 10-19, 1959.

LIN, C. S.; BINNS, M. R.; LEFKOVITCH, L. P. Stability analysis: where do we stand? **Crop Science**, Madison, v. 26, p. 894-900, Sept. 1986.

MAHALANOBIS, P. C. On the generalized distance in statistics. **Proceedings of the National Institute of Sciences of India**, Calcutta, v. 2, n. 1, p. 49-55, 1936.

MARANHA, F. G. C. B. **Estratificação ambiental para avaliação de genótipos de algodoeiro no estado de Mato Grosso**. 2005. 79 p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queirós, Piracicaba, 2005.

MULAMBA, N. N.; MOCK, J. J. Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. **Egyptian Journal of Genetic and Cytology**, Giza, v. 7, p. 40-51, 1978.

OLIVEIRA, A. B.; DUARTE, J. B.; PINHEIRO, J. B. Emprego da análise AMMI na avaliação da estabilidade produtiva em soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 38, n. 3, p. 357-364, mar. 2003.

OLIVEIRA, R. L. et al. Evaluation of maize hybrids and environmental stratification by the methods AMMI and GGE biplot. **Crop Breeding and Applied Biotechnology**, Londrina, v. 10, p. 247-253, 2010.

PESEK, J.; BAKER, R. J. Desired improvement in relation to selection indices. **Canadian Journal of Plant Science**, Ottawa, v. 49, p. 803-804, 1969.

RAMALHO, M. A. P. et al. Aplicações da **genética quantitativa no melhoramento de plantas autógamas**. Lavras: UFLA, 2012. 522 p.

R DEVELOPMENT CORE TEAM. **R**: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2011.

ROBERTSON, A. The sampling variance of the genetic correlation coefficient. **Biometrics**, Arlington, v. 15, n. 3, p. 469-485, Sept. 1959.

SANTOS, V. S. **Seleção de pré-cultivares de soja baseada em índices**. 2005. 104 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2005.

SAS INSTITUTE. **SAS/STAT use's guide**. Cary, 2000.

SMITH, H. F. A discriminant function for plant selection. **Annals Eugenics**, Local, v. 7, p. 240-250, 1936.

SMITH, O. S.; HALLAUER, A. R.; RUSSELL, W. A. Use of index selection in recurrent selection programs in maize. **Euphytica**, Dordrecht, v. 30, p. 611-618, 1981.

TAI, G. C. C. Index selection with desired gains. **Crop Science**, Madison, v. 17, p. 182-183, 1977.

TERASAWA JÚNIOR, F.; VENCOVSKY, R.; KOEHLER, H. Environment and genotype-environment interaction in maize breeding in Paraná, Brazil. **Crop Breeding Applied Biotechnology**, Viçosa, MG, v. 8, n. 1, p. 17-22, Dec. 2008.

TOLER, J. E. **Patters of genotypic performance over environmental arrays**. 1990. 154 p. Thesis (Ph.D.) – Clemson University, Clemson, 1990.

VENCOVSKY, R.; BARRIGA, P. Genética biométrica no fitomelhoramento. **Revista Brasileira de Genética**, Ribeirão Preto, v. 14, n. 2, p. 437-453, maio 1992.

WRICKE, G.; WEBER, W. E. **Quantitative genetics and selection in plant breeding**. New York: W. Gruyter, 1986. 406 p.

YAN, W. et al. Cultivar evaluation and mega-environment investigation based on the GGE biplot. **Crop Science**, Madison, v. 40, n. 3, p. 597-605, May/June 2000.

YAN, W. et al. GGE biplot vs. AMMI analysis of genotype-by-environment data. **Crop Science**, Madison, v. 47, n. 2, p. 643-655, Mar./Apr. 2007.

ZOBEL, R. W.; WRIGHT, M. J.; GAUCH, H. G. Statistical analysis of a yield trial. **Agronomy Journal**, Stanford, v. 80, n. 3, p. 388-393, May 1988.

ANEXOS

ANEXO A – Rotina utilizada para as análises realizada no software R.

```
mult <- read.table("ManovaGeral.txt",header=TRUE)
mult
# Analise de Variância Multivariada (MANOVA)
mult$NT <- as.factor(mult$NT)
mult$rep <- as.factor(mult$rep)
mult$LOCAL <- as.factor(mult$LOCAL)
attach(mult)
mult1<-manova(cbind(KG_HA,GR_ARD,PACAM,PQUEB,
PTOMB,DÇA1,DÇA2)~NT+rep+LOCAL+NT*LOCAL)
names(mult1)
mult.summ <- summary(mult1)
mult.summ$SS
#Obtenção da matriz da soma de quadrados e produtos residuais
E <- mult.summ$SS$Residuals
E
df<-df.residual(mult1)
df
cov<-E/df #Obtenção da matriz de variâncias e covariâncias residuais
cov
mcov <-matrix(cov,ncol=7,byrow=TRUE)
mcov
dados <-c(# inserir os dados de todas as parcelas de todos os ensaios, cada
#coluna é uma variável. As colunas são separadas por vírgulas)
dados
xdados <-matrix(dados,ncol=7,byrow=TRUE)
xdados
```

```
idt <-c(# inserir os ideótipos de todas as parcelas de todos os ensaios, cada
#coluna é uma variável. As colunas são separadas por vírgulas)
midt <-matrix(idt,ncol=7,byrow=TRUE)
midt
mdist<-xdados-midt # Obtenção da matriz de desvios
mdist
tmdist<-t(mdist)#Matriz de desvios transposta
tmdist
icov<-solve(mcov)#Inversa da matriz de variâncias e covariâncias
icov
idc<-diag(mdist%*%icov%*%tmdist)#Obtenção do ISDI
idc
idc1<-matrix(idc,ncol=1,byrow=TRUE)
idc1
mult <-cbind(mult,IDI=idc1)
mult
attach(mult)
mult2<-aov(IDI~NT+rep+LOCAL+NT*LOCAL)#Anova do ISDI
names(mult2)
resid<-residuals(mult2)
resid
xresid<-matrix(resid,ncol=1,byrow=TRUE)
xresid
shapiro.test(xresid)
summary(mult2)
library(GGEBiplotGUI)
mdados1 <- read.table("Matriz_Medprod2SCR1.txt",header=TRUE)
GGEBiplot(mdados1)#Análise GGEBiplot
```

ANEXO B – Rotina para validação cruzada, conforme Gabriel (2002), por meio do Software SAS v 9.0

```
DATA RAW;
  INPUT GEN $ ENV $ YLD;
  CARDS; /* Inserir as médias de cada genótipo em cada local para a variável
analizada */
;
PROC SORT DATA=RAW;
  BY ENV GEN;
PROC MEANS DATA = RAW NOPRINT;
  BY ENV GEN;
  VAR YLD;
  OUTPUT OUT = MEDIAS MEAN=YLDM;
PROC GLM DATA=MEDIAS NOPRINT;
  CLASS ENV;
  MODEL YLDM = ENV/SS4 ;
  OUTPUT OUT=OUTRES R=RESID;
DATA OUTRES;
  SET OUTRES;
PROC SORT DATA=OUTRES;
  BY GEN ENV;
PROC TRANSPOSE DATA=OUTRES OUT=OUTRES2;
  BY GEN;
  ID ENV;
  VAR RESID;
DATA MEDIAS;
SET MEDIAS;
```



```
PROC SORT DATA=MEDIAS;
BY GEN ENV;
PROC TRANSPOSE DATA=MEDIAS OUT=MEDIAS2;
BY GEN;
ID ENV;
VAR YLDM;
PROC IML;
USE MEDIAS2;
READ ALL INTO YLDM;
USE OUTRES2;
READ ALL INTO RESID;
g=nrow(RESID);
e=ncol(RESID);
GE=RESID;
XO=GE;
Y=GE;
g=nrow(y);
e=ncol(y);
start rowcodel(x,r,c);
return(x[setdif(1:nrow(x),r),setdif(1:ncol(x),c)]);
finish;
start pegavrow(x,r,c);
return(x[r,setdif(1:ncol(X),c)]);
finish;
start pegavcol(x,r,c);
return(x[setdif(1:nrow(X),r),c]);
finish;
start press_M(ARG1,ARG2,ARG3,ARG4);
```

```

ARG5=(1/(ARG1*ARG2))*sum((ARG3-ARG4)##2);
return(ARG5); finish press_M;
ASSY=(1/(g*e))*sum((XO)##2);
DR=(g-1)*e;
Wm=SHAPE(0,e,1);
PRESSM=shape(0,e,1);
PRESSM2=shape(0,e,1);
do m=1 to e-1;
XE=shape(0,g,e);
do r=1 to g;
do c=1 to e;
Y=XO;
Y=rowcodel(Y,r,c);
call SVD(U,D,V,Y);
D=diag(D);
dd=ginv(D);
U1=U` ;
Y=XO;
Yrt=pegavrow(Y,r,c);
Yc=pegavcol(Y,r,c);
XE[r,c]=Yrt*V[,1:m]*dd[1:m,1:m]*U1[1:m,]*Yc;
END;
END;
YY=YLDM[,1:ncol(YLDM)];
g=nrow(YY);
e=ncol(YY);
YYi=j(g,e,0);
do i=1 to g;

```

```

YYi[i,]=j(1,e,sum(YY[i,])/e);
end;
YYj=j(g,e,0);
do j=1 to e;
YYj[,j]=j(g,1,sum(YY[,j])/g);
end;
YY_barra=j(g,e,sum(YY)/(g*e));
yyj=YYj-yy_barra;
Ychapeu=XE+YYj+YY_barra;
CORR=corr(colvec(YY)||colvec(Ychapeu));
print CORR;
DM=g+e-2*m;
DR=DR-(g+e-(m-1)*2);
PRESSM[m]=press_M(g,e,GE,XE);
PRINT PRESSM;
PRESSM2[m]=press_M(g,e,YY,Ychapeu);
PRINT PRESSM2;
if m=1 then
Wm[m]=((ASSY-PRESSM[m])/DM)/(PRESSM[m]/DR);
else;
Wm[m]=((PRESSM[m-1]-PRESSM[m])/DM)/(PRESSM[m]/DR);
Print Wm;
end;
quit;
run;

```