



DEOCLÉCIO DOMINGOS GARBUGLIO

**METODOLOGIA DE ESTRATIFICAÇÃO
AMBIENTAL E ADAPTABILIDADE VIA
ANÁLISE DE FATORES ASSOCIADA AOS
EFEITOS GENOTÍPICOS E DE INTERAÇÃO
GENÓTIPOS POR AMBIENTES**

**LAVRAS - MG
2010**

DEOCLECIO DOMINGOS GARBUGLIO

**METODOLOGIA DE ESTRATIFICAÇÃO AMBIENTAL E
ADAPTABILIDADE VIA ANÁLISE DE FATORES ASSOCIADA AOS
EFEITOS GENOTÍPICOS E DE INTERAÇÃO GENÓTIPOS POR
AMBIENTES**

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

Prof. Dr. Daniel Furtado Ferreira

LAVRAS - MG

2010

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Garbuglio, Deoclécio Domingos.

Metodologia de estratificação ambiental e adaptabilidade via
análise de fatores associada aos efeitos genotípicos e de interação
genótipos por ambientes / Deoclécio Domingos Garbuglio. – Lavras
: UFLA, 2010.

105 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2010.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. Análise multivariada. 2. Milho. 3. Estabilidade. 4.
Adaptabilidade. I. Universidade Federal de Lavras. II. Título.

CDD – 631.53

DEOCLECIO DOMINGOS GARBUGLIO

**METODOLOGIA DE ESTRATIFICAÇÃO AMBIENTAL E
ADAPTABILIDADE VIA ANÁLISE DE FATORES ASSOCIADA AOS
EFEITOS GENOTÍPICOS E DE INTERAÇÃO GENÓTIPOS POR
AMBIENTES**

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 20 de maio de 2010

Prof. Dr. João Bosco dos Santos	UFLA
Dr Pedro Mário de Araújo	IAPAR
Dr Nelson da Silva Fonseca Junior	IAPAR
Dr Juarez Campolina Machado	IAPAR

Prof. Dr Daniel Furtado Ferreira
Orientador

LAVRAS – MG

2010

DEDICO

a minha mãe, meu irmão e meus avós que, em momento algum, deixaram de acreditar em mim e, tanto nos momentos bons como nos ruins, sempre me incentivaram a continuar em frente.

O pensamento científico não deseja persuadir, mas inspirar; inspirar um outro pensamento, por em movimento o pensar.

(José Wilson da Silva, 2009)

AGRADECIMENTOS

À Universidade Federal de Lavras

À CAPES, pela concessão da bolsa de estudos

Ao professor Daniel Furtado Ferreira, pela orientação, ensinamentos, apoio pessoal em vários momentos e pelo exemplo de competência nos campos da Estatística e Experimentação Agropecuária. Levarei comigo o privilégio de ter sido seu orientado.

Ao professor João Cândido de Souza, pelo apoio profissional e ensinamentos transmitidos, assim como aos demais docentes do Programa de Pós-Graduação em Genética e Melhoramento de Plantas da UFLA.

Ao pesquisador do IAPAR Nelson da Silva Fonseca Junior, pelos ensinamentos nos campos da experimentação agrícola e genética quantitativa, amizade e apoio, pessoal e profissional, nos vários momentos em que necessitei.

Aos pesquisadores Pedro Mário de Araújo e Antonio Carlos Gerage, do Instituto Agrônomo do Paraná - IAPAR, pelos ensinamentos e amizade, desde a graduação.

Ao grande amigo Gustavo Barnabé Biudes, pela amizade e apoio ao longo dos anos, desde o período de graduação e a todos os colegas da Pós-Graduação em GMP, pela agradável convivência e momentos de descontração ao longo do curso.

Aos técnicos agrícolas do Programa Milho do IAPAR, Antonio Alves Ferreira, Aparecido Sales de Carvalho, Mario Campos Kogima, Mauro Souza Santos e Vanildo Marques de Lima, pela competência e compromisso com o trabalho.

Meu muito obrigado !!!

SUMÁRIO

	Página
CAPÍTULO 1 Introdução Geral	7
1 INTRODUÇÃO	10
2 REFERENCIAL TEÓRICO	13
REFERÊNCIAS	31
CAPÍTULO 2 Metodologia FGGA: descrição e aplicação em dados de cultivares de milho	37
1 INTRODUÇÃO	40
2 MATERIAL E MÉTODOS	43
2.1 Dados utilizados nas análises	43
2.2 Análises de variância	44
2.3 Estimação dos efeitos genotípicos e de interação G×A	46
2.4 Aplicação da análise de fatores à matriz G+GA	48
2.5 Estratificação ambiental	53
2.6 Análise de adaptabilidade	54
3 RESULTADOS E DISCUSSÃO	56
4 CONCLUSÕES	70
REFERÊNCIAS	71
CAPÍTULO 3 Metodologia FGGA: comparação de métodos para estudo da interação genótipos × ambientes	75
1 INTRODUÇÃO	78
2 MATERIAL E MÉTODOS	80
3 RESULTADOS E DISCUSSÃO	83
3.1 Método AMMI	83
3.2 Método SREG-GGE Biplot	91
3.3 Método FGGA	97
4 CONCLUSÕES	101
REFERÊNCIAS	103

CAPÍTULO 1 Introdução Geral

RESUMO GERAL

O presente estudo foi realizado com o objetivo de propor uma metodologia utilizando a análise de fatores associada aos efeitos genotípicos e de interação genótipos \times ambientes (FGGA) buscando realizar, simultaneamente, análises de estratificação ambiental e adaptabilidade de cultivares de milho indicadas para cultivo no estado do Paraná e compará-las às repostas obtidas pelos métodos AMMI e SREG-GGE Biplot. A análise de adaptabilidade baseada na análise de fatores foi realizada graficamente, por meio dos escores em relação aos fatores, separando em quadrantes de adaptabilidade. A análise de estratificação ambiental foi realizada a partir das informações sobre a magnitude das cargas fatoriais finais, obtidas após rotações pelo método *varimax*. A técnica FGGA é eficiente no processo de estratificação ambiental e análise de adaptabilidade. Acima de 70% de variação captada nos primeiros autovalores, representa parte expressiva da variação contida na soma dos efeitos genotípicos e de interação $G \times A$. As condições climáticas atuaram como um dos fatores preponderantes no agrupamento de ambientes, no entanto, puderam ser detectados pela técnica FGGA. Nos métodos GGE e FGGA, o efeito genotípico atuou como um coeficiente de ajuste das interações $G \times A$ ao longo dos ambientes, implicando em altos percentuais de explicação, associados a um menor número de eixos. A estratificação ambiental pelos métodos GGE e FGGA apresentaram repostas similares enquanto, pelo método AMMI, não houve garantia de agrupamento de ambientes. As análises de adaptabilidade apresentaram poucas divergências de reposta pelos três métodos. Os genótipos P30F35, P30F53, P30R50, P30K64 e AS 1570 apresentaram altas produtividades, associadas à adaptabilidade geral. O método FGGA permitiu associar as diferenças de repostas de produtividade entre determinados conjuntos de ambientes e o impacto em localidades pertencentes ao mesmo conjunto ambiental (através de r_A), com o auxílio do nível de porção simples atuante da interação $G \times A$.

Palavra-chave: Análise multivariada. Milho. Estabilidade e adaptabilidade.

GENERAL ABSTRACT

The objective of the present study was to propose a methodology using factor analysis associated with genotypic and genotype \times environment interaction effects (FGGE) and to accomplish, simultaneously, analyses of environmental stratification and adaptability of maize cultivars indicated for planting in Paraná State. The analysis of adaptability based on factor analysis was accomplished graphically through the scores in relation to the factors, separating in adaptability quadrants. The analyses of environmental stratification were accomplished starting from the information on the magnitude of the final factor loadings, obtained after rotations through *varimax* method. The FGGE technique is efficient in the process of environmental stratification and adaptability analysis. Above 70% of retained variation on first eigenvalues represents expressive part into sum of the genotypic and G \times E interaction effects. The climatic conditions were preponderant in grouping of environments, however, could be detected by FGGE technique. In FGGA and GGE methods, the genotypic effect participated as an adjusting coefficient for G \times E interactions on environments, resulting in high percentiles of explanations, associated to a lower number of axis. The environmental stratification by FGGA and GGE methods showed similar responses, while by AMMI method there was no warranty of environmental grouping. The adaptability analysis showed low divergences of responses by the three methods. The genotypes P30F35, P30F53, P30R50, P30K64 and AS 1570 showed high yields associated with general adaptability. The FGGA method allowed to associate difference of yield responses in specific regions and the impact in locations belonging to the same environmental group (through r_E), with the level of simple portion of G \times E interaction.

Keywords: Multivariate analysis. Maize. Stability and adaptability.

1 INTRODUÇÃO

Nos programas de melhoramento genético das mais diversas espécies, a interação genótipos \times ambientes ($G \times A$) dificulta a seleção e a indicação de cultivares, em razão da inconsistência de desempenho dos genótipos em ambientes distintos.

No sentido biométrico, duas abordagens têm sido empregadas: i) a dos estudos de estabilidade e adaptabilidade de diferentes genótipos em que se procura particularizar as respostas de cada genótipo diante das variações ambientais, para identificar aqueles de adaptabilidade ampla ou específica e, ainda, os de comportamento previsível e ii) a relativa aos métodos de estratificação ambiental, que consiste na subdivisão de regiões heterogêneas em sub-regiões mais uniformes, onde se exclui qualquer interação $G \times A$ significativa ou, em outras situações, uma interação significativa com predominância de parte simples atuante, ou seja, que não venha a comprometer a indicação das cultivares.

Dentre os diversos métodos utilizados na quantificação dos efeitos da interação $G \times A$, os mais empregados amplamente são aqueles baseados em regressão linear única, regressão bissegmentada, regressão não-linear e métodos não-paramétricos.

Nos últimos anos, tem sido dada especial atenção às técnicas multivariadas. Uma dessas, chamada de método AMMI, considera modelos aditivos para os efeitos principais e multiplicativos para os efeitos da interação (Crossa, 1990). Os trabalhos precursores da análise AMMI são tributados conforme citações de Gollob (1968), Mandel (1971) e Gauch e Zobel (1996) e de De Cauwer e Ortiz (1998). Segundo Duarte e Vencovsky (1999), dentre as denominações iniciais recebidas pela metodologia estão: FANOVA, MI, ACP duplamente centrada e análise biplot. Segundo os autores, esta última

denominação é a mais inadequada, uma vez que biplot refere-se apenas a um tipo de gráfico e não a um determinado modelo em particular.

A aplicação do método AMMI em estudos de interação $G \times A$ de diversas espécies, como soja, feijão, milho e trigo, assim como a utilização de outras técnicas multivariadas, tem sido ampliada nos últimos anos, em virtude das facilidades que os pacotes computacionais proporcionam nas realizações de cálculos complexos, principalmente envolvendo modelos lineares e álgebra de matrizes.

No entanto, segundo Resende e Thompson (2004), o método AMMI apresenta pelo menos cinco grandes limitações: i) considera os efeitos de genótipo e de $G \times A$ como fixos; ii) é adequado apenas para dados balanceados; iii) não considera a variação espacial dentro de ensaios; iv) não considera a heterogeneidade de variância entre ensaios e v) não considera a utilização de diferentes números de repetições entre os ensaios.

Outro método que vem recebendo destaque é a análise de fatores que consiste, basicamente, em reduzir um número elevado de variáveis originais a um pequeno número de variáveis abstratas, também chamadas de fatores. Esta técnica vem sendo empregada com maior frequência em estudos da interação $G \times A$ para diferentes espécies de interesse agrônomo, como soja, feijão e milho, utilizando gráficos bidimensionais para análise de adaptabilidade dos genótipos, a partir do agrupamento de ambientes nos fatores extraídos.

Por meio de plotagem gráfica, o método GGE tem se mostrado muito promissor no que se refere à facilidade na obtenção de informações sobre a adaptabilidade de genótipos e à composição de megambientes, tendo por princípio o processo de regressão por sítios a partir dos efeitos genotípicos associados aos efeitos de interação $G \times A$.

Já a técnica de análise de fatores considera como entrada uma matriz de médias fenotípicas, o que pode levar a conclusões equivocadas, tanto para a

adaptabilidade das cultivares avaliadas quanto o processo de estratificação ambiental, devido ao ruído presente na interação $G \times A$.

Desse modo, buscando-se melhorar a precisão tanto nas estimativas de adaptabilidade quanto na estratificação ambiental, por meio de uma redução no ruído presente na interação $G \times A$, o presente estudo foi realizado com os seguintes objetivos:

- a) propor uma metodologia utilizando a análise de fatores associada aos efeitos genotípicos e de interação genótipos x ambientes (FGGE), simultaneamente, e realizar análises de estratificação ambiental e adaptabilidade de cultivares de milho indicados para cultivo no estado do Paraná;
- b) comparar a metodologia proposta com os métodos AMMI e GGE biplot, aplicados em dados de cultivares de milho indicadas para cultivo no estado do Paraná.

2 REFERENCIAL TEÓRICO

Nas fases finais do melhoramento de plantas, as novas cultivares são plantadas em várias localidades, em diferentes condições de clima, fertilidade de solo e também em diferentes safras (ACCIARESI; CHIDICHIMO, 1999; BECKER; LÉON, 1988), de modo a constituir as chamadas redes de avaliações de cultivares. A combinação das diferentes condições, definidas pelas localidades e safras, pode ser considerada como um único fator para as condições ambientais (FERREIRA et al., 2006).

Quando um mesmo genótipo é avaliado em diferentes ambientes, geralmente ocorre de o ambiente alterar, diferentemente do previsto pelo modelo matemático empregado, a resposta do mesmo caráter em genótipos distintos, ou seja, pode ocorrer uma interação entre fatores. Isso significa que os efeitos genéticos e ambientais não são independentes, uma vez que as respostas dos genótipos podem diferir com as variações ambientais (SOUZA JÚNIOR; VENCOVSKY, 1989).

Para Kang e Gauch Júnior (1996), em termos genéticos, a interação ocorre quando a contribuição dos alelos dos diferentes genes que controlam o caráter ou o nível de expressão dos mesmos não é coincidente entre os ambientes. Isso porque a expressão dos genes é influenciada e ou regulada pelo ambiente. Segundo Chaves (2001), a interação $G \times A$ resulta, então, da resposta diferencial dos genótipos à variação ambiental e não deve ser vista como um simples efeito estatístico e, sim, como um fenômeno biológico em suas aplicações no melhoramento de plantas. No entanto, pode-se dizer que é com ferramental estatístico, juntamente com conhecimento do fenômeno no âmbito biológico, que se podem tirar melhores informações para entender a interação $G \times A$ e propor alternativas para minimizá-la ou aproveitá-la em benefício dos agricultores.

Desse modo, constata-se que, se não houvesse a interação G×A, uma dada cultivar poderia se adaptar à maioria dos ambientes de cultivo (locais, safras, condições de clima e solo), de maneira que um único ensaio poderia ser a base para uma recomendação generalizada (DUARTE; VENCOVSKY, 1999). Partindo desse ponto de vista, a interação G×A não deve ser considerada apenas como um problema, mas como uma oportunidade a ser aproveitada.

Nos experimentos para se quantificar os efeitos da interação G×A, segundo Ferreira et al. (2006), os delineamentos em blocos completos casualizados e blocos incompletos são amplamente utilizados. Para os experimentos em blocos incompletos, os delineamentos em látice são empregados devido ao grande número de genótipos que devem ser testados. O modelo mais simples e comum para a análise estatística de um conjunto de dados em diferentes ambientes pode ser representado por meio de um modelo linear:

$$Y_{ij} = \mu + g_i + a_j + ga_{ij} + \bar{\varepsilon}_{ij}$$

em que Y_{ij} é a média do i -ésimo genótipo avaliado no j -ésimo ambiente, para $i = 1, 2, \dots, p$ e $j = 1, 2, \dots, q$; μ é a média geral; g_i é o efeito fixo de genótipos em que $\sum_{i=1}^p g_i = 0$; a_j é o efeito fixo do j -ésimo ambiente, em que $\sum_{j=1}^q a_j = 0$; ga_{ij} é o efeito da interação entre o i -ésimo genótipo no j -ésimo ambiente.

Admitindo-se, inicialmente, um conjunto de G genótipos avaliados em A ambientes (locais, safras), estes podem ser representados por meio de uma tabela de dupla entrada (Tabela 1), em que $\bar{Y}_{i\bullet}$ são as médias marginais de genótipos, $\bar{Y}_{\bullet j}$ são as médias marginais de ambientes e $\bar{Y}_{\bullet\bullet}$, a média geral.

Tabela 1 Estimativas das médias de p genótipos em cada um dos q ambientes

Genótipos	Ambientes				Médias marginais (\bar{Y}_i)
	1	2	...	q	
1	Y_{11}	Y_{12}	...	Y_{1q}	\bar{Y}_1
2	Y_{21}	Y_{22}	...	Y_{2q}	\bar{Y}_2
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
p	Y_{p1}	Y_{p2}	...	Y_{pq}	\bar{Y}_p
Médias (\bar{Y}_j)	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$...	$\bar{Y}_{.q}$	$\bar{Y}_{..}$

Tomando-se um exemplo hipotético, no qual 2 genótipos são avaliados em 2 ambientes (Figura 1), podem ser visualizados os diferentes tipos de expressão da interação $G \times A$. Nas Figuras 1a e 1b está representada a ausência de interação, ou seja, os genótipos apresentam o mesmo nível de resposta à variação ambiental (1a) ou, então, mantêm-se indiferentes a esta variação (1b). A Figura 1c representa a interação do tipo simples, em que esta é caracterizada pela não alteração das posições relativas dos genótipos avaliados, dentro de um conjunto de ambientes, tomados dois a dois, havendo porém, diferença da magnitude das respostas.

A interação do tipo complexa, representada na Figura 1d, ocorre quando a correlação entre o desempenho dos genótipos ao longo dos ambientes em estudo é baixa, fazendo com que a posição relativa dos genótipos seja alterada em virtude das diferentes respostas às variações ambientais (ROBERTSON, 1959), ou seja, há uma inversão significativa na classificação desses genótipos ao longo dos ambientes. Em situações reais, nas quais os melhoristas dispõem de um grande número de genótipos e de ambientes, o mais comum é que ocorram as interações do tipo simples e complexa em conjunto, sendo esta última a que mais atrapalha o melhorista, principalmente no momento da recomendação das cultivares.

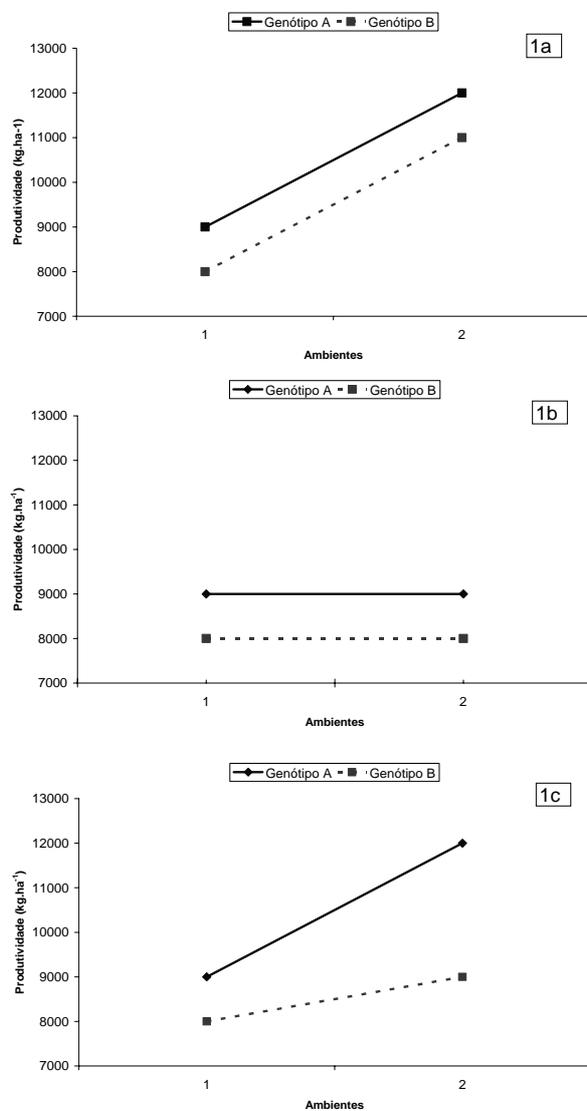


Figura 1 Padrões de comportamento dos genótipos em diferentes ambientes, destacando a ausência de interação (figuras 1a e 1b), bem como os tipos de interação simples (1c) e complexa (1d)

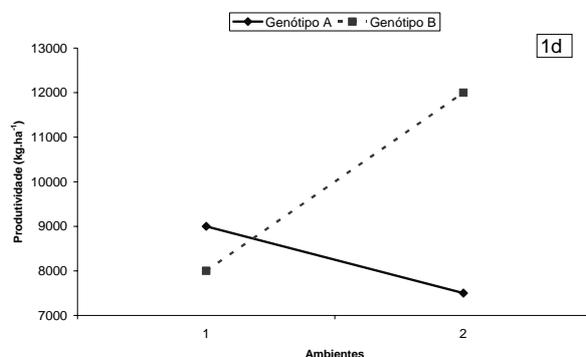


Figura 1, conclusão

Estudos a respeito da interação $G \times A$, embora de grande importância para o melhoramento, não oferecem informações detalhadas sobre o comportamento de cada genótipo, mediante as variações ambientais. Fatores bióticos, abióticos, estrutura física e química do solo são componentes que comumente identificam um ambiente. Estes fatores atuam sobre o genótipo que, conseqüentemente, irá produzir respostas diferentes em locais ou situações diferentes, influenciando, posteriormente, a estabilidade de produção no caso de plantas cultivadas (ALLIPRANDINI, 1992).

A partir das análises de adaptabilidade e estabilidade, torna-se possível a identificação de cultivares que apresentam comportamentos previsíveis e que sejam responsivos às variações ambientais. Atualmente, existe mais de uma dezena de métodos utilizados frequentemente no melhoramento de plantas para se avaliar a adaptabilidade e estabilidade fenotípica (CRUZ; REGAZZI; CARNEIRO, 2004). A diferença entre eles está nos parâmetros adotados para a sua avaliação, nos procedimentos biométricos empregados para avaliá-los ou na informação ou detalhamento de sua análise (CRUZ; CARNEIRO, 2003; VENCOVSKY; BARRIGA, 1992). No entanto, na aplicação de qualquer metodologia, o primeiro problema que se apresenta é a conceituação de estabilidade.

Para Mariotti et al. (1976), a estabilidade pode ser definida como a capacidade dos genótipos apresentarem desempenho o mais constante possível em função das variações ambientais, enquanto a adaptabilidade refere-se à capacidade dos genótipos de responderem vantajosamente à melhoria do ambiente. Este conceito de estabilidade é definido por Moraes (1980) como estabilidade de comportamento e não fenotípica, e é o que interessa ao melhorista, uma vez que está associado ao conceito de adaptabilidade, ou seja, a estabilidade de um dado genótipo determina, também, a confiabilidade nos parâmetros estimados de adaptabilidade.

Lin, Binns e Lefkovitch (1986) sugeriram três conceitos para estabilidade: i) a cultivar será considerada estável se sua variância entre ambientes for pequena; ii) a cultivar será estável se sua resposta aos ambientes for paralela ao desempenho médio de todas cultivares avaliadas e iii) será estável a cultivar que apresentar o quadrado médio do desvio da regressão baixo, próximo a zero, ou seja, alta confiabilidade na resposta estimada. Os três conceitos apresentados por Lin, Binns e Lefkovitch (1986) não levam em consideração as variações imprevisíveis do ambiente. Nesse sentido, Lin e Binns (1988) sugeriram o tipo 4, no qual a cultivar com maior estabilidade será aquela que apresentar menor quadrado médio da interação genótipos \times anos, ou seja, indivíduos estáveis frente às variações imprevisíveis.

Becker e Léon (1988) dividem a estabilidade em dois tipos: estática, associando aqueles genótipos que apresentam desempenho constante com as variações ambientais, e dinâmica, associando os genótipos que apresentam um comportamento previsível ao longo das variações ambientais.

Para que seja possível quantificar os níveis de interferência dos efeitos ambientais, diferentes métodos vêm sendo aplicados. Entre estes se destacam os não-paramétricos, os baseados em variância da interação cultivares \times ambientes, regressão linear única, regressão linear bissegmentada, regressão não-linear,

métodos multivariados e aqueles que integram métodos univariados e multivariados.

A regressão é o processo mais utilizado no estudo de estabilidade. Ela foi inicialmente proposta por Yates e Cochran (1938), porém, recebeu mais notoriedade por meio dos trabalhos de Finlay e Wilkinson (1963) e Eberhart e Russel (1966).

Os métodos propostos por Finlay e Wilkinson (1963) e Eberhart e Russel (1966) baseiam-se em análise de regressão linear unissegmentada, que medem a resposta de cada genótipo às variações ambientais. Sob os aspectos de análise e parâmetros estimados, as duas metodologias são bastante semelhantes, diferindo em alguns conceitos de estabilidade e adaptabilidade adotados.

No método de Finlay e Wilkinson (1963), para cada genótipo é computada uma regressão linear simples da variável dependente considerada, em relação a um índice ambiental, definido como a constante obtida pela diferença entre a média de todos os genótipos no ambiente e a média geral. Segundo Cruz, Regazzi e Carneiro (2004), nas análises de regressão, quando realizadas com dados previamente transformados para uma escala logarítmica, é possível induzir um alto grau de linearização. Os conceitos de adaptabilidade e estabilidade estão relacionados com os coeficientes de regressão linear e com a média do genótipo e nenhuma inferência é feita em relação ao ajuste das equações de regressão, pressupondo-se que a transformação logarítmica proporcione uma adequação satisfatória do modelo de regressão linear.

Eberhart e Russell (1966) expandiram o modelo proposto por Finlay e Wilkinson (1963), sob o aspecto de que tanto os coeficientes de regressão dos valores fenotípicos de cada genótipo em relação ao índice ambiental quanto os desvios desta regressão proporcionariam estimativas de parâmetros de estabilidade e adaptabilidade. Assim, pela metodologia destes autores, é considerado que um genótipo com coeficiente de regressão superior a 1,0 tem

comportamento consistentemente melhor em ambientes favoráveis, enquanto um que apresenta coeficiente de regressão inferior a 1,0 é tido como de desempenho relativamente melhor em ambientes desfavoráveis. A magnitude e a significância da variância dos desvios de regressão dão uma estimativa da previsibilidade do material genético. Desse modo, Eberhart e Russell (1966) consideram como genótipo ideal aquele que apresenta alta produção média, coeficiente de regressão igual a 1,0 e desvios de regressão tão pequenos quanto possíveis.

A diferença básica entre os modelos de Finlay e Wilkinson (1963) e Eberhart e Russell (1966) é que a primeira sugere uma transformação logarítmica dos dados, enquanto a segunda utiliza parâmetros de regressão e desvios para avaliar a estabilidade, a adaptabilidade e o padrão de resposta do genótipo ao ambiente.

Os modelos de regressão linear vêm recebendo críticas da comunidade científica (BECKER; LEON, 1998; TOLER; BURROWS, 1998), no entanto, são utilizados frequentemente. A principal crítica é baseada no fato de que o índice ambiental (I_j) não é independente da variável resposta (Y_{ij}), devido ao índice ser estimado a partir da resposta média ambiental. Porém, o efeito dessa dependência diminui com o aumento no número de genótipos. Uma segunda crítica está relacionada ao uso de estimadores viesados dos coeficientes de regressão devido à variável independente ser mensurada com erros (STORCK; VENCOSKY, 1994). O viés depende de p e da razão σ_e^2 / σ^2 , em que σ_e^2 é a variância ambiental e σ^2 é a variância residual. Porém, o viés reduz à medida que os efeitos ambientais aumentam. Outro problema que pode afetar a inferência é a violação da suposição de homogeneidade das variâncias do resíduo ambiental (FERREIRA et al., 2006).

Segundo Cruz, Regazzi e Carneiro (2004), as metodologias tradicionais são incapazes de reconhecer o genótipo ideal, uma vez que se avalia um único

coeficiente de regressão, o qual é estimado em uma única análise em que se consideram todos os ambientes testados. Nessas metodologias, o genótipo ideal corre o risco de ser descartado, pois, tendo uma dupla inclinação, seus desvios, que deveriam ser examinados nos diferentes ambientes, poderiam ser relativamente altos em relação à reta estimada.

Em relação à crítica quanto aos modelos de regressão linear simples serem baseados em um potencial padrão não linear, Verma, Chahal e Murty (1978) propuseram um método em que se buscou solucionar esse problema. Os ambientes são separados em dois grupos (favorável e desfavorável) e, então, é ajustado um modelo de regressão linear simples separadamente para cada parte. A divisão é feita com base no índice ambiental, que representa o desvio de cada média ambiental em relação à média geral. São considerados como ambientes desfavoráveis aqueles com índice negativo ou zero e, como ambientes favoráveis, aqueles com índices positivos. Recomenda-se, adicionalmente, que o subgrupo de ambientes favoráveis deva incluir, também, o ambiente de menor índice negativo para a continuidade das duas linhas de regressão (CRUZ; REGAZZI; CARNEIRO, 2004). Pode ser notado que esta classificação de ambientes é bastante simples e pode não ser satisfatório quando houver poucos genótipos (FERREIRA et al., 2006). Há, também, com esta aproximação erros aleatórios na classificação de ambientes quando o número de ambientes é pequeno. Por esta razão, Silva e Barreto (1985) propuseram um modelo de regressão linear bissegmentada em que os parâmetros de estabilidade e adaptabilidade seriam estimados por meio do ajuste de uma única equação.

Cruz, Torres e Vencovsky (1989) apresentaram uma extensão da metodologia proposta por Silva e Barreto (1985), tornando-a operacionalmente mais simples e com propriedades estatísticas mais adequadas aos propósitos do melhoramento. Esse modelo é descontínuo na junção dos dois segmentos de reta de ambientes desfavoráveis e favoráveis. Segundo Ferreira et al. (2006), essa

modificação foi adotada devido à existência de correlações residuais negativas entre os estimadores dos coeficientes de regressão dos ambientes favoráveis e desfavoráveis. Neste modelo, o genótipo tido como ideal apresenta alto desempenho médio de rendimento, alta estabilidade com bom ajustamento do modelo com, por exemplo, $\sigma_{di}^2 = 0$ e que seja pouco responsivo às condições ambientais adversas e capaz de responder positivamente quando as condições ambientais são melhoradas. Portanto, o genótipo ideal deve ter coeficiente de regressão menor do que 1,0 para ambientes desfavoráveis e maior do que 1,0 em ambientes favoráveis. Por outro lado, um genótipo indesejado apresenta coeficiente de regressão maior que 1,0 em ambientes desfavoráveis e menor que 1,0 em ambientes favoráveis.

Como no modelo de regressão linear simples, a crítica de que o índice ambiental não é independente da variável resposta (Y_{ij}) ainda permanece. Quanto aos vieses nos coeficientes de regressão, Storck e Vencovsky (1994) propuseram correções nos estimadores e nos teste de hipóteses. No entanto, se o número de ambientes é pequeno, o modelo é inapropriado e as inferências podem ser afetadas pela heterogeneidade das variâncias dos erros ambientais.

Crossa (1990) argumenta que a análise de regressão linear não é informativa se a linearidade falhar, é altamente dependente do grupo de genótipos e ambientes incluídos e tende a simplificar modelos de resposta, explicando a variação devido à interação em uma única dimensão, quando, na realidade, ela pode ser bastante complexa. Acrescenta-se aqui o fato de esses procedimentos, em geral, não informarem sobre as interações específicas de genótipos com ambientes (se positivas ou negativas), dificultando explorar vantajosamente os efeitos da interação (DUARTE; VENCOVSKY, 1999).

Toler (1990), buscando avaliar o padrão de comportamento genotípico em face do ambiente, por meio de um modelo e teste que mantivesse o rigor estatístico exigido nas aproximações por regressão, apresentou um modelo não

linear nos parâmetros em uma versão de regressão linear simples, análogo ao modelo de Eberhart e Russell (1966), e de regressão linear bissegmentado, análogo ao modelo de Cruz, Torres e Vencovsky (1989). Esses modelos são utilizados juntamente com o desempenho médio de cada genótipo, para agrupá-los de acordo com o seu padrão de resposta ao ambiente. Esse método fundamenta-se, basicamente, no fato de que o índice ambiental é um parâmetro a ser estimado e que os modelos são não-lineares nos parâmetros, exigindo métodos mais refinados de estimação.

Crossa (1990) sugere que a aplicação de métodos multivariados pode ser útil para melhor explorar as informações contidas nos dados. Desse modo, recomenda técnicas como a análise de componentes principais (ACP), a análise de agrupamentos e o procedimento *Additive Main Effects and Multiplicative Interaction*, ou AMMI. Vargas et al. (1999) citam que a análise AMMI é mais parcimoniosa que o modelo de análise de variância convencional em descrever a interação $G \times A$ e fornece melhor escopo para modelar e interpretar interação $G \times A$ que a regressão linear simples com base nas médias ambientais, pois a interação $G \times A$ pode ser modelada em mais de uma dimensão.

Zobel, Wright e Gauch Junior (1988) também reforçam a tese de que procedimentos univariados como a análise de variância têm limitações em detectar interação de fatores, mesmo em situações de elevada magnitude em termos de soma de quadrados. Como exemplos, citam casos em que a $SQ_{G \times A}$ atinge de 20% a 50% da SQ_{TOTAL} (de tratamentos: combinações de genótipos e ambientes) e, mesmo assim, o quadrado médio da interação $G \times A$ não atinge significância estatística (5% de probabilidade). Acrescentam que a regressão linear também explica, na maioria das vezes, apenas uma pequena porção da $SQ_{G \times A}$. Ademais, enquanto o método AMMI captou 71% dessa soma de quadrados, a análise pela regressão linear de Finlay e Wilkinson (1963) só o fez para 7,9%. Concluíram, então, que só uma análise apropriada permitiria captar

padrões agronômicos e estatísticos importantes presentes na interação G×A e que o método AMMI mostrou-se bastante promissor neste sentido.

A análise AMMI busca combinar, em um único modelo, componentes aditivos para os efeitos principais (genótipos e ambientes) e componentes multiplicativos para os efeitos da interação G×A (DUARTE; VENCOVSKY, 1999). Gauch e Zobel (1996) informam que esta análise pode ajudar tanto na identificação de genótipos de alta produtividade e largamente adaptados, como na realização do chamado zoneamento agronômico, com fins de recomendação regionalizada e seleção de locais de teste. Zobel, Wright e Gauch Junior (1988) sustentam que o método AMMI permite uma análise mais detalhada da interação G×A, garante a seleção de genótipos mais produtivos (capazes de capitalizar interações positivas com ambientes), propicia estimativas mais precisas das respostas genotípicas e possibilita uma fácil interpretação gráfica dos resultados da análise estatística.

Por meio da análise AMMI, a resposta média de um genótipo i num ambiente j é dada por:

$$Y_{ij} = \mu + g_i + a_j + \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \varepsilon_{ij}$$

$$\text{com } (ga)_{ij} \text{ modelado por } \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij}$$

Sob as restrições de identificabilidade: $\sum_i g_i = \sum_j a_j = \sum_i (ga)_{ij} = \sum_j (ga)_{ij} = 0$, além da média geral (μ) e do erro experimental médio (ε_{ij}), os demais termos do modelo resultam da chamada decomposição por valores singulares (DVS) da matriz de interações: $GA_{(gxa)} = [(g\hat{a})_{ij}]$. O procedimento para obter essa decomposição será tratado mais adiante. A matriz de interações é obtida como resíduo do ajuste aos efeitos principais, por ANOVA, aplicada à matriz de médias $Y_{(gxa)} = [Y_{ij}]$. Assim, λ_k é o k -ésimo valor singular de GA (escalar) e $\gamma_{k(gx1)}$ e $\alpha'_{k(1xa)}$ são os respectivos

vetores singulares (vetor coluna e vetor linha) associados a λ_k (GOOD, 1969; MANDEL, 1971; PIEPHO, 1995). Logo, γ_{ik} e α_{jk} são os elementos relacionados ao genótipo i e ao ambiente j dos vetores γ_k e α'_k , respectivamente. O índice k ($k=1,2, \dots, p$; em que $p = \min \{g-1, a-1\}$ é o posto da matriz GA), tomado até n no somatório ($n < p$), determina uma aproximação de mínimos quadrados para a matriz GA pelos n primeiros termos da DVS (GABRIEL, 1978; GOOD, 1969), deixando-se um resíduo adicional denotado por ρ_{ij} . Para $n = p$ não se tem mais a aproximação e sim uma decomposição exata da matriz, implicando em ρ_{ij} nulo.

No modelo, os componentes aditivos e multiplicativos podem ser ilustrados da seguinte forma:

$$Y_{ij} = \underbrace{\mu + g_i + a_j}_{\text{parte aditiva}} + \underbrace{\sum_{k=1}^p \lambda_k \gamma_{ik} \alpha_{jk}}_{(ga)_{ij} : \text{multiplicativa}} + \varepsilon_{ij}$$

$$Y_{ij} = \mu + g_i + a_j + \overbrace{\sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \sum_{k=n+1}^p \lambda_k \gamma_{ik} \alpha_{jk}}^{\square} + \varepsilon_{ij}$$

com : $n < p$ e $\sum_{k=n+1}^p \lambda_k \gamma_{ik} \alpha_{jk} = \rho_{ij}$.

Segundo Souza (2004), quando informações de variáveis ambientais externas como informações meteorológicas ou caracterização de solos estão disponíveis e também características genóticas ou fenotípicas como dados fisiológicos, precocidade, suscetibilidade a doenças, marcadores genéticos, dentre outras, elas podem ser correlacionadas com os escores genotípicos e ambientais estimados pelo AMMI. As informações dessas regressões podem ser sobrepostas no biplot AMMI, junto com escores genotípicos e ambientais (EEUWIJK, 1995), de forma que a interpretação da produtividade de grãos

explicada pela interação $G \times A$ é possível. Porém, as informações de variáveis ambientais externas não podem ser diretamente utilizadas no modelo AMMI.

Nesse sentido, trabalhos utilizando o modelo AMMI foram realizados por Gauch e Zobel (1996), que estudaram a interação entre ciclo de genótipos e fotoperíodo, na busca de uma modelagem que explicasse interação $G \times A$. Em outro trabalho, Oosterom et al. (1993) concluíram que a interação ciclo *versus* déficit hídrico foi responsável pela interação $G \times A$, observada em cevada (*Hordeum vulgare* L.), em ensaios em vários locais conduzidos na Síria e na África.

Posteriormente, outros estudos mostraram que a interação ciclo *versus* estresse hídrico e calor, em milheto [*Pennisetum glaucum* (L.) R. Br.], na Índia (OOSTEROM et al., 1996) e interações do ciclo *versus* estresse por baixa temperatura e altura de planta *versus* estresse hídrico, para trigo, na Itália, (ANNICCHIARICO; PERENZIN, 1994), foram responsáveis pela interação $G \times A$ observada. Outros exemplos nesta categoria foram citados por Gauch e Zobel (1996) e Eeuwijk (1995), propondo um método que impõe a covariância entre genótipos e ambientes no biplot GE, de forma que algumas causas de interação $G \times A$ possam ser visualizadas. Recentemente, esse procedimento foi adotado por Vargas et al. (1999), no estudo das causas da interação $G \times A$ em trigo.

Embora as estratégias possam diferir em amplos aspectos, é comum verificar que os diferentes métodos levam a conclusões semelhantes para um conjunto de dados específicos. Por exemplo, Baril et al. (1995) compararam regressão fatorial e os escores, com base na análise AMMI em batata-inglesa (*Solanum tuberosum* L.), de um conjunto de dados e obtiveram a mesma conclusão para a interação entre ciclo e estresse por frio ou calor, na explicação da interação $G \times A$ na produtividade de batata-inglesa.

Silva e Duarte (2006), avaliando diferentes métodos estatísticos de análise da interação $G \times A$, verificaram que, devido a uma associação relativamente fraca ($r_s = 0,52$) entre os métodos de Eberhart e Russel (1966) e AMMI, aliada à complementaridade de suas informações, seria possível indicar o seu uso combinado em estudos de estabilidade e adaptabilidade fenotípica. Os autores ainda inferem que métodos de estabilidade que se baseiam, exclusivamente, em coeficientes de regressão não se associam àqueles baseados na contribuição genotípica para a interação $G \times A$. Desse modo, é recomendável associá-los a um outro método com esse último princípio, ou introduzir medidas estatísticas dessa natureza. Utilizando o método de Eeuwijk (1995) de regressão pelos mínimos quadrados parciais e o método de regressão fatorial, Vargas et al. (1999) também chegaram a conclusões similares.

Uma modificação da análise AMMI foi apresentada por Yan et al. (2000), que busca agrupar o efeito de genótipo, que é um efeito aditivo, com a interação $G \times A$, efeito multiplicativo e submeter estes efeitos à análise de componentes principais, chamada de *Sites Regression* ou SREG.

A metodologia chamada de GGE biplot, segundo Souza (2004), tem sido reconhecida como revolucionária na abordagem com análise gráfica bidimensional para melhoramento e pesquisa genética, incluindo, também, análise de estabilidade e adaptabilidade na recomendação de cultivares, destacando-se das demais técnicas pela sua alta eficiência estatística na explicação da interação $G \times A$. Esta técnica integra a análise de variância com componentes principais e demonstra eficiência superior na explicação de uma maior proporção da soma de quadrados da interação, quando comparada com a análise de variância e regressão conjunta (YAN et al., 2000).

Souza (2004) menciona que o método GGE biplot pode auxiliar na interpretação de dados do tipo interação $G \times A$ para qualquer característica, sendo a de maior importância a produtividade; genótipos \times marcadores genéticos,

combinação de todos citados anteriormente, QTLs×ambientes, locais×fatores físicos, dialelos, genótipos hospedeiros×raças de patógenos, quaisquer outros dados de experimentos fatoriais que envolvem dois fatores ou matriz de dados bidimensionais.

Dentro do contexto de estudos da interação G×A por meio de técnicas multivariadas, outra metodologia que vem sendo utilizada nos últimos anos é a análise de fatores (AF). Conforme Murakami (2001) e Ferreira (2008), a técnica de AF foi inicialmente desenvolvida por psicólogos. Durante anos adquiriu linguagem e técnica própria e seu uso foi limitado em razão da complexidade dos cálculos envolvidos. Porém, com o desenvolvimento computacional, essa técnica tem sido retomada, abrangendo as mais diversas áreas.

A AF é uma técnica multivariada utilizada, entre outras formas, nos estudos de estratificação ambiental e que permite reduzir um número elevado de variáveis originais a um pequeno número de variáveis abstratas, também chamadas de fatores (CRUZ; CARNEIRO, 2003). Cada fator irá agrupar um conjunto de ambientes, altamente correlacionados entre si e fracamente correlacionados com os ambientes, agrupados nos demais fatores.

Como os estudos de adaptabilidade, estabilidade e de estratificação ambiental vêm sendo realizados de forma dissociada, Murakami e Cruz (2004) propuseram um método que contempla, simultaneamente, a análise de adaptabilidade e de estratificação ambiental, por meio do princípio da similaridade da performance fenotípica, baseada na técnica de análise de fatores (AF).

Na utilização da AF destinada à estratificação de ambientes, deve-se interpretar o significado de cada fator e avaliar o interesse por ele. São analisados os elementos ou cargas fatoriais, identificando, para cada variável (que neste caso representa um determinado ambiente), os fatores com coeficientes de ponderação próximos da unidade. O agrupamento de ambientes é

realizado, portanto, a partir das informações sobre a magnitude das cargas fatoriais finais (obtidas após rotações). Cargas fatoriais maiores ou iguais a 0,70, em valor absoluto, indicam ambientes com altas correlações e são agrupadas dentro de cada fator; cargas fatoriais com valores baixos ($\leq 0,5$) indicam que o respectivo ambiente não deva pertencer ao grupo e cargas fatoriais com valores intermediários não garantem nenhuma definição de agrupamento.

Antes de serem obtidas as cargas fatoriais que permitirão identificar o melhor agrupamento de ambientes, deve-se estabelecer o número de fatores a serem analisados. O número de fatores finais pode ser admitido como igual ao número de autovalores maiores ou iguais a 1,0. No entanto, nos casos em que a proporção da variabilidade explicada pelos autovalores maiores que 1,0 for baixa, podem ser considerados mais fatores, até que se atinja uma proporção adequada da variabilidade, geralmente próximo de 80% da variação total.

A análise de adaptabilidade fenotípica, baseada na análise de fatores, é realizada graficamente por meio de escores obtidos em relação aos fatores. De acordo com Murakami e Cruz (2004), nos quadrantes II e IV, estarão os genótipos com adaptabilidade específica aos ambientes, agrupados em cada um dos dois fatores utilizados na plotagem gráfica. No quadrante I, estarão os genótipos de adaptabilidade ampla, ou seja, apresentarão escores altos para os ambientes agrupados nos dois fatores, simultaneamente. No quadrante III estarão os genótipos de baixa performance, passíveis de não indicação de cultivo para os ambientes em estudo.

Garbuglio et al. (2007), estudando a adaptabilidade de genótipos de milho e realizando a estratificação em 22 ambientes no estado do Paraná, constataram que a adaptabilidade, analisada por meio de plotagens gráficas a partir dos escores fatoriais, apresentou respostas diferenciadas, quando comparada ao método de regressão bissegmentada, sendo que AF mostrou-se eficiente tanto

nos processos de estratificação ambiental quanto nos estudos de adaptabilidade dos genótipos de milho.

Mendonça et al. (2007), avaliando a eficiência de diferentes métodos de estratificação ambiental, a representatividade dos locais de avaliação e a adaptabilidade e estabilidade de genótipos de soja nos estados do Paraná e Santa Catarina, verificaram que a técnica de AF associada ao porcentual de parte simples da interação $G \times A$ foi mais seletiva no processo de estratificação ambiental do que o método tradicional de Lin (1982).

REFERÊNCIAS

ACCIARESI, H. A.; CHIDICHIMO, H. O. Genotype environment interaction in *Avena sativa* L.: employing AMMI and factorial correspondence models. **Pesquisa Agropecuária Brasileira**, Brasília, v. 34, n. 10, p. 1823-1830, out. 1999.

ALLIPRANDINI, L. F. **Estudo dos efeitos ambientais, estabilidade, adaptabilidade e ganho genético em linhagens de soja (*Glycine max* (L.) Merrill) no Estado do Paraná**. 1992. 122 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Universidade Estadual de Londrina, Londrina, 1992.

ANNICCHIARICO, P.; PERENZIN, M. Adaptation patterns and definition of macro-environments for selection and recommendation of common wheat genotypes in Italy. **Plant Breeding**, Berlin, v. 113, n. 3, p. 197-205, Nov. 1994.

BARIL, C. P. et al. Analysing genotype by environment interaction in Dutch potato variety trials using factorial regression. **Euphytica**, Wageningen, v. 82, n. 2, p. 149-155, Jan. 1995.

BECKER, H. C.; LEON, J. Stability analysis in plant breeding. **Plant Breeding**, Berlin, v. 101, n. 1, p. 1-23, Apr. 1988.

CHAVES, L. J. Interação de cultivares com ambientes. In: NASS, L. L.; VALOIS, A. C. C.; MELO, I. S.; VALADARES, M. C. **Recursos genéticos e melhoramento: plantas**. Rondonópolis: Fundação MT, 2001. p. 673-713.

CROSSA, J. Statistical analyses of multilocations trials. **Advances in agronomy**, San Diego, v. 44, p. 55-85, Dec. 1990.

CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2003. v. 2, 585 p.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2004. v. 1, 480 p.

CRUZ, C. D.; TORRES, R. A. de; VENCOVSKY, R. An alternative approach to the stability analysis proposed by Silva and Barreto. **Revista Brasileira de Genética**, Ribeirão Preto, v. 12, n. 2, p. 567-580, maio 1989.

DE CAUWER, I.; ORTIZ, R. Analysis of the genotype x environment interaction in *Musa* trials. **Experimental Agriculture**, Cambridge, v. 34, n. 2, p. 177-188, Apr. 1998.

DUARTE, J. B.; VENCOVSKY, R. **Interação genótipos x ambientes: uma introdução a análise "AMMI"**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 59 p.

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

EEUWIJK, F. A. van. Linear and bilinear models for the analysis of multienvironment trials: I. An inventory of models. **Euphytica**, Wageningen, v. 84, n. 1, p. 1-7, Feb. 1995.

FERREIRA, D. F. **Estatística multivariada**. Lavras: Ed. UFLA, 2008. 662 p.

FERREIRA, D. F. et al. Statistical models in agriculture: biometrical methods for evaluating phenotypic stability in plant breeding. **Cerne**, Lavras, v. 12, n. 4, p. 373-388, dez. 2006.

FINLAY, K. W.; WILKINSON, G. N. The analysis of adaptation in a plant breeding program. **Australian Journal of Agriculture Research**, East Melbourne, v. 14, n. 6, p. 742-754, Dec. 1963.

GABRIEL, K. R. Least squares approximation of matrices by additive and multiplicative models. **Journal of the Royal Statistical Society. Serie B. Methodological**, London, v. 40, n. 2, p. 186-196, May 1978.

GARBUGLIO, D. D. et al. Análise de fatores e regressão bissegmentada em estudo de estratificação ambiental e adaptabilidade em milho. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 2, p. 183-191, Feb. 2007.

GAUCH, H. G.; ZOBEL, R. W. AMMI analysis of yield trials. In: KANG, M. S.; GAUCH JÚNIOR, H. G. (Ed.). **Genotype-by-environment proved and under what conditions this can be most environment interaction**. Boca Raton: CRC, 1996. p. 1-40.

GOLLOB, H. F. A statistical models wich combines features of factor analytic and analysis of variance techniques. **Psychometrika**, v.33, n.1, p.73-115, 1968.

GOOD, I. J. Some applications of the singular decomposition of a matrix. **Technometrics**, Washington, v. 11, n. 4, p. 823-831, Nov. 1969.

KANG, M. S.; GAUCH JÚNIOR, H. G. **Genotype by environment interaction**. New York: CRC, 1996. 416 p.

LIN, C.S. Grouping genotypes by a cluster method directly related to genotype-environment interaction mean square. **Theoretical and Applied Genetics**, v.62, p.277-280, 1982.

LIN, C. S.; BINNS, M. R. A method of analyzing cultivar x location x year experiments: a new stability parameter. **Theoretical and Applied Genetics**, Berlin, v. 76, n. 3, p. 425-430, Sept. 1988

LIN, C. S.; BINNS, M. R.; LEFKOVITCH, L. P. Stability analysis: where do we stand? **Crop Science**, Madison, v. 26, n. 5, p. 894-900, Oct. 1986.

MANDEL, J. A. New analysis of variance model for non-additive data. **Technometrics**, Washington, v. 13, n. 1, p. 1-18, Feb. 1971.

MARIOTTI, J. A. et al. Analisis de estabilidad y adaptabilidad de genótipos de caña de azúcar. I. Interacciones dentro de una localidad experimental. **Revista Agronómica del Noroeste Argentino**, Tucuman, v. 13, n. 1/4, p. 405-412, jan. 1976.

MENDONÇA, O. et al. Análise de fatores e estratificação ambiental na avaliação da adaptabilidade e estabilidade em soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 11, p. 1567-1575, nov. 2007.

MORAIS, O. P. **Adaptabilidade, estabilidade de comportamento e correlações fenotípicas, genotípicas e de ambiente em variedades e linhagens de arroz (*Oryza sativa* L.)**. 1980. 70 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Universidade Federal de Viçosa, Viçosa, MG, 1980.

MURAKAMI, D. M. **Novas metodologias de análise de interação genótipos x ambientes**: análise combinada de estratificação, adaptabilidade e estabilidade e análise de representatividade ambiental. 2001. 155 p. Tese (Doutorado em Genética e Melhoramento de Plantas) – Universidade Federal de Viçosa, Viçosa, MG, 2001.

MURAKAMI, D. M.; CRUZ, C. D. Proposal of methodologies for environment stratification and analysis of genotype adaptability. **Crop Breeding and Applied Biotechnology**, Londrina, v. 4, n. 1, p. 7-11, jan. 2004.

OOSTEROM, E. J. van. et al. Genotype-by environment interactions of barley in the Mediterranean region. **Crop Science**, Madison, v. 33, n. 4, p. 669-674, July 1993.

PIEPHO, H. P. Robustness of statistical test for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trial. **Theoretical Applied of Genetics**, Berlin, v. 90, n. 3/4, p. 438-443, Mar. 1995.

RESENDE, M. D. V.; THOMPSON, R. Factor analytic multiplicative mixed models in the analysis of multiple experiments. **Revista de Matemática e Estatística**, São Paulo, v. 22, n. 2, p. 31-52, fev. 2004.

ROBERTSON, A. **Experimental design on the measurement of heritabilities and genetic correlations**: biometrical genetics. New York: Pergamon, 1959. 186 p.

SILVA, J. G. C.; BARRETO, J. N. Aplicação da regressão linear segmentada em estudos da interação genótipo x ambiente. In: SIMPÓSIO DE EXPERIMENTAÇÃO AGRÍCOLA, 1., 1985, Piracicaba. **Anais...** Piracicaba: ESALQ, 1985. p. 49-50.

SILVA, W. C. J.; DUARTE, J. B. Métodos estatísticos para estudo de adaptabilidade e estabilidade fenotípica em soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 1, p. 23-30, jan. 2006.

SOUZA JUNIOR, C. L.; VENCOVSKY, R. Covariância entre parentes na presença da interação genótipos x ambientes. In: SIMPÓSIO DE ESTATÍSTICA APLICADA A EXPERIMENTAÇÃO AGRONÔMICA, 3., 1989, Lavras. **Resumos...** Lavras: ESAL, 1989. p. 50-51.

SOUZA, H. N. **Análise AMMI e SREG da interação genótipos x em milho**. 2004. 106 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Viçosa, Viçosa, MG, 2004.

STORCK, L.; VENCOVSKY, R. Stability analysis based on a bi-segmented discontinuous model with measurement errors in the variables. **Revista Brasileira de Genética**, Ribeirão Preto, v. 17, n. 1, p. 75-81, jan. 1994.

TOLER, J. E. **Patterns of genotypic performance over environmental arrays**. 1990. 154 p. Thesis (PhD in Agronomy) - Clemson University, Clemson, 1990.

TOLER, J. E.; BURROWS, P. M. Genotype performance over environmental arrays: a non-linear grouping protocol. **Journal of Applied Statistics**, Abingdon, v. 25, n. 1, p. 131-143, Feb. 1998.

VARGAS, W. et al. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype-by-environment interaction. **Crop Science**, Madison, v. 39, n. 4, p. 955-967, July 1999.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sociedade Brasileira de Genética, 1992. 496 p.

VERMA, M. M.; CHAHAL, G. S.; MURTY, B. R. Limitations of conventional regression analysis: a proposed modification. **Theoretical and Applied Genetics**, Berlin, v. 53, n. 2, p. 89-91, Sept. 1978.

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

YATES, F.; COCHRAN, W. G. The analysis of group of experiments. **Journal of Agricultural Science**, Cambridge, v. 28, n. 4, p. 556-580, Oct. 1938.

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

**CAPÍTULO 2 Metodologia FGGA: descrição e aplicação em dados de
cultivares de milho**

RESUMO

O presente estudo foi realizado com o objetivo de propor uma metodologia utilizando a análise de fatores associada aos efeitos genotípicos e de interação genótipos \times ambientes (FGGA) e realizar, simultaneamente, análises de estratificação ambiental e adaptabilidade de cultivares de milho indicadas para cultivo no estado do Paraná. A análise de adaptabilidade baseada na análise de fatores foi realizada graficamente, por meio dos escores em relação aos fatores, separando em quadrantes de adaptabilidade. A análise de estratificação ambiental foi realizada a partir das informações sobre a magnitude das cargas fatoriais finais, obtidas após rotações pelo método *varimax*. A técnica FGGA é eficiente no processo de estratificação ambiental e análise de adaptabilidade. Acima de 70% de variação captada nos primeiros autovalores representa parte expressiva da variação contida na soma dos efeitos genotípicos e de interação $G \times A$. As condições climáticas atuaram como um dos fatores preponderantes no agrupamento de ambientes, no entanto, puderam ser detectados pela técnica FGGA. As cultivares com ampla adaptação e alta produtividade ao conjunto de ambientes testados foram P30F35, P30F53, P30K64, P30R50 e AS 1570.

Palavras-chave: Análise multivariada. Análise de fatores. Efeito genotípico. Interação genótipo \times ambiente.

ABSTRACT

The objective of the present study was to propose a methodology using the factor analysis associated with genotypic and genotype \times environment interaction effects (FGGE) and to accomplish, simultaneously, analyses of environmental stratification and adaptability of maize cultivars indicated for planting in Paraná State. The analysis of adaptability based on factor analysis was accomplished graphically through the scores in relation to the factors, separating in adaptability quadrants. The analyses of environmental stratification were accomplished starting from the information on the magnitude of the final factor loadings, obtained after rotations through *varimax* method. The FGGE technique is efficient in the process of environmental stratification and adaptability analysis. Above 70% of retained variation on first eigenvalues represents expressive part into sum of the genotypic and G \times E interaction effects. The climatic conditions were preponderant in grouping of environments, however, could be detected by the FGGE technique. The cultivars with wide adaptation and high yield in the group of tested environments were P30F35, P30F53, P30K64, P30R50 and AS 1570.

Keywords: Multivariate analysis. Factor analysis. Genotypic effect. Genotype \times environment interaction.

1 INTRODUÇÃO

A interação genótipos \times ambientes ($G \times A$) é um componente da variação fenotípica resultante do comportamento dos genótipos, diferente do previsto pelo modelo apresentado, quando submetidos a mais de um ambiente. Sua magnitude na expressão fenotípica do caráter pode reduzir a correlação entre fenótipo e genótipo, alterando a estimativa da variância genética e, por sua vez, parâmetros dependentes desta, como herdabilidade e ganho genético com a seleção (ROCHA; VELLO, 1999).

Uma particularização da interação $G \times A$ evidenciada por genótipos e ambientes pode ser realizada por meio de estudos sobre a adaptabilidade e a estabilidade fenotípica, assim como em estudos de estratificação ambiental. A adaptabilidade é a capacidade que um genótipo tem de aproveitar vantajosamente os efeitos ambientais, de maneira a assegurar alto nível de produtividade; já a estabilidade está relacionada com a manutenção da produtividade ou de sua previsibilidade com os ambientes diversos (ROCHA, 2002). A estratificação ambiental consiste na subdivisão de regiões heterogêneas em sub-regiões mais uniformes, nas quais se exclui qualquer interação $G \times A$ significativa ou, em outras situações, aceita-se $G \times A$ significativa porém com predominância de parte simples atuante, ou seja, que não venha a comprometer a recomendação das cultivares (CRUZ; CARNEIRO, 2003).

Estudos sobre a adaptabilidade e a estabilidade fenotípica são de suma importância, pois permitem particularizar os efeitos da interação $G \times A$ no âmbito de genótipos e de ambientes, identificando a contribuição relativa de cada um para a interação total. De acordo com Brasil (1990) e Duarte e Zimmermann (1995), é preciso dispor de metodologias estatísticas adequadas para se estimar e explorar a interação $G \times A$, permitindo até mesmo recomendações regionalizadas.

Embora várias metodologias tenham sido propostas ao longo dos últimos trinta anos, visando mitigar os desafios impostos pela interação $G \times A$, os pesquisadores atuais continuam buscando instrumentos estatísticos ainda mais efetivos nesta tarefa.

O tratamento estatístico dado à interação $G \times A$ nesta última década tem sofrido forte reorientação, seja pelos grandes avanços obtidos nos procedimentos estatísticos computacionais, seja pela pouca eficácia dos métodos usuais de análise da interação (DUARTE; VENCOVSKY, 1999). Essa mudança de rumos, nítida em publicações recentes, pode ser percebida numa rápida consulta ao livro *Genotype by Environment Interaction* (KANG; GAUCH JÚNIOR, 1996). Nos seus quatorze capítulos, escritos por autoridades internacionais no assunto, o enfoque tradicional univariado é substituído por uma abordagem que combina técnicas univariadas e multivariadas.

Dentre as técnicas multivariadas, a análise de fatores vem sendo empregada nos estudos agrônômicos há relativamente pouco tempo, devido à complexidade dos cálculos envolvidos. No entanto, a popularização dos computadores, nos últimos anos, vem permitindo maior difusão dessa técnica junto à comunidade científica.

Segundo Ferreira (2008), deve-se diferenciar a análise de fatores e a análise de componentes principais, embora ambas sejam propostas para estudar a estrutura de covariâncias ou de correlações populacionais. A explicação total das covariâncias ou das correlações na análise de componentes principais é feita utilizando-se todas as variáveis latentes. Na análise fatorial, é possível explicar todas as covariâncias ou correlações utilizando algumas poucas variáveis não observáveis ou latentes, que são denominadas fatores. A explicação total das covariâncias ou correlações pode ser feita por alguns poucos fatores, mas a explicação total das variâncias proporcionada por esses fatores é, em geral, menor do que a propiciada pelo mesmo número de componentes principais.

Desse modo, segundo Ferreira (2008), é possível concluir que o maior objetivo da análise de fatores é explicar as covariâncias, enquanto a dos componentes principais é explicar variâncias.

Além dos estudos de adaptabilidade e estratificação ambiental apresentados por diversos autores (GARBUGLIO et al., 2007; MENDONÇA et al., 2007; MURAKAMI; CRUZ, 2004), a técnica de análise de fatores tem sido adaptada a outros estudos do meio agrônomo, como na predição de ganhos por seleção (GRANATE et al., 2001); na associação entre características chave de qualidade em milho doce, representadas por 29 variáveis e tendo sua explicação a partir de seis fatores (JUVICK et al., 1993). Na cultura do feijoeiro, Acquah, Adams e Kelly (1992), estudando quinze variáveis relacionadas com a arquitetura da planta, obtiveram a extração de três fatores.

No entanto, a técnica originalmente proposta considera como entrada uma matriz de médias fenotípicas, o que pode levar a inferências equivocadas tanto na adaptabilidade das cultivares avaliadas quanto no processo de estratificação ambiental, devido ao ruído presente na interação $G \times A$, implicando também em um número elevado de fatores, dependendo da natureza das variáveis a serem consideradas no agrupamento.

Buscando melhorar a precisão dessas estimativas por meio de uma redução no ruído presente na interação $G \times A$, o presente estudo foi realizado com objetivo de propor uma metodologia utilizando a análise de fatores associada aos efeitos genotípicos e de interação $G \times A$ (FGGA) e, simultaneamente, realizar análises de estratificação ambiental e adaptabilidade de cultivares de milho indicadas para cultivo no estado do Paraná.

2 MATERIAL E MÉTODOS

2.1 Dados utilizados nas análises

Para a realização das análises propostas foram utilizados dados obtidos junto a Rede de Avaliação de Cultivares de Milho do Instituto Agrônômico do Paraná (IAPAR), referentes a 15 genótipos (Tabela 1) avaliados em 5 localidades naquele estado (Tabela 2), durante as safras 2005/2006 e 2006/2007, perfazendo um total de 10 ambientes. As análises foram realizadas utilizando-se os dados referentes à variável produtividade de grãos (corrigidos para 14% de umidade e transformados para $\text{kg}\cdot\text{ha}^{-1}$).

Tabela 1 Híbridos utilizados para coleta de dados de produtividade ($\text{kg}\cdot\text{ha}^{-1}$) em cinco localidades no estado do Paraná, safras 2005/2006 e 2006/2007.

Híbrido	Empresa	Ciclo	Classe
1 - AS 1570	Agroeste	Precoce	HS
2 - AS 1575	Agroeste	Precoce	HS
3 - BM 1120	Biomatrix	Superprecoces	HT
4 - DOW 2A525	DOW Agrosiences	Precoces	HS
5 - DOW 2B710	DOW Agrosiences	Precoces	HS
6 - IPR 119	IAPAR	Precoces	HD
7 - AG 5020	Monsanto-Agroceres	Precoces	HT
8 - AG 8021	Monsanto-Agroceres	Precoces	HS
9 - DKB 566	Monsanto-Dekalb	Precoces	HT
10 - P 30F35	Pioneer	Precoces	HS
11 - P 30F53	Pioneer	Precoces	HS
12 - P 30K64	Pioneer	Precoces	HS
13 - P 30R50	Pioneer	Precoces	HS
14 - SHS 4050	Santa Helena	Super-precoces	HD
15 - SHS 5070	Santa Helena	Super-precoces	HT

Tabela 2 Localidades (estado do Paraná) de implantação dos ensaios de produtividade pela Rede de Avaliação de Cultivares de Milho do IAPAR, Safras 2005/2006 e 2006/2007

Local	Altitude	Semeadura		Colheita	
		05/06	06/07	05/06	06/07
Campo Mourão	630 m	23/09/05	29/09/06	17/02/06	03/03/07
Londrina	576 m	20/09/05	03/10/06	24/02/06	27/02/07
Pato Branco	760 m	28/09/05	31/10/06	15/03/06	21/03/07
Ponta Grossa	975 m	27/09/05	24/10/06	18/04/06	19/04/07
Wenceslau Braz	836 m	11/10/05	27/09/06	21/03/06	08/03/07

2.2 Análises de variância

Os ensaios foram realizados em delineamento de blocos casualizados, com três repetições. A área experimental de cada parcela foi constituída de duas linhas de 5 m, com 80 cm entre linhas, 20 cm entre plantas, tendo-se obtido 25 plantas por linha após desbaste. Os controles de pragas e de plantas daninhas foram realizados quando necessário.

Os dados foram submetidos às análises de variância individuais e, posteriormente, depois de verificada a relação entre o maior e o menor quadrado médio residual, foi realizada a análise de variância conjunta.

$$Y_{ijkm} = \mu + G_i + S_j + L_k + (B/S)/L_{jkm} + GS_{ij} + GL_{ik} + SL_{jk} + GSL_{ijk} + \varepsilon_{ijkm}$$

em que μ : média geral; G_i : efeito do i -ésimo genótipo ($i = 1, 2, \dots, p$); S_j : efeito da j -ésima safra ($j = 1, 2, \dots, s$); L_k : efeito do k -ésimo local ($k=1, 2, \dots, l$); GS_{ij} , GL_{ik} , LS_{jk} : efeitos das interações de primeira ordem entre genótipos e safras, genótipos e locais, locais e safras, respectivamente; GSL_{ijk} : efeito da interação tripla entre genótipos, safras e locais; $(B/S)/L_{jkm}$: efeito de blocos dentro de safras dentro de locais; ε_{ijkm} : erro aleatório. Todos os efeitos, exceto o erro, foram considerados como fixos (Tabela 3).

Tabela 3 Resumo do esquema da análise de variância conjunta considerando todos os efeitos como fixos (CRUZ; REGAZZI; CARNEIRO, 2004).

F.V	GL	QM	F
B(L×S)	(r-1)(sl)	Q1	Q1/Q9
Safras (S)	s-1	Q2	Q2/Q9
Locais (L)	l-1	Q3	Q3/Q9
Genótipos (G)	p-1	Q4	Q4/Q9
G × S	(p-1)(s-1)	Q5	Q5/Q9
G × L	(p-1)(l-1)	Q6	Q6/Q9
S × L	(s-1)(l-1)	Q7	Q7/Q9
G × S × L	(p-1)(s-1)(l-1)	Q8	Q8/Q9
Resíduo	(r-1)(p-1)sl	Q9	-----

Uma vez detectada interação significativa entre safras, locais e suas interações, cada local dentro de cada safra será considerado como um ambiente independente e, neste caso, serão analisados dez ambientes (Tabela 4), com todos os efeitos considerados como fixos. O modelo reduzido passará a ser

$$Y_{ij} = \mu + g_i + a_j + (ga)_{ij} + \varepsilon_{ij}$$

em que μ : média geral; g_i : efeito de genótipos ($i=1,2,\dots,p$); a_j : efeitos de ambientes ($j=1,2,\dots,q$); $(ga)_{ij}$: efeito da interação genótipos \times ambientes; ε_{ij} : erro aleatório.

Tabela 4 Análise de variância com base no modelo reduzido, considerando todos os efeitos como fixos.

F.V	GL	QM	F	E(QM)
B/A	q(r-1)	Q1	Q1/Q5	$\sigma_e^2 + p\phi_B$
Genótipos (G)	p-1	Q2	Q2/Q5	$\sigma_e^2 + pq\phi_G$
Ambientes (A)	q-1	Q3	Q3/Q5	$\sigma_e^2 + rp\phi_A$
G × A	(p-1)(q-1)	Q4	Q4/Q5	$\sigma_e^2 + r\phi_{GA}$
Resíduo	q(r-1)(p-1)	Q5	-----	σ_e^2

Foi estimado, também, o percentual de porção simples (PS) da interação $G \times A$, de acordo com adaptação do método de Robertson (1959), assim como a correlação entre as 45 combinações de pares de ambientes.

$$PS\% = \frac{\frac{1}{2}(\sqrt{Q_{2a}} - \sqrt{Q_{2b}})^2}{\left[\frac{1}{2}(\sqrt{Q_{2a}} - \sqrt{Q_{2b}})^2 \right] + \left[(1-r)\sqrt{Q_{2a} \cdot Q_{2b}} \right]} \cdot 100$$

em que PS% é o percentual de parte simples da interação $G \times A$ e Q_{2a} e Q_{2b} são os quadrados médios entre genótipos nos ambientes a e b, respectivamente.

2.3 Estimação dos efeitos genotípicos e de interação $G \times A$

Nos moldes apresentados na análise AMMI, buscando a estimação do efeito genotípico e de interação $G \times A$, a partir de uma análise de variância comum aplicada à matriz de médias ($Y_{(g \times a)}$) em que o modelo que descreve o desempenho fenotípico de uma dada cultivar i é representado por $Y_{ij} = \mu + g_i + a_j + (ga)_{ij} + \bar{\varepsilon}_{ij}$. Desse modo, foram estimados seus efeitos e aplicada a análise fatorial. Esses procedimentos estão descritos na sequência.

A solução por quadrados mínimos ordinários para o sistema de equações normais correspondente, sob as restrições de identificabilidade $\sum_i g_i = \sum_j a_j = 0$, é dada por:

média geral $\hat{\mu} = \bar{Y}_{..}$; efeito genotípico: $\hat{g}_i = \bar{Y}_{i.} - \bar{Y}_{..}$; efeito de ambiente

$$\hat{a}_j = \bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot}; \text{ considerando-se que: } \bar{Y}_{\cdot\cdot} = \sum_{i,j} Y_{ij} / ga; \quad \bar{Y}_{i\cdot} = \sum_j Y_{ij} / a; \quad \text{e}$$

$$\bar{Y}_{\cdot j} = \sum_i Y_{ij} / g$$

A aproximação de quadrados mínimos (\hat{Y}_{ij}) e seu respectivo resíduo, correspondente ao termo geral da interação $(g\hat{a})_{ij}$, ambos invariantes, são dados por $Y_{ij} = Y_{i\cdot} + Y_{\cdot j} - Y_{\cdot\cdot}$; portanto, $(g\hat{a})_{ij} = Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot}$.

A partir desse último resultado foi construída a matriz de interações (GA), bem como a matriz que inclui os efeitos genotípicos somados aos efeitos de interação (G+GA) que foi utilizada nessa proposta.

gen _p \ amb _q	1	2	...	q
1	$\begin{pmatrix} (g\hat{a})_{11} & (g\hat{a})_{12} & \cdots & (g\hat{a})_{1q} \\ (g\hat{a})_{21} & (g\hat{a})_{22} & \cdots & (g\hat{a})_{2q} \\ \vdots & \vdots & & \vdots \\ (g\hat{a})_{p1} & (g\hat{a})_{p2} & \cdots & (g\hat{a})_{pq} \end{pmatrix}$			
2				
...				
p				

$= \mathbf{GA}_{(pxq)}$

G _p \ A _q	1	2	...	q
1	$\begin{pmatrix} g_1 + (g\hat{a})_{11} & g_1 + (g\hat{a})_{12} & \cdots & g_1 + (g\hat{a})_{1q} \\ g_2 + (g\hat{a})_{21} & g_2 + (g\hat{a})_{22} & \cdots & g_2 + (g\hat{a})_{2q} \\ \vdots & \vdots & & \vdots \\ g_p + (g\hat{a})_{p1} & g_p + (g\hat{a})_{p2} & \cdots & g_p + (g\hat{a})_{pq} \end{pmatrix}$			
2				
...				
p				

$= \mathbf{G+GA}_{(pxq)}$

2.4 Aplicação da análise de fatores à matriz G+GA

O modelo fatorial pressupõe que cada variável pode ser expressa por uma combinação linear de fatores comuns, além do fator específico, sendo dado por:

$$\begin{aligned}x_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\x_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2 \\&\dots \\x_h &= \ell_{h1}F_1 + \ell_{h2}F_2 + \dots + \ell_{hm}F_m + \varepsilon_h\end{aligned}$$

ou

$$x_j = \sum_{k=1}^m \ell_{jk} F_k + \varepsilon_j$$

para $m < h$; em que ℓ_{jk} é a carga fatorial para a j -ésima variável associada ao k -ésimo vetor; F_k é o k -ésimo fator comum; ε_j é o fator específico associado à j -ésima variável; sendo $h=p$ para genótipos ou $h=q$ para ambientes.

Para a utilização da análise de fatores em estudos de adaptabilidade e estratificação ambiental, deve-se destacar o fato de que $X_1, X_2 \dots X_h$ representam uma única variável como, por exemplo, produtividade de grãos, porém, avaliada em cada um dos “ q ” ambientes em que os genótipos foram avaliados ou dos “ p ” genótipos referentes aos q ambientes. No caso em questão, essas variáveis representam as colunas da matriz constituída pela soma dos efeitos genotípicos e de interação apresentada anteriormente.

A carga fatorial do fator F_k é dada pela correlação entre o fator e a variável correspondente, ou seja:

$$r_{x_j F_k} = \frac{Cov(x_j, F_k)}{\sqrt{V(x_j)V(F_k)}} = \ell_{jk}$$

Na análise de fatores são impostas as seguintes restrições em relação aos fatores comuns:

Média nula: $\bar{F}_k = 0$, para $k=1, 2, \dots, m$

Variância unitária: $Var(F_k) = 1$

Covariância nula: $Cov(F_k, F_{k'}) = 0$, ($k \neq k'$), propriedade que garante que os fatores sejam ortogonais

Também são apresentadas as seguintes propriedades relativas aos fatores específicos:

Média nula: 0

sendo $x_j = \sum_{k=1}^m \ell_{jk} F_k + \varepsilon_j$ e $\bar{x}_j = 0$ e $\bar{F}_k = 0$, portanto, $\bar{\varepsilon} = 0$

Variância:

A variância de um fator específico é dada por: $Var(\varepsilon_j) = \sigma_{\varepsilon_j}^2$, esta variância é denominada especificidade ou unicidade.

Covariância entre fatores específicos: $Cov(\varepsilon_j, \varepsilon_{j'}) = 0$ com ($j \neq j'$). Assim os fatores específicos são também não correlacionados.

Covariância entre fatores específicos e comuns:

$Cov(\varepsilon_j, F_k) = 0$ para $j=1, 2, \dots, v$ e $k=1, 2, \dots, m$

ou seja, os vetores de fatores específicos são ortogonais aos vetores de fatores comuns.

Comunalidade e especificidade:

A comunalidade é a medida da eficiência da representação de uma variável x_j por uma parte comum, também envolvida nas demais variáveis

analisadas. Souza (1988 citado por CRUZ; CARNEIRO, 2003) relata que têm sido aceitos como razoáveis valores de comunalidade superiores a 0,64, por ser equivalente a um coeficiente de correlação entre x_j e z_j superior a 0,80.

Define-se por comunalidade o componente quadrático dado por:

$$h_j^2 = \sum_{k=1}^m \ell_{jk}^2 = 1 - \sigma_{ej}^2$$

Para cada variável x_j tem-se que $h_j^2 + \sigma_{ej}^2 = 1$, em que h_j^2 , denominado comunalidade, representa a proporção da variância da variável padronizada “j” devida a fatores comuns. Considera-se, na análise de fatores, que cada variável possa ser explicada por uma parte comum ($\sum_{k=1}^m \ell_{jk} F_k$), determinada por “m” fatores (F_1, F_2, \dots, F_m) e por outra parte específica (ε_j). A forma funcional do modelo de análise de fatores é dada por:

$$X = Y\Lambda + \varepsilon$$

em que

X: originalmente, é a matriz ($g \times a$) de observações que apresenta valores obtidos para cada genótipo e variável analisada, destacando-se o fato de que a variável é representada por uma mesma característica medida em vários ambientes; Y: matriz ($g \times m$) de fatores comuns; Λ : matriz ($m \times a$) de cargas fatoriais; ε : matriz ($g \times a$) de fatores específicos. Para cada vetor de variáveis tem-se:

$$\tilde{x}_j = Y\tilde{\ell}_j + \tilde{\varepsilon}_j$$

em que \tilde{x}_j : vetor de observações da variável no ambiente j, de dimensão $g \times 1$;

Y: matriz de fatores comuns de dimensão $g \times m$; $\tilde{\ell}_j$: vetor de cargas fatoriais,

de dimensão $m \times 1$; $\tilde{\varepsilon}_j$: vetor de fatores específicos, associado à variável quantificada no j -ésimo ambiente, de dimensão $g \times 1$.

No presente trabalho, a matriz X de observações será substituída pela matriz de efeitos G+GA. Procurou-se, na análise de fatores, o estabelecimento de combinações lineares que permitissem a interpretação quanto aos grupos de ambientes com o mesmo padrão de discriminação genotípica, capazes de reter o máximo das associações originalmente disponíveis entre os ambientes estudados. Como, nesta análise, procura-se relacionar conjunto de variáveis observáveis, representadas na matriz X , com outras que, em parte, são denominadas fatores comuns, representadas na matriz Y , e em outras, por fator específico, representada na matriz ε , tem-se também associação entre as respectivas matrizes de variâncias e covariâncias, como ilustrado a seguir:

$$R = \Lambda' \Lambda + \Psi$$

em que R : matriz ($a \times a$) de correlações entre os desempenhos dos genótipos em cada par de ambientes considerados. Por outro lado, pode ser obtido:

$$R_c = R - \Psi = \Lambda' \Lambda$$

A matriz R_c é denominada matriz de correlações reduzidas do modelo de análise fatorial e apresenta os mesmos elementos da matriz de correlação original à exceção da sua diagonal, cujos elementos são reduzidos de σ_{ej}^2 e chamados de comunalidade. Segundo Cruz e Carneiro (2003), o problema fundamental da análise de fatores é o cálculo da matriz de cargas Λ e da diagonal de especificidades Ψ , cujos elementos não-nulos são σ_{ej}^2 , a partir da matriz R conhecida.

Uma maneira de obter as cargas fatoriais é por meio da fatoração de R_c , usando o teorema da decomposição espectral, em que qualquer matriz simétrica R_c pode ser escrita como:

$$R_c = R - \Psi = \Lambda' \Lambda = BAB' = \sum_{j=1}^h \lambda_j \tilde{\alpha}_j \tilde{\alpha}_j'$$

sendo

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_h$ os autovalores de R_c ;

$\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_h$ os autovetores correspondentes.

De outra forma, tem-se: $R_c = \Lambda' \Lambda = (B' A^{1/2})(A^{1/2} B)$. Assim,

$$\Lambda = BA^{1/2}$$

Na prática, não se dispõe de R_c , mas de R e utiliza-se um número $m < h$ de fatores comuns. O objetivo é explicar o máximo da variação em x_j com o menor número possível de fatores, tornando ε_j mínimo.

$$\Lambda = BA^{1/2}$$

sendo

$B = [\tilde{\alpha}_1 \quad \tilde{\alpha}_2 \quad \dots \quad \tilde{\alpha}_m]$: matriz de cargas fatoriais.

$$B = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{m2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1h} & \alpha_{2h} & \dots & \alpha_{hm} \end{bmatrix}_{h \times m}$$

$$A^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & & \phi \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ \phi & & & \sqrt{\lambda_m} \end{bmatrix}_{m \times m}$$

em que α_{jk} : elemento do autovetor $\tilde{\alpha}_k$ ($k = 1, 2, \dots, m$) que corresponde ao termo de ponderação da variável k , na análise de componentes principais;

$$l_{jk} = \sqrt{\lambda_k} \alpha_{jk}$$

2.5 Estratificação ambiental

Na utilização da análise de fatores destinada à estratificação ambiental, os resultados foram interpretados para cada fator e do potencial de seu interesse. Foram analisados os elementos ou cargas fatoriais identificando, para cada variável, os fatores com coeficientes de ponderação próximos da unidade. Neste caso, cada variável foi representada por um determinado ambiente, porém, ao invés das médias dos genótipos dentro de cada ambiente, a matriz de dados foi composta pelos efeitos genotípicos somados aos efeitos de interação $G \times A$.

O agrupamento dos ambientes foi realizado a partir das informações sobre a magnitude das cargas fatoriais finais, obtidas após rotações. Cargas fatoriais maiores ou iguais a 0,70, em valor absoluto, indicam ambientes com altas correlações e foram agrupados dentro de um mesmo fator. Cargas fatoriais finais de valores baixos ($\leq 0,50$) indicavam que o respectivo ambiente não deveria pertencer ao grupo. Cargas fatoriais com valores intermediários não garantiam nenhuma definição de agrupamento. No caso dessa não-definição foram utilizados os valores de PS% e correlação entre ambientes para se fazer inferências sobre o agrupamento.

Antes de serem obtidas as cargas fatoriais que permitiram identificar o melhor agrupamento de ambientes, foi estabelecido o número de fatores a serem considerados no modelo. O número de fatores finais pôde ser admitido como igual ao número de autovalores maiores ou iguais a 1,0. No entanto, nos casos em que a proporção da variabilidade explicada pelos autovalores maiores que 1,0 for baixa, podem ser considerados mais fatores até que se atinja, pelo menos, 80% da variação total.

Outro aspecto que foi considerado é o fato de que as cargas fatoriais não permitem clara definição dos ambientes mais similares, recomendando-se obter novos valores, denominados cargas fatoriais finais, que são obtidas por rotação via recursos computacionais.

Em determinados pacotes estatísticos, a análise é feita assumindo que os fatores são não-correlacionados, com base no modelo de fatores ortogonais. A extração das cargas fatoriais é feita, então, pelo método dos componentes principais e os fatores são estabelecidos pelo método de rotação *varimax* (GOULD, 1996), com um máximo de 50 rotações, em que este permite que os eixos de uma dada projeção representem a direção de máxima variância possível.

2.6 Análise de adaptabilidade

A análise de adaptabilidade baseada na análise de fatores é realizada graficamente por meio dos escores em relação aos fatores. Para isso, são traçados eixos paralelos, tomando-se a média dos escores, de modo a estabelecer quatro quadrantes. Nos quadrantes II e IV ficaram situados aqueles genótipos com adaptabilidade específica ao grupo de ambientes determinado pelo fator. No quadrante I ficaram situados aqueles genótipos de adaptabilidade ampla e, no quadrante III, os genótipos ruins, de baixo desempenho e passíveis de descarte ou não indicação para cultivo.

Em situações em que apenas dois fatores foram suficientes para explicar mais de 80% da variação total, determinando apenas duas sub-regiões, a análise de adaptabilidade foi feita com base somente nas informações de um único gráfico. Quando mais de duas sub-regiões foram formadas, teve-se o número de gráficos correspondentes à combinação de “n” regiões, tomadas duas a duas (C_2^n), de forma que as indicações de genótipos passaram a ser realizadas de modo mais eficiente do que os métodos tradicionais, pois, no método baseado em análise de fatores, não se consideram apenas ambientes favoráveis e desfavoráveis, como nos métodos de regressão linear uni ou bissegmentada, mas grupos de ambientes estabelecidos de acordo com suas similaridades.

Os escores fatoriais foram obtidos sabendo-se que $X = \hat{Y}\Lambda$ ou $X' = \Lambda'\hat{Y}'$, de modo que $\hat{Y}' = (\Lambda\Lambda')^{-1}\Lambda X' = \beta X'$, sendo \hat{Y}' : matriz de dimensão $m \times p$ em que cada linha apresenta os escores do k-ésimo fator (1,2,...,m); Λ : matriz de dimensão $m \times q$ das cargas fatoriais rotacionadas finais; X: originalmente é a matriz de dimensão $p \times q$ de observações, que será substituída pela matriz G+GA; β : matriz de dimensão $m \times q$ de coeficientes de ponderação dos caracteres para obtenção dos escores dos fatores.

3 RESULTADOS E DISCUSSÃO

Por meio das análises de variância individuais, foi constatado que os genótipos apresentaram comportamento distinto ($p < 0,01$) dentro de cada ambiente de estudo (Tabela 5). Os coeficientes de variação ($CV_e\%$) para produtividade oscilaram entre 4,66% (Ponta Grossa, safra 2006/2007) e 7,94% (Pato Branco, safra 2005/2006), com média de 6,54%, indicando alta precisão experimental do conjunto de ensaios. As médias de produtividade ficaram compreendidas entre 8617 $kg \cdot ha^{-1}$ (Pato Branco, safra 2005/2006) e 12.064 $kg \cdot ha^{-1}$ (Ponta Grossa, safra 2006/2007), com média geral 10.973 $kg \cdot ha^{-1}$, refletindo o alto potencial produtivo das cultivares avaliadas.

Tabela 5 Quadrados médios de genótipos (QMG) e de resíduo (QMR), médias de produtividade e coeficiente de variação experimental em diferentes ambientes do Estado do Paraná, nas safras 2005/2006 e 2006/2007.

Ambiente	¹ QMG		¹ QMR	² Médias	C.V.(%)
1. LONDRINA safra 05/06	20,8420	**	4,8646	12030	5,80
2. C MOURAO safra 05/06	59,0733	**	5,5035	9426	7,87
3. P BRANCO safra 05/06	54,4262	**	4,6796	8617	7,94
4. W BRAZ safra 05/06	38,0138	**	7,4585	11744	7,35
5. P GROSSA safra 05/06	39,2482	**	6,6821	12017	6,80
6. LONDRINA safra 06/07	39,2588	**	4,1698	10457	6,18
7. C MOURAO safra 06/07	62,4776	**	3,8435	11449	5,41
8. P BRANCO safra 06/07	34,8636	**	6,6244	10763	7,56
9. W BRAZ safra 06/07	33,9652	**	4,2414	11161	5,83
10. P GROSSA safra 06/07	64,3626	**	3,1634	12064	4,66

¹: QM x 10⁵; ²: Médias, em $kg \cdot ha^{-1}$

** : significativo, a 1% de probabilidade, pelo teste F

Em recente revisão apresentada por Troyer e Wellin (2009), ficou constatado que, ao longo do tempo, nos Estados Unidos, a seleção natural e artificial praticada pelos melhoristas, com foco na adaptabilidade culminou no

aumento de rendimento tanto das linhagens parentais, quanto dos híbridos, o que refletiu na redução da heterose. Por outro lado, essas linhagens e híbridos mais produtivos apresentaram tolerância a estresses bióticos e abióticos com maior frequência (DUVICK, 1999; TROYER, 2006). No Brasil, a exploração cada vez mais intensiva da cultura do milho em novos sistemas de produção, semeadura em sucessão à soja e obedecendo ao zoneamento agrícola, estão fazendo com que o cultivo do milho seja realizado praticamente o ano todo, o que tem contribuído para o aparecimento mais intenso de pragas e doenças, exigindo cultivares mais adaptadas a essas condições.

Uma primeira análise conjunta foi realizada considerando como diferentes fatores os efeitos de safras e locais (Tabela 6). Foi constatada alta significância ($p < 0,01$) para todos os fatores individuais e interações. Considerando o efeito significativo das interações safras \times locais e Genótipos \times locais \times safras, optou-se por realizar uma nova análise conjunta em que cada local, dentro de cada safra, fosse tratado com um ambiente independente (Tabela 7). Como houve um simples agrupamento de fatores (safras e locais em ambientes), os quadrados médios para blocos, genótipos e erro se mantiveram, assim como a média geral e, conseqüentemente, o valor de 6,52% para o CV_e. Foram constatadas diferenças significativas entre ambientes e para a interação genótipos \times ambientes, indicando que os genótipos apresentaram respostas diferentes do modelo em cada ambiente de avaliação.

Outros autores constataram que a variação nas produtividades de genótipos de arroz (RAMOS; SANCHES; COTES, 2009) e trigo (YAN et al., 2000), em diferentes ambientes, ocorreu, principalmente, devido a efeitos ambientais, que chegaram a compor de 70% a 80% da soma de quadrados. Estes valores indicam que os conjuntos de genótipos avaliados apresentaram pouca variação de produtividade devido a causas genéticas, o que não ocorre no presente estudo com genótipos de milho. Nas cultivares comerciais de milho há

grande variação entre e dentro de classes genéticas (HS, HT, HD e VPA), resultado dos primeiros intercâmbios de germoplasma melhorado entre instituições públicas nacionais e, atualmente, por parte de empresas privadas, devido às introduções de genótipos temperados, oriundos, principalmente, da América do Norte e genótipos tropicais introduzidos da Ásia.

Tabela 6 Quadrados médios obtidos pela análise de variância conjunta para a variável produtividade de grãos ($\text{kg}\cdot\text{ha}^{-1}$) de milho, avaliados em diferentes locais e safras no Estado do Paraná nas safras 2005/2006/e 2006/2007.

FV	GL	¹SQ	¹QM	
Rep(Locais×Saфра)	20	151,1661	7,5583	
Safras	1	190,7519	190,7519	**
Locais	4	3038,2921	759,5730	**
Genótipos	14	3985,380	284,6700	**
Genótipos × Safras	14	599,8176	42,8441	**
Genótipos × Locais	56	903,4695	16,1334	**
Safras × Locais	4	2399,2002	599,8000	**
Genótipos × Locais × Safras	56	762,7739	13,6210	**
erro	280	1434,4636	5,1231	

CV%: 6,52

Média: 10973 $\text{kg}\cdot\text{ha}^{-1}$

¹: QM x 10⁵

** : significativo, a 1% de probabilidade, pelo teste F

No presente estudo, 46,5% da soma de quadrados devem-se aos efeitos ambientais (Tabela 7), enquanto 33,5% são devido aos efeitos genotípicos e 18,7% à interação G×A que, somados, totalizam 52,2% da variação de produtividade. Em termos de variância, foi constatado que 42,4% devem-se à variância entre ambientes e 41,9%, devido à soma das variâncias entre genótipos e à interação G×A. Este resultado é importante na definição de uma nova metodologia para a qual sejam considerados, conjuntamente, os efeitos

genotípicos e de interação, de modo a testar sua potencialidade para análises simultâneas de adaptabilidade e de estratificação ambiental.

Tabela 7. Quadrados médios obtidos pela análise de variância conjunta, assumindo locais e safras como ambientes e estimativas da variância experimental ($\hat{\sigma}_e^2$) e componentes quadráticos de ambientes ($\hat{\phi}_A$), genotípico ($\hat{\phi}_G$) e da interação G×A($\hat{\phi}_{G \times A}$).

FV	GL	¹ SQ	¹ QM	
Blocos/Ambientes	20	151,1661	7,5583	
Genótipos (G)	14	3985,3802	284,6700	**
Ambientes (A)	9	5628,2443	625,3605	**
G×A	126	2266,0609	17,9846	**
Resíduo	280	1434,4636	5,1231	
$\hat{\sigma}_e^2$	5,1231			
$\hat{\phi}_A$	13,7831			
$\hat{\phi}_G$	9,3182			
$\hat{\phi}_{G \times A}$	4,2872			
Média: 10973 kg.ha ⁻¹				
CV%: 6,52				

¹: QM x 10⁵

** : significativo, a 1% de probabilidade, pelo teste F

A ideia básica de uma análise de estratificação ambiental se fundamenta na subdivisão de regiões heterogêneas em sub-regiões mais uniformes, buscando reunir ambientes de classificação genotípica similar e que seja independente da qualidade ambiental.

Pelos dados da Tabela 8 constata-se que os dois primeiros autovalores maiores que 1,0 absorveram 77,7% da variação. Em diversos estudos envolvendo técnicas multivariadas tem-se considerado que acima de 70% da variação captada nos primeiros eixos constitui parte expressiva da variação

(RAMOS; SANCHES; COTES, 2009; YAN et al., 2000). Crossa et al. (1991) comentam que uma característica marcante dos modelos multivariados é que eles buscam capitalizar grande parte do chamado “padrão” nos primeiros eixos principais, relegando aos eixos subsequentes, cada vez menos “padrão” e mais “ruído”. Portanto, quanto mais variação for captada nos primeiros eixos, preferencialmente os dois primeiros, menos “ruídos” associados haverá, melhorando a qualidade das estimativas.

Na análise de fatores tem sido utilizado, empiricamente, um número de eixos que absorvam, pelo menos, 80% da variação para se determinar o número final de fatores que serão utilizados no agrupamento (CRUZ; CARNEIRO, 2003; GARBUGLIO et al., 2007; MENDONÇA et al., 2007; MURAKAMI; CRUZ, 2004). No presente caso, os três primeiros eixos captaram 84,6% da variação, fixando-se, assim, um número final de três fatores para a obtenção das cargas fatoriais após rotação.

As comunalidades (h_j^2), ou seja, as medidas de eficiência na representação das variáveis por uma parte comum ficaram acima de 0,8, indicando alta eficiência do processo de fatoração para o agrupamento dos ambientes.

O segundo fator agrupou os ambientes Wenceslau Braz 2005/2006 e Campo Mourão 2005/2006 e 2006/2007. Dentro da mesma safra (2005/2006), foi constatado que Campo Mourão e Wenceslau Braz apresentaram um percentual de parte simples (PS%) de 69,1% e uma correlação significativa ($p < 0,01$) de 0,84 (Tabela 9), indicando que os níveis de incremento ou redução de produtividade do conjunto de cultivares, de modo geral, foram similares de um ambiente para outro. A alocação do ambiente Campo Mourão 2006/2007 no fator 2 pode ser devido às altas correlações ($r > 0,7$) com os demais ambientes agrupados neste fator.

Tabela 8 Estratificação ambiental por meio da análise de fatores, com 15 genótipos e 10 ambientes no estado do Paraná, nas safras 2005/2006 e 2006/2007.

Ambientes	Ano Agrícola	Cargas fatoriais após rotação			Comunalidades
		Fator 1	Fator 2	Fator 3	
Londrina	2005/2006	0,4668	0,6746	0,1712	0,89
Campo Mourão	2005/2006	0,2513	0,8042	0,4000	0,82
Pato Branco	2005/2006	0,1347	0,2695	0,9352	0,85
Wenceslau Braz	2005/2006	0,1883	0,8751	0,3294	0,91
Ponta Grossa	2005/2006	0,4057	0,6780	0,4809	0,98
Londrina	2006/2007	0,8396	0,1587	0,2049	0,92
Campo Mourão	2006/2007	0,4563	0,8052	-0,0493	0,84
Pato Branco	2006/2007	0,8729	0,3286	-0,0050	0,80
Wenceslau Braz	2006/2007	0,8474	0,3216	0,1330	0,94
Ponta Grossa	2006/2007	0,7115	0,4819	0,2866	0,93

Autovalor	Porcentagem Acumulada
6,49	64,9
1,27	77,7
0,70	84,6
0,50	89,6
0,40	93,6
0,24	96,1
0,16	97,7
0,12	99,0
0,09	99,9
0,01	100,0

Tabela 9 Porcentagem de parte simples da interação G×A (PS%) e estimativas de correlações entre ambientes (r) e quadrados médios de G×A.

A1	A2	r	PS%	¹ QM G×A	A1	A2	r	PS%	¹ QM G×A
1	2	0,70 •	49,4	16,3621 **	3	10	0,45	26,0	20,4602 **
1	3	0,44	25,2	20,8890 **	4	5	0,80 •	64,9	11,5896 **
1	4	0,75 •	50,2	9,3552 *	4	6	0,42	27,8	16,1396 **
1	5	0,62+	42,9	20,6812 **	4	7	0,75 •	56,8	12,8632 **
1	6	0,49	35,1	16,0146 **	4	8	0,43	27,8	25,8589 **
1	7	0,63+	42,7	18,6714 **	4	9	0,50	29,7	17,9241 **
1	8	0,63+	41,0	17,8658 **	4	10	0,60+	37,3	14,6428 **
1	9	0,71 •	46,0	11,2966 *	5	6	0,44	42,1	26,4701 **
1	10	0,62+	38,0	15,0938 **	5	7	0,72 •	47,2	17,3593 **
2	3	0,60+	43,1	20,5543 **	5	8	0,59+	36,7	24,3034 **
2	4	0,84 •	69,1	9,5148 *	5	9	0,68 •	48,6	17,7795 **
2	5	0,80 •	54,8	12,9953 **	5	10	0,81 •	62,0	11,2949 *
2	6	0,45	42,4	25,2789 **	6	7	0,57+	50,7	19,8438 **
2	7	0,72 •	47,1	17,1034 **	6	8	0,68 •	58,9	14,6495 **
2	8	0,50	29,6	29,3120 **	6	9	0,69 •	51,8	9,8930
2	9	0,47	31,5	27,1463 **	6	10	0,72 •	54,9	9,5239
2	10	0,71 •	50,8	15,6765 **	7	8	0,63+	39,5	20,8653 *
3	4	0,56+	34,0	14,9967 **	7	9	0,58+	38,8	21,0256 **
3	5	0,66 •	49,0	18,2813 **	7	10	0,71 •	49,4	15,1493 **
3	6	0,35	23,1	18,5108 **	8	9	0,83 •	62,5	8,3972
3	7	0,26	18,0	35,0463 **	8	10	0,76 •	53,6	11,7284 *
3	8	0,22	14,5	34,9909 **	9	10	0,71 •	46,1	11,2478
3	9	0,32	17,9	24,6602 **					

•,+: diferente de zero a 1% e 5%, pelo teste t, respectivamente; **: significativo, a 1% e 5% de probabilidade, pelo teste F; ¹: QM x 10⁵

A1 e A2: refere-se à codificação dos ambientes testados sendo para safra 2005/2006, 1: Londrina; 2: Campo Mourão; 3: Pato Branco; 4: Wenceslau Braz; 5: Ponta Grossa e para safra 2006/2007, 6: Londrina; 7: Campo Mourão; 8: Pato Branco; 9: Wenceslau Braz; 10: Ponta Grossa.

Dentro do conjunto de ambientes agrupados no fator 1, constatou-se que algumas das quatro localidades da safra 2006/2007 apresentaram interação $G \times A$ não significativa ($p > 0,05$) quando analisadas aos pares (Tabela 9). Apenas Campo Mourão, na mesma safra, apresentou significância da interação $G \times A$ associada a baixos valores de PS% ($PS\% < 50$), o que pode explicar, novamente, a sua alocação no fator 2, em vez de 1.

Garbuglio et al. (2007) realizaram a estratificação ambiental em 22 ambientes no estado do Paraná, constatando que, das 231 combinações, apenas 23 apresentaram interação $G \times A$ com predominância de parte simples ($PS\% > 50$) e correlações acima de 0,50. Constatou-se, ainda, que, pelo método tradicional de Lin (1982), os grupos foram formados por ambientes em que a porção simples foi mais pronunciada, ou seja, as combinações entre ambientes, dentro de um mesmo grupo, apresentaram $PS\% > 50$ e $r > 0,5$, com exceção de um grupo com 47,8% de PS, porém, com $r = 0,71$.

As condições climáticas na safra 2005/2006 não foram tão favoráveis para o desenvolvimento da cultura do milho em algumas regiões do Paraná, principalmente pela distribuição irregular das chuvas. As precipitações foram bastante elevadas nas fases de plantio e desenvolvimento inicial da cultura em setembro e outubro, a ponto de prejudicar o início da semeadura na região Centro-Sul (GERAGE; SHIOGA; ARAUJO, 2006). Contudo em novembro, dezembro e início de janeiro, as precipitações foram abaixo da média e a distribuição de chuvas muito irregular nas regiões sudoeste (Pato Branco), centro-sul (Ponta Grossa) e oeste (Campo Mourão) (GERAGE; SHIOGA; ARAUJO, 2006). A região norte foi exceção, pois apresentou condições melhores de precipitação neste período, permitindo que o experimento de Londrina apresentasse a maior média de produtividade (Tabela 10), seguido do experimento de Ponta Grossa (centro-sul) onde houve precipitações localizadas na área do experimento (GERAGE; SHIOGA; ARAUJO, 2006).

Tabela 10 Estimativas de produtividade média entre e dentro de dez ambientes e ranqueamento (rk) das quinze cultivares de milho no estado do Paraná.

	(a) 1	rk	(a) 2	rk	(a) 3	rk	(a) 4	rk	(a) 5	rk		
P 30F35	12624	4	10542	3	9826	4	13945	1	13235	2		
P 30K64	12566	5	10295	5	9761	5	12403	5	13112	3		
AS 1570	12934	2	10525	4	8840	8	11658	10	14125	1		
P 30F53	12304	8	11087	2	10032	2	12307	6	12830	5		
P 30R50	12377	7	12080	1	9962	3	12912	2	12383	6		
BM 1120	11996	9	9721	7	8663	9	11946	7	12880	4		
AS 1575	12423	6	9288	9	7992	11	11578	11	11197	12		
DKB 566	13458	1	9457	8	8573	10	12589	3	12270	7		
AG 5020	11016	13	10204	6	7784	12	10515	12	11555	11		
AG 8021	12754	3	9214	10	10069	1	11845	8	11834	9		
2A525	11471	12	8315	11	9192	7	11665	9	12076	8		
2B710	11532	11	7395	15	9215	6	12422	4	11613	10		
SHS 5070	10971	14	7987	12	6511	14	9947	15	10046	15		
SHS 4050	11551	10	7706	13	7048	13	10019	14	10957	13		
IPR 119	10476	15	7578	14	5791	15	10402	13	10149	14		
	(a) 6	rk	(a) 7	rk	(a) 8	rk	(a) 9	rk	(a) 10	rk	(b) Média Geral	rk
P 30F35	12512	2	12595	3	13117	1	12740	2	14844	1	12598 a	1
P 30K64	12538	1	14472	1	11906	2	12928	1	13095	4	12307 a	2
AS 1570	10798	5	12981	2	11622	4	12316	3	13728	2	11953 a	3
P 30F53	11666	3	11938	7	10199	10	11681	6	13024	5	11707 a	4
P 30R50	11213	4	11972	6	8871	15	11950	5	12299	8	11602 a	5
BM 1120	8904	15	11920	8	10995	6	10468	11	13224	3	11072 b	6
AS 1575	9859	9	12366	5	11293	5	12019	4	12432	7	11045 b	7
DKB 566	10685	6	10583	10	10727	8	10428	12	11175	11	10995 b	8
AG 5020	10519	8	12592	4	11674	3	11417	7	12601	6	10988 b	9
AG 8021	9788	10	10154	12	10946	7	10513	10	11084	12	10820 b	10
2A525	9573	11	10183	11	10066	12	10587	9	12121	9	10525 c	11
2B710	10657	7	10772	9	10120	11	9749	14	11203	10	10468 c	12
SHS 5070	9229	14	9867	14	10686	9	10186	13	10207	14	9564 c	13
SHS 4050	9437	13	9943	13	9504	14	9619	15	9272	15	9506 c	14
IPR 119	9471	12	9395	15	9715	13	10819	8	10652	13	9445 c	15

(a): codificação dos ambientes testados, sendo para safra 2005/2006, 1: Londrina; 2: Campo Mourão; 3: Pato Branco; 4: Wenceslau Braz; 5: Ponta Grossa e para safra 2006/2007, 6: Londrina; 7: Campo Mourão; 8: Pato Branco; 9: Wenceslau Braz; 10: Ponta Grossa.

(b): Agrupamento de médias, pelo teste de Scott-Knott, a 1% de probabilidade

O experimento implantado na localidade de Pato Branco sofreu estresse hídrico nas duas safras, tendo, em 2005/2006, o estresse atingido, principalmente, os estádios mais críticos da cultura (GERAGE; SHIOGA; ARAUJO, 2006), resultando em rendimento apenas satisfatório ($8.617 \text{ kg}\cdot\text{ha}^{-1}$), além de r e PS% baixos. Este fato resultou no agrupamento isolado de Pato Branco, 2005/2006, no Fator 3, e este fator recebeu a denominação, através da metodologia FGGA, como “*Seca*”.

É possível constatar que a metodologia foi eficiente no sentido de discriminar os conjuntos de ambientes dentro de safras e que a condição climática também foi preponderante no agrupamento de ambientes. Na safra 2006/2007, as condições climáticas foram favoráveis para o desenvolvimento da cultura (SHIOGA; GERAGE; ARAUJO, 2007), o que resultou em altas correlações entre ambientes, com exceção de Campo Mourão e não significância da interação $G \times A$ entre alguns pares de ambientes, o que pode ter facilitado o agrupamento. Neste sentido o fator 1 pode ser denominado pela metodologia como “*Condições edafoclimáticas boas*”.

A safra 2005/2006, caracterizada por estresses em fases críticas do desenvolvimento da cultura, permitiu o agrupamento de somente duas localidades da mesma safra, sendo o fator 2 denominado como “*Condições edafoclimáticas ruins*”. As localidades de Londrina e Ponta Grossa, safra 2005/2006, apresentaram as melhores condições climáticas, o que as diferiu muito das demais, podendo ser a causa do seu não agrupamento. Suas cargas fatoriais obtidas foram de 0,6746 (Londrina) e 0,6780 (Ponta Grossa), muito próximas de 0,7, porém, não há garantia de seus agrupamentos no fator 2, devendo ser consideradas como ambientes distintos dos demais.

No caso do melhoramento genético, esse tipo de estratificação baseada nas covariâncias e correlações entre ambientes, levando-se em consideração os efeitos genotípicos somados aos efeitos de interação $G \times A$, permitiu detectar

melhor a magnitude dos efeitos de interação e, no caso de sua existência, considera que a parte atuante seja de natureza simples, quando possível o agrupamento, permitindo ao melhorista uma indicação mais generalizada ou mais regionalizada.

Para as análises de adaptabilidade por dispersão gráfica, a partir dos escores fatoriais finais (Figuras 1, 2 e 3), foram utilizados os três fatores que agruparam 8 dos 10 ambientes.

Os genótipos P30F35, P30F53, P30R50, P30K64 e AS 1570 apresentaram adaptabilidade ampla ao conjunto de ambientes agrupados nos fatores 1 (Condições edafoclimáticas boas) e 2 (Condições edafoclimáticas ruins), pelo fato de estarem posicionados no quadrante I (QI). Pelo teste de Scott-Knott, a 1% de probabilidade (Tabela 10), essas cinco cultivares constituíram o grupo de maior produtividade. Isolando-se os fatores de Condições edafoclimáticas boas e Seca (Fator 3), que agruparam cinco ambientes, e os fatores de Condições edafoclimáticas ruins e Seca, que agruparam quatro ambientes, constata-se, novamente, que P30F35 e P30K64 se mantêm em QI, além de apresentarem as maiores médias de produtividade ($12.598 \text{ kg.ha}^{-1}$ e $12.307 \text{ kg.ha}^{-1}$, respectivamente). As cultivares P 30F53 e P 30R50, quando avaliadas em Pato Branco 2005/2006, no Fator 3, foi constatada baixa adaptabilidade a esse ambiente. A única cultivar a apresentar adaptabilidade específica a esse ambiente foi SHS 5070, no entanto, apresentou baixa adaptabilidade aos demais ambientes.

As cultivares AS 1575 e AG 5020 apresentaram adaptabilidade específica aos ambientes agrupados nos fatores 1 e 3, devido à suas alocações no Q4 da Figura 1 e Q1 da figura 2. DKB 566, AG 8021 e BM 1120 apresentaram adaptabilidade específica aos ambientes agrupados nos fatores 2 e 3, por se posicionarem no Q2 das Figuras 1 e 2, e Q1 da figura 3. Para os ambientes agrupados no fator 2, constatou-se adaptabilidade específica das cultivares DOW

2B710 e DOW 2A525. As cultivares que apresentaram adaptabilidade mais baixa ao conjunto de ambientes testados foram SHS 4050 e IPR 119, por se posicionarem em Q3 das três figuras.

Tomando-se como referência a Figura 1, constata-se que os 77,7% da variação captada nos dois primeiros fatores permite clara distinção dos genótipos com ampla adaptabilidade, adaptabilidade específica e baixa adaptabilidade. Isso pode ser confirmado, também, pela média geral dos dez ambientes testados e pelo agrupamento de médias pelo teste de Scott-Knott.

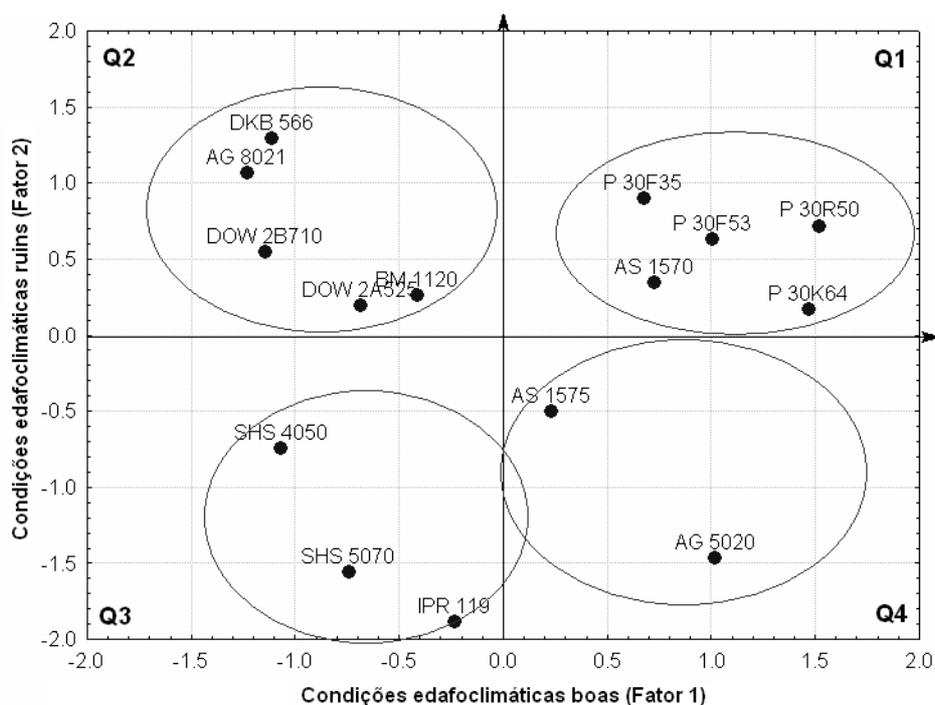


Figura 1 Análise gráfica da adaptabilidade de quinze cultivares de milho, a partir dos escores obtidos, considerando sete ambientes contidos nos fatores de Condições edafoclimáticas boas (Fator 1) e Condições edafoclimáticas ruins (Fator 2).

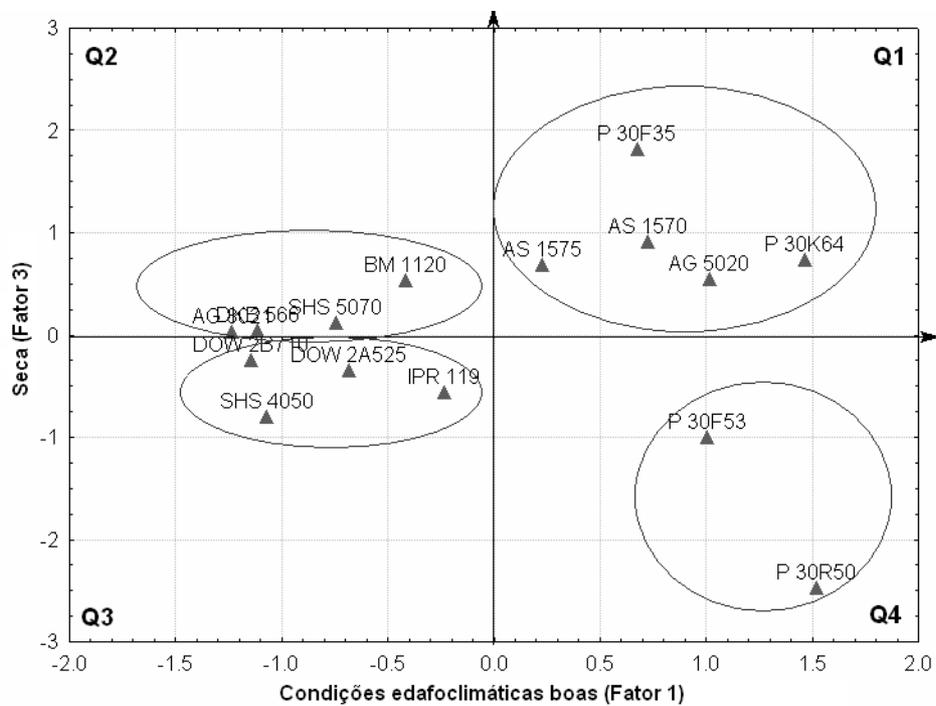


Figura 2 Análise gráfica da adaptabilidade de quinze cultivares de milho, a partir dos escores obtidos, considerando cinco ambientes contidos nos fatores de Condições edafoclimáticas boas (Fator 1) e Seca (Fator 3).

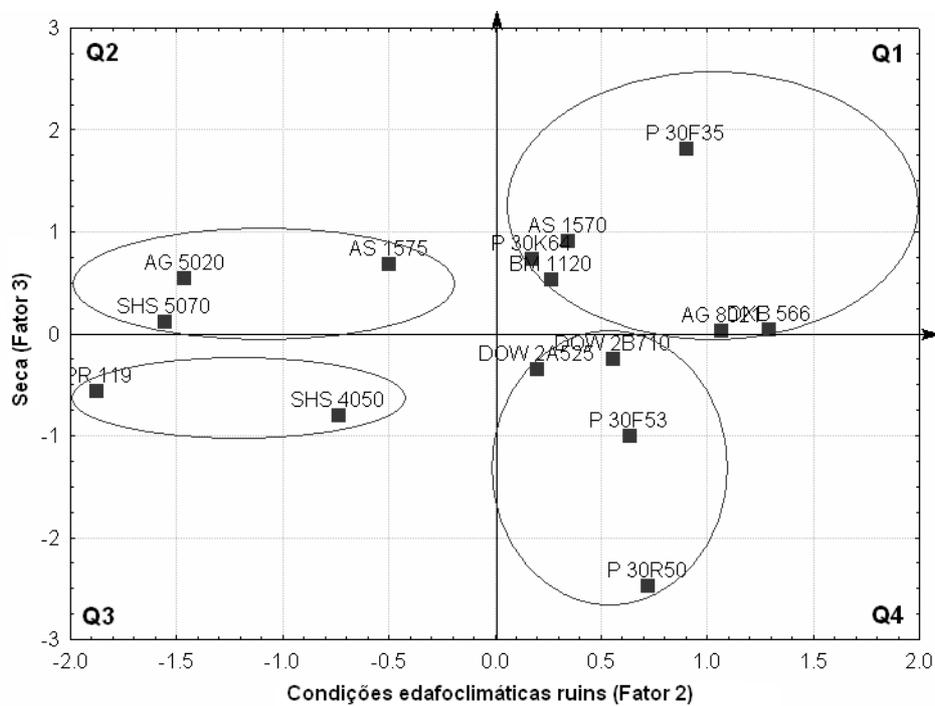


Figura 3 Análise gráfica da adaptabilidade de quinze cultivares de milho, a partir dos escores obtidos, considerando quatro ambientes contidos nos fatores de Condições edafoclimáticas ruins (Fator 2) e Seca (Fator 3).

As cinco cultivares mais produtivas alocadas em Q1, as três menos produtivas em Q3 e as demais alocadas em Q2 e Q4 mantiveram oscilações de produtividade conforme suas adaptabilidades específicas. Este fato também foi observado por Garbuglio et al. (2007), que utilizaram, para análise de adaptabilidade, três eixos que absorveram 73,7% da variação, em vez dos cinco fatores finais, com 81,8% da variação, concluindo que a análise de fatores foi eficiente no estudo de adaptabilidade.

Este fato vai ao encontro das observações anteriores feitas para outras técnicas multivariadas, para as quais mais de 70% captados nos primeiros eixos, representam parte expressiva da variação (RAMOS; SANCHES; COTES, 2009;

YAN et al., 2000; ZOBEL; WRIGHT; GAUCH JUNIOR, 1988), não sendo necessária a utilização de valores mais elevados, como 80% utilizado empiricamente na análise de fatores original.

4 CONCLUSÕES

- 1) A técnica de análise de fatores associada a efeitos genotípicos somados aos efeitos de interação $G \times A$ (FGGA) é eficiente no processo de estratificação ambiental e análise de adaptabilidade.
- 2) Acima de 70% de variação captada nos primeiros autovalores representa parte expressiva da variação contida na soma dos efeitos genotípicos e de interação $G \times A$.
- 3) As condições climáticas atuaram como um dos fatores preponderantes no agrupamento de ambientes, no entanto, puderam ser detectados pela técnica FGGA.
- 4) As cultivares com ampla adaptação e alta produtividade no conjunto de ambientes testados foram P30F35, P30F53, P30K64, P30R50 e AS 1570.

REFERÊNCIAS

- ACQUAAH, G.; ADAMS, M. W.; KELLY, J. D. A factor analysis of plant variables associated with architecture and seed size in dry bean. **Euphytica**, Wageningen, v. 60, n. 3, p. 171-177, Apr. 1992.
- BRASIL, E. M. **Comparação de métodos no estudo da interação de genótipos com ambientes em milho (*Zea mays* L.)**. 1990. 181 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Universidade Federal de Goiás, Goiânia, 1990.
- CROSSA, J. et al. AMMI adjustment for statistical analysis of an international wheat yield trial. **Theoretical and Applied Genetics**, Berlin, v. 81, n. 1, p. 27-37, Jan. 1991.
- CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2003. v. 2, 585 p.
- CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2004. v. 1, 480 p.
- DUARTE, J. B.; VENCOVSKY, R. **Interação genótipos x ambientes: uma introdução a análise "AMMI"**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 59 p.
- DUARTE, J. B.; ZIMMERMANN, M. J. O. Correlation among yield stability parameters in common bean. **Crop Science**, Madison, v. 35, n. 3, p. 905-912, June 1995.

DUVICK, D. N. Heterosis: feeding people and protecting natural resources. In: COORS, J. G.; PANDEY, S. (Ed.). **The genetics and exploitation of heterosis in crops**. Madison: ASA, 1999. p. 19-29.

FERREIRA, D. F. **Estatística multivariada**. Lavras: Ed. UFLA, 2008. 662 p.

GARBUGLIO, D. D. et al. Análise de fatores e regressão bissegmentada em estudo de estratificação ambiental e adaptabilidade em milho. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 2, p. 183-191, Feb. 2007.

GERAGE, A. C.; SHIOGA, P. S.; ARAUJO, P. M. **Avaliação estadual de cultivares de milho safra 2005/2006**. Londrina: IAPAR, n.150, 2006. 55p.

GOULD, S. J. **The mismeasure of man**. New York: W. W. Norton, 1996. 443 p.

GRANATE, M. J. et al. A análise de fatores na predição de ganhos por seleção em milho (*Zea mays* L.). **Acta Scientiarum**, Maringá, v. 23, n. 5, p. 1271-1279, set. 2001.

JUVIK, J. A. et al. Kernel changes in a *shrunk-2* maize population associated with selection for increased field emergence. **Journal on the American Society of Horticultural Science**, Alexandria, v. 118, n. 1, p. 135-140, Jan. 1993.

KANG, M. S.; GAUCH JÚNIOR, H. G. **Genotype by environment interaction**. New York: CRC, 1996. 416 p.

LIN, C. S. Grouping genotypes by a cluster method directly related to genotype-environment interaction mean-square. **Theoretical and Applied Genetics**, Berlin, v. 62, n. 3, p. 277-280, Sept. 1982.

MENDONÇA, O. et al. Análise de fatores e estratificação ambiental na avaliação da adaptabilidade e estabilidade em soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 11, p. 1567-1575, nov. 2007.

MURAKAMI, D. M.; CRUZ, C. D. Proposal of methodologies for environment stratification and analysis of genotype adaptability. **Crop Breeding and Applied Biotechnology**, Londrina, v. 4, n. 1, p. 7-11, jan. 2004.

RAMOS, L. M.; SANCHES, A.; COTES, J. M. Teste multiambientais na seleção de genótipos de arroz utilizando o modelo de regressão nos sítios ou locais. **Ciência Rural**, Santa Maria, v. 39, n. 1, p. 52-57, jan. 2009.

ROBERTSON, A. **Experimental design on the measurement of heritabilities and genetic correlations. biometrical genetics**. New York: Pergamon Press, 1959.186p.

ROCHA, M. M. **Seleção de linhagens experimentais de soja para adaptabilidade e estabilidade fenotípica**. 2002. 173 p. Tese (Doutorado em Genética e Melhoramento) – Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2002.

ROCHA, M. M.; VELLO, N. A. Interação genótipos e locais para rendimento de grãos de linhagens de soja com diferentes ciclos de maturação. **Bragantia**, Campinas, v. 58, n. 1, p. 69-81, jun. 1999.

SHIOGA, P. S.; GERAGE, A. C.; ARAUJO, P. M. **Avaliação estadual de cultivares de milho safra 2006/2007**. Londrina: IAPAR, n.152, 2007. 53p.

TROYER, A. F. Adaptedness and heterosis in corn and mule hybrids. **Crop Science**, Madison, v. 46, n. 2, p. 529-543, Mar. 2006.

TROYER, A. F.; WELLIN, E. J. Heterosis decreasing in hybrids: yield test inbreds. **Crop Science**, Madison, v. 49, n. 6, p. 1969-1976, Nov. 2009.

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

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

**CAPÍTULO 3 Metodologia FGGA: comparação de métodos para estudo da
interação genótipos × ambientes**

RESUMO

Este trabalho foi realizado com o objetivo de avaliar três métodos estatísticos multivariados para análise de adaptabilidade e estratificação ambiental simultaneamente, utilizando dados de cultivares de milho indicadas para cultivo no estado do Paraná. Nos métodos GGE e FGGA, o efeito genotípico atuou como um coeficiente de ajuste das interações $G \times A$ ao longo dos ambientes, implicando em altos percentuais de explicação, associados a um menor número de eixos. A estratificação ambiental pelos métodos GGE e FGGA apresentou respostas similares, enquanto pelo método AMMI não houve garantia de agrupamento de ambientes. As análises de adaptabilidade apresentaram poucas divergências de resposta, pelos três métodos. Os genótipos P30F35, P30F53, P30R50, P30K64 e AS 1570 apresentaram altas produtividades associadas à adaptabilidade geral. O método FGGA permitiu associar as diferenças de respostas de produtividade entre determinados conjuntos de ambientes e o impacto em localidades pertencentes ao mesmo conjunto ambiental (através de rA), com o auxílio do nível de porção simples atuante da interação $G \times A$.

Palavras-chave: Análise multivariada. Análise de fatores. Efeito genotípico. Milho.

ABSTRACT

The objective of this work was to evaluate three statistical multivariate methods for adaptability and environmental stratification, simultaneously, using data from maize cultivars indicated for planting in Paraná State. In FGGA and GGE methods, the genotypic effect has an adjust effect in the $G \times E$ interactions over environments, resulting in high percentages of explanations, associated to a lower number of axis. The environmental stratification by FGGA and GGE methods showed similar responses, while by AMMI method there was no warranty of environmental grouping. The adaptability analysis showed low divergences patterns of responses by the three methods. The genotypes P30F35, P30F53, P30R50, P30K64 and AS 1570 showed high yields associated with general adaptability. The FGGA method allowed to associate difference of yield responses in specific regions and the impact in locations belonging to the same environmental group (through r_E), with the level of simple portion of $G \times E$ interaction.

Keywords: Multivariate analysis. Factor analysis. Genotypic effect. Maize.

1 INTRODUÇÃO

Os ensaios finais entre cultivares são conduzidos, anualmente, em todo o mundo, em uma ampla gama de ambientes, por várias empresas de melhoramento e produção de sementes, públicas e privadas. O ponto chave é identificar cultivares superiores, em termos de produtividade, associada a uma ampla adaptabilidade e alta estabilidade desta produção, frente às oscilações ambientais cada vez mais comuns nos novos cenários agrícolas.

Um segundo ponto, igualmente importante, refere-se a compreender o comportamento dessas cultivares em determinadas regiões agrícolas de interesse e determinar se essas regiões podem ser subdivididas dentro de diferentes mega-ambientes (YAN et al., 2000) ou sub-regiões mais uniformes, em que se exclui qualquer interação $G \times A$ significativa ou, em outras situações, uma interação significativa com predominância de parte simples atuante, ou seja, que não venha a comprometer a recomendação das cultivares (CRUZ; CARNEIRO, 2003).

Existem diversas metodologias destinadas à avaliação da interação $G \times A$ em que a escolha do método dependerá dos dados experimentais, especialmente do número de ambientes disponíveis, da precisão requerida e do tipo de informação desejada. Nos últimos anos, as técnicas multivariadas têm ganhado espaço neste tipo de estudo, em virtude da popularização dos computadores e de modernos pacotes estatísticos, permitindo que os complexos cálculos envolvendo álgebra de matrizes e modelos lineares sejam realizados em segundos.

Uma dessas técnicas, chamada de método AMMI, considera modelos aditivos para os efeitos principais (genótipos e ambientes) e multiplicativos para os efeitos da interação $G \times A$ (CROSSA; GAUCH; ZOBEL, 1990). Desse modo, a resposta média de um genótipo i em um ambiente j é dada por:

$$Y_{ij} = \mu + g_i + a_j + \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \varepsilon_{ij},$$
 com $(ga)_{ij}$ modelado por $\sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \varepsilon_{ij}$. Sob as restrições de identificabilidade, além da média geral (μ) e do erro experimental médio (ε_{ij}), os demais termos do modelo resultam da chamada decomposição por valores singulares (DVS) da matriz de interações: $GA_{(g \times a)} = [(\hat{g}\hat{a})_{ij}]$ (DUARTE; VENCOVSKY, 1999).

Outro método multiplicativo que vem ganhando espaço é o chamado SREG - GGE Biplot. O modelo multiplicativo SREG (CROSSA; CORNELIUS, 1997) é semelhante ao modelo AMMI na modelagem da interação $G \times A$, com a diferença de que os efeitos principais dos genótipos são considerados junto com o efeito da interação $G \times A$ que, no AMMI, são estimados como efeitos aditivos (YAN et al., 2000). Desse modo, o biplot originado por esse método contém os efeitos genotípicos somados aos efeitos de interação, sendo denominado de GGE biplot, ao contrário do método AMMI, para o qual o biplot é baseado na DVS da matriz de interações $G \times A$, podendo ser denominado de GE Biplot.

A técnica de análise de fatores assemelha-se à técnica de componentes principais, no sentido de que ambas sejam propostas para estudar a estrutura de covariâncias ou de correlações populacionais (FERREIRA, 2008). No entanto, a explicação total das covariâncias ou das correlações na análise de componentes principais é feita utilizando-se todas as variáveis latentes, enquanto, na análise fatorial, é possível explicar todas as covariâncias ou correlações, utilizando algumas poucas variáveis não observáveis ou latentes, que são denominadas fatores (FERREIRA, 2008).

A técnica originalmente proposta considera como entrada uma matriz de médias fenotípicas, o que pode levar a inferências equivocadas, tanto para a adaptabilidade das cultivares avaliadas quanto no processo de estratificação ambiental, devido ao ruído presente na interação $G \times A$. Isso implica também em

um número elevado de fatores, dependendo da natureza das variáveis a serem consideradas no agrupamento.

A proposta derivada da análise de fatores, denominada FGGA, considera como entrada uma matriz contendo os efeitos genotípicos somados aos efeitos de interação $G \times A$, buscando, assim, melhorar a eficiência no processo de fatoração, visando à estratificação ambiental e à análise de adaptabilidade, por meio de uma redução no ruído presente na interação $G \times A$.

Este trabalho foi realizado com o objetivo de comparar três métodos estatísticos multivariados para análise de adaptabilidade e estratificação ambiental simultaneamente, utilizando dados de cultivares de milho indicadas para cultivo no estado do Paraná.

2 MATERIAL E MÉTODOS

Foram utilizados dados de quinze cultivares de milho indicadas para cultivo no estado do Paraná, sendo estas nove híbridos simples (AS 1570, AS 1575, DOW 2A525, DOW 2B710, AG 8021, P30F35, P30F53, P30K64 e P30R50), quatro híbridos triplos (BM 1120, DKB 566 e SHS 5070) e dois híbridos duplos (IPR 119 e SHS 4050). Os dados utilizados nas análises são referentes à variável produtividade de grãos (corrigidos para 14% de umidade e transformados para $\text{kg} \cdot \text{ha}^{-1}$), obtida a partir de experimentos conduzidos em delineamento de blocos casualizados com três repetições por ambiente. A parcela experimental utilizada foi a de duas linhas de 5 m de comprimento, espaçadas 80 cm entre linhas, deixando-se 25 plantas por linha após desbaste. Considerou-se como ambiente a combinação de local e safra. Desse modo, pela combinação dos cinco locais (Londrina, Campo Mourão, Wenceslau Braz, Ponta Grossa e Pato Branco) e das duas safras (2005/2006 e 2006/2007), foram obtidos dez ambientes (A1, A2, ..., A10).

Após as análises de variância individuais e conjunta, foram realizadas as análises multivariadas com ênfase na adaptabilidade e na estratificação ambiental, sendo estas a técnica SREG-GGE Biplot (YAN et al., 2000), o método AMMI (ZOBEL; WRIGHT; GAUCH JUNIOR, 1988) e a proposta do presente trabalho FGGA.

O modelo multiplicativo SREG (regressão por sítios ou locais), como citado anteriormente, é semelhante ao modelo utilizado no método AMMI. A média de um genótipo i em um ambiente j é comumente descrita por meio de um modelo linear: $Y_{ij} = \mu + g_i + a_j + \phi_{ij}$, em que μ é a média geral; g_i é o efeito aditivo de genótipos; a_j é o efeito aditivo de ambientes e ϕ_{ij} é o efeito de interação entre o genótipo i e o ambiente j . A deleção de g_i e/ou a_j , ou o conjunto $\mu + g_i + a_j$, permite que a variação explicada pelos termos deletados sejam absorvidos para dentro do termo ϕ_{ij} , ou seja, submetendo somente o componente ϕ_{ij} a DVS, sem qualquer deleção, resulta no modelo AMMI.

O que ocorre no GGE-biplot é utilizar o modelo SREG, obtido após eliminação do componente g_i e submetendo ϕ_{ij} a DVS (YAN et al., 2000). Desse modo, $\phi_{ij} = Y_{ij} - \mu - a_j = \sum_{n=1}^r \xi_{in}^* \eta_{jn}^*$, em que $\xi_{in}^* = \lambda_n^{0,5} \xi_{in}$ e $\eta_{jn}^* = \lambda_n^{0,5} \eta_{jn}$ são utilizados para se obter uma escala simétrica entre os escores dos genótipos e dos ambientes para a construção do biplot. Dentro do modelo, λ_n é o valor singular do CP_n (Componente principal “n”), em que o seu valor ao quadrado corresponde à soma de quadrados explicada pelo CP_n ; ξ_{in} e η_{jn} são os escores do i -ésimo genótipo e do j -ésimo ambiente para o CP_n .

O modelo FGGA, a partir de uma matriz contendo os efeitos g_i , somados aos efeitos $ga_{(ij)}$, foi adotado, sendo estimados os fatores que, combinados de forma linear, explicam cada variável estudada. O modelo fatorial (FGGA) é dado por:

$$\begin{aligned}
x_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\
x_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2 \\
&\dots \\
x_h &= \ell_{h1}F_1 + \ell_{h2}F_2 + \dots + \ell_{hm}F_m + \varepsilon_h \\
\text{ou } x_j &= \sum_{k=1}^m \ell_{jk}F_k + \varepsilon_j
\end{aligned}$$

sendo $m < h$; em que ℓ_{jk} é a carga fatorial para a j -ésima variável associada ao k -ésimo vetor; F_k é o k -ésimo fator comum; ε_j é o fator específico associado à j -ésima variável; para $h=p$ para genótipos ou $h=q$ para ambientes.

Para a utilização da análise de fatores no estudo de adaptabilidade e estratificação ambiental empregado, deve-se destacar o fato de que $X_1, X_2 \dots X_h$ representam uma única variável, como, por exemplo, produtividade de grãos, porém, avaliada em cada um dos j ambientes em que os genótipos foram avaliados ou dos i genótipos referentes aos j ambientes. Assim, os efeitos de genótipos adicionados aos efeitos da interação genótipos por ambientes foram utilizados como variáveis. Ora tinham-se os ambientes, relativos aos valores dessa soma, representando as h variáveis e ora tinham-se os genótipos.

Para a estratificação ambiental, o agrupamento dos ambientes foi realizado a partir das informações sobre a magnitude das cargas fatoriais finais, obtidas após rotações. Cargas fatoriais maiores ou iguais a 0,70, em valor absoluto, indicaram ambientes com altas correlações, que foram agrupados dentro de um mesmo fator. Cargas fatoriais finais de valores baixos ($\leq 0,50$) indicaram que o respectivo ambiente não deveria pertencer ao grupo. Cargas fatoriais com valores intermediários não garantiam nenhuma definição de agrupamento.

A análise de adaptabilidade baseada na análise de fatores foi realizada graficamente, por meio do uso dos escores em relação aos fatores. Nos quadrantes II e IV foram situados aqueles genótipos com adaptabilidade específica à região determinada pelo fator. No quadrante I foram situados os genótipos de adaptabilidade ampla e, no quadrante III, os genótipos ruins, de baixo desempenho e passíveis de descarte ou não indicados para cultivo.

Em situações nas quais apenas dois fatores eram suficientes para explicar mais de 80% da variação total, determinando apenas duas sub-regiões, a análise de adaptabilidade foi feita com base nas informações de um único gráfico. No entanto, como constatado no capítulo anterior, valores acima de 70% foram suficientes para explicar a variação, o que é indicado também por outros autores com diferentes técnicas multivariadas (GARBUGLIO et al., 2007; RAMOS; SANCHES; COTES, 2009; YAN et al., 2000; ZOBEL; WRIGHT; GAUCH JUNIOR, 1988). Porém, não há estudos demonstrando o que ocorre no agrupamento dos ambientes, em virtude da redução do percentual de variação para a determinação no número final de fatores. Desse modo, para o agrupamento de ambientes, foi considerado que o número final de fatores seria igual ao número de autovalores que apresentem pelo menos 80% da variação e, para análise de adaptabilidade, foram utilizados os fatores que apresentassem conjuntamente 70% da variação.

3 RESULTADOS E DISCUSSÃO

3.1 Método AMMI

A interação G×A foi significativa, a 1% de probabilidade (Tabela 1), ou seja, os genótipos apresentaram respostas diferenciadas ao longo dos ambientes

de avaliação, o que pode dificultar a recomendação de cultivares para a região abrangida pelo estudo.

Por meio da análise AMMI, espera-se capturar grande parte do padrão estrutural da $SQ_{G \times A}$ nos primeiros componentes. Em concordância com o modelo AMMI, a interação $G \times A$ original pôde ser decomposta em nove componentes (posto da matriz $G \times A$), dos quais os quatro primeiros componentes foram altamente significativos ($p < 0,01$). Segundo regra de Gollob, isto levaria à escolha do modelo AMMI4, em alusão ao fato de que seriam necessários 4 eixos principais para explicar a interação de forma significativa. Neste caso, seria possível a construção de 6 gráficos envolvendo os 4 CPs, combinados 2 a 2. Considerando que a seleção de eixos é um critério liberal (PIEPHO, 1995), uma alternativa seria a aplicação do teste F_R aplicado aos resíduos AMMI, o qual Piepho (1995) defende como um dos mais robustos. Por meio deste critério, seleciona-se o modelo AMMI a partir do IPCA em que o resíduo AMMI torna-se não significativo. Neste caso, o modelo adotado seria AMMI3, em que este explica 70,2% da variação na interação $G \times A$ (Tabela 2). No entanto, não há um consenso sobre qual seria a proporção mínima da soma de quadrados de $G \times A$ que deve ser acumulada pelo primeiro componente principal para a construção do Biplot.

Carbonell et al. (2004) mencionam que é conveniente uma análise visual prévia. Segundo Duarte e Vencovsky (1999), os biplots podem ser apresentados desde que o primeiro eixo acumule proporções entre 27,1% e 71%. No presente estudo, optou-se por construir um biplot entre médias e escores obtidos no primeiro componente, que absorveu 33,3% da variação e um segundo biplot envolvendo os escores do primeiro e do segundo componentes, que absorveram 56,1% da variação.

Tabela 1 Quadrados médios obtidos pela análise de variância conjunta com desdobramento da interação original G×A pelo método AMMI.

FV	GL	^(a) QM	F _G	GL	^(a) QM _{res AMMI}	F _R
Genótipo	14	2846,7 **				
Ambiente	9	6253,6 **				
G x A	126	179,85 **				
IPCA-1	22	114,28 **	104		48,46 **	
IPCA-2	20	86,13 **	84		39,49 **	
IPCA-3	18	59,08 **	66		34,14	
IPCA-4	16	52,01 *	50		28,42	
IPCA-5	14	43,27	36		22,65	
IPCA-6	12	28,5	24		19,73	
IPCA-7	10	25,28	14		15,76	
IPCA-8	8	24,46	6		4,16	
IPCA-9	6	4,16	0		0	

^(a): QM x 10⁴

Para os modelos SREG-GGE e FGGA, foi constatado que os dois primeiros componentes absorveram 79,3% e 77,7% da variação, respectivamente (Tabela 2) e, incluindo mais um componente, estes passaram a explicar 85,9% e 84,6% da variação, respectivamente. Estes altos percentuais de explicação, associados a um menor número de eixos, podem ser devido ao fato de o efeito genotípico atuar como um coeficiente de ajuste das interações G×A ao longo dos ambientes. Desse modo, para o processo de estimação dos autovalores, via componentes principais, as variâncias de cada variável (neste caso ambientes) e as covariâncias entre variáveis são maximizadas (Figura 1). Porém, isso ocorre de forma não proporcional, uma vez que o efeito genotípico é constante entre ambientes e variável dentro de ambientes, o que pode ter refletido no aumento da eficiência na captação da variação para genótipos e ambientes estudados simultaneamente.

Tabela 2 Porcentagens de explicação acumulada para diferentes métodos empregados de análise multivariada.

Eixos	% Explicação Acumulada		
	AMMI	GGE Biplot	FGGA
1	33,3	67,3	64,9
2	56,1	79,3	77,7
3	70,2	85,9	84,6
4	81,2	90,4	89,6
5	89,2	93,6	93,6
6	93,7	96,2	96,1
7	97,1	97,8	97,7
8	99,7	99,0	99,0
9	100,0	99,9	99,9

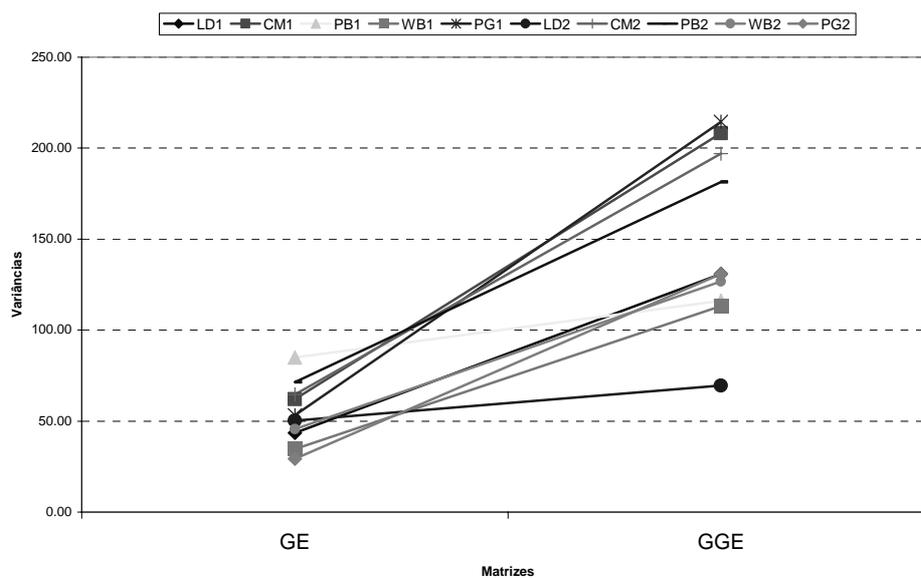


Figura 1 Estimativas de variâncias para produtividade ($\text{kg}\cdot\text{ha}^{-1}$) dentro de ambientes, utilizando as matrizes GE e GGE. Dados obtidos de quinze híbridos de milho avaliados em cinco locais no estado do Paraná (LD – Londrina; CM - Campo Mourão; PG - Ponta Grossa; WB - Wenceslau Braz; PB - Pato Branco) em duas safras (1 - 2006/2007 e 2 - 2005/2006).

Uma das abordagens utilizadas frequentemente no método AMMI é a construção do biplot envolvendo médias×IPCA1, chamado AMMI1 (CARBONELL et al., 2004; MAIA et al., 2006; OLIVEIRA; DUARTE; PINHEIRO, 2003). Neste caso, o enfoque complementar às características de adaptabilidade e estabilidade fenotípica, como a produtividade média dos genótipos testados, são imprescindíveis para recomendações de uma nova cultivar.

Pelo AMMI1(Figura 2), dos quinze híbridos avaliados, nove obtiveram médias superiores à média geral (10973 kg.ha⁻¹). Dentre esses híbridos, os que apresentaram interações relativamente baixas com determinados ambientes e tendo suas classificações determinadas, basicamente, pelos efeitos genotípicos, foram P30K64, AS 1570, P30F53, P30R50 e AS 1575, em ordem decrescente de produtividade. P30F35 obteve a maior média de produtividade (12.598 kg.ha⁻¹), no entanto, não apresentou baixos efeitos de interações específicas com quaisquer dos ambientes testados. Os ambientes de maiores produtividades (PG1, LD2 e PG2) não se destacaram quanto à estabilidade, por estarem distantes da origem. Os ambientes CM2 e LD1 apresentaram, em relação ao conjunto, baixas produtividades médias (9.426 kg.ha⁻¹ e 10.457 kg.ha⁻¹), porém, alta estabilidade, considerando o primeiro IPCA, o que pode facilitar a seleção de genótipos nesses ambientes.

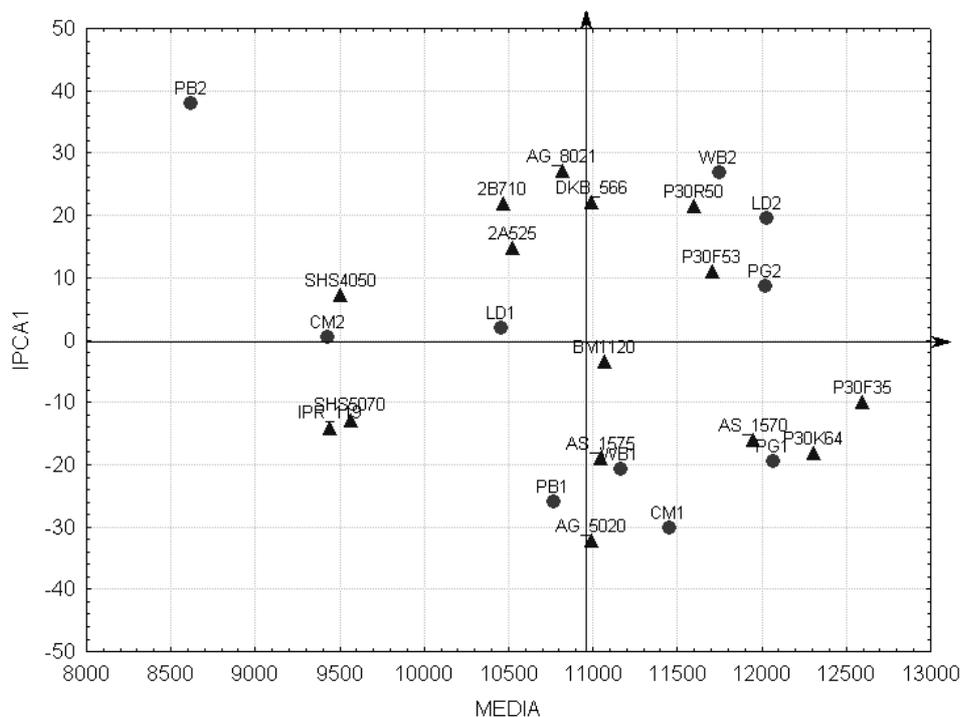


Figura 2 Biplot AMMI para dados de produtividade de grãos ($\text{kg}\cdot\text{ha}^{-1}$) de quinze híbridos de milho em cinco locais no estado do Paraná (LD – Londrina; CM - Campo Mourão; PG - Ponta Grossa; WB - Wenceslau Braz; PB - Pato Branco) em duas safras (1 - 2006/2007 e 2 - 2005/2006).

As relações adaptativas podem ser constatadas facilmente em um biplot AMMI2 (Figura 3), observando-se os sinais dos escores para cada par de genótipos e ambientes (MAIA et al., 2006). Dessa forma, genótipos e ambientes com escores de mesmo sinal (-,- ou +,+) devem interagir positivamente, enquanto sinais opostos (-,+ ou +,-) indicam interação negativa (DUARTE; VENCOVSKY, 1999).

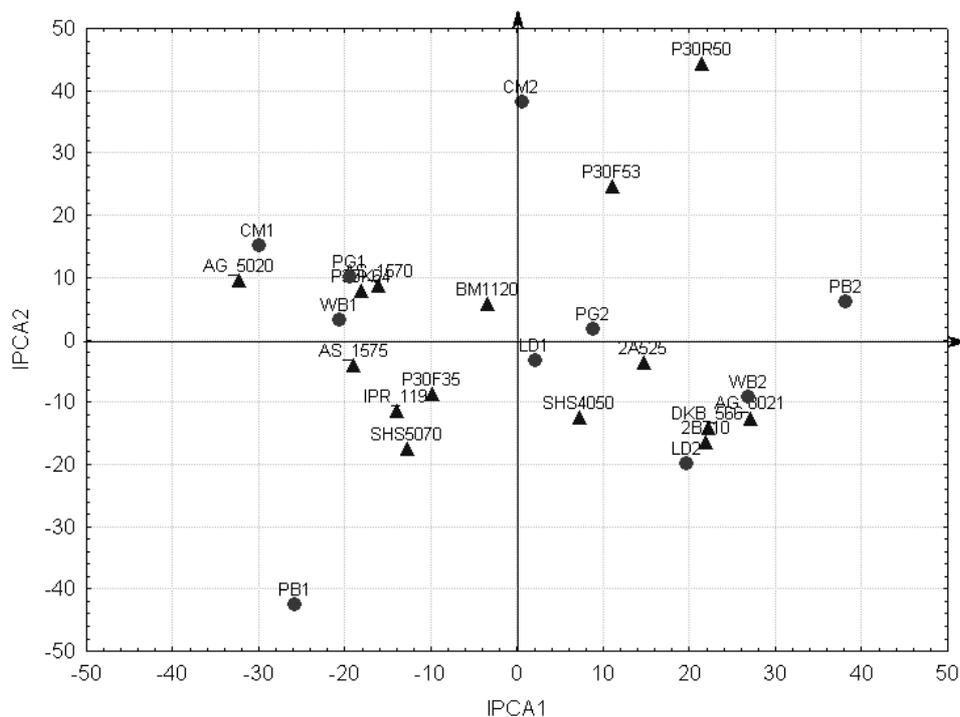


Figura 3 Biplot AMMI2 para dados de produtividade de grãos ($\text{kg}\cdot\text{ha}^{-1}$) de quinze híbridos de milho em cinco locais no estado do Paraná (LD – Londrina; CM - Campo Mourão; PG - Ponta Grossa; WB - Wenceslau Braz; PB - Pato Branco) em duas safras (1 - 2006/2007 e 2 - 2005/2006).

O único genótipo que contribuiu pouco para a interação cruzada foi BM 1120, apresentando produtividade pouco acima da média geral ($11.072 \text{ kg}\cdot\text{ha}^{-1}$). Os demais apresentaram altas magnitudes para o efeito de interação $G \times A$. Um grupo formado pelos genótipos P30K64 e AS 1570 apresentou interação positiva com os ambientes PG1 e WB1 e estabilidade média, assim como os híbridos 2B710 e DKB566, com ambiente LD2, e AG 8021 com WB2. Essas informações podem ser válidas, no caso de recomendações específicas. Porém, tomando como base a safra 2005/2006, deve ser considerado o forte estresse hídrico que atingiu as áreas produtoras de milho, tornando a recomendação cautelosa, caso seja um peso maior ao nível de estabilidade e não ao potencial

produtivo médio desses híbridos. Esta afirmação também é válida para a recomendação de genótipos caracterizados como instáveis, como P30F53 e P30R50, que têm elevados potenciais produtivos.

Um grupo que apresentou estabilidade geral média é formado pelos genótipos AS 1575, P30F35, IPR 119, SHS 5070, SHS 4050 e 2A525. Os ambientes mais estáveis foram LD1 e PG2, próximos à origem dos eixos. Porém, os genótipos IPR 119, SHS 4050 e SHS 5070, por apresentarem baixa produtividade, são passíveis de não indicação de cultivo dentro do conjunto estudado. CM2, apesar de distante da origem de IPCA2, foi caracterizado como de média estabilidade, por estar posicionado na origem de IPCA1, que te, a maior variação captada de $G \times A$ associada a alto padrão. Dentre os ambientes que mais inflacionaram a interação $G \times A$, destacam-se CM1, PB1 e PB2.

Uma das falhas do método AMMI é restringir o nível de estabilidade de um dado genótipo à sua faixa de produtividade. Sendo assim, um genótipo com baixo potencial produtivo, mas que o mantenha, mesmo na melhoria de condições ambientais, é caracterizado como estável, como, por exemplo, o genótipo SHS 5070. Por outro lado, genótipos de altas produtividades, como $12.000 \text{ kg} \cdot \text{ha}^{-1}$, podem ser caracterizados como instáveis, se sua produção cair menos 10%, como no caso do genótipo P30F35.

Carbonell et al. (2004), analisando os níveis de estabilidade de diferentes cultivares de feijoeiro, por meio dos métodos AMMI e desvio máximo de rendimento ou DMR (LINN; BINNS, 1988 citado por CARNEIRO, 1998), constataram que o método DMR identificou invariavelmente as cultivares mais estáveis e também mais produtivas, um fato que não foi verificado sempre pelo método AMMI. A razão apontada pelos autores é o fato de que a técnica DMR identifica cultivares mais estáveis como sendo aquelas com menor desvio da maior média de produtividade em cada ambiente estudado. Dessa forma, o procedimento fica amplamente relacionado com uma recente definição para

adaptação apresentada por Cecarelli (1996), em que são consideradas adaptadas aquelas cultivares/linhagens que apresentem as maiores produções econômicas.

3.2 Método SREG-GGE Biplot

Por meio do método GGE-biplot, a produtividade de um genótipo i em um determinado ambiente j será, aproximadamente, o produto do escore de CP1 do genótipo i versus o CP1 do ambiente j , mais o produto do escore do CP2 do genótipo i versus o escore do CP2 do ambiente j . Geometricamente, isto será o comprimento do vetor ambiente (distância absoluta da origem ao marcador do genótipo), multiplicado pelo comprimento do vetor do genótipo (distância absoluta da origem ao marcador do ambiente) e multiplicado pelo cosseno do ângulo entre eles (KROONENBERG, 1995).

Segundo Yan et al. (2000), essa propriedade permite que as seguintes informações sejam visualizadas: i) a similaridade e dissimilaridade entre ambiente e as diferentes respostas provocadas nos genótipos; ii) a similaridade e a dissimilaridade entre genótipos e respostas nos ambientes; iii) a natureza positiva ou negativa e a magnitude de interação entre qualquer genótipo e qualquer ambiente.

Na Figura 4, os genótipos localizados nos vértices do polígono formado têm a maior distância da origem de todos os genótipos dentro do setor delimitado por ele, sendo classificados como os mais responsivos. Estes, sendo os melhores ou piores genótipos, em algumas ou todas as localidades, podem ser utilizados para identificar possíveis mega-ambientes. Os genótipos dentro do polígono foram menos responsivos para os ambientes localizados dentro dos setores.

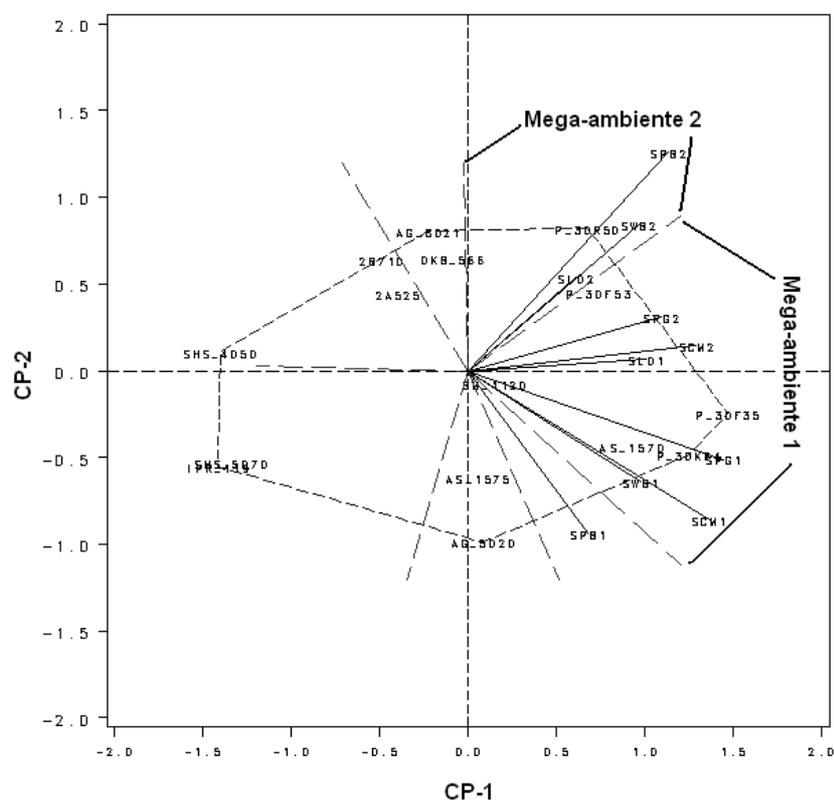


Figura 4 Plotagem gráfica referente à análise SREG-GGE Biplot para quinze híbridos de milho avaliados em cinco locais no estado do Paraná (LD – Londrina; CM - Campo Mourão; PG - Ponta Grossa; WB - Wenceslau Braz; PB - Pato Branco) em duas safras (1 - 2006/2007 e 2 - 2005/2006).

Os genótipos localizados nos vértices foram P30K64, P30F35, P30R50, AG 5020 (escores CP1 positivos), AG 8021, SHS 4050 e IPR 119 (CP1 negativos). Ao conectar os marcadores desses genótipos situados nos extremos obtém-se a formação do polígono. A partir de retas perpendiculares traçadas para cada lado do polígono, partindo da origem, os ambientes são divididos em dois grandes setores, podendo ter um ou mais genótipos vencedores nos vértices. O primeiro setor contém o maior conjunto de ambientes, sendo estes PG2, CM2, LD1, PG1, WB1, CM1 e PB1, tendo P30K64 e P30F35 como genótipos

vencedores. Um segundo setor, formado pelos ambientes PB2, WB2 e LD2, todos referentes à safra 2005/2006, foi identificado, tendo P30R50 como genótipo vencedor.

De modo geral, o método GGE Biplot separou o conjunto de ambientes conforme a safra, indicando a existência de dois mega-ambientes. Esse resultado era esperado, uma vez que o conjunto de localidades testadas busca representar a área agrícola de milho no estado do Paraná, em duas safras distintas, apesar dos altos níveis de interação $G \times A$ presentes. A alocação de CM2 e PG2 com as demais localidades da safra 2006/2007 pode estar relacionada às condições, como chuvas localizadas nas áreas desses experimentos durante a safra 2005/2006 (GERAGE; SHIOGA; ARAUJO, 2006), refletindo em bons resultados de produtividade. Outro fator favorável à CM2 é o plantio mais cedo realizado naquela localidade.

Calculando-se o coeficiente de correlação simples de Pearson entre os escores do CP1 e os efeitos de genótipos (médias de produtividade em todos os ambientes), o valor obtido foi de 0,99 (dados não apresentados), significativo ($p < 0,01$). Essa quase perfeita correlação de escores de CP1 e médias, também constatada por Souza (2004), confirma o que sugeriram Yan et al. (2001), ou seja, quando G corresponde a 40% da soma de quadrados, ou que esta seja superior à $G \times A$, a correlação entre G e CP1 apresenta valores elevados ($r > 0,90$). No presente estudo, 33,5% da SQ_{total} devem-se a efeitos genotípicos, enquanto 18,7% à interação $G \times A$, como apresentado no capítulo anterior. Analisando, ainda, as estimativas dos componentes quadráticos, percebe-se que as proporções são equivalentes (28,66% para $\hat{\phi}_G$ e 13,19% para $\hat{\phi}_{G \times A}$). Por outro lado, quando ocorre baixa correlação, esta pode ser devido a uma alta magnitude da parte complexa da interação $G \times A$, principalmente quando se analisam dados de vários anos conjuntamente e, geralmente, acompanhados de quantidades similares de soma de quadrados explicadas por CP1 e CP2 (YAN; HUNT,

2001). Portanto, genótipos com escores mais elevados para CP1 terão maiores médias de produtividade e serão mais bem identificados em ambientes com escores também elevados para CP1. Desse modo, ressalta-se a afirmação de Yan e Hunt (2001) de que o efeito de genótipo, também chamado de efeito principal de genótipo, acaba sendo resultado da própria interação $G \times A$.

A modelagem da técnica faz com que todos os ambientes possuam escores de mesmo sinal para CP1, assim, de acordo com Crossa e Cornelius (1997), estes representarão a diferença proporcional de produtividade nos vários ambientes, devido à interação $G \times A$ simples, não apresentando qualquer porção complexa atuante. O CP2, por sua vez, resume as fontes de variação que levam à interação complexa. Os ambientes e os genótipos podem assumir valores positivos e negativos, indistintamente. Desse modo, a porção complexa entre os melhores genótipos leva à diferenciação dos mega-ambientes.

É possível constatar que a localidade de Pato Branco, nas duas safras, apresentou os maiores níveis de interação $G \times A$ complexa, também constatado pelo método AMMI.

No GGE biplot os genótipos são avaliados quanto à adaptabilidade, a partir de estimativas aproximadas dadas pelos escores de CP1, e estabilidade quanto aos escores do CP2. Assim, genótipos produtivos e estáveis deverão possuir escores elevados para CP1, porém, próximos de zero para CP2. Esses genótipos são mais facilmente identificados em ambientes com elevados escores de CP1 e próximos de zero em CP2.

Os genótipos que mais se aproximaram desse conceito foram P30F35 e P30K64. Pelo método AMMI, os resultados indicaram, para esses mesmos genótipos, adaptabilidade média ou específica a um conjunto pequeno de ambientes. Uma das causas dessa diferença pode estar associada à baixa explicação dos dois primeiros IPCAs, mais especificamente do IPCA 1, que contém o maior percentual de padrão associado à variação captada nesse eixo.

O genótipo BM 1120, apesar da alta estabilidade, obteve rendimento produtivo médio, como constatado pelo método AMMI. Na contínua busca por novos genótipos produtivos e estáveis em ambientes que facilitem a identificação, constatou-se que Londrina (safra 2006/2007) e Campo Mourão (safra 2005/2006) são adequados, por possuírem considerável escore para CP1 e baixo escore para CP2. Pelo método AMMI foi ressaltada a estabilidade desses mesmos locais, apesar das baixas médias de produtividade. No entanto, deve ser ressaltado que a safra 2006/2007 apresenta uma melhor definição de posicionamento, em virtude das melhores condições climáticas que ocorreram naquele período. Portanto, Londrina seria a melhor escolha.

No mega-ambiente 1, além das cultivares vencedoras, destacaram-se AS 1570 e P30F53. Cultivares alocadas neste setor podem ter suas indicações de plantio mais generalizadas, uma vez que sete dos dez ambientes testados compõem este mega-ambiente. P30R50 apresentou adaptabilidade específica às demais localidades da safra 2005/2006, agrupadas no mega-ambiente 2, resultado verificado, em parte, pelo método AMMI. Porém, por AMMI2, seria considerado como instável, divergindo do método GGE. Os demais genótipos localizados nos vértices, como SHS 4050 e IPR 119, além de SHS 5070, apresentaram baixo rendimento e, pelo fato de estarem distantes dos marcadores de ambientes, refletiu em baixa adaptabilidade no conjunto de dez ambientes testados.

Este fato também foi detectado pelo método AMMI, mas somente para os genótipos IPR 119 e SHS 5070. Os demais setores, por possuírem genótipos delimitadores próximos uns dos outros, ainda que distantes da origem, não tiveram ambientes localizados em seus setores formados. Isso é devido à similaridade de produtividades nos ambientes avaliados. Desse modo as perpendiculares entre eles poderiam ser desconsideradas na análise ou, nem mesmo, plotadas.

O efeito de genótipo no método GGE é considerado como multiplicativo, em termos de interação $G \times A$ e, partindo do princípio de que os escores de CP1 para locais/ambientes tendem a apresentar mesmo sinal, CP1 passa a representar a interação $G \times A$ simples (YAN et al., 2000). Devido ao fato de os escores de CP1 para genótipos serem altamente correlacionados com os efeitos de genótipos (médias), para fins práticos, os escores podem substituir os efeitos de genótipos. No entanto, deve ser salientado que, conceitualmente, são muito diferentes.

Como apresentado anteriormente, o efeito genotípico é constante em qualquer ambiente, porém, as predições de rendimento, a partir de CP1, no método GGE biplot, para um dado genótipo, não são constantes. Essas predições variam ao longo dos ambientes em uma proporção direta aos escores do CP1 para locais/ambientes. Yan et al. (2000) acreditam que essa proporcionalidade na resposta de rendimento dos genótipos seja mais lógica e biologicamente plausível do que o conceito de efeitos principais aditivos. No entanto, a única propriedade que sustenta esse conceito é que os locais/ambientes, que facilitam a identificação de genótipos, com grandes efeitos principais, também são simultaneamente indicados.

Yan et al. (2000) destacam, como outra importante propriedade da análise GGE, a diferenciação entre respostas proporcionais ou desproporcionais de cultivares e suas implicações nas interações $G \times A$ simples e complexa. O entendimento dessas interações pode ser obtido relacionando-se os escores genotípicos de CP1 (interação simples) e CP2 (interação complexa) ou por meio da covariância ambiental (SOUZA, 2004). O método AMMI, apesar de ser representado por um biplot, requer, por parte do melhorista, um alto conhecimento de estatística avançada, para que se possam extrair todas as interpretações possíveis para genótipos, ambientes e suas interações, não

permitindo, ainda, a diferenciação de respostas proporcionais e desproporcionais de genótipos/cultivares.

O biplot bidimensional, com base em SREG2, além de respostas visuais mais diretas, sempre usa um número intermediário de graus de liberdade e também explica uma magnitude intermediária da soma de quadrados $G+G \times A$ entre AMMI1 e AMMI2. Desse modo, GGE biplot sempre estará próximo de um dos modelos, porém, podendo apresentar maior percentual de explicação nos primeiros eixos. Souza (2004), utilizando um método simplificado sugerido por Gauch e Zobel (1996) para estimar o “padrão” e o “ruído” na avaliação de cultivares de milho em Minas Gerais, constatou que SREG2 foi o melhor modelo para safra 1998/1999 em relação a AMMI2, pois explicou a maior proporção da soma de quadrados da interação $G \times A$. Os autores ainda verificaram que, na safra 1998/1999, o método GGE explicou 53,4% da $SQG \times A$, contra 50,8% da análise AMMI, enquanto na safra 1999/2000 GGE explicou 52,3% e AMMI explicou 45,4%.

3.3 Método FGGA

Por meio do método FGGA, os ambientes foram agrupados em três fatores (apresentado no capítulo 2, Tabela 8). Em razão da elevada porção complexa atuante discutida no capítulo anterior, não foi possível reduzir o número de ambientes testados, mas sim identificar uma macrorregião de cultivo de milho no estado do Paraná. O agrupamento ocorreu em razão de safras com predominância dos efeitos climáticos como estresse por falta de água em fases críticas da cultura.

A localidade de Pato Branco (2005/2006) foi, novamente, agrupada de forma isolada no fator 3, em razão da elevada porção complexa atuante, podendo estar relacionada ao forte estresse hídrico que ocorreu naquela safra, atingindo

os estádios de pré-florescimento e enchimento de grãos. Durante a safra 2006/2007, para esta mesma localidade, as condições climáticas não foram tão favoráveis, mas permitiram um bom desenvolvimento da cultura (SHIOGA; GERAGE; ARAUJO, 2007). Pelo método GGE fica nítida a separação desses dois ambientes dos demais, dentro dos seus mega-ambientes, tanto pela distância da origem quanto de CP1.

O fator 1 no método FGGA, de modo geral, agrupou ambientes da safra 2006/2007, enquanto o fator 2 agrupou ambientes da safra 2005/2006. No fator 2, as localidades de Londrina e Ponta Grossa apresentaram cargas fatoriais de 0,6746 e 0,6780, respectivamente, muito próximas de 0,7, o que poderia confirmá-las no fator 2, porém, não há garantia total de seus agrupamentos. Pelo método GGE, LD2, apesar de estar alocada no mega-ambiente 2, está muito próxima da perpendicular, apresentando, em relação a WB2, $r=0,75$ ($p<0,01$) e PS% de 50,2, contra 0,62 ($p<0,05$) e PS% de 42,9%, o que pode ter influenciando no seu posicionamento, uma vez que CP1 relaciona-se estreitamente com PS% atuante na interação $G \times A$. Pelos métodos AMMI1 e AMMI2, é possível verificar os níveis de estabilidade dos ambientes em razão de suas distâncias ao IPCA1 (no caso de AMMI1) e da origem (no caso de AMMI2). Porém, torna-se subjetivo o estabelecimento de grupos de ambientes similares, principalmente quando aparentes adaptações específicas, exemplificadas nos casos de P30R50 e P30F53 com PG2, LD2 e WB2 por AMMI1 não se confirmam, quando há incremento das informações capturadas pelos segundo eixo (AMMI2). Duarte e Vencovsky (1999) também encontraram respostas similares utilizando dados de produtividade de feijoeiro obtidos em cinco ambientes.

Pelo fato de as análises GGE e FGGA permitirem inferências diretas e fortemente relacionadas às porções simples (CP1) e complexa (CP2) da interação $G \times A$, no caso do método GGE, e às covariâncias e correlações, no

caso do método FGGA, buscou-se considerar o agrupamento de ambientes somente por essas duas metodologias, restringindo o método AMMI para avaliação de genótipos.

Por meio das análises GGE e FGGA é possível separar dois grandes conjuntos de ambientes no estado do Paraná. No entanto, em ambos os conjuntos predominou a captação da interação $G \times A$, devido ao efeito de safras, como constatado anteriormente. Na formação dos conjuntos, as diferenças de posicionamento de determinados locais, tomando como referência diferentes fatores ou mega-ambientes, devem-se à base da metodologia utilizada. No método GGE, essa diferenciação de posicionamento é atribuída ao PS% atuante, enquanto no método FGGA predominam os efeitos de PS% mais a correlação entre ambientes.

Portanto, a associação entre os efeitos de r_A e PS% no método FGGA permite que os ambientes contidos dentro de cada fator também representem mega-ambientes para a indicação de culturas, desde que dentro de uma mesma safra. Ainda, partindo do efeito de r_A , pode-se inferir que, na formação dos grupos de ambientes, são levados em consideração os possíveis aumentos ou reduções de produtividade que ocorreram em determinados nichos de produção (analisados por pares de ambientes).

A análise gráfica de adaptabilidade pelo método FGGA (representada pela figura 1 do capítulo 2) indicou que P30F35, P30F53, P30R50, P30K64 e AS 1570 apresentam ampla adaptabilidade ao conjunto de ambientes agrupados nos fatores 1 e 2. Em relação ao método AMMI1, verificaram-se baixas magnitudes dos efeitos de interação com ambientes por parte de P30F53, P30R50, P30K64 e AS 1570, além de altas produtividades. P30F35 não foi incluído, apesar da maior média de produtividade, por apresentar magnitudes elevadas dos efeitos de interações com quaisquer dos ambientes testados. Nesse caso, foi incluído o genótipo AS 1575, por apresentar pequeno efeito da

interação específica com o ambiente WB1. No entanto, considerando associar estabilidade geral alta com produtividade média elevada, o melhor genótipo a ser incluído seria BM 1120 em vez de AS 1575, confirmado pelos métodos GGE e FGGA.

Pelo método GGE, P30F35 e P30K64 foram os genótipos vencedores no mega-ambiente 1, composto por sete dos dez ambientes testados. Dentro do polígono foram incluídos P30F53, AS 1570 e BM1120, enquanto, no mega-ambiente 2, P30R50 destacou-se como genótipo vencedor.

Os três métodos confirmaram a maioria dos mesmos genótipos com ampla adaptabilidade, ocorrendo algumas divergências quando se buscou analisar as adaptabilidades específicas relacionadas a determinados grupos de ambientes formados pelos métodos GGE e FGGA. No caso do genótipo P30R50, por GGE, este teria uma adaptabilidade para três locais da safra 2005/2006 (Wenceslau Braz, Pato Branco e Londrina). Porém, considerando o nível de produtividade, constata-se que o genótipo ocupou a 2^a, a 3^a e a 7^a posições, respectivamente em cada um desses ambientes (dados apresentados na tabela 10, capítulo 2). Nas demais localidades desta mesma safra e na safra 2006/2007, somente em Pato Branco foi verificada baixa produtividade deste genótipo, enquanto nas demais localidades ele esteve entre as seis maiores produtividades, na maioria das vezes. Este fato pode ajudar na confirmação de adaptabilidade geral constatada pelo método FGGA.

Em relação à adaptabilidade específica, os três métodos apresentaram boa concordância, no entanto, para os métodos GGE e AMMI, uma análise visual, com base nos dados gerais de produtividade, se faz necessária, para um correto direcionamento de cultivares.

Ainda, por meio desses métodos, especificamente AMMI2, os genótipos DKB 566, 2B710 e AG 8021 permaneceram próximos, estando 2A525 um pouco mais distante. Pelo método AMMI detectou-se baixa interação destes

genótipos com LD2 e WB2. Pelo método GGE estes mesmo genótipos ficaram próximos à delimitação do polígono e das perpendiculares, o que pode direcioná-los, a critério do melhorista, para os ambientes agrupados no mega-ambiente 2. Porém, deve ser levada em consideração a porcentagem de porção complexa atuante, uma vez que estão próximos de CP2. Por meio do método FGGA confirma-se o agrupamento desses genótipos nos ambientes referentes à safra 2005/2006, alocados no fator 2, neste caso WB2 e CM2.

Os genótipos AS 1575 e AG 5020 podem ser considerados com adaptabilidade específica ao mega-ambiente 1 ou fator 1, nos casos dos métodos GGE e FGGA. Por meio do gráfico AMMI1 verificou-se que esses genótipos estão próximos de WB1, CM1 e PB1, e por meio de AMMI2 permaneceram próximos de WB1 e CM1, distanciando-se de PB1 e aproximando-se de PG1. Como mencionado anteriormente, o método AMMI apresenta algumas diferenças de resposta à medida que se incorporam novos eixos na análise. No entanto, é possível constatar a adaptação desses genótipos com as localidades da safra 2006/2007, confirmando as respostas obtidas pelos demais métodos.

Pelas três metodologias, os genótipos SHS 4050, SHS 5070 e IPR 119, dentro do conjunto de genótipos e ambientes estudados, foram os que apresentaram os desempenhos menos satisfatórios em termos de produtividade e adaptabilidade.

4 CONCLUSÕES

- 1) Nos métodos GGE e FGGA, o efeito genotípico atuou como um coeficiente de ajuste das interações $G \times A$ ao longo dos ambientes, implicando em altos percentuais de explicação, associados a um menor número de eixos.

- 2) A estratificação ambiental pelos métodos GGE e FGGA apresentaram respostas similares, enquanto, pelo método AMMI, não houve garantia de agrupamento de ambientes.
- 3) As análises de adaptabilidade apresentaram poucas divergências de resposta pelos três métodos.
- 4) Os genótipos P30F35, P30F53, P30R50, P30K64 e AS 1570 apresentaram altas produtividades, associadas à adaptabilidade geral.
- 5) O método FGGA permitiu associar as diferenças de respostas de produtividade em determinadas regiões e o impacto em localidades pertencentes ao mesmo conjunto ambiental (através de r_A), com o nível de porção simples atuante da interação GA.

REFERÊNCIAS

CARBONELL, S. A. M. et al. Common bean cultivars and lines interaction with environments. **Scientia Agrícola**, Piracicaba, v. 61, n. 2, p. 169-177, mar. 2004.

CARNEIRO, P. C. S. **Novas metodologias de análise de adaptabilidade e estabilidade de comportamento**. 1998. 155 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Universidade Federal de Viçosa, Viçosa, MG, 1998.

CECARELLI, S. Adaptation to low/high input cultivation. **Euphytica**, Wageningen, v. 92, n. 1/2, p. 203-214, Jan. 1996.

CROSSA, J.; CORNELIUS, P. L. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. **Crop Science**, Madison, v. 37, n. 2, p. 405-415, Mar. 1997.

CROSSA, J.; GAUCH, H. G.; ZOBEL, R. W. Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. **Crop Science**, Madison, v. 30, n. 3, p. 493-500, May 1990.

CRUZ, C. D.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: UFV, 2003. v. 2, 585 p.

DUARTE, J. B.; VENCOSKY, R. **Interação genótipos x ambientes: uma introdução a análise "AMMI"**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 59 p.

FERREIRA, D. F. **Estatística multivariada**. Lavras: Ed. UFLA, 2008. 662 p.

GARBUGLIO, D. D. et al. Análise de fatores e regressão bissegmentada em estudo de estratificação ambiental e adaptabilidade em milho. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, n. 2, p. 183-191, Fev. 2007.

GAUCH, H. G.; ZOBEL, R. W. AMMI analysis of yield trials. In: KANG, M. S.; GAUCH JÚNIOR, H. G. (Ed.). **Genotype-by-environment proved and under what conditions this can be most environment interaction**. Boca Raton: CRC, 1996. p. 1-40.

GERAGE, A. C.; SHIOGA, P. S.; ARAUJO, P. M. **Avaliação estadual de cultivares de milho safra 2005/2006**. Londrina: IAPAR, n.150, 2006. 55p.

KROONENBERG, P. M. **Introduction to biplots for GxE tables**. Brisbane: University of Queensland, 1995.

MAIA, M. C. C. et al. Adaptabilidade e estabilidade de linhagens experimentais de soja selecionadas para caracteres agronômicos através de método uni e multivariado. **Bragantia**, Campinas, v. 65, n. 2, p. 215-226, jun. 2006.

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.

PIEPHO, H. P. Robustness of statistical test for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trial. **Theoretical Applied of Genetics**, Berlin, v. 90, n. 3/4, p. 438-443, Mar. 1995.

RAMOS, L. M.; SANCHES, A.; COTES, J. M. Teste multiambientais na seleção de genótipos de arroz utilizando o modelo de regressão nos sítios ou locais. **Ciência Rural**, Santa Maria, v. 39, n. 1, p. 52-57, jan. 2009.

SHIOGA, P. S.; GERAGE, A. C.; ARAUJO, P. M. **Avaliação estadual de cultivares de milho safra 2006/2007**. Londrina: IAPAR, n.152, 2007. 53p.

SOUZA, H. N. **Análise AMMI e SREG da interação genótipos x em milho**. 2004. 106 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Universidade Federal de Viçosa, Viçosa, MG, 2004.

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

YAN, W. et al. Two types of GGE biplots for analyzing multi-environment trial data. **Crop Science**, Madison, v. 41, n. 3, p. 656-663, May 2001.

YAN, W.; HUNT, L. A. Interpretation of genotype x environment interaction for winter wheat yield in ontario. **Crop Science**, Madison, v. 41, n. 1, p. 19-25, Jan. 2001.

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