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

Teste estatístico para contribuição de genótipos e ambientes na
matriz de interação GE

Mirian Fernandes Carvalho Araújo

Dissertação apresentada para obtenção do título de Mestre
em Agronomia. Área de concentração: Estatística e Experimentação Agronômica

Piracicaba
2008

Mirian Fernandes Carvalho Araújo
Licenciada em Matemática

**Teste estatístico para contribuição de genótipos e ambientes na matriz de
interação GE**

Orientador:
Prof. Dr. **CARLOS TADEU DOS SANTOS DIAS**

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

**Piracicaba
2008**

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Araújo, Mirian Fernandes Carvalho

Teste estatístico para contribuição de genótipos e ambientes na matriz de interação GE
/ Mirian Fernandes Carvalho Araújo. - - Piracicaba, 2008.
113 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura Luiz de Queiroz, 2008.
Bibliografia.

1. Análise de dados 2. Distribuição qui-quadrada 3. Genética estatística 4. Genótipos
Título

CDD 575.10212
A663t

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

DEDICATÓRIA

*“Deus é a minha salvação,
terei confiança e não temerei.”* (Isaías 12:2a)

Ao meu querido esposo **Lúcio Borges de Araújo**,
por toda ajuda, pelo incentivo e o amor sempre constante. Te amo muito!

Aos meus pais **Jair Vilarinho Carvalho** e **Lindomar Vilela Fernandes Carvalho**,
por todo amor construído a base de confiança e respeito.

A minha irmã **Léia Fernandes Carvalho**,
pela amizade e cumplicidade que sempre tivemos.

AGRADECIMENTOS

A Deus, autor e consumidor da minha fé, por sua eterna fidelidade e sem o qual nada podemos fazer.

Ao professor Dr. Carlos Tadeu dos Santos Dias pela orientação cultivada pela amizade, apoio e ajuda à elaboração deste trabalho. A sua esposa Elvina e seus filhos Vitor e Laura, pelos momentos prazerosos passados juntos.

A ESALQ/USP pela estrutura física e humana disponível.

A CNPq pela concessão de bolsas de estudos.

A Evani, D. Edna e Lucimara que me acolheram em sua casa de braços abertos e fez brotar um amor fraternal que nunca mais se acabará. Vocês são muito especiais.

Aos grandes amigos e irmãos, Marcos e Lilian Shimizu, pelo constante amor, pelos maravilhosos momentos juntos e por serem parte da nossa família em Piracicaba. A seus filhos, Lucas e Marcos, pelos momentos de descontração. Deus os abençoe ricamente.

Aos amigos e irmãos em Cristo da Igreja Evangélica na Paulista pela amizade e apoio sempre presente em todos os momentos.

Ao conselho do programa de Pós-graduação em Estatística e Experimentação Agronômica e os professores Dr. César Gonçalves de Lima, Dra. Clarice Garcia Borges Demétrio, Dr. Décio Barbin, Dr. Gerson Barreto Mourão, Dra. Roseli Aparecida Leandro, Dra. Sônia Maria De Stefano Piedade, pelos cuidados na formação.

Aos funcionários do Departamento de Ciências Exatas da ESALQ/USP, Solange de Assis Paes Sabadin e Eduardo Bonilha, pelos auxílios permanentes, em especial a Luciane Brajão pela amizade e ajuda sempre que preciso.

Ao colega de curso e terno amigo, Osmar Jesus Macedo, pela alegria de sua companhia e colaboração em tantos momentos. Um carinho todo especial estendido a sua esposa e filhos.

A todos os colegas de estudo pela convivência e respeito, em especial a Lucymary, Vanderly e Wilson que estiveram sempre presentes nos momentos alegres e também complicados da sala de estudos, bem como na ajuda da parte técnica do trabalho. A vocês, um enorme abraço.

Um agradecimento especial reservo as colegas Marina, Renata e Simone, pelo companheirismo, carinho e por várias vezes enxugarem minhas lágrimas com palavras de

amor e incentivo. Agora, não mais colegas, mas grandes amigas.

E não poderia deixar de agradecer imensamente aos novos e velhos amigos recém chegados, Marcelino Popó, Renato, Tiago e Patrícia, pelas risadas nos muitos momentos de descontração e que apesar do pouco tempo juntos, mas bem aproveitados, vocês vão ficar marcados para sempre em minha memória de uma maneira muito especial. Sucesso nessa nova caminhada.

A todos que cooperaram direta ou indiretamente na realização deste trabalho, muito obrigada.

SUMÁRIO

RESUMO	8
ABSTRACT	9
LISTA DE FIGURAS	10
LISTA DE TABELAS	16
1 INTRODUÇÃO	18
2 REVISÃO DE LITERATURA	20
2.1 A interação $G \times E$ e o enfoque genético	20
2.2 A interação $G \times E$ e o enfoque estatístico	24
2.3 Decomposição da interação $G \times E$	29
2.3.1 Decomposição da soma dos graus de liberdade da interação $G \times E$ mais os graus de liberdade do fator G	29
2.3.2 Decomposição da soma dos graus de liberdade da interação $G \times E$ mais os graus de liberdade do fator E	31
2.4 Autovalores e Autovetores de uma matriz e a Decomposição por Valores Singulares (DVS)	34
2.5 O modelo <i>AMMI</i>	35
2.5.1 O desdobramento da soma de quadrados da interação $G \times E$ por DVS	40
2.6 Representação gráfica em “ <i>biplot</i> ”	41
2.7 A seleção do modelo <i>AMMI</i>	44
2.7.1 Determinação do número ótimo de termos multiplicativos no modelo <i>AMMI</i> . .	44
2.7.2 Testes estatísticos para os componentes da soma de quadrados da interação $G \times E$	44
2.7.3 Avaliação preditiva por validação cruzada	50
2.7.4 Predição de respostas genotípicas por ambiente	51
2.8 Distribuição qui-quadrado não-central	53
3 MATERIAL E MÉTODOS	54
3.1 Características dos dados	54
3.2 Estudo de simulação	54
3.2.1 Análise de variância conjunta	56
3.3 Contribuição de cada genótipo e ambiente para a interação	57

3.4 Proposta de teste estatístico para a contribuição de genótipos e ambientes para a interação	58
4 RESULTADOS E DISCUSSÃO	62
4.1 Estudo de simulação	62
4.2 Aplicação em dados reais	93
4.2.1 Análise de variância conjunta	93
5 CONCLUSÕES	102
REFERÊNCIAS	103
ANEXOS	108

RESUMO

Teste estatístico para contribuição de genótipos e ambientes na matriz de interação GE

O presente trabalho teve por objetivos propor um método para testar a contribuição de cada genótipo e ambiente para a interação genótipos $\times$ ambientes em ensaios multi-ambientais através de um teste F e implementar uma rotina computacional para a realização da análise de dados segundo o teste proposto. O estudo avalia quatro conjuntos de dados, cada um com diferentes números de genótipos dentro de ambientes com quatro blocos. Para um dos conjuntos, simulou-se as somas de quadrados das linhas (genótipos) e colunas (ambientes) da matriz de interação genótipos $\times$ ambientes (**GE**) gerando 500, 5000 e 10000 experimentos para verificar a distribuição empírica. Os resultados indicaram um ajuste à distribuição qui-quadrado não-central para as linhas e colunas da matriz de interação **GE**, verificados também pelo teste de Kolmogorov-Smirnov e o gráfico QQplot. Na aplicação do teste F proposto aos quatro conjuntos de dados, identificou-se os genótipos e ambientes que contribuíram mais para a interação genótipos $\times$ ambientes. Dessa forma, os melhoristas podem selecionar bons genótipos e ambientes nos seus estudos.

Palavras-chaves: Interação genótipos $\times$ ambientes; Contribuição de genótipos e ambientes; Distribuição qui-quadrado não-central; Teste F modificado

ABSTRACT

Statistical test for contribution in the interaction matrix of genotypes and environments

The objective of the present work was to propose a method for testing the contribution of each element in a genotypes $\times$ environments interaction using multi-environment analyses by means of an F test and implementation of a computational routine to analyze the data according to the test proposed. The study evaluated four data sets, each with a different number of genotypes and environments, in a block design with four repetitions. In one group, the sum of squares within rows (genotypes) and columns (environments) of the genotypes $\times$ environments (**GE**) matrix was simulated, generating 500, 5000 and 10000 experiments to verify the empirical distribution. Results indicate a non-central chi-squared distribution for rows and columns of the **GE** interaction matrix, which was also verified by the Kolmogorov-Smirnov test and QQplot graph. Application of the F test to the four data sets identified the genotypes and environments that contributed the most to the genotypes $\times$ environments interaction. In this way, geneticists can select good genotypes and environments in their studies.

Keywords: Genotypes $\times$ environments interaction; Contribution of genotypes and environments; Non-central chi-squared distribution; Modified F test

LISTA DE FIGURAS

Figura 1 - Comportamento de dois genótipos (G_1 e G_2) em duas condições ambientais (A_1 e A_2) com ausência de interação	21
Figura 2 - Comportamento de dois genótipos (G_1 e G_2) em duas condições ambientais (A_1 e A_2) com interação simples ou quantitativa	22
Figura 3 - Comportamento de dois genótipos (G_1 e G_2) em duas condições ambientais (A_1 e A_2) com interação cruzada ou qualitativa	22
Figura 4 - Histogramas para as SQ dos genótipos 1, 2, 3 e 4 com 500 experimentos simulados . .	62
Figura 5 - Histogramas para as SQ dos genótipos 5, 6, 7 e 8 com 500 experimentos simulados . .	62
Figura 6 - Histogramas para as SQ dos genótipos 9, 10, 11 e 12 com 500 experimentos simulados	63
Figura 7 - Histogramas para as SQ dos genótipos 13, 14, 15 e 16 com 500 experimentos simulados	63
Figura 8 - Histogramas para as SQ dos genótipos 17, 18, 19 e 20 com 500 experimentos simulados	63
Figura 9 - Histogramas para as SQ dos genótipos 1, 2, 3 e 4 com 5000 experimentos simulados . .	64
Figura 10 -Histogramas para as SQ dos genótipos 5, 6, 7 e 8 com 5000 experimentos simulados . .	64
Figura 11 -Histogramas para as SQ dos genótipos 9, 10, 11 e 12 com 5000 experimentos simulados	64
Figura 12 -Histogramas para as SQ dos genótipos 13, 14, 15 e 16 com 5000 experimentos simulados	65
Figura 13 -Histogramas para as SQ dos genótipos 17, 18, 19 e 20 com 5000 experimentos simulados	65
Figura 14 -Histogramas para as SQ dos genótipos 1, 2, 3 e 4 com 10000 experimentos simulados .	65
Figura 15 -Histogramas para as SQ dos genótipos 5, 6, 7 e 8 com 10000 experimentos simulados .	66
Figura 16 -Histogramas para as SQ dos genótipos 9, 10, 11 e 12 com 10000 experimentos simulados	66
Figura 17 -Histogramas para as SQ dos genótipos 13, 14, 15 e 16 com 10000 experimentos simulados	66
Figura 18 -Histogramas para as SQ dos genótipos 17, 18, 19 e 20 com 10000 experimentos simulados	67

	11
Figura 19 -Histogramas para as SQ dos ambientes 1, 2, 3 e 4 com 500 experimentos simulados . .	67
Figura 20 -Histogramas para as SQ dos ambientes 5, 6, 7 e 8 com 500 experimentos simulados . .	68
Figura 21 -Histogramas para as SQ dos ambientes 9, 10, 11 e 12 com 500 experimentos simulados	68
Figura 22 -Histogramas para as SQ dos ambientes 13, 14, 15 e 16 com 500 experimentos simulados	68
Figura 23 -Histogramas para as SQ dos ambientes 17, 18, 19 e 20 com 500 experimentos simulados	69
Figura 24 -Histogramas para as SQ dos ambientes 21, 22, 23 e 24 com 500 experimentos simulados	69
Figura 25 -Histogramas para as SQ dos ambientes 25, 26, 27 e 28 com 500 experimentos simulados	69
Figura 26 -Histogramas para as SQ dos ambientes 29, 30, 31 e 32 com 500 experimentos simulados	70
Figura 27 -Histogramas para as SQ dos ambientes 33 e 34 com 500 experimentos simulados . . .	70
Figura 28 -Histogramas para as SQ dos ambientes 1, 2, 3 e 4 com 5000 experimentos simulados .	70
Figura 29 -Histogramas para as SQ dos ambientes 5, 6, 7 e 8 com 5000 experimentos simulados .	71
Figura 30 -Histogramas para as SQ dos ambientes 9, 10, 11 e 12 com 5000 experimentos simulados	71
Figura 31 -Histogramas para as SQ dos ambientes 13, 14, 15 e 16 com 5000 experimentos simulados	71
Figura 32 -Histogramas para as SQ dos ambientes 17, 18, 19 e 20 com 5000 experimentos simulados	72
Figura 33 -Histogramas para as SQ dos ambientes 21, 22, 23 e 24 com 5000 experimentos simulados	72
Figura 34 -Histogramas para as SQ dos ambientes 25, 26, 27 e 28 com 5000 experimentos simulados	73
Figura 35 -Histogramas para as SQ dos ambientes 29, 30, 31 e 32 com 5000 experimentos simulados	73
Figura 36 -Histogramas para as SQ dos ambientes 33 e 34 com 5000 experimentos simulados . . .	73
Figura 37 -Histogramas para as SQ dos ambientes 1, 2, 3 e 4 com 10000 experimentos simulados .	74

	12
Figura 38 -Histogramas para as SQ dos ambientes 5, 6, 7 e 8 com 10000 experimentos simulados .	74
Figura 39 -Histogramas para as SQ dos ambientes 9, 10, 11 e 12 com 10000 experimentos simulados	74
Figura 40 -Histogramas para as SQ dos ambientes 13, 14, 15 e 16 com 10000 experimentos simulados	75
Figura 41 -Histogramas para as SQ dos ambientes 17, 18, 19 e 20 com 10000 experimentos simulados	75
Figura 42 -Histogramas para as SQ dos ambientes 21, 22, 23 e 24 com 10000 experimentos simulados	75
Figura 43 -Histogramas para as SQ dos ambientes 25, 26, 27 e 28 com 10000 experimentos simulados	76
Figura 44 -Histogramas para as SQ dos ambientes 29, 30, 31 e 32 com 10000 experimentos simulados	76
Figura 45 -Histogramas para as SQ dos ambientes 33 e 34 com 10000 experimentos simulados . .	76
Figura 46 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 1, 2, 3 e 4 com 500 experimentos simulados	79
Figura 47 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 5, 6, 7 e 8 com 500 experimentos simulados	79
Figura 48 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 9, 10, 11 e 12 com 500 experimentos simulados	80
Figura 49 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 13, 14, 15 e 16 com 500 experimentos simulados	80
Figura 50 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 17, 18, 19 e 20 com 500 experimentos simulados	80
Figura 51 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 1, 2, 3 e 4 com 5000 experimentos simulados	81
Figura 52 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 5, 6, 7 e 8 com 5000 experimentos simulados	81
Figura 53 -Gráficos QQplot qui-quadrado para as SQ dos genótipos 9, 10, 11 e 12 com 5000 experimentos simulados	81

Figura 54 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 13, 14, 15 e 16 com 5000 experimentos simulados	82
Figura 55 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 17, 18, 19 e 20 com 5000 experimentos simulados	82
Figura 56 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 1, 2, 3 e 4 com 10000 experimentos simulados	82
Figura 57 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 5, 6, 7 e 8 com 10000 experimentos simulados	83
Figura 58 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 9, 10, 11 e 12 com 10000 experimentos simulados	83
Figura 59 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 13, 14, 15 e 16 com 10000 experimentos simulados	83
Figura 60 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 17, 18, 19 e 20 com 10000 experimentos simulados	84
Figura 61 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 1, 2, 3 e 4 com 500 experimentos simulados	84
Figura 62 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 5, 6, 7 e 8 com 500 experimentos simulados	84
Figura 63 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 9, 10, 11 e 12 com 500 experimentos simulados	85
Figura 64 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 13, 14, 15 e 16 com 500 experimentos simulados	85
Figura 65 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 17, 18, 19 e 20 com 500 experimentos simulados	85
Figura 66 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 21, 22, 23 e 24 com 500 experimentos simulados	86
Figura 67 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 25, 26, 27 e 28 com 500 experimentos simulados	86
Figura 68 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 29, 30, 31 e 32 com 500 experimentos simulados	86

Figura 69 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 33 e 34 com 500 experimentos simulados	87
Figura 70 - Gráficos QQplot qui-quadrado as SQ dos ambientes 1, 2, 3 e 4 com 5000 experimentos simulados	87
Figura 71 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 5, 6, 7 e 8 com 5000 experimentos simulados	87
Figura 72 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 9, 10, 11 e 12 com 5000 experimentos simulados	88
Figura 73 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 13, 14, 15 e 16 com 5000 experimentos simulados	88
Figura 74 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 17, 18, 19 e 20 com 5000 experimentos simulados	88
Figura 75 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 21, 22, 23 e 24 com 5000 experimentos simulados	89
Figura 76 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 25, 26, 27 e 28 com 5000 experimentos simulados	89
Figura 77 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 29, 30, 31 e 32 com 5000 experimentos simulados	89
Figura 78 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 33 e 34 com 5000 experimentos simulados	90
Figura 79 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 1, 2, 3 e 4 com 10000 experimentos simulados	90
Figura 80 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 5, 6, 7 e 8 com 10000 experimentos simulados	90
Figura 81 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 9, 10, 11 e 12 com 10000 experimentos simulados	91
Figura 82 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 13, 14, 15 e 16 com 10000 experimentos simulados	91
Figura 83 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 17, 18, 19 e 20 com 10000 experimentos simulados	91

Figura 84 -Gráficos QQplot qui-quadrado para as SQ dos ambientes 21, 22, 23 e 24 com 10000	
experimentos simulados	92
Figura 85 -Gráficos QQplot qui-quadrado para as SQ dos ambientes 25, 26, 27 e 28 com 10000	
experimentos simulados	92
Figura 86 -Gráficos QQplot qui-quadrado para as SQ dos ambientes 29, 30, 31 e 32 com 10000	
experimentos simulados	92
Figura 87 -Gráficos QQplot qui-quadrado para as SQ dos ambientes 33 e 34 com 10000 experimentos	
simulados	93

LISTA DE TABELAS

Tabela 1 - Representação dos dados médios de g genótipos avaliados em e ambientes para um caráter genérico Y	25
Tabela 2 - Esquema da ANOVA com teste F obtida pela decomposição da interação $G \times E$ mais o fator G	31
Tabela 3 - Esquema da ANOVA com teste F obtida pela decomposição da interação $G \times E$ mais o fator E	34
Tabela 4 - Análise de variância conjunta obtida pelo sistema de Gollob	46
Tabela 5 - Esquema da ANOVA para experimentos de um mesmo grupo de genótipos avaliado em e ambientes com k blocos	57
Tabela 6 - Esquema da ANOVA com teste F para obtenção de genótipos que contribuem significativamente para a interação $G \times E$	59
Tabela 7 - Esquema da ANOVA com teste F para obtenção de ambientes que contribuem significativamente para a interação $G \times E$	61
Tabela 8 - Estimativas dos parâmetros de não-centralidade ($\hat{\theta}$) e valor- p do teste de Kolmogorov-Smirnov para aderência à distribuição qui-quadrado não-central da soma de quadrados dos genótipos para 500, 5000 e 10000 experimentos simulados	77
Tabela 9 - Estimativas dos parâmetros de não-centralidade ($\hat{\theta}$) e valor- p do teste de Kolmogorov-Smirnov para aderência à distribuição qui-quadrado não-central da soma de quadrados dos ambientes para 500, 5000 e 10000 experimentos simulados	78
Tabela 10 - ANOVA do Conjunto 1 com 20 genótipos de trigo avaliados em 34 ambientes com 4 blocos	93
Tabela 11 - Teste F , aplicado ao Conjunto 1, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$	94
Tabela 12 - Teste F , aplicado ao Conjunto 1, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$	95
Tabela 13 - ANOVA do Conjunto 2 com 16 genótipos de milho avaliados em 24 ambientes com 4 blocos	96

Tabela 14 -Teste F , aplicado ao Conjunto 2, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$	96
Tabela 15 -Teste F , aplicado ao Conjunto 2, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$	97
Tabela 16 -ANOVA do Conjunto 3 com 9 genótipos de milho avaliados em 20 ambientes com 4 blocos	98
Tabela 17 -Teste F , aplicado ao Conjunto 3, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$	98
Tabela 18 -Teste F , aplicado ao Conjunto 3, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$	99
Tabela 19 -ANOVA do Conjunto 4 com 8 genótipos de milho avaliados em 59 ambientes com 4 blocos	99
Tabela 20 -Teste F , aplicado ao Conjunto 4, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$	100
Tabela 21 -Teste F , aplicado ao Conjunto 4, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$	100

1 INTRODUÇÃO

Um dos princípios básicos no estudo da herança de caracteres métricos é o de que o valor de um indivíduo (fenótipo) é determinado por dois componentes básicos: a composição genética (genótipo) e uma contribuição do ambiente específico em que aquele indivíduo se encontra (CHAVES, 2001).

Na realização de um experimento, quando o mesmo grupo de genótipos é avaliado em diferentes ambientes, geralmente ocorre, de o ambiente alterar diferentemente o mesmo caráter em diferentes genótipos, ou seja, pode ocorrer uma interação entre fatores. Assim, a interação entre genótipos e ambientes (genótipos $\times$ ambientes) é resultante da resposta diferencial de cada genótipo à variação ambiental.

A interação genótipos $\times$ ambientes é uma das principais dificuldades encontradas pelos melhoristas durante sua atividade seletiva dos melhores genótipos. Por outro lado, se não houvesse interação genótipos $\times$ ambientes, uma variedade adaptar-se-ia à maioria dos ambientes de cultivo (locais, anos ou estações), de maneira que um único ensaio seria suficiente para fornecer resultados universais.

A presença da interação é essencial para a eficiência do processo podendo afetar positiva ou negativamente a sua finalização. O problema resulta da deficiência dos genótipos manterem um padrão consistente de desempenho nos ambientes. Logo, a interação não é apenas um problema, mas também uma oportunidade a ser aproveitada.

Várias metodologias têm sido propostas para o estudo da interação, bem como, na seleção de genótipos consistentes e de elevada produtividade. Entre os métodos tradicionais, encontram-se o zoneamento agrônômico que permite recomendações regionalizadas e métodos de regressão linear simples e múltipla que identificam os genótipos que contribuem menos para a interação, podendo ser recomendados de forma generalizada (PIOVESAN, 2007). Entretanto, as metodologias baseadas em regressão linear, como proposta por Eberhart e Russell (1966), possuem limitações e têm sido muito criticada quanto à falta de linearidade.

Na busca por novas metodologias de exploração da interação genótipos $\times$ ambientes, as técnicas multivariadas vêm ganhando muita importância e aplicabilidade nos últimos

anos. Dentre essas técnicas, recomenda-se a análise de componentes principais (ACP), a análise de agrupamentos e a análise dos modelos *AMMI*.

A análise dos modelos *AMMI* combina, num único modelo, componentes aditivos para os efeitos principais (genótipos e ambientes) e componentes multiplicativos para os efeitos da interação genótipos $\times$ ambientes. Gauch Jr e Zobel (1996) informam que essa análise pode ajudar tanto na identificação de genótipos de alta produtividade e largamente adaptados, como na realização do zoneamento agrônômico. Além disso, o método *AMMI* permite uma análise mais detalhada da interação genótipos $\times$ ambientes, 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.

A interação genótipos $\times$ ambientes é representada pelos termos multiplicativos do modelo *AMMI*. Por essa razão, um objetivo inicial é determinar o número ótimo de termos multiplicativos a serem retidos no modelo *AMMI* com a finalidade de descrever o padrão relacionado à interação. A seleção do número ótimo desses termos pode ser feita através de testes de hipóteses como os apresentados por Gollob (1968), Mandel (1971), Cornelius (1993), Johnson e Graybill (1972) entre outros, ou através de validação cruzada como apresentada por Dias e Krzanowski (2003).

Os melhoristas de plantas compreendem que a interação genótipos $\times$ ambientes é de suma importância para a obtenção de variedades superiores. Entretanto, há a necessidade de entender a contribuição de cada genótipo e ambiente para a interação, uma vez que as avaliações do comportamento diferencial dos genótipos em função da diversidade ambiental podem ocasionar mudanças na posição relativa dos genótipos como, por exemplo, a produtividade, ou até mesmo na magnitude das suas diferenças.

Os objetivos do presente trabalho são propor um método para testar a contribuição de cada genótipo e ambiente para a interação genótipos $\times$ ambientes através de um teste F e implementar uma rotina computacional para a realização da análise de dados, mediante a contribuição de cada genótipo e ambiente para a interação.

A proposta de um teste F como foco de análise neste trabalho, deve-se a importância por parte dos melhoristas em conhecer a contribuição de cada genótipo e ambiente para a interação, contribuindo para a seleção de bons genótipos e ambientes nos seus estudos.

2 REVISÃO DE LITERATURA

2.1 A interação $G \times E$ e o enfoque genético

Em experimentação agronômica, é freqüente a condução de experimentos em locais representativos de uma determinada região ou país, sendo repetidos um número suficiente de vezes de forma a representar as alternativas do clima de toda a região ou país. Por exemplo, a EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária) possui programas para ensaios nacionais de culturas como a do milho.

No processo de melhoramento genético, a influência que o ambiente exerce sobre cada genótipo tem merecido atenção especial por parte dos melhoristas de plantas, visto que as interações entre genótipos e ambientes podem interferir positiva ou negativamente nos resultados finais (GONZALEZ, 1988).

A interação de genótipos com ambientes, segundo Chaves (2001), resulta 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 ferramenta estatístico juntamente com conhecimento do fenômeno ao nível biológico que se pode tirar melhores informações para entender a complexidade da interação.

Uma população que ocupa uma área geográfica com grandes variedades ambientais, estará sujeita a pressões de seleção diferentes, ao longo de sua extensão espacial. Ocorrendo variabilidade de adaptação dos indivíduos, serão beneficiados aqueles indivíduos mais adaptados na condição ambiental particular. Esta adaptação específica a determinados ambientes é a causa evolutiva básica da interação de genótipos e ambientes.

Foram ilustrados por Allard e Bradshaw (1964) e ampliados aqui, alguns tipos de interações em que são considerados dois genótipos em dois ambientes. Nas Figuras 1, 2 e 3 são apresentadas três situações básicas que trazem diferentes conseqüências para o melhoramento. Em dois ambientes A_1 e A_2 são avaliados dois genótipos G_1 e G_2 . Assume-se que o ambiente A_1 é mais favorável para a manifestação do caráter genérico Y, que pode ser qualquer variável métrica. Na Figura 1 há ausência de interação, ou seja, a mudança das condições ambientais afeta igualmente o comportamento dos genótipos e a diferença entre eles permanece constante nos dois ambientes. Observa-se nas Figuras 1c, 1d e 1e que os

genótipos podem coincidir e portanto, terem o mesmo comportamento nos dois ambientes ou apresentarem diferenças de um ambiente para outro, mas com ausência da interação. Nas Figuras 1 e 2, exceto Figura 1c, a mudança de ambiente afeta desigualmente a manifestação do caráter para os dois genótipos, ou seja, a diferença entre os genótipos varia entre ambientes. Na Figura 2, o efeito de cada ambiente não modifica a classificação dos genótipos, sendo o G_1 superior a G_2 em todas condições. Neste caso a interação é denominada simples ou quantitativa. A Figura 3, apresenta uma mudança na classificação dos genótipos. Para este tipo, a interação é denominada cruzada ou qualitativa.

Figura 1 - Comportamento de dois genótipos (G_1 e G_2) em duas condições ambientais (A_1 e A_2) com ausência de interação

Nas Figuras 1 e 2, as implicações para o melhoramento é que um mesmo genótipo G_1 é melhor adaptado às duas condições ambientais e uma seleção baseada na média dos ambientes beneficiará sempre o melhor genótipo. Na Figura 3, a seleção baseada na média dos ambientes não é capaz de satisfazer o conjunto dos ambientes podendo levar a uma seleção de genótipos mal adaptados a uma situação particular. Nas Figuras 1b e

2b ocorrem sinergismos e nas Figuras 1a e 2a antagonismo entre genótipos e ambientes, ao passar do ambiente A_1 para o ambiente A_2 .

Figura 2 - Comportamento de dois genótipos (G_1 e G_2) em duas condições ambientais (A_1 e A_2) com interação simples ou quantitativa

Figura 3 - Comportamento de dois genótipos (G_1 e G_2) em duas condições ambientais (A_1 e A_2) com interação cruzada ou qualitativa

Para as três figuras anteriores o papel de G_1 e G_2 pode ser permutado com A_1 e A_2 , obtendo-se interpretações semelhantes. Quando se consideram vários genótipos avaliados em vários ambientes, a combinação de situações como as das Figuras 1, 2 e 3 formam um emaranhado de situações, difícil de ser interpretado, exigindo métodos adequados de análise da interação genótipos $\times$ ambientes.

As variáveis ambientais também podem ser classificadas em dois tipos: previsíveis e imprevisíveis (ALLARD e BRADSHAW, 1964). As variáveis previsíveis seriam as

características gerais do clima e solos que ocorrem de maneira sistemática ou que estão sob controle do homem. Já as variáveis imprevisíveis correspondem às flutuações climáticas tais como quantidade e distribuição de chuvas, temperatura e outros fatores que não podem ser controlados pelo homem.

A interação genótipos $\times$ ambientes é uma fonte de variação fenotípica, que na maioria dos casos é inseparável da variância ambiental (FALCONER, 1987). Na prática, para verificar a significância da interação de genótipos com ambientes, é necessário repetir o experimento várias vezes, pois se o experimento for realizado somente em um ambiente, poderá ocorrer uma superestimação dos ganhos genéticos (CROSSA, 1990).

Existe uma concordância geral entre melhoristas de plantas de que interação genótipos $\times$ ambientes têm um importante significado para a obtenção de variedades superiores. Entretanto, de acordo com Allard (1971) é muito difícil encontrar concordâncias sobre o que se deve conhecer em relação à interação genótipos $\times$ ambientes e como utilizá-la.

A natureza da interação genótipos $\times$ ambientes também deve ser considerada e não somente a verificação de sua existência (VENCOVSKY e BARRIGA, 1992). Assim a natureza pode ser simples e complexa. A interação de natureza simples indica a presença de genótipos adaptados em um grande número de ambientes, sendo possível fazer uma recomendação generalizada de cultivares. A interação com natureza complexa mostra que existem genótipos adaptados a apenas alguns ambientes, o que traz uma complicação ao pesquisador, quando da recomendação de cultivar.

A existência de interação genótipos $\times$ ambientes, produz uma barreira de dificuldades aos melhoristas na identificação de genótipos superiores, tanto no processo de seleção quanto no processo de recomendação de cultivares. Essa interação indica que o comportamento dos genótipos nos experimentos depende principalmente das condições ambientais a que são submetidos. Assim a resposta obtida de um genótipo, em comparação a outro, é variável, sendo que essas variações se apresentam devido a mudança de ambientes (OLIVEIRA; DUARTE; PINHEIRO, 2003; KANG, 1998).

Assim, a interação de genótipos $\times$ ambientes deve ser considerada, não como um problema ou um fator indesejável, cujo efeito deve ser minimizado, mas deve ser enfrentada como um fenômeno biológico natural, que deve ser bem conhecido para melhor aproveitá-lo no processo de seleção (CHAVES, 2001). Logo os genótipos que interagem po-

sitivamente com os ambientes podem fazer a diferença entre um bom e um ótimo cultivar (DUARTE; VENCovsky, 1999).

2.2 A interação $G \times E$ e o enfoque estatístico

A existência de interação genótipos $\times$ ambientes têm sido reconhecida há longo tempo de acordo com Freeman (1973), sendo a referência mais antiga feita por Fisher e Mackenzie em 1923, a qual precede a análise de variância conjunta (ANOVA). Desde então, muitos trabalhos têm sido feitos para análises estatísticas da interação genótipos $\times$ ambientes, seja por estatísticos, agrônomos, melhoristas e geneticistas.

Os métodos mais simples que utilizam componentes de variância para a interpretação dos resultados quando existem interações, incluem estudos de regressão, métodos baseados em análises modificadas e métodos envolvendo variáveis ambientais externas.

Estudos realizados sobre interação genótipos $\times$ ambientes mostraram que a regressão foi utilizada pela primeira vez por Yates e Cochran (1938), analisando grupos de experimentos. Neste caso, o grau de associação entre diferenças varietais pôde ser verificado pelo cálculo da regressão dos rendimentos das variedades isoladas sobre os rendimentos médios de todas as variedades, mostrando assim que a regressão explicava grande parte da interação em uma série de experimentos de cevada. O método de regressão foi também usado por Perkins e Jinks (1968) para estimação de parâmetros em um modelo genético aplicado à biometria.

Outro método usado é aquele no qual os dados são organizados em uma tabela de dupla entrada em que o processo de investigação da interação é feito através da ANOVA. Esta análise envolve vários experimentos, e por isso é possível determinar a magnitude da interação, através da razão do quadrado médio da interação ($QM_{G \times E}$) pelo quadrado médio do resíduo (QM_{Res}). A detecção de significância para a interação não esclarece, contudo, as implicações que estas possam ter sobre o melhoramento, de forma que, estudos de detalhamento deste componente da variação são em geral necessários.

Supondo-se que os genótipos foram avaliados nos diversos ambientes com r repetições em delineamento experimental que permita a estimativa da variação residual em cada ambiente, o modelo mais simples e comum para a análise estatística de um conjunto de dados é dado por (ANOVA para grupos de experimentos):

$$Y_{ij} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ij} \quad (1)$$

sendo que:

Y_{ij} : é a resposta média do i -ésimo genótipo no j -ésimo ambiente;

μ : é uma constante comum às respostas (normalmente a média geral);

g_i : é o efeito do i -ésimo genótipo ($i = 1, 2, \dots, g$);

e_j : é o efeito do j -ésimo ambiente ($j = 1, 2, \dots, e$);

$(ge)_{ij}$: é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente; e

ε_{ij} : é o erro experimental médio, assumido independente e $\varepsilon_{ij} \sim N(0, \sigma^2/r)$.

A forma básica de organização dos dados para análise da interação genótipos $\times$ ambientes está representada na Tabela 1, na qual estão representadas as médias dos g tratamentos (genótipos) sobre r repetições, derivadas de um experimento em delineamento apropriado. Quando todos os genótipos são avaliados em todos os ambientes, com o mesmo número de repetições por experimento, diz-se que os dados são balanceados.

Tabela 1 - Representação dos dados médios de g genótipos avaliados em e ambientes para um caráter genérico Y

Genótipos	Ambientes				Média ($\bar{Y}_{i.}$)
	1	2	$\dots$	e	
1	Y_{11}	Y_{12}	$\dots$	Y_{1e}	$\bar{Y}_{1.}$
2	Y_{21}	Y_{22}	$\dots$	Y_{2e}	$\bar{Y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
g	Y_{g1}	Y_{g2}	$\dots$	Y_{ge}	$\bar{Y}_{g.}$
Médias ($\bar{Y}_{.j}$)	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	$\dots$	$\bar{Y}_{.e}$	$\bar{Y}_{..}$

A análise de variância feita sob o modelo (1) com base nos dados da Tabela 1 é bastante simples, sendo calculados os efeitos principais pelas marginais da Tabela 1 e a interação como desvios do modelo, em que $\hat{ge}_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$ são os correspondentes elementos da Tabela 1 para o estudo da interação. O erro experimental ε_{ij} é calculado para cada experimento, utilizando-se o resíduo na análise conjunta.

Na solução do modelo (1) visando encontrar os estimadores dos parâmetros, pelo processo de mínimos quadrados, utiliza-se o sistema de equações normais. O modelo matricial é da seguinte forma:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{pmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1e} \\ y_{21} \\ y_{22} \\ \vdots \\ y_{2e} \\ \vdots \\ y_{g1} \\ y_{g2} \\ \vdots \\ y_{ge} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \mu \\ g_1 \\ g_2 \\ \vdots \\ g_g \\ e_1 \\ e_2 \\ \vdots \\ e_e \\ (ge)_{11} \\ (ge)_{12} \\ \vdots \\ (ge)_{ge} \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1e} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2e} \\ \vdots \\ \varepsilon_{g1} \\ \varepsilon_{g2} \\ \vdots \\ \varepsilon_{ge} \end{pmatrix}$$

Têm-se que a matrix $\mathbf{X}$ não é de posto coluna completo e, portanto $(\mathbf{X}'\mathbf{X})$ não possui inversa. Uma possível solução paramétrica do sistema de equações normais $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$ é dada por:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$$

em que $(\mathbf{X}'\mathbf{X})^{-}$ é uma inversa generalizada, uma vez que o sistema é indeterminado, admitindo-se as condições marginais:

$$\sum_{i=1}^g g_i = \sum_{j=1}^e e_j = \sum_{i=1}^g (ge)_{ij} = \sum_{j=1}^e (ge)_{ij} = \sum_{i=1}^g \sum_{j=1}^e (ge)_{ij} = 0 \quad (2)$$

Assim, a solução do sistema de equações normais é:

$$\hat{\mu} = \frac{\sum_{ij} Y_{ij}}{ge} = \bar{Y}_{..};$$

$$\hat{g}_i = \frac{\sum_j Y_{ij}}{e} - \hat{\mu} = \bar{Y}_{i.} - \bar{Y}_{..};$$

$$\hat{e}_j = \frac{\sum_i Y_{ij}}{g} - \hat{\mu} = \bar{Y}_{.j} - \bar{Y}_{..};$$

$$(\hat{ge})_{ij} = Y_{ij} - \frac{\sum_j Y_{ij}}{e} - \frac{\sum_i Y_{ij}}{g} + \frac{\sum_{ij} Y_{ij}}{ge} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}.$$

Esses são estimadores não viesados dos parâmetros μ , g_i , e_j e $(ge)_{ij}$, sujeitos às condições marginais dadas acima. Se essas ou quaisquer outras condições não forem impostas sobre os parâmetros, então os estimadores dos parâmetros individuais são viesados.

A aproximação de mínimos quadrados $\hat{Y}_{ij}$ (ou valor predito) e seu respectivo resíduo, correspondente ao termo geral de interação $(\hat{ge})_{ij}$, são dados por:

$$\hat{Y}_{ij} = \frac{\sum_j Y_{ij}}{e} + \frac{\sum_i Y_{ij}}{g} - \frac{\sum_{ij} Y_{ij}}{ge} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..}$$

$$\hat{e}_{ij} = Y_{ij} - \hat{Y}_{ij}$$

Agora, pode-se construir a matriz de interações $\mathbf{GE}_{(g \times e)} = (\hat{ge}_{ij})$, como segue:

$$\mathbf{GE}_{(g \times e)} = \begin{pmatrix} \hat{ge}_{11} & \hat{ge}_{12} & \cdots & \hat{ge}_{1e} \\ \hat{ge}_{21} & \hat{ge}_{22} & \cdots & \hat{ge}_{2e} \\ \cdots & \cdots & \cdots & \cdots \\ \hat{ge}_{g1} & \hat{ge}_{g2} & \cdots & \hat{ge}_{ge} \end{pmatrix}$$

As somas de quadrados devido aos efeitos principais e interação, referentes aos dados de médias como na Tabela 1, são obtidas por:

$$SQ_{Total} = \sum_{i=1}^g \sum_{j=1}^e Y_{ij}^2 - ge\bar{Y}_{..}^2, \text{ com } ge - 1 \text{ graus de liberdade;}$$

$$SQ_G = e \sum_{i=1}^g \bar{Y}_{i.}^2 - ge\bar{Y}_{..}^2, \text{ com } g - 1 \text{ graus de liberdade;}$$

$$SQ_E = g \sum_{j=1}^e \bar{Y}_{.j}^2 - ge\bar{Y}_{..}^2, \text{ com } e - 1 \text{ graus de liberdade;}$$

$$SQ_{G \times E} = SQ_{Total} - SQ_G - SQ_E, \text{ com } (g - 1)(e - 1) \text{ graus de liberdade.}$$

Os quadrados médios correspondentes são obtidos dividindo-se cada soma de quadrados pelos respectivos graus de liberdade. O quadrado médio do resíduo ou do erro médio (QM_{Res}) é calculado pela média ponderada dos QM 's do resíduo (obtido nas ANOVA's

individuais de experimentos), assumidos homogêneos, usando seus respectivos graus de liberdade como pesos, ou seja:

$$QM_{Res} = \frac{\sum_j SQ_{Resj}}{\sum_j GL_{Resj}},$$

em que $j = 1, 2, \dots, e$ ambientes ou experimentos.

Uma consideração importante a ser feita para realização do teste F e interpretação dos resultados da análise conjunta da variância, diz respeito à natureza fixa ou aleatória dos efeitos do modelo de análise (CHAVES, 2001). A natureza fixa ou aleatória da interação é determinada pelos efeitos principais. Se genótipos e ambientes são fixos, a interação será fixa. Se pelo menos um dos fatores for aleatório, a interação será aleatória. Na análise de ensaios multi-ambientais, consideram-se, em geral, os efeitos de genótipos como fixos e os efeitos de ambientes como aleatórios, de tal forma que o efeito da interação genótipos $\times$ ambientes é aleatória nesse caso.

O estudo da interação genótipos $\times$ ambientes possibilita a identificação de cultivares mais adaptados a determinadas regiões, onde as mesmas poderão expressar o seu potencial genético. Assim, estudos sobre a magnitude de tais interações podem ser úteis na regionalização de cultivares, objetivando indicar áreas onde as mesmas possam expressar o máximo que as condições ambientais particulares permitam, com respeito a respostas de genótipos, e possibilitar a exploração de efeitos específicos de adaptação para determinadas regiões.

No que diz respeito a adaptabilidade e estabilidade de cada genótipo, tais fenômenos não devem ser considerados iguais, apesar de estarem relacionados entre si. A adaptabilidade refere-se à capacidade de os genótipos aproveitarem vantajosamente o estímulo do ambiente e a estabilidade diz respeito à capacidade de os genótipos mostrarem comportamento altamente previsível em razão do estímulo do ambiente.

O conceito de estabilidade foi classificado por Lin, Binns e Lefkovicth (1986) em três tipos. No tipo 1, o genótipo é considerado estável se sua variância entre os ambientes for pequena. Ela pode ser medida pela variância de cada genótipo nos diferentes ambientes e é útil para características que devem ser mantidas, tal como resistência a patógenos e pragas (NUNES, 2000). Na estabilidade tipo 2, o genótipo é considerado estável se sua resposta ao ambiente for paralela à resposta média de todos os materiais avaliados no experimento, o que

ocorre quando o genótipo possui interações mínimas com o ambiente. A estabilidade do tipo 3 é aquela na qual o genótipo seria considerado estável se o quadrado médio dos desvios de regressão for pequeno, classificando-o como de alta confiabilidade de resposta.

2.3 Decomposição da interação $G \times E$

O efeito da interação genótipos $\times$ ambientes visa verificar, estatisticamente, o efeito do genótipo na presença dos ambientes ou o efeito do ambiente na presença dos genótipos. Quando se verifica de forma significativa evidências na presença da interação do genótipo com o ambiente, conclui-se que há um efeito diferenciado do genótipo na presença de pelo menos um ambiente ou vice-versa. Neste caso, o procedimento adequado para interpretar o efeito da interação $G \times E$ é através de sua decomposição, que consiste em estudar o efeito de cada fator na presença de cada nível do outro fator.

Quando o efeito da interação é significativo e, também, verifica-se significância para algum efeito principal, deve-se interpretar e concluir somente os resultados referentes ao efeito da interação, ignorando os resultados obtidos para os efeitos principais. Este fato é justificado através das hipóteses testadas para os efeitos principais, que sob a condição de interação significativa perde o seu significado, uma vez que a esperança de Y_{ij} fica $\mu + g_i + e_j + (ge)_{ij}$.

De acordo com o objetivo da pesquisa, o método tradicional de decomposição da interação $G \times E$ em relação a cada efeito será descrito a seguir (BARBIN, 2003; PIMENTEL-GOMES, 2000).

2.3.1 Decomposição da soma dos graus de liberdade da interação $G \times E$ mais os graus de liberdade do fator G

Essa decomposição corresponde a obter o efeito do fator G (genótipo) dentro do fator E (ambiente), representado por $G(E)$. Suponha que o delineamento seja aleatorizado em blocos. Assim, o modelo adotado é:

$$Y_{ijk} = \mu + b_{k(j)} + g_{i(j)} + e_j + \varepsilon_{ijk}$$

em que:

Y_{ijk} : é o valor observado do i -ésimo genótipo no j -ésimo ambiente e no k -ésimo bloco, com $i = 1, 2, \dots, g$, $j = 1, 2, \dots, e$ e $k = 1, 2, \dots, r$;

μ : é uma constante comum às respostas (normalmente a média geral);

$b_{k(j)}$: é o efeito do k -ésimo bloco dentro do j -ésimo ambiente;

$g_{i(j)}$: é o efeito do i -ésimo genótipo na presença do j -ésimo ambiente;

e_j : é o efeito do j -ésimo ambiente; e

ε_{ijk} : é erro experimental associado ao i -ésimo genótipo, no j -ésimo ambiente e no k -ésimo bloco assumido ser independente e $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

O $g_{i(j)}$ é definido como:

$$g_{i(j)} = \mu_{i(j)} - \mu_{.j}, \text{ tal que } \sum_{i=1}^g g_{i(j)} = 0, \text{ para todo } j.$$

Sendo que $\mu_{i(j)}$ refere-se à média populacional do i -ésimo genótipo na presença do j -ésimo ambiente e $\mu_{.j}$ refere-se à média populacional do j -ésimo ambiente.

As hipóteses testadas são:

- i) $H_{01} : e_j = 0$, para todo ambiente j
 $H_{a1} : e_j \neq 0$, para algum ambientes j

ou ainda, H_{01} e H_{a1} podem ser escritas da seguinte maneira:

$$\begin{aligned} H_{01} & : \mu_{.1} = \dots = \mu_{.e} \\ H_{a1} & : \text{ pelo menos } \mu_{.j} \neq \mu_{.j'}, \text{ para } j \neq j' \text{ (} j' = 1, \dots, e \text{)} \end{aligned}$$

- ii) $H_{02} : g_{i(j)} = 0$, para todo genótipo i na presença do ambiente j
 $H_{a2} : g_{i(j)} \neq 0$, para algum genótipo i na presença do ambiente j

ou ainda, H_{02} e H_{a2} podem ser escritas da seguinte maneira:

$$\begin{aligned} H_{02} & : \mu_{1(j)} = \dots = \mu_{g(j)}, \text{ na presença do ambiente } j \\ H_{a2} & : \text{ pelo menos } \mu_{i(j)} \neq \mu_{i'(j)}, \text{ para genótipo } i \neq i' \text{ na presença do ambiente } j. \end{aligned}$$

A soma de quadrados do genótipo dentro do ambiente é dada por:

$$SQ_{G(E_j)} = \frac{1}{k} \sum_{i=1}^g y_{i(j)}^2 - \frac{y_{.j}^2}{gk}, \text{ com } j = 1, \dots, e \text{ e } (g-1) \text{ graus de liberdade associados,}$$

sendo que

$$y_{i(j).} = \sum_{k=1}^r y_{i(j)k} \text{ o total do } i\text{-ésimo genótipo dentro do } j\text{-ésimo ambiente.}$$

$$y_{.j.} = \sum_{i=1}^g \sum_{k=1}^r y_{ijk} \text{ o total do } j\text{-ésimo ambiente.}$$

Verifica-se que,

$$F_{G(E_j)} = \frac{QM_{G(E_j)}}{QM_{Res}} \sim F_{(\alpha, (g-1), e(g-1)(r-1))}, \text{ para todo } j \text{ sob } H_{02}.$$

Ao nível α de significância, rejeita-se H_{02} quando,

$$F_{G(E_j)} = \frac{QM_{G(E_j)}}{QM_{Res}} \geq F_{(\alpha, (g-1), e(g-1)(r-1))}, \text{ para } j = 1, \dots, e.$$

O esquema da ANOVA, com aplicação do teste F , considerando a decomposição dos $(g-1)(e-1)$ graus de liberdade da interação $G \times E$ com $(g-1)$ graus de liberdade dos genótipos, para verificar o efeito de $G(E)$ é dado na Tabela 2.

Tabela 2 - Esquema da ANOVA com teste F obtida pela decomposição da interação $G \times E$ mais o fator G

Fonte de Variação	GL	SQ	QM	F
B d. E	$e(r-1)$	$SQ_{B \text{ d. } E}$	$QM_{B \text{ d. } E}$	
E	$(e-1)$	SQ_E	QM_E	
$G(E_1)$	$(g-1)$	$SQ_{G(E_1)}$	$QM_{G(E_1)}$	$QM_{G(E_1)}/QM_{Res}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$G(E_e)$	$(g-1)$	$SQ_{G(E_e)}$	$QM_{G(E_e)}$	$QM_{G(E_e)}/QM_{Res}$
$G(E)$	$e(g-1)$	$SQ_{G(E)}$	$QM_{G(E)}$	
Resíduo	$e(g-1)(r-1)$	SQ_{Res}	QM_{Res}	
Total	$ger-1$	SQ_{Total}		

GL: Graus de liberdade; SQ: Soma de quadrados; QM: Quadrado médio; B d. E: Blocos dentro do ambiente.

Têm-se que a $SQ_{G(E)} = \sum_{j=1}^e SQ_{G(E_j)}$, com $e(g-1)$ graus de liberdade e, observa-se que $SQ_{G(E)} = SQ_{G \times E} + SQ_G$, com $[(g-1)(e-1) + (g-1)] = [e(g-1)]$ graus de liberdade. Vê-se claramente por esta decomposição que o efeito não é puro, no que diz respeito à interação, pois à interação adiciona-se SQ_G .

2.3.2 Decomposição da soma dos graus de liberdade da interação $G \times E$ mais os graus de liberdade do fator E

Essa decomposição corresponde a obter o efeito do fator E (ambiente) dentro do fator G (genótipo), representado por $E(G)$. Suponha que o delineamento também seja

aleatorizado em blocos. Assim, o modelo adotado é:

$$Y_{ijk} = \mu + b_{k(j)} + g_i + e_{j(i)} + \varepsilon_{ijk}$$

em que:

Y_{ijk} : é o valor observado do i -ésimo genótipo no j -ésimo ambiente e no k -ésimo bloco, com $i = 1, 2, \dots, g$, $j = 1, 2, \dots, e$ e $k = 1, 2, \dots, r$;

μ : é uma constante comum às respostas (normalmente a média geral);

$b_{k(j)}$: é o efeito do k -ésimo bloco dentro do j -ésimo ambiente;

g_i : é o efeito do i -ésimo genótipo;

$e_{j(i)}$: é o efeito do j -ésimo ambiente na presença do i -ésimo genótipo; e

ε_{ijk} : é erro experimental associado ao i -ésimo genótipo, no j -ésimo ambiente e no k -ésimo bloco assumido ser independente e $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

O $e_{j(i)}$ é definido como:

$$e_{j(i)} = \mu_{j(i)} - \mu_i, \text{ tal que } \sum_{j=1}^e e_{j(i)} = 0, \text{ para todo } i.$$

Sendo que $\mu_{j(i)}$ refere-se à média populacional do j -ésimo ambiente na presença do i -ésimo genótipo e μ_i refere-se à média populacional do i -ésimo genótipo.

As hipóteses testadas são:

- i) $H_{01} : g_i = 0$, para todo i
 $H_{a1} : g_i \neq 0$, para algum i

ou ainda, H_{01} e H_{a1} podem ser escritas da seguinte maneira:

$$H_{01} : \mu_1 = \dots = \mu_g.$$

$$H_{a1} : \text{ pelo menos } \mu_i \neq \mu_{i'}, \text{ para } i \neq i' \ (i' = 1, \dots, g)$$

- ii) $H_{02} : e_{j(i)} = 0$, para todo ambiente j na presença do genótipo i
 $H_{a2} : e_{j(i)} \neq 0$, para algum ambiente j na presença do genótipo i

ou ainda, H_{02} e H_{a2} podem ser escritas da seguinte maneira:

H_{02} : $\mu_{1(i)} = \dots = \mu_{e(i)}$, na presença de i

H_{a2} : pelo menos $\mu_{j(i)} \neq \mu_{j'(i)}$, para ambiente $j \neq j'$ na presença do genótipo i .

A soma de quadrados do ambiente dentro do genótipo é dada por:

$$SQ_{E(G_i)} = \frac{1}{k} \sum_{j=1}^e y_{(i)j}^2 - \frac{y_{i..}^2}{ek}, \text{ com } i = 1, \dots, g \text{ e } (e-1) \text{ graus de liberdade associados,}$$

sendo que

$$y_{(i)j.} = \sum_{k=1}^r y_{(i)jk} \text{ o total do } j\text{-ésimo ambiente dentro do } i\text{-ésimo genótipo.}$$

$$y_{i..} = \sum_{j=1}^e \sum_{k=1}^r y_{ijk} \text{ o total do } i\text{-ésimo genótipo.}$$

Verifica-se que,

$$F_{E(G_i)} = \frac{QM_{E(G_i)}}{QM_{Res}} \sim F_{(\alpha, (e-1), e(g-1)(r-1))}, \text{ para todo } i \text{ sob } H_{02}.$$

Ao nível α de significância, rejeita-se H_{02} quando,

$$F_{E(G_i)} = \frac{QM_{E(G_i)}}{QM_{Res}} \geq F_{(\alpha, (e-1), e(g-1)(r-1))}, \text{ para } i = 1, \dots, g.$$

O esquema da ANOVA , com aplicação do teste F , considerando a decomposição dos $(g-1)(e-1)$ graus de liberdade da interação $G \times E$ com $(e-1)$ graus de liberdade dos ambientes, para verificar o efeito de $E(G)$ é dado na Tabela 3.

Têm-se que a $SQ_{E(G)} = \sum_{i=1}^g SQ_{E(G_i)}$, com $g(e-1)$ graus de liberdade e, observa-se que $SQ_{E(G)} = SQ_{G \times E} + SQ_E$, com $[(g-1)(e-1) + (e-1)] = [g(e-1)]$ graus de liberdade. Vê-se também claramente que por esta decomposição o efeito não é puro, no que diz respeito à interação, pois à interação adiciona-se SQ_E .

Tabela 3 - Esquema da ANOVA com teste F obtida pela decomposição da interação $G \times E$ mais o fator E

Fonte de Variação	GL	SQ	QM	F
B d. E	$e(r-1)$	$SQ_{B \text{ d. } E}$	$QM_{B \text{ d. } E}$	
G	$(g-1)$	SQ_G	QM_A	
$E(G_1)$	$(e-1)$	$SQ_{E(G_1)}$	$QM_{E(G_1)}$	$QM_{E(G_1)}/QM_{Res}$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$E(G_g)$	$(e-1)$	$SQ_{E(G_g)}$	$QM_{E(G_g)}$	$QM_{E(G_g)}/QM_{Res}$
$E(G)$	$g(e-1)$	$SQ_{E(G)}$	$QM_{E(G)}$	
Resíduo	$e(g-1)(r-1)$	SQ_{Res}	QM_{Res}	
Total	$ger-1$	SQ_{Total}		

GL: Graus de liberdade; SQ: Soma de quadrados; QM: Quadrado médio; B d. E: Blocos dentro do ambiente.

2.4 Autovalores e Autovetores de uma matriz e a Decomposição por Valores Singulares (DVS)

Em uma matriz quadrada $\mathbf{A}$ pode ser calculado um escalar λ e um vetor não nulo $\mathbf{x}$, de tal forma que:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (3)$$

em que λ é chamado de autovalor de $\mathbf{A}$ e $\mathbf{x}$ de autovetor de $\mathbf{A}$ associado a λ . Para encontrar λ e $\mathbf{x}$ associados à matriz $\mathbf{A}$, toma-se o sistema:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \quad (4)$$

sendo que $\mathbf{0}$ é um vetor nulo e $\mathbf{I}$ é uma matriz identidade de ordem p . O sistema dado em (4) admite solução não trivial se, e somente se, o determinante da matriz $(\mathbf{A} - \lambda\mathbf{I})$ for igual a zero, ou seja:

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

Se $|\mathbf{A} - \lambda\mathbf{I}| \neq 0$, então a matriz $(\mathbf{A} - \lambda\mathbf{I})$ possui inversa e $\mathbf{x} = \mathbf{0}$ é a única solução.

A equação $|\mathbf{A} - \lambda\mathbf{I}| = 0$ é um polinômio de grau p conhecido como polinômio característico ou equação característica e suas raízes são os autovalores de $\mathbf{A}$. Assim, após a determinação dos autovalores λ_i , com $i = 1, 2, \dots, p$, retorna-se à expressão (3) e calcula-se o autovetor associado a cada autovalor λ_i .

A decomposição por valores singulares (DVS) é uma técnica da álgebra matricial, de propriedades matemáticas muito fortes. A DVS é a base dos métodos mais precisos para a resolução de problemas de mínimos quadrados, para a determinação de posto de matrizes, do espaço-imagem ($R(\cdot)$) e do espaço nulo ($N(\cdot)$) de matrizes e de solução de vários

problemas envolvendo normas euclidianas ($\|\cdot\|$) (ARAÚJO, 2005).

Uma aplicação da DVS consiste na construção de uma aproximação para uma matriz real qualquer, de posto p , por p matrizes ortogonais entre si e de posto unitário (SOUZA, 1988). A decomposição exata para uma matriz real $\mathbf{A}_{(g \times e)}$, por DVS, é dada por: $\mathbf{A} = \sum_k \lambda_k \boldsymbol{\gamma}_k \boldsymbol{\alpha}'_k$; em que λ_k é o k -ésimo valor singular de $\mathbf{A}$, com $\boldsymbol{\gamma}_k$ e $\boldsymbol{\alpha}'_k$ vetores singulares (coluna e linha, respectivamente) associados ($k = 1, 2, \dots, p$; com $p \leq \min\{g, e\}$). Fazendo-se o somatório de um número n de termos, inferior a p , obtém-se uma aproximação de posto n para a matriz $\mathbf{A}$. Além disso, a matriz resultante representa a melhor aproximação de posto n para $\mathbf{A}$.

Arranjando-se os p valores singulares, em ordem decrescente, numa matriz diagonal ($\boldsymbol{\Lambda}_p$) e os respectivos vetores singulares $\boldsymbol{\gamma}_k$ e $\boldsymbol{\alpha}'_k$, em duas matrizes $\mathbf{U}_{g \times p}$ e $\mathbf{V}'_{p \times e}$, respectivamente, a expressão anterior pode ainda ser escrita da seguinte forma: $\mathbf{A} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{V}'$. Assim, $\mathbf{U}$ terá os p vetores $\boldsymbol{\gamma}_k$ em suas colunas e $\mathbf{V}'$ os p vetores $\boldsymbol{\alpha}'_k$ em suas linhas.

O uso da DVS, conforme Duarte e Vencovsky (1999), para obter a aproximação de uma matriz por outra de posto inferior, representa a aplicação adotada pela análise *AMMI* para compreender a estrutura subjacente aos dados da matriz de interações $\mathbf{GE}$, como definido na seção anterior. Ou seja, $\mathbf{GE}$ assumirá o papel da matriz $\mathbf{A}$.

2.5 O modelo *AMMI*

São muitas as metodologias de análise e interpretação para a interação genótipos $\times$ ambientes proveniente de um grupo de cultivares testados em vários ambientes. Uma alternativa é a utilização da chamada análise *AMMI* (*Additive Main effects and Multiplicative Interaction*), que significa modelo de efeitos principais aditivos e interação multiplicativa.

A abordagem *AMMI* combina análise de variância (ANOVA) e análise de componentes principais (ACP) em um único modelo, aditivo em relação aos efeitos principais de genótipos e ambientes e multiplicativo no detalhamento do efeito da interação. Tais análises (ANOVA e ACP) buscam resumir os dados contidos em uma matriz de interação genótipos $\times$ ambientes em um modelo. Ambas as análises modelam uma matriz $g \times e$ por um vetor de g parâmetros de genótipo e um vetor de e parâmetros de ambiente, ou seja, uma matriz é modelada por um par de vetores. A única diferença, segundo Gauch Jr (1992),

é que na ANOVA as estimativas dos efeitos dos genótipos e dos ambientes são somadas à média geral, enquanto que na ACP os efeitos são multiplicados entre si e depois são somados à média geral.

A ACP foi desenvolvida por Pearson em 1902 e a ANOVA foi proposta por Fisher em 1918 (MIRANDA, 2004). Assim, a ACP é mais antiga e considerada mais eficiente, entretanto é a ANOVA que se tornou mais popular devido à simplicidade dos cálculos e à fácil interpretação do modelo aditivo. A pouca utilização da ACP foi devido à exigência de grandes recursos computacionais, ausente quando a teoria foi proposta e pela dificuldade de interpretação do modelo multiplicativo.

Os primeiros a aplicar a ANOVA e a ACP, separadamente, em um experimento com 12 variedades de batata em crescimento com 6 fertilizantes (ambientes) e 3 repetições foram Fisher e Mackenzie (1923). Para este experimento em particular, eles acharam que a ACP, em comparação com a ANOVA, ajustou-se melhor aos dados. As razões para este melhor ajuste dado pela ACP é devido, em primeiro lugar, aos cálculos triviais da ANOVA, enquanto que a ACP requer cálculos iterativos extensos. Fisher e Mackenzie empenharam-se em três caminhos computacionais diferentes para encontrar e confirmar a solução da ACP, mesmo sendo uma matriz simples 12×3 . Em segundo lugar, os parâmetros aditivos da ANOVA eram mais facilmente compreendidos pelos agrônomos do que os parâmetros multiplicativos da ACP. Estes autores aplicaram ACP aos dados de produção original, mas não para a interação. Conseqüentemente, eles não desenvolveram a análise *AMMI*. Em 1918 ambos os componentes do modelo *AMMI* estavam disponíveis, mas o simples passo de combiná-los em uma única análise não foi dado até 1952.

A análise da interação a partir da regressão linear (RL) foi usada por Yates e Cochran (1938). Finlay e Wilkinson (1963) mais adiante popularizaram este método. Foram propostas muitas variações de regressão, incluindo cálculo das regressões complementares de ambientes nas médias de genótipos e substituindo uma medida física atual do ambiente para a média do ambiente. Historicamente, estas análises tiveram a vantagem de ser fácil de implementar até mesmo sem computadores.

Artigos utilizando metodologia semelhante ao modelo *AMMI*, de acordo com Gauch Jr (1992), foram publicados por Williams (1952) e Pike e Silverberg (1952). O primeiro foi publicado em um jornal estatístico e analisava uma experiência de química, já o segundo

artigo foi publicado em um jornal de engenharia e interessava-se pelo uso de computadores analógicos para balísticas e outras aplicações. Trabalhos envolvendo o modelo *AMMI* permaneceram parados por duas décadas e só foram retomados a partir de 1968 com Gollob. Provavelmente não é mera coincidência que o interesse ativo em modelos *AMMI* coincidiu com a disponibilidade de computadores digitais.

Foi introduzido por Gollob (1968) e Mandel (1961, 1969, 1971) o modelo FANOVA com fator analítico para estudar a interação em tabelas de dupla entradas completas. Entretanto, Zobel, Wright e Gauch Jr (1988), Gauch (1988), Gauch Jr e Zobel (1988) renomearam o modelo FANOVA como “Additive Main Effects and Multiplicative Interactions (*AMMI*) model”. Gauch e Zobel (1990) apresentaram o algoritmo *EM – AMMI* para imputar dados perdidos. Gauch (1990), Gauch e Furnas (1991) apresentaram o MAT-MODEL, um programa de microcomputador, para ajustar o modelo *AMMI*, *EM – AMMI* e modelos relacionados.

Comparando a ANOVA, a ACP, a RL e o modelo *AMMI*, Zobel, Wright e Gauch Jr (1988) verificaram que a ANOVA não detectou a significância da interação genótipos $\times$ ambientes, a ACP não identificou e separou os componentes principais de genótipos e de ambientes, a RL (metodologia mais usada) reteve uma grande parcela da soma de quadrados da interação genótipos $\times$ ambientes, unindo o efeito da interação genótipos $\times$ ambientes com os efeitos principais e a análise *AMMI* mostrou uma alta e significativa interação genótipos $\times$ ambientes com sentido agrônomo, permitindo separar os ambientes quanto à latitude e os genótipos quanto à precocidade.

Como mencionado anteriormente, o modelo de análise *AMMI* pressupõe componentes aditivos para os efeitos principais de genótipos e ambientes e componentes multiplicativos para o efeito de interação. Assim, a resposta média sobre r repetições ou blocos do i -ésimo genótipo no j -ésimo ambiente é representada por:

$$Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \varepsilon_{ij}, \quad (5)$$

em que:

Y_{ij} : é a resposta média do i -ésimo genótipo no j -ésimo ambiente;

μ : é uma constante comum às respostas (normalmente a média geral);

g_i : é o efeito do i -ésimo genótipo ($i = 1, 2, \dots, g$);

e_j : é o efeito do j -ésimo ambiente ($j = 1, 2, \dots, e$);

$(ge)_{ij} = \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij}$: é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente;

ρ_{ij} é o ruído; e

ε_{ij} : é o erro experimental médio, assumido independente e $\varepsilon_{ij} \sim N(0, \sigma^2/r)$.

Assume-se as restrições dadas em (2) e, além da média geral (μ) e do erro experimental médio (ε_{ij}), os demais termos do modelo resultam da decomposição por valores singulares (DVS) da matriz de interações: $\mathbf{GE}_{(g \times e)} = (\hat{g}\hat{e}_{ij})$. A matriz de interações é obtida como resíduo do ajuste aos efeitos principais, por ANOVA, aplicada à matriz de médias $\mathbf{Y}_{(g \times e)} = [Y_{ij}]$. Assim, λ_k é o k -ésimo valor singular de $\mathbf{GE}$, $\boldsymbol{\gamma}_{k(g \times 1)}$ e $\boldsymbol{\alpha}'_{k(1 \times e)}$ são os respectivos vetores singulares (vetor coluna e vetor linha) associados a λ_k (GOOD, 1969; MANDEL, 1971; PIEPHO, 1995). Correspondências entre DVS e ACP são como seguem:

λ_k é o k -ésimo valor singular ou a raiz quadrada do k -ésimo autovalor da matriz $(\mathbf{GE})(\mathbf{GE})'$ (ou $(\mathbf{GE})'(\mathbf{GE})$);

γ_{ik} é o i -ésimo elemento do k -ésimo autovetor de $(\mathbf{GE})(\mathbf{GE})'$ associado a λ_k^2 ;

α_{jk} é o j -ésimo elemento do k -ésimo autovetor de $(\mathbf{GE})'(\mathbf{GE})$ associado com λ_k^2 .

O índice k pode variar de 1 a p , sendo $p = \min\{(g-1), (e-1)\}$ o posto da matriz $\mathbf{GE}$. Tomando até n no somatório ($n < p$), determina uma aproximação de mínimos quadrados para a matriz $\mathbf{GE}$ pelos n primeiros termos da DVS (GOOD, 1969; GABRIEL, 1978); 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.

Pode-se escrever os componentes aditivos e multiplicativos do modelo *AMMI* da seguinte forma:

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

em que a parte multiplicativa pode ser desdobrada da seguinte forma:

$$Y_{ij} = \mu + g_i + e_j + \underbrace{\sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk}}_{\text{padrão}} + \underbrace{\sum_{k=n+1}^p \lambda_k \gamma_{ik} \alpha_{jk}}_{\text{ruído}} + \varepsilon_{ij}$$

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

A abordagem *AMMI* não busca recuperar toda a $SQ_{G \times E}$, mas apenas a parcela mais fortemente determinada por genótipos e ambientes (linhas e colunas da matriz **GE**, respectivamente), ou seja, o padrão de resposta (parte determinística ou sistemática). Pelas propriedades da DVS, $\lambda_1^2 \geq \lambda_2^2 \geq \dots \geq \lambda_p^2$ de forma que as primeiras parcelas captam proporções maiores da variação contida na matriz de interações. Por isso, o método *AMMI* é visto como um procedimento capaz de separar padrão e ruído na análise da $SQ_{G \times E}$: $\sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk}$ e $\sum_{k=n+1}^p \lambda_k \gamma_{ik} \alpha_{jk}$, respectivamente (WEBER; WRICKE; WESTERMANN, 1996), com um modelo parcimonioso, isto é, um modelo simples em termos de parâmetros a serem estimados.

Por padrão entende-se a lei geral de formação da matriz e que governa a interação $G \times E$, ou pelos n primeiros termos da DVS ou eixos da ACP. A porção ruído, presente em **GE**, mas não fortemente determinada por genótipos e ambientes, é descartada buscando-se melhorar a capacidade preditiva do modelo. Tais aspectos são vantajosos no modelo *AMMI* em comparação com os modelos mais tradicionais. Ademais, a análise *AMMI* considera os efeitos da interação de forma isolada e multiplicativa.

Algumas vantagens enumeradas por Zobel, Wright e Gauch Jr (1988) para o método *AMMI* foram:

1. permite uma análise mais detalhada da interação $G \times E$;
2. garante a seleção de genótipos, capitalizando suas interações positivas com os ambientes;
3. propicia estimativas mais precisas das respostas genotípicas;
4. possibilita fácil interpretação gráfica dos resultados, nos chamados “*biplot*” (representação gráfica simultânea dos genótipos e ambientes (GABRIEL, 1971)).

2.5.1 O desdobramento da soma de quadrados da interação $G \times E$ por DVS

A decomposição exata da matriz $\mathbf{GE}_{(g \times e)}$ por DVS, pode ser escrita como uma soma de p matrizes de posto unitário da seguinte forma:

$$\begin{aligned}\mathbf{GE} &= \mathbf{U}\mathbf{\Lambda}\mathbf{V}' = \sum_{k=1}^p \lambda_k \gamma_k \boldsymbol{\alpha}'_k; \text{ com } p = \min\{(g-1), (e-1)\} \\ \mathbf{GE} &= \lambda_1 \gamma_1 \boldsymbol{\alpha}'_1 + \lambda_2 \gamma_2 \boldsymbol{\alpha}'_2 + \cdots + \lambda_p \gamma_p \boldsymbol{\alpha}'_p = \sum_{i,j} (\hat{ge})_{ij}^2 = SQ_{G \times E} \\ \sum_k \lambda_k^2 &= \sum_{i,j} (\hat{ge})_{ij}^2 = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_p^2 = \sum_k SQ_{IPCA_k} = SQ_{G \times E}\end{aligned}$$

A soma de quadrados dos elementos da matriz $\mathbf{GE}$ reproduz exatamente a $SQ_{G \times E}$, que é desdobrada em componentes relativos a cada termo da DVS, ou a cada eixo $IPCA_k$ de interação da ACP em que o termo $IPCA$ vem de “*Incremental Principal Component Analysis*”, ou seja, $IPCA_k$ é o efeito de interação do componente principal associado ao k -ésimo eixo. Obtém-se esse desdobramento por meio dos autovalores de $(\mathbf{GE})(\mathbf{GE})'$ ou de $(\mathbf{GE})'(\mathbf{GE})$. A soma de quadrados relativa ao k -ésimo termo ou eixo de interação equivale ao quadrado do autovalor correspondente: $SQ_{G \times E[IPCA_k]} = \lambda_k^2$.

Utilizando-se de forma cumulativa os sucessivos termos do desdobramento, obtém-se aproximações cada vez melhores para a $SQ_{G \times E}$ original. Como $\lambda_1^2 \geq \lambda_2^2 \geq \cdots \geq \lambda_p^2$, é possível que com poucos primeiros termos já sejam suficientes para descrever uma alta proporção dessa soma de quadrados. Isso pode permitir a seleção de um modelo parcimonioso e com boa capacidade para descrever a variabilidade devida à interação. Logo, o objetivo da análise *AMMI* é descrever a interação $G \times E$ por um número reduzido de eixos ($n = 1$ a 3), resultando num modelo informativo, mas que tenha associados poucos graus de liberdade. Assim, espera-se que o modelo seja capaz de: (i) captar o padrão subjacente à interação, ou seja, a possível lei geral que governa o fenômeno e (ii) descartar os ruídos presentes nos dados e sem interesse agrônomo, os quais só prejudicam a capacidade preditiva do modelo. Essa separação em padrão e ruído pode também ser expressa em termos de soma de quadrados, da seguinte forma:

$$SQ_{G \times E} = \sum_{k=1}^p \lambda_k^2 = \left(\sum_{k=1}^n \lambda_k^2 \right) + \left(\sum_{k=n+1}^p \lambda_k^2 \right) = (SQ_{G \times E[\text{padrão}]}) + (SQ_{G \times E[\text{ruídos}]})$$

Dependendo do número de termos retidos para descrever o padrão da interação, o modelo é denotado: *AMMI0*, *AMMI1*, ..., *AMMIF*, gerando uma família de modelos *AMMI*. Por exemplo, em *AMMI0* nenhum eixo é ajustado, ou seja, o modelo é aditivo com

os fatores principais (genótipo e ambiente) e, portanto, sem interação. Em *AMMI1* apenas o primeiro eixo de interação multiplicativa é ajustado e assim por diante até o *AMMI_F*, o modelo completo de médias de caselas, com $n = p$. Vale ressaltar que se todos os eixos dessa decomposição (ou ACP) forem retidos, o modelo consumirá muitos graus de liberdade, apesar do ajuste perfeito à matriz de médias. O objetivo, entretanto, é resumir grande parte da interação $G \times E$ em apenas poucos eixos ($SQ_{G \times E[\text{padrão}]}$), resultando na seleção de um modelo *AMMI* reduzido, que descarta um resíduo adicional ($SQ_{G \times E[\text{ruídos}]}$).

2.6 Representação gráfica em “*biplot*”

O método *AMMI* possibilita representar, num único gráfico, os efeitos de interação para cada genótipo e cada ambiente. O gráfico, apresentado por Gabriel (1971), denominado “*biplot*”, baseia-se na aproximação da decomposição por valores singulares (DVS) de uma matriz, por outra de posto inferior. Para permitir a construção do gráfico, o posto da matriz aproximada deverá ser igual a um ou dois, resultando num “*biplot*” em uma ou duas dimensões, respectivamente.

Os gráficos “*biplot*” são úteis para diagnosticar o melhor modelo estatístico para um determinado conjunto de dados (GAUCH Jr, 1992). Eles também auxiliam a compreender e definir os ambientes ou mega-ambientes de um experimento, como também identificar fatores que influenciam os rendimentos. Kempton (1984) utilizou os gráficos “*biplot*” pela primeira vez em dados de produtividade e sua interpretação foi feita nos modelos de ANOVA, RL, ACP e *AMMI*. Este autor considera o “*biplot*” uma ferramenta estatística útil para interações complexas de genótipos $\times$ ambientes.

Na análise *AMMI* o uso do “*biplot*” é garantido pela DVS aplicada à matriz de interações, que será representada, na maioria das aplicações, por meio de sua aproximação de posto um ou dois. Seja a seguinte aproximação DVS, de posto 2, para **GE** (matriz de posto conhecido p):

$$\mathbf{GE}_{(g \times e)} = \sum_{k=1}^2 \lambda_k \boldsymbol{\gamma}_k \boldsymbol{\alpha}'_k = \tilde{\mathbf{U}} \tilde{\mathbf{\Lambda}} \tilde{\mathbf{V}}'$$

em que $\tilde{\mathbf{U}}_{(g \times 2)}$ tem em suas colunas apenas os 2 primeiros autovetores $\boldsymbol{\gamma}_{(g \times 1)}$ de $(\mathbf{GE})(\mathbf{GE})'$; $\tilde{\mathbf{V}}'_{(2 \times e)}$ tem em suas linhas os 2 primeiros autovetores $\boldsymbol{\alpha}'_{(1 \times e)}$ de $(\mathbf{GE})'(\mathbf{GE})$; e $\tilde{\mathbf{\Lambda}}_{(2)}$ é a matriz diagonal com os primeiros valores singulares não nulos: λ_1 e λ_2 de $(\mathbf{GE})(\mathbf{GE})'$ ou

$(\mathbf{GE})'(\mathbf{GE})$. A expressão pode também ser escrita como: $\mathbf{GE}_{(g \times e)} = (\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}^{1/2})(\tilde{\mathbf{\Lambda}}^{1/2}\tilde{\mathbf{V}}')$ ou, ainda, por: $\mathbf{GE}_{(g \times e)} = (\tilde{\mathbf{G}})(\tilde{\mathbf{H}}')$ em que, $\tilde{\mathbf{G}}_{(g \times 2)} = (\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}^{1/2})$ e $\tilde{\mathbf{H}}'_{(2 \times e)} = \tilde{\mathbf{\Lambda}}^{1/2}\tilde{\mathbf{V}}'$. A matriz $\tilde{\mathbf{G}}$ terá, então, g vetores denominados marcadores de linhas, cada um com 2 componentes (as coordenadas dos genótipos para cada eixo singular selecionado). Da mesma forma, $\tilde{\mathbf{H}}$ terá e marcadores de colunas, cada um com 2 componentes, correspondendo às coordenadas de ambientes para os mesmos 2 eixos. Dada a ortogonalidade dos eixos singulares, pode-se plotar os g genótipos e os e ambientes num sistema cartesiano de 2 eixos perpendiculares a fim de obter uma visualização gráfica.

Vê-se claramente, que o padrão subjacente aos dados da matriz de interações $\mathbf{GE}$ passa a ser expresso, inclusive graficamente, como o resultado da multiplicação de um termo relativo aos genótipos (efeitos de linhas) e outro relativo aos ambientes (efeitos de colunas). Ambos os termos expressos em unidades de raiz quadrada da unidade de Y_{ij} . Portanto, fazendo o produto interno de vetores, isto é, multiplicando uma linha i de $\tilde{\mathbf{G}}$ por uma coluna j de $\tilde{\mathbf{H}}'$, obtém-se a estimativa $AMMI_2$ para a interação do genótipo i com o ambiente j , ou seja: $\sum_{k=1}^2 \lambda_k \gamma_k \alpha'_k$. Dessa forma, possibilita capitalizar na seleção de genótipos, interações positivas de genótipos e ambientes específicos, o que corresponderá, no gráfico, a combinações de genótipos e ambientes com coordenadas (ou escores) de mesmo sinal. Assim, genótipos e ambientes com escores de mesmo sinal devem interagir positivamente evidenciando um sinergismo adaptativo a ser aproveitado na seleção. Já aqueles com sinais opostos devem interagir negativamente, sugerindo um certo antagonismo, ou seja, uma combinação desfavorável de genótipo e ambiente. Dessa forma, informações fundamentais ao processo tornam-se disponíveis, aumentando assim, a probabilidade de sucesso na recomendação de cultivares.

Um esquema que resume o raciocínio subjacente à representação em “*biplot*”, no caso da análise $AMMI$, pode ser escrito como:

$$\begin{aligned} \mathbf{GE}_{(g \times e)} &= \tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{V}}' = \overbrace{\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}^{1/2}}^{\tilde{\mathbf{G}}} \overbrace{\tilde{\mathbf{\Lambda}}^{1/2}\tilde{\mathbf{V}}'}^{\tilde{\mathbf{H}}'} = \tilde{\mathbf{G}}\tilde{\mathbf{H}}' \\ \mathbf{GE}_{(g \times e)} &= \tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{V}}' = \tilde{\mathbf{G}}\tilde{\mathbf{H}}' = \sum_{k=1}^2 (\lambda_k^{1/2} \gamma_k) (\lambda_k^{1/2} \alpha'_k) \\ \tilde{\mathbf{G}}_{(g \times 2)} &\Rightarrow \text{marcadores de linhas (efeitos de genótipos)} \end{aligned}$$

$$\tilde{\mathbf{H}}'_{(2 \times e)} \Rightarrow \text{marcadores de colunas (efeitos de ambientes)}$$

Vale ressaltar que o termo “*biplot*” não se refere a qualquer modelo estatístico particular, mas apenas a um tipo de gráfico contendo duas categorias de pontos ou marcadores, porém, muito rico em informações de relação entre e dentro dos marcadores, sendo chamado de análise *biplot* por Yan e Kang (2002). Neste caso, uma delas referindo-se aos genótipos e a outra aos ambientes, daí o termo “bi”.

Os gráficos em “*biplot*” são construídos em duas dimensões, o eixo das abscissas representa os efeitos principais (médias de genótipos e de ambientes) e o das ordenadas representa os escores de genótipos e ambientes para o referido eixo de interação.

A interpretação de um “*biplot*” quanto à interação $G \times E$ é feita observando a magnitude e o sinal dos escores de genótipos e ambientes para o(s) eixo(s) de interação. Assim, escores baixos (próximos de zero) são próprios de genótipos e ambientes que contribuíram pouco ou quase nada para a interação, caracterizando-os como estáveis. Por exemplo, num “*biplot*” *AMMI1* a estabilidade é avaliada observando as ordenadas (primeiro eixo de interação - $IPCA_1$). A estabilidade de um genótipo é indicativo de suas respectivas amplitudes adaptativas, ou seja, genótipos estáveis apresentam-se amplamente adaptados aos ambientes de teste. Entretanto, genótipos estáveis devem apresentar uma performance desejável para que as cultivares sejam recomendadas, o que é avaliado através de suas médias (efeitos principais).

As informações de um “*biplot*” *AMMI1* foram resumidas por Alvarez e Eyhéabide (1996), classificando os genótipos e ambientes em quatro grupos: 1) com rendimentos médios elevados (acima da média) e de baixa interação $G \times E$ (estáveis); 2) com baixos rendimentos (inferiores à média) e estáveis; 3) com rendimentos elevados, porém, apresentando maior interação $G \times E$ (instáveis); e 4) com baixos rendimentos e instáveis.

Com base no método de reamostragem “bootstrap” no modelo *AMMI*, Lavoranti (2003) concluiu que a interpretação pelos gráficos “*biplot*” dos escores “bootstrap” *AMMI* foi suficiente para descrever o comportamento da interação padrão para o grupo de genótipos e ambientes trabalhados, sendo que suas relações adaptativas foram melhores interpretadas com auxílio das médias e interações preditas pelos modelos *AMMI* selecionados.

Em função dessas informações, a análise *AMMI* é apontada como de grande utilidade não só na identificação de genótipos de alta produtividade e ampla adaptação, mas

também para a realização do chamado zoneamento agrônômico, incluindo a seleção de locais “chaves” (locais de teste, como referem-se Gauch Jr e Zobel (1996)).

2.7 A seleção do modelo *AMMI*

2.7.1 Determinação do número ótimo de termos multiplicativos no modelo *AMMI*

Dois procedimentos básicos têm sido desenvolvidos para determinar o número ótimo de termos multiplicativos a serem retidos nos componentes da interação genótipos $\times$ ambientes.

Um dos procedimentos básicos utiliza o método de validação cruzada no qual os dados são aleatoriamente divididos para modelagem de dados e para validação do modelo. O modelo *AMMI* é ajustado para modelar os dados e o quadrado médio do erro de predição são determinados dos dados de validação. Uma crítica a esse procedimento é que o melhor modelo preditivo computado de um subconjunto de dados pode não ser o melhor modelo quando todos os dados são considerados (CORNELIUS e CROSSA, 1999). No uso de validação cruzada em dados de experimentos multi-ambientais, deve-se ajustar os dados para número de repetições diferentes dentro de ambientes.

Outro procedimento para determinar o melhor modelo preditivo truncado é usar testes de hipóteses sobre o k -ésimo componente, $H_{0k} : \lambda_k = 0$, usando o conjunto de dados completos. Esses testes são baseados em somas de quadrados seqüenciais explicadas por termos multiplicativos.

2.7.2 Testes estatísticos para os componentes da soma de quadrados da interação $G \times E$

Várias pesquisas têm sido feitas para definição do número de eixos a serem retidos, a fim de descrever o padrão relacionado à interação (GOLLOB, 1968; MANDEL, 1971; GAUCH, 1988; entre outros). Como em ACP, todos os critérios de testes envolvidos, pelo menos indiretamente, levam em consideração a proporção da $SQ_{G \times E}$ acumulada até o n -ésimo eixo, isto é, $\sum_{k=1}^n \lambda_k^2 / SQ_{G \times E}$ (DIAS e KRZANOWSKI, 2003).

Um procedimento usual consiste na determinação dos graus de liberdade associados à parcela da $SQ_{G \times E}$ relacionada a cada membro da família de modelos *AMMI*. Isto permite calcular o quadrado médio para cada componente e obtém-se um teste F para

avaliar a significância de cada componente em relação ao QM_{Res} . Desde que se tenha um desdobramento ortogonal da soma de quadrados da interação, assume-se que o quociente entre o quadrado médio de um dos componentes da interação e o quadrado médio do resíduo segue uma aproximação F com os correspondentes graus de liberdade (desde que garantida uma aproximação normal para o resíduo da variável resposta original). A seleção de um modelo ótimo baseia-se na significância do teste F para os sucessivos termos da interação. O resíduo $AMMI$, coletado nos últimos termos da $SQ_{G \times E}$, também pode ser testado para confirmar sua significância (DIAS E KRZANOWSKI, 2003).

São mencionados por Gauch Jr e Zobel (1996) vários métodos de atribuição dos graus de liberdade aos componentes do modelo $AMMI$, dentre os mais populares são o de Gollob (1968) e o de Mandel (1971). Entretanto, os autores advertem sobre divergências entre estes métodos, já que recaem sobre considerações de natureza teórica e prática. A aproximação de Gollob (1968) é de fácil aplicação e é simplesmente definido por: $GL_{IPCA_k} = g + e - 1 - 2k$. Outros métodos requerem extensas simulações e por isso, esse sistema tenha sido grandemente utilizado.

O número de graus de liberdade associados a cada parcela λ_k^2 é definido por Mandel (1971) como: $GL_{IPCA_k} = E[\lambda_k^2]/\sigma^2$, em que $E[.]$ denota o operador esperança matemática e σ^2 a variância populacional. Entretanto, deve-se fazer simulações que resultam em três tabelas relacionadas ao número de graus de liberdade. Krzanowski (1979) dá algumas versões exatas dessas tabelas. Estas tabelas, entretanto, não são exaustivas e isto reduz a utilidade prática do método. Em comparação com Gollob (1968), o sistema de Mandel (1971) geralmente resulta em uma diminuição não linear dos graus de liberdade para sucessivos termos da interação, em que tais graus de liberdade podem ainda ser fracionários.

O trabalho de Mandel (1971) representa uma das melhores descrições da análise $AMMI$, mas o autor adverte que não é seguro aplicar o teste F se os autovalores λ_k^2 da matriz $(\mathbf{GE})(\mathbf{GE})'$ (ou $(\mathbf{GE})'(\mathbf{GE})$) não têm distribuição qui-quadrado (χ^2) ou desde que a soma de quadrados não são variáveis aleatórias independentes que sugerem distribuição qui-quadrado.

A abordagem proposta por Mandel (1971) foi considerada exata por muitos anos e por isso, a mais correta. Atualmente esta proposta têm sido muito criticada e o sistema mais utilizado no momento é a aproximação de Gollob (1968). As razões destas críticas centram nas pressuposições feitas por Mandel nas suas simulações.

O sistema de Gollob segundo Gauch Jr (1992), utiliza a definição estrutural clássica de graus de liberdade, isto é, o número de parâmetros estimados menos o número de restrições. Do modelo *AMMI* (equação 5), cada eixo de interação ($IPCA_k$) tem um valor singular λ_k , g elementos para o vetor γ_k e e elementos para o vetor α_k , totalizando $[g + e + 1]$ parâmetros estimados. Após a retirada dos efeitos principais, os autovalores ficam sujeitos a duas restrições: $\sum_i \gamma_{ik} = 0$ e $\sum_j \alpha_{jk} = 0$. Escalonando-os como vetores de comprimento unitário, compõem-se outras duas restrições: $\sum_i \gamma_{ik}^2 = 1$ e $\sum_j \alpha_{jk}^2 = 1$. Finalmente, requerendo que cada eixo $IPCA_k$, para $k > 1$, seja ortogonal a todos $[k - 1]$ eixos anteriores, impõe-se $2(k - 1)$ restrições adicionais: $\sum_i \gamma_{in} \gamma_{ik} = 0$ e $\sum_j \alpha_{jn} \alpha_{jk} = 0$, para $k = 1, 2, \dots, n - 1$ (DUARTE e VENCOSKY, 1999). Daí, conforme já enunciado, o eixo $IPCA_k$ terá:

$$GL_{IPCA_k} = g + e + 1 - [2 + 2 + 2(k - 1)] = g + e - 1 - 2k$$

Usando-se o sistema de Gollob, o quadro completo da análise de variância conjunta, obtida a partir de médias, é como mostrado na Tabela 4 a seguir:

Tabela 4 - Análise de variância conjunta obtida pelo sistema de Gollob

Fonte de Variação	GL	SQ
G	$(g - 1)$	SQ_G
E	$(e - 1)$	SQ_E
G×E	$(g - 1)(e - 1)$	$SQ_{G \times E}$
$IPCA_1$	$g + e - 1 - (2 \times 1)$	λ_1^2
$IPCA_2$	$g + e - 1 - (2 \times 2)$	λ_2^2
$IPCA_3$	$g + e - 1 - (2 \times 3)$	λ_3^2
$IPCA_4$	$g + e - 1 - (2 \times 4)$	λ_4^2
...	...	...
$IPCA_p$	$g + e - 1 - (2 \times p)$	λ_p^2
Resíduo	$ge(r - 1)$	SQ_{Res}
Total	$ger - 1$	SQ_{Total}

GL: Graus de liberdade; SQ: Soma de quadrados.

A robustez de alguns testes alternativos para selecionar um modelo *AMMI* em relação às suposições de homogeneidade e normalidade dos erros foi investigada por Piepho (1995). O autor comenta que o teste F é bastante liberal em selecionar mais termos multiplicativos quando se aplica o critério de Gollob (1968). Dos quatro métodos estudados, incluindo o de Gollob (1968), o teste F_R proposto por Cornelius, Seyedsadr e Crossa (1992) mostrou-se o mais robusto.

A estatística F_R com n termos multiplicativos no modelo é como segue:

$$F_{R,n} = \frac{SQ_{G \times E} - \sum_{k=1}^n \lambda_k^2}{f_2 QM_{Res}}; \text{ com } f_2 = (g-1-n)(e-1-n).$$

Sob hipótese nula de que não haja mais do que n termos determinando a interação, o numerador ($SQ_{G \times E}$ para o resíduo $AMMI$) é, aproximadamente, uma variável qui-quadrado (PIEPHO, 1995). Então, a estatística F_R tem uma distribuição F aproximada com f_2 graus de liberdade no numerador e graus de liberdade do resíduo do modelo no denominador (GL_{Res}).

Assim, um resultado significativo pelo teste sugere que ao menos um ou mais termos multiplicativos devem ser adicionados aos n já ajustados. Assim, F_R pode ser visto como um teste de significância dos $n+1$ primeiros termos da interação (similar ao teste de falta de ajuste na regressão linear). Quando $n = 0$, ou seja, quando nenhum termo multiplicativo é ajustado, o teste é equivalente ao F para a interação $G \times E$ global, na ANOVA, que é um teste exato. Vale lembrar que, os graus de liberdade do numerador de F_R é igual aos graus de liberdade para toda interação menos os graus de liberdade atribuídos por Gollob (1968) aos n primeiros termos. Portanto, conclui-se que a aplicação de F_R é equivalente ao teste do resíduo $AMMI$ para a interação $G \times E$, como sugerido anteriormente, ou ao teste para falta de ajuste na análise de regressão.

Em discussão sobre a obtenção dos graus de liberdade para componentes multiplicativos de um modelo $AMMI$ por Gauch Jr (1992), concluiu-se que sempre que simulações rigorosas parecerem desnecessárias, a recomendação geral é que se use o sistema de Gollob, de forma a permitir um teste F aproximado. Quando houver uma divisão clara entre grandes componentes determinando o padrão e pequenos componentes ricos em ruídos, o autor sugere que uma atribuição igualitária de graus de liberdade ($GL_{IPCA_k} = [(g-1)(e-1)]/p$) é interessante para os primeiros componentes, já que geralmente há pouco interesse em partições adicionais do resíduo (rico em ruídos). Ressalta-se ainda que, em certas questões de pesquisas, pode ser necessário a atribuição exata dos graus de liberdade a cada termo multiplicativo.

Na análise de um experimento de dois fatores sem repetição, a partir do modelo matemático: $Y_{ij} = \mu + \tau_i + \beta_j + \lambda_1 \alpha_i \gamma_j + \varepsilon_{ij}$, Johnson e Graybill (1972) proporam um teste estatístico baseado na razão de verossimilhança para testar a hipótese $H_0 : \lambda_1 = 0$ versus a

hipótese alternativa $H_a : \lambda_1 \neq 0$. A estatística é dada por:

$$\Delta^* = \frac{\sum_i \sum_j (ge)_{ij}^2 - \lambda_1}{\sum_i \sum_j (ge)_{ij}^2}$$

em que $(ge)_{ij}$ é o i -ésimo e j -ésimo elemento da matriz de interações **GE** e λ_1 é a raiz quadrada do maior (ou primeiro) autovalor da matriz $(\mathbf{GE})(\mathbf{GE})'$.

Para uma probabilidade de erro Tipo I igual a α , a hipótese $H_0 : \lambda_1 = 0$ é rejeitada em favor de $H_a : \lambda_1 \neq 0$ se $\Delta^* < K^*$, em que K^* é tal que $P_{H_0}[\Delta^* < K^*] = \alpha$ para um dado α . Dizer que $\Delta^* < K^*$ é equivalente a dizer que:

$$\frac{\lambda_1}{\lambda_1 + \lambda_2 + \dots + \lambda_g} > K$$

em que K é tal que

$$P_{H_0} \left[\frac{\lambda_1}{\lambda_1 + \lambda_2 + \dots + \lambda_g} > K \right] = \alpha$$

para um dado α .

O teste de Johnson e Graybill (1972) foi usado por Boik (1986) para o caso onde havia mais de uma observação por casela. Neste caso, o resíduo foi adicionado ao denominador da estatística de Johnson e Graybill (1972). Boik chamou este teste de razão máxima estudentizada e construiu tabelas para seu uso.

O mesmo teste foi derivado por Hegemann e Johnson (1976), neste caso, para o segundo autovalor da forma $\lambda_2/[\lambda_2 + \dots + \lambda_g]$ e incluiu algumas tabelas para o seu uso. Finalmente, Yochmowitz e Cornell (1978) apresentaram a generalização para o teste no caso da k -ésima razão dada por:

$$\Delta_k^* = \frac{\lambda_k}{\lambda_k + \dots + \lambda_g}.$$

Foi proposto por Cornelius, Seyedsadr e Crossa (1992) um teste para dados com repetições. Admitiram $\mu_1 = E(\sum_{k=1}^n \lambda_k^2/\sigma^2)$ e $\mu_2' = Var(\sum_{k=1}^n \lambda_k^2/\sigma^2)$. Além disso, seja s^2 e f a estimativa de σ^2 e os graus de liberdade, respectivamente. As expressões seguintes são distribuições F aproximadas e podem ser usadas como testes aproximados para significância do k -ésimo termo multiplicativo.

1. Seja $q_1 = \mu_2' + \mu_1^2 + (f-4)\mu_1$, $q_2 = (f-2)\mu_2' + 2\mu_1^2$, $b = q_1\mu_1/q_2$, $a = 1 + [(f-2)q_1]/q_2$, então:

$$F_1 = \frac{n \sum_{k=1}^n \lambda_k^2/b}{f s^2/a}$$

é aproximadamente uma distribuição F com $2b$ e $2a$ graus de liberdade.

2. Seja $f_1 = 2\mu_1^2/\mu_2'$, então:

$$F_2 = \frac{n \sum_{k=1}^n \lambda_k^2}{\mu_1 s^2}$$

é aproximadamente uma distribuição F com f_1 e f graus de liberdade.

Ambas aproximações da distribuição F são obtidas indiretamente. A primeira usa o fato de que se aX/b tem distribuição F com $2b$ e $2a$ graus de liberdade, então $X+1$ tem a mesma distribuição de $1/B$, em que B é uma variável aleatória Beta com parâmetros a e b . Os valores para a e b , parâmetros da distribuição Beta, decorrem de tomar $X = n \sum_{k=1}^n \lambda_k^2 / f s^2$ e adotam o primeiro e segundo momentos de $X + 1$ igual ao momento correspondente de $1/B$. A segunda aproximação usa essencialmente o mesmo procedimento usado por Johnson (1976) para deduzir o valor crítico de contrastes na interação. Investigações posteriores, mas não publicadas, sugeriram que os testes F_1 e F_2 controlavam a taxa de erro Tipo I para modelos *AMMI*.

Baseados nos teoremas de Goodman e Haberman (1990), Cornelius (1993) estendeu os testes F_1 e F_2 denotando-os de F_{GH1} e F_{GH2} . O autor admitiu os valores u_1 para a esperança e u_2 para o desvio padrão do maior autovalor de uma matriz central de Wishart com dimensão e graus de liberdade específicos (p e n para o primeiro autovalor a ser testado, $p-1$ e $n-1$ para o segundo, etc., em que $p = \min(g-1, e-1)$ e $n = \max(g-1, e-1)$). Quando ocorrer $p \leq 19$ e $n \leq 99$, os valores da esperança e desvio padrão podem ser obtidos na tabela de Mandel (1971) que fora construída por simulação de Monte Carlo. Cornelius (1980) forneceu fórmulas, obtidas por análise de regressão dos resultados de Mandel, que aproximam dos valores da tabela e podem ser usadas em programas computacionais. Entretanto, são válidos somente dentro das variações para p e n na tabela de Mandel. Assim, seja s^2 o quadrado médio do resíduo sobre a média da casela com f graus de liberdade, isto é, $s^2 = s_e^2/r$ em que s_e^2 é a raiz quadrada do quadrado médio do resíduo e r o número de repetições dentro do ambiente. Admitindo-se $v_1 = u_2^2 + u_1^2 + (f-4)u_1$ e $v_2 = (f-2)u_2^2 + 2u_1^2$ e considerando também, $h_1 = 2v_1u_1/v_2$, $h_2 = 2u_1^2/u_2^2$ e $g = 2 + [2(f-2)v_1]/v_2$. Então, os testes são dados por:

$$F_{GH1} = \frac{g\hat{\lambda}^2}{h_1 f s^2}, \text{ com } h_1 \text{ e } g \text{ graus de liberdade e}$$

$$F_{GH2} = \frac{\hat{\lambda}^2}{u_1 s^2}, \text{ com } h_2 \text{ e } f \text{ graus de liberdade,}$$

em que $\hat{\lambda}^2$ é um autovalor particular testado. Os testes F_{GH1} e F_{GH2} tem aproximadamente distribuição F .

2.7.3 Avaliação preditiva por validação cruzada

Procedimentos estatísticos computacionalmente intensivos vêm sendo amplamente aplicados. Dentre eles encontram-se os baseados em reamostragem como jackknife, bootstrap e validação cruzada. Na análise *AMMI*, Gauch (1988) introduziu a avaliação preditiva, por meio de validação cruzada, para fins de seleção do modelo (*AMMI0*, *AMMI1*, *AMMI2*, ..., ou *AMMIF*).

A avaliação pelo teste F é comentada por Gauch Jr e Zobel (1988) ser “*predictiva*”, na medida em que busca um modelo que explique uma grande parte da variação mostrada por todos os dados observados (com coeficiente de determinação elevado). Entretanto, argumentam que tal abordagem mostra-se menos eficiente para selecionar modelos parcimoniosos e capazes de desprezar ruídos de menor interesse.

Por outro lado, um critério “*preditivo*” de avaliação prioriza a capacidade de um modelo aproximar suas predições a dados não incluídos na análise (simulando respostas futuras ainda não mensuradas). Como ilustração dos dois enfoques, os autores exemplificam que se uma SQ_{Trat} for constituída de 70% de padrão e 30% de ruídos, o objetivo de uma análise estatística não deveria ser recuperar 100% dessa soma de quadrados. Mas, preferencialmente, apenas 70% e, especialmente, aqueles 70% que representam o padrão. Recuperar um padrão implica num aumento na precisão das predições, mas recuperar ruídos diminui essa precisão cada vez mais. Ademais, um modelo que seletivamente recupera o padrão e relega ruídos a um resíduo desconsiderado na predição de respostas, pode resultar em melhor precisão do que os próprios dados. Esse é o princípio subjacente à proposta de Gauch (1988) para a seleção do modelo *AMMI*.

Dessa forma, os dados de repetições para cada combinação de genótipos e ambientes são divididos através da validação cruzada por um critério aleatório, em dois subconjuntos: (i) dados para o ajuste de modelo *AMMI* e (ii) dados de validação. As respostas preditas por um determinado modelo da família *AMMI* são confrontadas com os respectivos dados de validação, calculando-se as diferenças, entre esses valores. Em seguida, obtém-se a soma de quadrados dessas diferenças, dividindo-se o resultado pelo número delas. Extrairdo

a raiz quadrada desse resultado, tem-se o que os autores chamaram de raiz quadrada da diferença preditiva média (RMSPD). Segundo Crossa et al. (1991), esse procedimento deve ser repetido pelo menos dez vezes, obtendo-se uma média dos resultados para cada membro da família de modelos. Um pequeno valor de RMSPD indica um bom sucesso preditivo do modelo. O modelo escolhido é, então, utilizado para analisar os dados de todas as r repetições, conjuntamente, numa análise definitiva.

Embora o procedimento de Gauch vêm sendo amplamente usado, Piepho (1994) sugere algumas modificações, tais como que a obtenção do valor médio de RMSPD seja obtido a partir de 1000 randomizações diferentes, ao invés de 10 como sugerem Crossa et al. (1991). O autor ainda propõe uma modificação para a partição aleatória dos dados (modelagem e validação) quando esses são provenientes de experimentos delineados em blocos. Nesse caso, recomenda-se que sejam feitos sorteios de blocos inteiros dentro de experimentos e não de parcelas para cada combinação de genótipo e ambiente (tratamento).

Assim, fica preservada a estrutura original de blocos. Contudo, apesar da coerência lógica desse tipo de proposta, estudos confirmando sua eficácia ainda não são suficientemente disponíveis. Gauch Jr e Zobel (1996) sugerem ainda que se faça o conjunto de dados de validação sempre com uma só observação para cada tratamento. Acredita-se, com isso, que seja mais provável, a partir de $r - 1$ dados, encontrar um modelo que mais se aproxime do ideal para analisar o conjunto completo dos r dados por tratamento.

Vale ressaltar que ao avaliar o modelo por validação cruzada, a análise *AMMI* deve partir das observações individuais propriamente ditas (dados de cada repetição dentro de experimentos). Em contrapartida, se o modelo for avaliado por um teste F , a análise pode ser feita a partir das médias dos genótipos nos ambientes (experimentos), desde que se disponha dos quadrados médios residuais, obtidos nas análises de variâncias de cada experimento, como é usual proceder na análise conjunta de experimento.

2.7.4 Predição de respostas genotípicas por ambiente

A partir da seleção do melhor modelo que descreva a estrutura dos dados, ou seja, *AMMI0*, *AMMI1*, *AMMI2*, ... ou *AMMIF*, pode-se prever a resposta de cada

genótipo em cada ambiente ($\hat{Y}_{ij}$) baseado no modelo escolhido:

$$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk}$$

em que: n é o número de eixos ou componentes principais selecionados para descrever o padrão da interação $G \times E$ sendo $n \leq p$, em que p refere-se a escolha do modelo completo com todos os termos multiplicativos.

Mediante a seleção do modelo, a expressão acima assume as seguintes formas, conforme o modelo escolhido:

Modelo <i>AMMI</i> selecionado	Expressão da resposta esperada do genótipo i no ambiente j ($\hat{Y}_{ij}$)
<i>AMMI0</i>	$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..}$
<i>AMMI1</i>	$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + \lambda_1 \gamma_{i1} \alpha_{j1}$
<i>AMMI2</i>	$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + \lambda_1 \gamma_{i1} \alpha_{j1} + \lambda_2 \gamma_{i2} \alpha_{j2}$
<i>AMMI3</i>	$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + \lambda_1 \gamma_{i1} \alpha_{j1} + \lambda_2 \gamma_{i2} \alpha_{j2} + \lambda_3 \gamma_{i3} \alpha_{j3}$
...	...
<i>AMMIF</i>	$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + \lambda_1 \gamma_{i1} \alpha_{j1} + \lambda_2 \gamma_{i2} \alpha_{j2} + \dots + \lambda_p \gamma_{ip} \alpha_{jp}$

Se *AMMIF* for o modelo selecionado (modelo completo com todos os termos multiplicativos), a resposta esperada equivale à média da casela relativa ao genótipo i e no ambiente j , ou seja, a média das r repetições do genótipo naquele ambiente, o que pode ser facilmente verificado por:

$$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + (\hat{ge})_{ij};$$

pois $\sum_{k=1}^p \lambda_k \gamma_{ik} \alpha_{jk} = (\hat{ge})_{ij}$, mas tem-se que $(\hat{ge})_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$. Logo,

$$\hat{Y}_{ij} = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} + (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = Y_{ij}.$$

Um modelo *AMMI* não leva em conta somente a interação, pois sempre contempla também os efeitos principais. É possível não incluir nenhum termo de interação (*AMMI0*), quando nenhum deles se mostrar significativo. Entretanto, na maior parte das pesquisas em genética e melhoramento de plantas, a interação $G \times E$ é significativa, de forma que pelo menos um termo multiplicativo normalmente deve ser incluído na predição das respostas.

Para dados não repetidos, Gauch Jr e Zobel (1988) mostram que uma análise *AMMI* associada à validação cruzada para a seleção do modelo é preditivamente mais precisa

do que as médias de tratamentos obtidas de duas repetições. Esse benefício aumenta com o tamanho do ensaio (aumentando o número de genótipos e ambientes) e quanto maior for o nível de ruído presente nos dados.

2.8 Distribuição qui-quadrado não-central

Na teoria de probabilidade e estatística, se $U_1, U_2, \dots, U_v$ são variáveis aleatórias independentes e normalmente distribuídas, e $\delta_1, \delta_2, \dots, \delta_v$ são constantes, então a distribuição de

$$\sum_{j=1}^v (U_j + \delta_j)^2 \quad (6)$$

depende de $\delta_1, \delta_2, \dots, \delta_v$ somente através das suas somas de quadrados. Esta é a chamada distribuição qui-quadrado (χ^2) não-central com v graus de liberdade e parâmetro de não-centralidade $\theta = \sum_{j=1}^v \delta_j^2$, representada por $\chi_v^2(\theta)$.

A função densidade probabilidade é expressa como (JOHNSON, KOTZ e BALAKRISHNAN, 1995):

$$f(u; v, \theta) = \frac{\exp\{-1/2(\theta + u)\}}{2^{v/2}} \sum_{j=0}^{\infty} \left(\frac{\theta}{4}\right)^j \frac{u^{(v/2)+j-1}}{j! \Gamma(1/2v + j)},$$

em que $\Gamma[.]$ refere-se à função gama.

Uma estimativa do parâmetro de não-centralidade é apresentada por Hogg e Craig (1978) como:

$$\hat{\theta} = \frac{\sum_{j=1}^v U_j^2}{\sigma^2}$$

3 MATERIAL E MÉTODOS

3.1 Características dos dados

Os dados a serem utilizados são os mesmos utilizados por Cornelius e Crossa (1999) e Dias e Krzanowski (2003). Foram obtidos pelo CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo) em experimentos realizados em vários países, caracterizando assim, experimentos multi-ambientais. Foram utilizados genótipos de milho e trigo sendo que em todos os experimentos utilizou-se o delineamento aleatorizado em blocos. Cada conjunto tem a seguinte descrição:

Conjunto 1: 20 genótipos de trigo, sendo que um genótipo é do tipo trigo “durum” e os outros 19 são do tipo trigo “bread”. Cada genótipo foi avaliado em 34 ambientes com 4 blocos;

Conjunto 2: 16 genótipos de milho avaliados em 24 ambientes com 4 blocos;

Conjunto 3: 9 genótipos de milho avaliados em 20 ambientes com 4 blocos;

Conjunto 4: 8 genótipos de milho avaliados em 59 ambientes com 4 blocos.

3.2 Estudo de simulação

Inicialmente, estudou-se por simulação, utilizando o módulo IML do sistema computacional SAS (2003), o comportamento das somas de quadrados das linhas e colunas da matriz de interação genótipos $\times$ ambientes (**GE**), visando a aplicação de um teste para os efeitos de genótipos e ambientes.

Supõe-se que $\mathbf{Y} = [y_{ij}]$ é a tabela de dupla entrada para um conjunto de dados escolhido, e que se ajuste a análise de variância (ANOVA):

$$Y_{ij} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ij}, \quad (7)$$

em que:

Y_{ij} : é a resposta média do i -ésimo genótipo no j -ésimo ambiente;

μ : é uma constante comum às respostas (normalmente a média geral);

g_i : é o efeito do i -ésimo genótipo ($i = 1, 2, \dots, g$);

e_j : é o efeito do j -ésimo ambiente ($j = 1, 2, \dots, e$);

$(ge)_{ij}$: é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente; e

ε_{ij} : é o erro experimental médio, assumido independente e $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Considere uma barra sobre a letra para denotar média, com ‘.’ indicando o sobescrito sobre o qual o somatório ocorre. Dessa forma, os estimadores dos parâmetros do modelo (7) são:

$$\begin{aligned}\hat{\mu} &= \frac{\sum_{ij} Y_{ij}}{ge} = \bar{Y}_{..} & (\hat{ge})_{ij} &= Y_{ij} - \frac{\sum_j Y_{ij}}{e} - \frac{\sum_i Y_{ij}}{g} + \frac{\sum_{ij} Y_{ij}}{ge} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \\ \hat{g}_i &= \frac{\sum_j Y_{ij}}{e} - \hat{\mu} = \bar{Y}_{i.} - \bar{Y}_{..} & \hat{e}_j &= \frac{\sum_i Y_{ij}}{g} - \hat{\mu} = \bar{Y}_{.j} - \bar{Y}_{..}\end{aligned}$$

Assim, com base nos dados originais é possível estimar $\hat{\mu}$, $\hat{g}_i$, $\hat{e}_j$ e $(\hat{ge})_{ij}$ gerando números aleatórios de uma distribuição normal com média 0 e variância σ^2 para o termo ε_{ij} (σ^2 é também estimado pelo QM_{Res}). Pode-se também gerar repetições que imitem o padrão dos dados originais, mas os quais têm embutido uma estrutura de variabilidade dos mesmos. Geraram-se 500, 5000 e 10000 experimentos para o Conjunto 1, dos quais obteve-se as somas de quadrados dos elementos das linhas e colunas da matriz de interação **GE**.

No software R (2008), utilizou-se a função “*fitdistr*” para estimar o parâmetro de não-centralidade da distribuição qui-quadrado não-central. Esta função faz a estimação de máxima verossimilhança, utilizando um processo de otimização, sendo que a estimativa inicial utilizada foi a média dos dados simulados. Utilizou-se ainda a função “*ks.test*” para verificar a aderência das somas de quadrados dos dados simulados a distribuição qui-quadrado não-central através do teste de Kolmogorov-Smirnov. A estatística do teste de Kolmogorov-Smirnov é definida por:

$$D = \sup_x |F_n(x) - F(x)|,$$

em que $\sup_x |\cdot|$ refere-se ao supremo de x representando a maior distância, medida no sentido vertical, entre as imagens das funções $F_n(x)$ e $F(x)$, sendo $F(x)$ a função de distribuição acumulada definida para um conjunto de n observações independentes $X_1, \dots, X_n$ e $F_n(x)$ a função de distribuição empírica definida a partir das estatísticas de ordem $X_{(1)}, \dots, X_{(n)}$, da seguinte forma:

$$F_n(x) = 0, \text{ se } x < X_{(1)};$$

$$F_n(x) = \frac{i}{n}, \text{ se } X_{(i)} \leq x < X_{(i+1)} \text{ com } i = 1, \dots, n-1;$$

$$F_n(x) = 1, \text{ se } X_{(n)} \leq x.$$

Por último, construiu-se o gráfico QQplot para visualização do ajuste das somas de quadrados das linhas e colunas da matriz de interação **GE** a distribuição qui-quadrado não-central.

3.2.1 Análise de variância conjunta

Com o objetivo de verificar se existe a interação entre genótipos e ambientes, realiza-se uma análise de variância conjunta que envolve o estudo de todos os genótipos em todos os ambientes, sendo que em cada ambiente utilizou-se um delineamento aleatorizado em blocos. Esses dados serão representados pelo seguinte modelo matemático:

$$Y_{ijk} = \mu + b_{k(j)} + g_i + e_j + (ge)_{ij} + \varepsilon_{ijk} \quad (8)$$

sendo que:

Y_{ijk} : é o valor observado do i -ésimo genótipo no j -ésimo ambiente e no k -ésimo bloco, com $i = 1, 2, \dots, g$, $j = 1, 2, \dots, e$ e $k = 1, 2, \dots, r$;

μ : é uma constante comum às respostas (normalmente a média geral);

$b_{k(j)}$: é o efeito do k -ésimo bloco dentro do j -ésimo ambiente;

g_i : é o efeito do i -ésimo genótipo;

e_j : é o efeito do j -ésimo ambiente;

$(ge)_{ij}$: é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente;

ε_{ijk} : é erro experimental associado ao i -ésimo genótipo, no j -ésimo ambiente e no k -ésimo bloco assumido ser independente e $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

Na Tabela 5 apresenta-se o esquema da análise de variância para o modelo (8):

Tabela 5 - Esquema da ANOVA para experimentos de um mesmo grupo de genótipos avaliado em e ambientes com k blocos

Fonte de Variação	GL	QM
Blocos dentro do ambiente (B d. E)	$e(r - 1)$	$QM_{B \text{ d. } E}$
Genótipos (G)	$(g - 1)$	QM_G
Ambientes (E)	$(e - 1)$	QM_E
Interação (G×E)	$(g - 1)(e - 1)$	$QM_{G \times E}$
Resíduo	$e(g - 1)(r - 1)$	QM_{Res}
Total	$ger - 1$	

GL: Graus de liberdade; QM: Quadrado médio.

3.3 Contribuição de cada genótipo e ambiente para a interação

A manifestação fenotípica dos caracteres métricos resulta do efeito conjunto do genótipo, do ambiente específico e da interação destes fatores. As avaliações do comportamento diferencial dos genótipos em função da diversidade ambiental, pode ocasionar mudanças na posição relativa dos genótipos ou mesmo na magnitude das suas diferenças (MAIA, 2004). Diferenças em adaptação de genótipos em populações resultam de diferenças de constituição gênica para os caracteres importantes nesta adaptação. A reação diferencial às mudanças ambientais pode-se dar desde os mecanismos de regulação gênica até caracteres morfológicos finais.

No que compete à contribuição dos diferentes ambientes para a interação, sua variância é atribuída a uma porção considerada fixa e uma porção imprevisível que é aleatória. A porção previsível faz parte da variação e pode ser medida em magnitude e em direção, seu caráter sistemático facilita razoável controle. Já a porção imprevisível faz parte da variação experimental não controlada, por ter natureza indeterminada, ou seja, de sentido e direção incertos.

Estudos sobre a interação genótipos $\times$ ambientes devem pormenorizar o comportamento tanto dos genótipos disponíveis, em estudos de adaptabilidade e estabilidade, quanto dos ambientes estudados. Para Chaves (2001) o efeito do ambiente é um confundimento de efeitos fixos e aleatórios. Assim, o ambiente mais previsível terá uma maior possibilidade de aproveitamento dos efeitos da interação, sendo ideal que os efeitos de ambientes pudessem ser tomados como fixos ou, pelo menos, que a porção fixa da interação pudesse ser

medida e os resultados utilizados na recomendação de cultivares.

3.4 Proposta de teste estatístico para a contribuição de genótipos e ambientes para a interação

Seja a matriz de interação $\mathbf{GE}_{(g \times e)} = (\hat{ge}_{ij})$ dada da seguinte forma:

$$\mathbf{GE}_{(g \times e)} = \begin{pmatrix} \hat{ge}_{11} & \cdots & \hat{ge}_{1e} \\ \hat{ge}_{21} & \cdots & \hat{ge}_{2e} \\ \cdots & \cdots & \cdots \\ \hat{ge}_{g1} & \cdots & \hat{ge}_{ge} \end{pmatrix}$$

Como o objetivo é verificar a contribuição de cada genótipo e ambiente para a interação $G \times E$, as hipóteses testadas são:

- i) H_{01} : o i -ésimo genótipo não contribui para a interação $G \times E$, para $i = 1, \dots, g$
 H_{a1} : o i -ésimo genótipo contribui para a interação $G \times E$, para $i = 1, \dots, g$

Seja $G_i(G \times E)$ a representação do efeito do i -ésimo genótipo para a interação.

A soma de quadrados devido ao efeito desse genótipo é obtida por:

$$SQ_{G_i(G \times E)} = \sum_{j=1}^e (\hat{ge}_{ij})^2 \quad (i = 1, 2, \dots, g), \text{ com } \frac{(g-1)(e-1)}{g} \text{ graus de liberdade associados,}$$

e o quadrado médio correspondente é obtido dividindo-se cada soma de quadrados pelos respectivos graus de liberdade, ou seja:

$$QM_{G_i(G \times E)} = \frac{SQ_{G_i(G \times E)}}{\frac{(g-1)(e-1)}{g}} = \frac{g(SQ_{G_i(G \times E)})}{(g-1)(e-1)}, \quad i = 1, \dots, g.$$

Observa-se que a $SQ_{G_1(G \times E)} + \cdots + SQ_{G_g(G \times E)} = SQ_{G \times E}$, com $(g-1)(e-1)$ graus de liberdade associados.

Tem-se que a soma de quadrados devido ao efeito de genótipo dividido pela variância segue distribuição qui-quadrado não-central com parâmetro de não-centralidade θ_{1i} , ou seja:

$$\frac{SQ_{G_i(G \times E)}}{\sigma^2} \sim \chi^2_{\frac{(g-1)(e-1)}{g}}(\theta_{1i})$$

$$\text{em que } \theta_{1i} = \frac{\sum_{j=1}^e (\hat{ge}_{ij})^2}{\sigma^2}.$$

Assim, verifica-se que:

$$F_{G_i(G \times E)}^* = \frac{QM_{G_i(G \times E)}}{QM_{Res}} \sim F_{(\alpha, [(g-1)(e-1)]/g, e(g-1)(r-1))}(\theta_{1i}) \text{ para todo genótipo } i \text{ sob } H_{01}.$$

Seja a hipótese $H_0 : \theta_{1i} = 0$ versus $H_a : \theta_{1i} > 0$, que é equivalente a hipótese H_{01} versus H_{a1} , tem-se que a estatística $F_{G_i(G \times E)}^*$ sob H_0 tem distribuição F central com $[(g-1)(e-1)]/g$ e $e(g-1)(r-1)$ graus de liberdade (MOSER, 1996).

Ao nível α de significância, rejeita-se H_0 ou H_{01} quando,

$$F_{G_i(G \times E)}^* = \frac{QM_{G_i(G \times E)}}{QM_{Res}} \geq F_{(\alpha, [(g-1)(e-1)]/g, e(g-1)(r-1))}(\theta_{1i}) \text{ para } i = 1, \dots, g.$$

Através do teste F , aplicado a ANOVA, levando em consideração a decomposição dos $(g-1)(e-1)$ graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/g$ graus de liberdade para os genótipos, resulta no seguinte esquema:

Tabela 6 - Esquema da ANOVA com teste F para obtenção de genótipos que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F
B d. E	$e(r-1)$	$SQ_{B \text{ d. } E}$	$QM_{B \text{ d. } E}$	
G	$(g-1)$	SQ_G	QM_G	
E	$(e-1)$	SQ_E	QM_E	
$G \times E$	$(g-1)(e-1)$	$SQ_{G \times E}$	$QM_{G \times E}$	
$G_1(G \times E)$	$\frac{(g-1)(e-1)}{g}$	$SQ_{G_1(G \times E)}$	$QM_{G_1(G \times E)}$	$\frac{QM_{G_1(G \times E)}}{QM_{Res}}$
...	...	...	...	...
$G_g(G \times E)$	$\frac{(g-1)(e-1)}{g}$	$SQ_{G_g(G \times E)}$	$QM_{G_g(G \times E)}$	$\frac{QM_{G_g(G \times E)}}{QM_{Res}}$
Resíduo	$e(g-1)(r-1)$	SQ_{Res}	QM_{Res}	
Total	$ger-1$	SQ_{Total}		

GL: Graus de liberdade; SQ: Soma de quadrados; QM: Quadrado médio;

$G_i(G \times E)$: é o efeito do i -ésimo genótipo para a interação, com $i = 1, \dots, g$.

- ii) H_{02} : o j -ésimo ambiente não contribui para a interação $G \times E$, para $j = 1, \dots, e$
 H_{a2} : o j -ésimo ambiente contribui para a interação $G \times E$, para $j = 1, \dots, e$

Seja $E_j(G \times E)$ a representação do efeito do j -ésimo ambiente para a interação.

A soma de quadrados devido ao efeito desse ambiente é obtida por:

$$SQ_{E_j(G \times E)} = \sum_{i=1}^g (\hat{g}e_{ij})^2 \quad (j = 1, 2, \dots, e), \text{ com } \frac{(g-1)(e-1)}{e} \text{ graus de liberdade associados,}$$

e o quadrado médio correspondente é obtido dividindo-se cada soma de quadrados pelos respectivos graus de liberdade, ou seja:

$$QM_{E_j(G \times E)} = \frac{SQ_{E_j(G \times E)}}{\frac{(g-1)(e-1)}{e}} = \frac{e(SQ_{E_j(G \times E)})}{(g-1)(e-1)}, \quad j = 1, \dots, e.$$

Observa-se que a $SQ_{E_1(G \times E)} + \dots + SQ_{E_e(G \times E)} = SQ_{G \times E}$, com $(g-1)(e-1)$ graus de liberdade associados.

Tem-se que a soma de quadrados devido ao efeito de ambiente dividido pela variância segue distribuição qui-quadrado não-central com parâmetro de não-centralidade θ_2 , ou seja:

$$\frac{SQ_{E_j(G \times E)}}{\sigma^2} \sim \chi^2_{\frac{(g-1)(e-1)}{e}}(\theta_{2j})$$

$$\text{em que } \theta_{2j} = \frac{\sum_{i=1}^g (\hat{g}e_{ij})^2}{\sigma^2}.$$

Assim verifica-se que:

$$F_{E_j(G \times E)}^* = \frac{QM_{E_j(G \times E)}}{QM_{Res}} \sim F_{(\alpha, [(g-1)(e-1)]/e, e(g-1)(r-1))}(\theta_{2j}) \text{ para todo ambiente } j \text{ sob } H_{02}.$$

Seja a hipótese $H_0 : \theta_{2j} = 0$ versus $H_a : \theta_{2j} > 0$, que é equivalente a hipótese H_{02} versus H_{a2} , tem-se que a estatística $F_{E_j(G \times E)}^*$ sob H_0 tem distribuição F central com $[(g-1)(e-1)]/e$ e $e(g-1)(r-1)$ graus de liberdade (MOSER, 1996).

Ao nível α de significância, rejeita-se H_0 ou H_{02} quando,

$$F_{E_j(G \times E)}^* = \frac{QM_{E_j(G \times E)}}{QM_{Res}} \geq F_{(\alpha, [(g-1)(e-1)]/e, e(g-1)(r-1))}(\theta_{2j}) \text{ para } j = 1, \dots, e.$$

Através do teste F , aplicado a ANOVA, levando em consideração a decomposição dos $(g-1)(e-1)$ graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/e$ graus de liberdade para os ambientes, resulta no seguinte esquema:

Tabela 7 - Esquema da ANOVA com teste F para obtenção de ambientes que contribuem significativamente para a interação $G \times E$

Fonte de variação	GL	SQ	QM	F
B d. E	$e(r - 1)$	$SQ_{B \text{ d. } E}$	$QM_{B \text{ d. } E}$	
G	$(g - 1)$	SQ_G	QM_G	
E	$(e - 1)$	SQ_E	QM_E	
$G \times E$	$(g - 1)(e - 1)$	$SQ_{G \times E}$	$QM_{G \times E}$	
$E_1(G \times E)$	$\frac{(g - 1)(e - 1)}{e}$	$SQ_{E_1(G \times E)}$	$QM_{E_1(G \times E)}$	$\frac{QM_{E_1(G \times E)}}{QM_{Res}}$
...	...	...	...	...
$E_e(G \times E)$	$\frac{(g - 1)(e - 1)}{e}$	$SQ_{E_e(G \times E)}$	$QM_{E_e(G \times E)}$	$\frac{QM_{E_e(G \times E)}}{QM_{Res}}$
Resíduo	$e(g - 1)(r - 1)$	SQ_{Res}	QM_{Res}	
Total	$ger - 1$	SQ_{Total}		

GL: Graus de liberdade; SQ: Soma de quadrados; QM: Quadrado médio;

$E_j(G \times E)$: é o efeito do j -ésimo ambiente para a interação, com $j = 1, \dots, e$.

A proposta dos graus de liberdade igualitários para genótipos e ambientes, vai de encontro à proposta de Gauch Jr (1992) para obtenção dos graus de liberdade dos componentes multiplicativos de um modelo *AMMI*.

Claramente o teste aqui proposto para interação de dois fatores pode ser aplicado a qualquer conjunto de dados de um ensaio fatorial, não necessitando ser apenas de um ensaio multi-ambiente.

4 RESULTADOS E DISCUSSÃO

4.1 Estudo de simulação

A simulação de dados foi realizada para o Conjunto 1 que contém 20 genótipos de trigo, sendo que um genótipo é do tipo trigo “durum” e os outros 19 são do tipo trigo “bread”. Cada genótipo foi avaliado em 34 ambientes com 4 blocos.

As Figuras de 4 a 18 apresentam os histogramas dos 500, 5000 e 10000 experimentos simulados para as somas de quadrados (SQ) de cada linha (genótipos) da matriz de interação **GE** com as linhas contínuas representando o ajuste de uma distribuição qui-quadrado não-central teórica e as barras verticais representando a distribuição empírica dos dados de somas de quadrados simulados para genótipos. Observa-se pelas Figuras de 4 a 18 que as distribuições das somas de quadrados dos genótipos estão bem próximas de uma distribuição qui-quadrado não-central com 31,35 graus de liberdade e ainda, a medida que o número de experimentos simulados aumenta, essa aproximação se torna melhor.

Figura 4 - Histogramas para as SQ dos genótipos 1, 2, 3 e 4 com 500 experimentos simulados

Figura 5 - Histogramas para as SQ dos genótipos 5, 6, 7 e 8 com 500 experimentos simulados

Figura 6 - Histogramas para as SQ dos genótipos 9, 10, 11 e 12 com 500 experimentos simulados

Figura 7 - Histogramas para as SQ dos genótipos 13, 14, 15 e 16 com 500 experimentos simulados

Figura 8 - Histogramas para as SQ dos genótipos 17, 18, 19 e 20 com 500 experimentos simulados

Figura 9 - Histogramas para as SQ dos genótipos 1, 2, 3 e 4 com 5000 experimentos simulados

Figura 10 - Histogramas para as SQ dos genótipos 5, 6, 7 e 8 com 5000 experimentos simulados

Figura 11 - Histogramas para as SQ dos genótipos 9, 10, 11 e 12 com 5000 experimentos simulados

Figura 12 - Histogramas para as SQ dos genótipos 13, 14, 15 e 16 com 5000 experimentos simulados

Figura 13 - Histogramas para as SQ dos genótipos 17, 18, 19 e 20 com 5000 experimentos simulados

Figura 14 - Histogramas para as SQ dos genótipos 1, 2, 3 e 4 com 10000 experimentos simulados

Figura 15 - Histogramas para as SQ dos genótipos 5, 6, 7 e 8 com 10000 experimentos simulados

Figura 16 - Histogramas para as SQ dos genótipos 9, 10, 11 e 12 com 10000 experimentos simulados

Figura 17 - Histogramas para as SQ dos genótipos 13, 14, 15 e 16 com 10000 experimentos simulados

Figura 18 - Histogramas para as SQ dos genótipos 17, 18, 19 e 20 com 10000 experimentos simulados

As Figuras de 19 a 45 apresentam os histogramas dos 500, 5000 e 10000 experimentos simulados para as somas de quadrados (SQ) de cada coluna (ambientes) da matriz de interação **GE** com a linha contínua representando a distribuição qui-quadrado não-central teórica e as barras verticais representando a distribuição empírica dos dados de somas de quadrados simulados para ambientes. Observa-se pelas Figuras de 19 a 45 que as distribuições das somas de quadrados dos ambientes estão bem próximas de uma distribuição qui-quadrado não-central com aproximadamente 18,44 graus de liberdade e ainda, a medida que o número de experimentos simulados aumenta, essa aproximação se torna mais evidente.

Figura 19 - Histogramas para as SQ dos ambientes 1, 2, 3 e 4 com 500 experimentos simulados

Figura 20 - Histogramas para as SQ dos ambientes 5, 6, 7 e 8 com 500 experimentos simulados

Figura 21 - Histogramas para as SQ dos ambientes 9, 10, 11 e 12 com 500 experimentos simulados

Figura 22 - Histogramas para as SQ dos ambientes 13, 14, 15 e 16 com 500 experimentos simulados

Figura 23 - Histogramas para as SQ dos ambientes 17, 18, 19 e 20 com 500 experimentos simulados

Figura 24 - Histogramas para as SQ dos ambientes 21, 22, 23 e 24 com 500 experimentos simulados

Figura 25 - Histogramas para as SQ dos ambientes 25, 26, 27 e 28 com 500 experimentos simulados

Figura 26 - Histogramas para as SQ dos ambientes 29, 30, 31 e 32 com 500 experimentos simulados

Figura 27 - Histogramas para as SQ dos ambientes 33 e 34 com 500 experimentos simulados

Figura 28 - Histogramas para as SQ dos ambientes 1, 2, 3 e 4 com 5000 experimentos simulados

Figura 29 - Histogramas para as SQ dos ambientes 5, 6, 7 e 8 com 5000 experimentos simulados

Figura 30 - Histogramas para as SQ dos ambientes 9, 10, 11 e 12 com 5000 experimentos simulados

Figura 31 - Histogramas para as SQ dos ambientes 13, 14, 15 e 16 com 5000 experimentos simulados

Figura 32 - Histogramas para as SQ dos ambientes 17, 18, 19 e 20 com 5000 experimentos simulados

Figura 33 - Histogramas para as SQ dos ambientes 21, 22, 23 e 24 com 5000 experimentos simulados

Figura 34 - Histogramas para as SQ dos ambientes 25, 26, 27 e 28 com 5000 experimentos simulados

Figura 35 - Histogramas para as SQ dos ambientes 29, 30, 31 e 32 com 5000 experimentos simulados

Figura 36 - Histogramas para as SQ dos ambientes 33 e 34 com 5000 experimentos simulados

Figura 37 - Histogramas para as SQ dos ambientes 1, 2, 3 e 4 com 10000 experimentos simulados

Figura 38 - Histogramas para as SQ dos ambientes 5, 6, 7 e 8 com 10000 experimentos simulados

Figura 39 - Histogramas para as SQ dos ambientes 9, 10, 11 e 12 com 10000 experimentos simulados

Figura 40 - Histogramas para as SQ dos ambientes 13, 14, 15 e 16 com 10000 experimentos simulados

Figura 41 - Histogramas para as SQ dos ambientes 17, 18, 19 e 20 com 10000 experimentos simulados

Figura 42 - Histogramas para as SQ dos ambientes 21, 22, 23 e 24 com 10000 experimentos simulados

Figura 43 - Histogramas para as SQ dos ambientes 25, 26, 27 e 28 com 10000 experimentos simulados

Figura 44 - Histogramas para as SQ dos ambientes 29, 30, 31 e 32 com 10000 experimentos simulados

Figura 45 - Histogramas para as SQ dos ambientes 33 e 34 com 10000 experimentos simulados

Nas Tabelas 8 e 9 são apresentadas as estimativas de máxima verossimilhança dos parâmetros de não-centralidade da distribuição qui-quadrado. Nestas mesmas tabelas

também são apresentados os valores do nível descritivo p do teste de Kolmogorov-Smirnov para verificar se as somas de quadrados das linhas e colunas da matriz de interação **GE** seguem uma distribuição qui-quadrado não-central com 31, 35 e 18, 44 graus de liberdade, respectivamente.

Tabela 8 - Estimativas dos parâmetros de não-centralidade ($\hat{\theta}$) e valor- p do teste de Kolmogorov-Smirnov para aderência à distribuição qui-quadrado não-central da soma de quadrados dos genótipos para 500, 5000 e 10000 experimentos simulados

	500		5000		10000	
	$\hat{\theta}$	valor- p	$\hat{\theta}$	valor- p	$\hat{\theta}$	valor- p
G_1	108,6197	0,3855	108,5908	0,5126	108,5437	0,1027
G_2	51,1050	0,6455	50,9525	0,6705	50,8489	0,2220
G_3	43,4694	0,9898	42,3223	0,8190	42,5279	0,2833
G_4	37,9770	0,7969	38,0530	0,4874	38,0294	0,1671
G_5	76,5203	0,6897	76,4220	0,6329	76,1384	0,1337
G_6	44,9578	0,9307	44,3966	0,3889	44,4034	0,3644
G_7	96,0936	0,5712	96,5530	0,2131	96,7086	0,0207
G_8	74,2493	0,5406	73,6738	0,7236	73,8949	0,8790
G_9	52,2325	0,5714	52,6581	0,4144	52,6225	0,2367
G_{10}	76,6308	0,7140	76,3332	0,3238	76,4783	0,7882
G_{11}	52,7447	0,2375	52,2932	0,7642	52,0944	0,5212
G_{12}	107,4176	0,9997	109,9130	0,6118	109,3350	0,2460
G_{13}	56,4111	0,5103	56,6207	0,6776	56,5189	0,4606
G_{14}	109,7172	0,1839	109,7279	0,1899	108,9042	0,1850
G_{15}	36,4324	0,9930	37,0387	0,7508	37,1842	0,1471
G_{16}	96,2576	0,4699	96,7074	0,8570	97,0332	0,4719
G_{17}	116,6450	0,5952	116,2115	0,3081	116,6613	0,1173
G_{18}	58,7156	0,7050	59,0081	0,0657	58,8805	0,2883
G_{19}	104,8892	0,8935	106,5185	0,1296	106,1936	0,2299
G_{20}	272,5516	0,6567	274,6099	0,5655	274,4027	0,2886

Da Tabela 8, nota-se que os valores estimados dos parâmetros de não-centralidade, para cada genótipo dentro da interação, não diferem a medida que se aumenta o número de simulações, mostrando assim que um número de 500 simulações foi bom para se estimar o parâmetro de não-centralidade da distribuição qui-quadrado. No entanto, o $\hat{\theta}$ varia de um genótipo para outro, indicando assim, que cada genótipo deve ser testado com uma particular distribuição F não-central, no sentido de verificar sua contribuição para a interação. Ainda nesta tabela, observa-se pelo valor- p que quase todas somas de quadrados devido aos genótipos dentro da interação genótipos $\times$ ambientes seguem uma distribuição qui-quadrado não-central. Este fato não foi observado, ao nível de 5% de significância, somente para o genótipo 7 e com 10000 simulações, estando perfeitamente dentro da margem

de erro do teste.

Tabela 9 - Estimativas dos parâmetros de não-centralidade ($\hat{\theta}$) e valor- p do teste de Kolmogorov-Smirnov para aderência à distribuição qui-quadrado não-central da soma de quadrados dos ambientes para 500, 5000 e 10000 experimentos simulados

	500		5000		10000	
	$\hat{\theta}$	valor- p	$\hat{\theta}$	valor- p	$\hat{\theta}$	valor- p
E_1	47,9933	0,8274	47,5419	0,7600	47,4448	0,4527
E_2	18,5578	0,9650	18,3636	0,2099	18,4021	0,4401
E_3	18,2636	0,8389	18,2692	0,8900	18,2803	0,7319
E_4	88,0627	0,6028	88,7763	0,6361	88,8459	0,3685
E_5	105,3483	0,0549	106,6328	0,6540	105,7354	0,7280
E_6	77,3272	0,8555	76,1497	0,6180	75,9230	0,5808
E_7	95,7689	0,6009	94,4437	0,8425	94,3937	0,6344
E_8	23,8334	0,7683	24,8442	0,8794	24,5371	0,1582
E_9	58,2305	0,3515	57,5403	0,3071	57,6803	0,8469
E_{10}	33,6378	0,9780	34,3815	0,4765	34,1411	0,6399
E_{11}	25,3091	0,3087	25,5565	0,7432	25,6057	0,6102
E_{12}	52,3734	0,5243	52,0422	0,6592	52,2213	0,1165
E_{13}	14,9108	0,2778	13,8305	0,8009	14,0507	0,9983
E_{14}	44,0804	0,4584	44,8162	0,9841	44,7146	0,1158
E_{15}	33,6823	0,9214	34,1133	0,3457	34,2616	0,5984
E_{16}	45,8735	0,8312	46,2242	0,1991	46,2753	0,7545
E_{17}	16,0549	0,9607	16,0598	0,9916	15,5132	0,4634
E_{18}	96,8405	0,7747	97,4067	0,6447	97,5725	0,2953
E_{19}	60,8326	0,9048	60,5904	0,6443	60,7523	0,1201
E_{20}	84,9604	0,9860	86,8706	0,7878	87,2916	0,8719
E_{21}	21,2416	0,7060	21,1500	0,7160	21,2328	0,1283
E_{22}	51,7497	0,9734	51,9349	0,8794	51,7137	0,8167
E_{23}	30,1671	0,9103	30,8733	0,8477	30,9801	0,8761
E_{24}	46,2759	0,4099	47,5078	0,3844	47,0886	0,4266
E_{25}	58,3453	0,6626	59,5726	0,1733	59,4737	0,3564
E_{26}	50,0072	0,6161	49,2312	0,6664	49,1291	0,9217
E_{27}	38,9532	0,7406	38,4890	0,8133	38,4574	0,0744
E_{28}	29,7554	0,9957	28,9504	0,2773	28,9457	0,6254
E_{29}	76,0706	0,6518	76,1735	0,7977	76,0228	0,3372
E_{30}	8,0419	0,9800	8,2774	0,8936	8,4957	0,2298
E_{31}	14,1124	0,5245	14,2522	0,4211	13,9414	0,0969
E_{32}	30,4566	0,6012	30,4168	0,8199	30,5236	0,7806
E_{33}	80,6634	0,9807	81,3923	0,8113	81,7776	0,8583
E_{34}	95,4098	0,8885	95,7457	0,6438	96,0478	0,9026

Observando a Tabela 9, têm-se as mesmas conclusões para as somas de quadrados devido a ambientes dentro da interação genótipos $\times$ ambientes, ou seja, que os valores dos parâmetros de não-centralidade, para cada ambiente dentro da interação, não difere a medida que se aumenta o número de simulações, mostrando assim que um número de 500 simulações foi um bom número para se estimar o parâmetro de não-centralidade da dis-

tribuição qui-quadrado e que todas somas de quadrados devido aos ambientes dentro da interação genótipos $\times$ ambientes seguem uma distribuição qui-quadrado não-central.

As mesmas observações feitas anteriormente com relação a distribuição de probabilidades das somas de quadrados devido a genótipos e ambientes dentro da interação genótipos $\times$ ambientes, também podem ser verificadas através dos gráficos QQplot nas Figuras de 46 a 87. Nota-se que em todos os gráficos os pontos estão dispostos próximo de linha reta, o que leva a concluir que estas somas de quadrados apresentam um bom ajuste à distribuição qui-quadrado não-central com parâmetro de não-centralidade θ , até mesmo a soma de quadrados referente ao genótipo 7 que havia sido rejeitado pelo teste de Kolmogorov-Smirnov com 10000 simulações.

Figura 46 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 1, 2, 3 e 4 com 500 experimentos simulados

Figura 47 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 5, 6, 7 e 8 com 500 experimentos simulados

Figura 48 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 9, 10, 11 e 12 com 500 experimentos simulados

Figura 49 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 13, 14, 15 e 16 com 500 experimentos simulados

Figura 50 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 17, 18, 19 e 20 com 500 experimentos simulados

Figura 51 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 1, 2, 3 e 4 com 5000 experimentos simulados

Figura 52 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 5, 6, 7 e 8 com 5000 experimentos simulados

Figura 53 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 9, 10, 11 e 12 com 5000 experimentos simulados

Figura 54 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 13, 14, 15 e 16 com 5000 experimentos simulados

Figura 55 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 17, 18, 19 e 20 com 5000 experimentos simulados

Figura 56 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 1, 2, 3 e 4 com 10000 experimentos simulados

Figura 57 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 5, 6, 7 e 8 com 10000 experimentos simulados

Figura 58 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 9, 10, 11 e 12 com 10000 experimentos simulados

Figura 59 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 13, 14, 15 e 16 com 10000 experimentos simulados

Figura 60 - Gráficos QQplot qui-quadrado para as SQ dos genótipos 17, 18, 19 e 20 com 10000 experimentos simulados

Figura 61 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 1, 2, 3 e 4 com 500 experimentos simulados

Figura 62 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 5, 6, 7 e 8 com 500 experimentos simulados

Figura 63 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 9, 10, 11 e 12 com 500 experimentos simulados

Figura 64 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 13, 14, 15 e 16 com 500 experimentos simulados

Figura 65 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 17, 18, 19 e 20 com 500 experimentos simulados

Figura 66 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 21, 22, 23 e 24 com 500 experimentos simulados

Figura 67 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 25, 26, 27 e 28 com 500 experimentos simulados

Figura 68 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 29, 30, 31 e 32 com 500 experimentos simulados

Figura 69 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 33 e 34 com 500 experimentos simulados

Figura 70 - Gráficos QQplot qui-quadrado as SQ dos ambientes 1, 2, 3 e 4 com 5000 experimentos simulados

Figura 71 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 5, 6, 7 e 8 com 5000 experimentos simulados

Figura 72 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 9, 10, 11 e 12 com 5000 experimentos simulados

Figura 73 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 13, 14, 15 e 16 com 5000 experimentos simulados

Figura 74 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 17, 18, 19 e 20 com 5000 experimentos simulados

Figura 75 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 21, 22, 23 e 24 com 5000 experimentos simulados

Figura 76 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 25, 26, 27 e 28 com 5000 experimentos simulados

Figura 77 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 29, 30, 31 e 32 com 5000 experimentos simulados

Figura 78 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 33 e 34 com 5000 experimentos simulados

Figura 79 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 1, 2, 3 e 4 com 10000 experimentos simulados

Figura 80 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 5, 6, 7 e 8 com 10000 experimentos simulados

Figura 81 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 9, 10, 11 e 12 com 10000 experimentos simulados

Figura 82 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 13, 14, 15 e 16 com 10000 experimentos simulados

Figura 83 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 17, 18, 19 e 20 com 10000 experimentos simulados

Figura 84 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 21, 22, 23 e 24 com 10000 experimentos simulados

Figura 85 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 25, 26, 27 e 28 com 10000 experimentos simulados

Figura 86 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 29, 30, 31 e 32 com 10000 experimentos simulados

Figura 87 - Gráficos QQplot qui-quadrado para as SQ dos ambientes 33 e 34 com 10000 experimentos simulados

4.2 Aplicação em dados reais

4.2.1 Análise de variância conjunta

A Tabela 10 corresponde a análise de variância conjunta (ANOVA) efetuada com o Conjunto 1 com 20 genótipos de trigo avaliados em 34 ambientes com 4 blocos. Verifica-se que o efeito de blocos dentro de ambiente (*B d. E*), o efeito de genótipos (*G*), o efeito de ambientes (*E*) e o efeito da interação genótipos $\times$ ambientes ($G \times E$) são significativos ao nível de 5% de significância e suas somas de quadrados (SQ) correspondem a 4,32%, 1,49%, 72,7% e 9,96%, respectivamente, da soma de quadrados total.

Tabela 10 - ANOVA do Conjunto 1 com 20 genótipos de trigo avaliados em 34 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	<i>F</i>	valor- <i>p</i>
B d. E	102	257862519	2528064	7,14	<0,0001
G	19	89066441	4687707	13,23	<0,0001
E	33	4333925428	131331074	370,67	<0,0001
$G \times E$	627	594108485	947541	2,67	<0,0001
Resíduo	1938	686646195	354307		
Total	2719	5961609068			

Através do teste *F*, aplicado a ANOVA, levando em consideração a decomposição dos $(g - 1)(e - 1) = (20 - 1)(34 - 1) = 627$ graus de liberdade da interação $G \times E$ com $[(g - 1)(e - 1)]/g = [(20 - 1)(34 - 1)]/20 = 31,35$ graus de liberdade para os genótipos, resulta na Tabela 11.

Tabela 11 - Teste F , aplicado ao Conjunto 1, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- p
$G_1(G \times E)$	31,35	38419779	1225511,30	3,4589006	<0,0001
$G_2(G \times E)$	31,35	18000775	574187,39	1,6205947	0,0164
$G_3(G \times E)$	31,35	15051875	480123,62	1,3551077	0,0910
$G_4(G \times E)$	31,35	13438384	428656,59	1,2098465	0,1970
$G_5(G \times E)$	31,35	26944156	859462,70	2,4257598	<0,0001
$G_6(G \times E)$	31,35	15771832	503088,74	1,4199248	0,0619
$G_7(G \times E)$	31,35	34115953	1088228,20	3,0714306	<0,0001
$G_8(G \times E)$	31,35	26152950	834224,88	2,3545282	<0,0001
$G_9(G \times E)$	31,35	18638713	594536,31	1,6780278	0,0109
$G_{10}(G \times E)$	31,35	27052040	862903,98	2,4354725	<0,0001
$G_{11}(G \times E)$	31,35	18457831	588766,55	1,6617431	0,0122
$G_{12}(G \times E)$	31,35	38927562	1241708,50	3,5046158	<0,0001
$G_{13}(G \times E)$	31,35	20045716	639416,78	1,8046990	0,0042
$G_{14}(G \times E)$	31,35	38666042	1233366,60	3,4810713	<0,0001
$G_{15}(G \times E)$	31,35	13098543	417816,37	1,1792509	0,2278
$G_{16}(G \times E)$	31,35	34294875	1093935,40	3,0875388	<0,0001
$G_{17}(G \times E)$	31,35	41353606	1319094,30	3,7230305	<0,0001
$G_{18}(G \times E)$	31,35	20927058	667529,76	1,8840455	0,0022
$G_{19}(G \times E)$	31,35	37625900	1200188,20	3,3874282	<0,0001
$G_{20}(G \times E)$	31,35	97124895	3098082,80	8,7440729	<0,0001

Nota-se que somente os genótipos 3, 4, 6 e 15 são não significativos ao nível de 5% de significância, ou seja, tais genótipos não contribuem significativamente para a interação genótipos $\times$ ambientes. Os demais genótipos são significativos e contribuem mais para a interação. Sendo assim, tais genótipos podem ser descartados pelos melhoristas, uma vez que eles apresentam respostas heterogêneas quanto aos ambientes.

Agora, levando em consideração a decomposição dos mesmos 627 graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/e = [(20-1)(34-1)]/34 \approx 18,44$ graus de liberdade para os ambientes com aplicação do teste F na ANOVA, resulta na Tabela 12.

Observa-se que os ambientes 2, 3, 8, 11, 13, 17, 21, 28, 30 e 31 são os que não contribuem para a interação genótipos $\times$ ambientes, sendo esses ambientes que podem ser escolhidos pelos melhoristas por serem ambientes de resposta homogênea aos genótipos para essa variável. Os demais ambientes contribuem significativamente, ao nível de 5% de significância, para a interação.

Tabela 12 - Teste F , aplicado ao Conjunto 1, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- p
$E_1(G \times E)$	18,44	16838211,0	913076,83	2,5770810	0,0002
$E_2(G \times E)$	18,44	6501297,7	352542,46	0,9950209	0,4628
$E_3(G \times E)$	18,44	6448827,2	349697,17	0,9869903	0,4724
$E_4(G \times E)$	18,44	31408385,0	1703166,00	4,8070400	<0,0001
$E_5(G \times E)$	18,44	37599852,0	2038907,50	5,7546415	<0,0001
$E_6(G \times E)$	18,44	26887597,0	1458019,60	4,1151354	<0,0001
$E_7(G \times E)$	18,44	33399579,0	1811141,50	5,1117914	<0,0001
$E_8(G \times E)$	18,44	8687710,9	471103,94	1,3296505	0,1566
$E_9(G \times E)$	18,44	20359469,0	1104022,20	3,1160081	<0,0001
$E_{10}(G \times E)$	18,44	12093933,0	655811,38	1,8509714	0,0150
$E_{11}(G \times E)$	18,44	9053531,4	490941,10	1,3856391	0,1258
$E_{12}(G \times E)$	18,44	18406587,0	998124,34	2,8171203	<0,0001
$E_{13}(G \times E)$	18,44	4906271,5	266049,81	0,7509028	0,7630
$E_{14}(G \times E)$	18,44	15826596,0	858220,51	2,4222538	<0,0001
$E_{15}(G \times E)$	18,44	12162492,0	659529,07	1,8614642	0,0142
$E_{16}(G \times E)$	18,44	16377350,0	888085,98	2,5065465	<0,0001
$E_{17}(G \times E)$	18,44	5595639,5	303431,81	0,8564103	0,6357
$E_{18}(G \times E)$	18,44	34465014,0	1868916,20	5,2748558	<0,0001
$E_{19}(G \times E)$	18,44	21472712,0	1164389,50	3,2863895	<0,0001
$E_{20}(G \times E)$	18,44	30798200,0	1670077,80	4,7136514	<0,0001
$E_{21}(G \times E)$	18,44	7480479,7	405640,05	1,1448843	0,2997
$E_{22}(G \times E)$	18,44	18406653,0	998127,92	2,8171305	<0,0001
$E_{23}(G \times E)$	18,44	10940424,0	593260,61	1,6744272	0,0356
$E_{24}(G \times E)$	18,44	16740586,0	907782,95	2,5621395	0,0003
$E_{25}(G \times E)$	18,44	21066198,0	1142345,70	3,2241727	<0,0001
$E_{26}(G \times E)$	18,44	17427185,0	945014,82	2,6672233	0,0001
$E_{27}(G \times E)$	18,44	13639858,0	739641,45	2,0875746	0,0043
$E_{28}(G \times E)$	18,44	10275510,0	557204,67	1,5726624	0,0569
$E_{29}(G \times E)$	18,44	27022757,0	1465348,90	4,1358215	<0,0001
$E_{30}(G \times E)$	18,44	3011783,5	163318,40	0,4609522	0,9752
$E_{31}(G \times E)$	18,44	4996776,7	270957,59	0,7647545	0,7471
$E_{32}(G \times E)$	18,44	10857237,0	588749,69	1,6616955	0,0378
$E_{33}(G \times E)$	18,44	28897327,0	1567000,20	4,4227237	<0,0001
$E_{34}(G \times E)$	18,44	34056454,0	1846761,50	5,2123258	<0,0001

A Tabela 13 corresponde a ANOVA efetuada com o Conjunto 2 com 16 genótipos de milho avaliados em 24 ambientes com 4 blocos. Verifica-se que o efeito de blocos dentro de ambiente ($B d. E$), o efeito de genótipos (G), o efeito de ambientes (E) e o efeito da interação genótipos $\times$ ambientes ($G \times E$) são significativos ao nível de 5% de significância e suas somas de quadrados (SQ) correspondem a 20,4%, 2,03%, 55,6% e 11,96%, respectivamente, da soma de quadrados total.

Tabela 13 - ANOVA do Conjunto 2 com 16 genótipos de milho avaliados em 24 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	valor- <i>p</i>
B d. E	72	3065470183	42575975	30,56	<0,0001
G	15	304645583	20309706	14,58	<0,0001
E	23	8355949394	363302148	260,81	<0,0001
G×E	345	1797966686	5211498	3,74	<0,0001
Resíduo	1080	1504417193	1392979		
Total	1535	15028449039			

Através do teste F , aplicado a ANOVA, levando em consideração a decomposição dos $(g - 1)(e - 1) = (16 - 1)(24 - 1) = 345$ graus de liberdade da interação $G \times E$ com $[(g - 1)(e - 1)]/g = [(16 - 1)(24 - 1)]/16 \approx 21,56$ graus de liberdade para os genótipos, resulta na Tabela 14.

Tabela 14 - Teste F , aplicado ao Conjunto 2, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- <i>p</i>
$G_1(G \times E)$	21,56	71599087	3320537,4	2,3837672	0,0004
$G_2(G \times E)$	21,56	99958533	4635758,1	3,3279457	<0,0001
$G_3(G \times E)$	21,56	168626237	7820347,2	5,6141176	<0,0001
$G_4(G \times E)$	21,56	126055632	5846058,3	4,1968033	<0,0001
$G_5(G \times E)$	21,56	109059825	5057847,0	3,6309574	<0,0001
$G_6(G \times E)$	21,56	90085384	4177872,9	2,9992363	<0,0001
$G_7(G \times E)$	21,56	117225015	5436522,4	3,9028032	<0,0001
$G_8(G \times E)$	21,56	110027437	5102721,7	3,6631723	<0,0001
$G_9(G \times E)$	21,56	172730013	8010667,3	5,7507457	<0,0001
$G_{10}(G \times E)$	21,56	149907452	6952229,7	4,9909081	<0,0001
$G_{11}(G \times E)$	21,56	92199304	4275909,7	3,0696156	<0,0001
$G_{12}(G \times E)$	21,56	112873972	5234734,9	3,7579428	<0,0001
$G_{13}(G \times E)$	21,56	58513866	2713686,5	1,9481175	0,0059
$G_{14}(G \times E)$	21,56	107917084	5004850,3	3,5929118	<0,0001
$G_{15}(G \times E)$	21,56	106151931	4922988,1	3,5341441	<0,0001
$G_{16}(G \times E)$	21,56	105035914	4871230,8	3,4969883	<0,0001

Nota-se que todos os genótipos são significativos ao nível de 5% de significância e, portanto, todos contribuem significativamente para a interação genótipos $\times$ ambientes.

Agora, levando em consideração a decomposição dos mesmos 345 graus de liberdade da interação $G \times E$ com $[(g - 1)(e - 1)]/e = [(16 - 1)(24 - 1)]/24 \approx 14,37$ graus de liberdade para os ambientes com aplicação do teste F na ANOVA, resulta na Tabela 15.

Tabela 15 - Teste F , aplicado ao Conjunto 2, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- p
$E_1(G \times E)$	14,37	55565422	3865420,7	2,7749313	0,0004
$E_2(G \times E)$	14,37	51116700	3555944,4	2,5527626	0,0011
$E_3(G \times E)$	14,37	31738003	2207861,1	1,5849925	0,0746
$E_4(G \times E)$	14,37	64888735	4513998,9	3,2405365	<0,0001
$E_5(G \times E)$	14,37	55280460	3845597,2	2,7607003	0,0004
$E_6(G \times E)$	14,37	100226291	6972263,7	5,0052903	<0,0001
$E_7(G \times E)$	14,37	268517980	18679512,0	13,4097590	<0,0001
$E_8(G \times E)$	14,37	32780602	2280389,7	1,6370598	0,0615
$E_9(G \times E)$	14,37	462315191	32161057,0	23,0879710	<0,0001
$E_{10}(G \times E)$	14,37	54322077	3778927,1	2,7128388	0,0005
$E_{11}(G \times E)$	14,37	50397371	3505904,0	2,5168393	0,0014
$E_{12}(G \times E)$	14,37	52354311	3642039,1	2,6145688	0,0008
$E_{13}(G \times E)$	14,37	69400027	4827828,0	3,4658300	<0,0001
$E_{14}(G \times E)$	14,37	28748176	1999873,1	1,4356808	0,1269
$E_{15}(G \times E)$	14,37	53542558	3724699,7	2,6739097	0,0006
$E_{16}(G \times E)$	14,37	71116118	4947208,2	3,5515314	<0,0001
$E_{17}(G \times E)$	14,37	32769005	2279583,0	1,6364806	0,0616
$E_{18}(G \times E)$	14,37	36566300	2543742,6	1,8261172	0,0293
$E_{19}(G \times E)$	14,37	44491325	3095048,7	2,2218921	0,0054
$E_{20}(G \times E)$	14,37	12761891	887783,7	0,6373275	0,8392
$E_{21}(G \times E)$	14,37	14199903	987819,4	0,7091417	0,7707
$E_{22}(G \times E)$	14,37	25672412	1785906,9	1,2820775	0,2093
$E_{23}(G \times E)$	14,37	41565902	2891541,0	2,0757967	0,0102
$E_{24}(G \times E)$	14,37	87629925	6095994,8	4,3762291	<0,0001

Observa-se que somente os ambientes 3, 8, 14, 17, 20, 21 e 22 não contribuem para a interação genótipos $\times$ ambientes, sendo estes os ambientes mais estáveis aos genótipos, enquanto que os demais ambientes contribuem significativamente, ao nível de 5% de significância, para a interação.

A Tabela 16 corresponde a ANOVA efetuada com o Conjunto 3 com 9 genótipos de milho avaliados em 20 ambientes com 4 blocos. Verifica-se que o efeito de blocos dentro de ambiente ($B d. E$), o efeito de genótipos (G), o efeito de ambientes (E) e o efeito da interação genótipos $\times$ ambientes ($G \times E$) são significativos ao nível de 5% de significância e suas somas de quadrados (SQ) correspondem a 6,88%, 4,62%, 57,29% e 14,46%, respectivamente, da soma de quadrados total.

Tabela 16 - ANOVA do Conjunto 3 com 9 genótipos de milho avaliados em 20 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	valor- p
B d. E	60	118813053,8	1980217,6	3,28	<0,0001
G	8	79828574,7	9978571,8	16,55	<0,0001
E	19	989593771,8	52083882,7	86,40	<0,0001
$G \times E$	152	249704161,7	1642790,5	2,73	<0,0001
Resíduo	480	289366499,0	602847,0		
Total	719	1727306061,0			

Através do teste F , aplicado a ANOVA, levando em consideração a decomposição dos $(g-1)(e-1) = (9-1)(20-1) = 152$ graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/g = [(9-1)(20-1)]/9 \approx 16,89$ graus de liberdade para os genótipos, resulta na Tabela 17.

Tabela 17 - Teste F , aplicado ao Conjunto 3, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- p
$G_1(G \times E)$	16,89	35147999	2081131,50	3,4521727	<0,0001
$G_2(G \times E)$	16,89	17268730	1022490,60	1,6961034	0,0405
$G_3(G \times E)$	16,89	10872797	643784,06	1,0679064	0,3829
$G_4(G \times E)$	16,89	42526664	2518026,20	4,1768918	<0,0001
$G_5(G \times E)$	16,89	26242317	1553821,40	2,5774728	0,0006
$G_6(G \times E)$	16,89	15653251	926837,25	1,5374339	0,0779
$G_7(G \times E)$	16,89	22544078	1334846,70	2,2142384	0,0037
$G_8(G \times E)$	16,89	53853715	3188706,80	5,2894142	<0,0001
$G_9(G \times E)$	16,89	25594611	1515470,40	2,5138562	0,0008

Tem-se para o Conjunto 3 que somente os genótipos 3 e 6 são não significativos ao nível de 5% de significância, sendo assim, os demais genótipos são os que contribuem mais para a interação genótipos $\times$ ambientes.

Agora, levando em consideração a decomposição dos mesmos 152 graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/e = [(9-1)(20-1)]/20 = 7,6$ graus de liberdade para os ambientes com aplicação do teste F na ANOVA, resulta na Tabela 18.

Tabela 18 - Teste F , aplicado ao Conjunto 3, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- p
$E_1(G \times E)$	7,6	13515978,0	1778418,10	2,9500329	0,0037
$E_2(G \times E)$	7,6	4141059,8	544876,29	0,9038386	0,5092
$E_3(G \times E)$	7,6	18493555,0	2433362,50	4,0364521	0,0002
$E_4(G \times E)$	7,6	5068124,9	666858,55	1,1061823	0,3576
$E_5(G \times E)$	7,6	5877124,1	773305,80	1,2827566	0,2528
$E_6(G \times E)$	7,6	7048106,0	927382,37	1,5383382	0,1455
$E_7(G \times E)$	7,6	5005552,9	658625,38	1,0925252	0,3668
$E_8(G \times E)$	7,6	72882752,0	9589835,80	15,9075810	<0,0001
$E_9(G \times E)$	7,6	4798500,1	631381,60	1,0473333	0,3984
$E_{10}(G \times E)$	7,6	8310938,0	1093544,50	1,8139672	0,0762
$E_{11}(G \times E)$	7,6	23728096,0	3122117,90	5,1789569	<0,0001
$E_{12}(G \times E)$	7,6	9093533,4	1196517,60	1,9847786	0,0499
$E_{13}(G \times E)$	7,6	18976722,0	2496937,00	4,1419093	0,0001
$E_{14}(G \times E)$	7,6	377195,3	49630,96	0,0823276	0,9995
$E_{15}(G \times E)$	7,6	11592125,0	1525279,70	2,5301278	0,0120
$E_{16}(G \times E)$	7,6	2537423,7	333871,54	0,5538248	0,8071
$E_{17}(G \times E)$	7,6	8212795,4	1080631,00	1,7925464	0,0802
$E_{18}(G \times E)$	7,6	16303900,0	2145250,00	3,5585321	0,0006
$E_{19}(G \times E)$	7,6	9205089,7	1211196,00	2,0091271	0,0470
$E_{20}(G \times E)$	7,6	4535590,5	596788,23	0,9899499	0,4409

Observa-se que os ambientes 2, 4, 5, 6, 7, 9, 10, 14, 16, 17 e 20 são os que não contribuem para a interação genótipos $\times$ ambientes, enquanto que os demais ambientes contribuem significativamente, ao nível de 5% de significância, para a interação.

A Tabela 19 corresponde a ANOVA efetuada com o Conjunto 4 com 8 genótipos de milho avaliados em 59 ambientes com 4 blocos. Verifica-se que o efeito de blocos dentro de ambiente ($B d. E$), o efeito de genótipos (G), o efeito de ambientes (E) e o efeito da interação genótipos $\times$ ambientes ($G \times E$) são significativos ao nível de 5% de significância e suas somas de quadrados (SQ) correspondem a 6,06%, 1,66%, 76,01% e 5,28%, respectivamente, da soma de quadrados total.

Tabela 19 - ANOVA do Conjunto 4 com 8 genótipos de milho avaliados em 59 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	valor- p
B d. E	177	301817526	1705184	3,86	<0,0001
G	7	82527757	11789680	26,66	<0,0001
E	58	3787154892	65295774	147,67	<0,0001
$G \times E$	406	263183718	648236	1,47	<0,0001
Resíduo	1239	547865274	442183		
Total	1887	4982549167			

Através do teste F , aplicado a ANOVA, levando em consideração a decomposição dos $(g-1)(e-1) = (8-1)(59-1) = 406$ graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/g = [(8-1)(59-1)]/8 = 50,75$ graus de liberdade para os genótipos, resulta na Tabela 20.

Tabela 20 - Teste F , aplicado ao Conjunto 4, para obtenção de genótipos que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	valor- p
$G_1(G \times E)$	50,75	32732736	644980,01	1,4586255	0,0208
$G_2(G \times E)$	50,75	34895741	687600,80	1,5550126	0,0083
$G_3(G \times E)$	50,75	31377221	618270,37	1,3982215	0,0356
$G_4(G \times E)$	50,75	34604938	681870,70	1,5420539	0,0094
$G_5(G \times E)$	50,75	28872097	568908,31	1,2865889	0,0878
$G_6(G \times E)$	50,75	24990248	492418,68	1,1136072	0,2737
$G_7(G \times E)$	50,75	38238525	753468,47	1,7039726	0,0017
$G_8(G \times E)$	50,75	37472213	738368,73	1,6698245	0,0025

Nota-se que somente os genótipos 5 e 6 não contribuem para a interação genótipos $\times$ ambientes, enquanto que os demais genótipos são significativos ao nível de 5% de significância e contribuem mais para a interação.

Agora, levando em consideração a decomposição dos mesmos 406 graus de liberdade da interação $G \times E$ com $[(g-1)(e-1)]/e = [(8-1)(59-1)]/59 \approx 6,88$ graus de liberdade para os ambientes com aplicação do teste F na ANOVA, resulta na Tabela 21:

Tabela 21 - Teste F , aplicado ao Conjunto 4, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$

(continua)					
Fonte de Variação	GL	SQ	QM	F	valor- p
$E_1(G \times E)$	6,88	1315190,2	191123,70	0,4322272	0,8797
$E_2(G \times E)$	6,88	4347385,0	631762,85	1,4287348	0,1908
$E_3(G \times E)$	6,88	5369858,5	780348,89	1,7647628	0,0920
$E_4(G \times E)$	6,88	3847128,9	559065,53	1,2643294	0,2653
$E_5(G \times E)$	6,88	7829528,9	1137788,70	2,5731147	0,0129
$E_6(G \times E)$	6,88	8744661,6	1270775,90	2,8738660	0,0059
$E_7(G \times E)$	6,88	2903547,1	421944,03	0,9542285	0,4625
$E_8(G \times E)$	6,88	4017131,3	583770,32	1,3201994	0,2378
$E_9(G \times E)$	6,88	3472897,1	504682,09	1,1413410	0,3346
$E_{10}(G \times E)$	6,88	4491941,1	652769,77	1,4762420	0,1729
$E_{11}(G \times E)$	6,88	2905901,6	422286,19	0,9550023	0,4620
$E_{12}(G \times E)$	6,88	2236501,3	325008,80	0,7350090	0,6300
$E_{13}(G \times E)$	6,88	3294071,6	478695,13	1,0825714	0,3719
$E_{14}(G \times E)$	6,88	3609742,0	524568,42	1,1863140	0,3079
$E_{15}(G \times E)$	6,88	7807391,7	1134571,70	2,5658395	0,0131

Tabela 21 - Teste F , aplicado ao Conjunto 4, para obtenção de ambientes que contribuem significativamente para a interação $G \times E$

Fonte de Variação	GL	SQ	QM	F	(conclusão)
					valor- p
$E_{16}(G \times E)$	6,88	1868065,9	271467,70	0,6139255	0,7420
$E_{17}(G \times E)$	6,88	29088277,0	4227114,10	9,5596393	<0,0001
$E_{18}(G \times E)$	6,88	3515808,8	510918,03	1,1554436	0,3260
$E_{19}(G \times E)$	6,88	5773918,2	839066,93	1,8975540	0,0678
$E_{20}(G \times E)$	6,88	12286633,0	1785496,00	4,0379079	0,0002
$E_{21}(G \times E)$	6,88	1512292,2	219766,60	0,4970032	0,8343
$E_{22}(G \times E)$	6,88	2946767,6	428224,85	0,9684326	0,4519
$E_{23}(G \times E)$	6,88	1284283,9	186632,39	0,4220701	0,8863
$E_{24}(G \times E)$	6,88	644492,0	93657,71	0,2118074	0,9819
$E_{25}(G \times E)$	6,88	1954132,4	283974,90	0,6422106	0,7184
$E_{26}(G \times E)$	6,88	3608892,3	524444,94	1,1860348	0,3080
$E_{27}(G \times E)$	6,88	1856906,6	269846,03	0,6102581	0,7450
$E_{28}(G \times E)$	6,88	1391280,8	202181,20	0,4572338	0,8629
$E_{29}(G \times E)$	6,88	2464782,1	358182,62	0,8100317	0,5770
$E_{30}(G \times E)$	6,88	1474710,2	214305,18	0,4846522	0,8434
$E_{31}(G \times E)$	6,88	1135401,4	164996,75	0,3731410	0,9159
$E_{32}(G \times E)$	6,88	2467670,2	358602,31	0,8109809	0,5762
$E_{33}(G \times E)$	6,88	12961882,0	1883623,30	4,2598232	0,0001
$E_{34}(G \times E)$	6,88	8672098,8	1260231,10	2,8500188	0,0063
$E_{35}(G \times E)$	6,88	2739227,7	398065,11	0,9002262	0,5041
$E_{36}(G \times E)$	6,88	5331768,1	774813,59	1,7522447	0,0946
$E_{37}(G \times E)$	6,88	6736266,2	978915,53	2,2138223	0,0317
$E_{38}(G \times E)$	6,88	4723904,1	686478,67	1,5524749	0,1470
$E_{39}(G \times E)$	6,88	2580797,7	375042,02	0,8481594	0,5457
$E_{40}(G \times E)$	6,88	9139837,8	1328203,00	3,0037377	0,0042
$E_{41}(G \times E)$	6,88	2569609,5	373416,15	0,8444825	0,5487
$E_{42}(G \times E)$	6,88	2531778,3	367918,51	0,8320495	0,5589
$E_{43}(G \times E)$	6,88	1821980,1	264770,50	0,5987798	0,7545
$E_{44}(G \times E)$	6,88	5943722,2	863742,88	1,9533587	0,0594
$E_{45}(G \times E)$	6,88	6446685,6	936833,63	2,1186538	0,0400
$E_{46}(G \times E)$	6,88	3171265,5	460848,93	1,0422121	0,3991
$E_{47}(G \times E)$	6,88	8980210,8	1305006,00	2,9512775	0,0048
$E_{48}(G \times E)$	6,88	6755165,1	981661,92	2,2200333	0,0313
$E_{49}(G \times E)$	6,88	1167061,5	169597,61	0,3835458	0,9000
$E_{50}(G \times E)$	6,88	1964439,4	285472,72	0,6455980	0,7156
$E_{51}(G \times E)$	6,88	2286134,1	332221,46	0,7513205	0,6262
$E_{52}(G \times E)$	6,88	2243040,0	325959,01	0,7371579	0,6381
$E_{53}(G \times E)$	6,88	2206637,9	320669,05	0,7251946	0,6483
$E_{54}(G \times E)$	6,88	5361318,7	779107,89	1,7619563	0,0926
$E_{55}(G \times E)$	6,88	1920834,5	279136,05	0,6312675	0,7276
$E_{56}(G \times E)$	6,88	1461669,8	212410,14	0,4803666	0,8465
$E_{57}(G \times E)$	6,88	3907132,0	567785,19	1,2840490	0,2553
$E_{58}(G \times E)$	6,88	3589662,0	521650,39	1,1797149	0,3117
$E_{59}(G \times E)$	6,88	4502396,7	654289,17	1,4796782	0,1716

Observa-se que os ambientes 5, 6, 15, 17, 20, 33, 34, 37, 40, 45, 47 e 48 são significativos ao nível de 5% de significância e são os que contribuem para a interação genótipos $\times$ ambientes. Enquanto que a maioria dos ambientes não contribuem para a interação.

5 CONCLUSÕES

Os resultados obtidos nessa pesquisa permitem extrair as seguintes conclusões:

- a) As somas de quadrados das linhas e colunas da matriz de interação genótipos $\times$ ambientes seguem uma distribuição qui-quadrado não central;
- b) Uma combinação entre o método de Kolmogorov-Smirnov e os gráficos histograma e QQplot foi importante para verificar que todas as somas de quadrados seguem a distribuição qui-quadrado não-central;
- c) O teste proposto para detectar quais genótipos e ambientes contribuem mais para a interação é de fácil aplicação e interpretação;
- d) O teste proposto para detectar quais genótipos e ambientes contribuem mais para a interação não requer grandes quantidades de cálculos e são facilmente implementados em computador;
- e) O teste aqui proposto pode ser aplicado a qualquer estrutura fatorial envolvendo dois fatores.

SUGESTÕES PARA NOVAS PESQUISAS

Uma generalização do teste proposto para testar a contribuição dos níveis dos fatores para a interação de qualquer ordem, deve ser o próximo passo dessa pesquisa.

REFERÊNCIAS

- ALLARD, R.W. **Princípios do melhoramento genético das plantas**. Rio de Janeiro: USAID/Edgard Blucher, 1971. 381p.
- ALLARD, R.W. BRADSHAW, A.D. Implications of genotype-environmental interactions in applied plant breeding. **Crop Science**, Madison, v.4, n.5, p.503-508, 1964.
- ALVAREZ, M.P.; EYHERÁBIDE, G.H. Estabilidad del rendimiento de cultivares de híbridos comerciales de maíz en el área de la EEA Pergamino. **Revista de Tecnología Agropecuaria**, Pergamino, v.1, n.2, p.17-21, 1996.
- ARAÚJO, L.B. **Métodos de correção de autovalores e regressão isotônica nos modelos AMMI**. 2005. 75p. Dissertação (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2005.
- BARBIN, D. **Planejamento e análise estatística de experimentos agronômicos**. Arapongas: Midas, 2003. 194p.
- BOIK, R.J. Testing the rank of a matrix with applications to the analysis of interaction in ANOVA. **Journal of the American Statistical Association**, Boston, v.81, p.243-248, 1986.
- 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.
- CORNELIUS, P.L. Functions approximating Mandel's tables for the means and standard deviations of the first three roots of a Wishart matrix. **Technometrics**, Alexandria, v.22, p.613-616, 1980.
- CORNELIUS, P.L. Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. **Crop Science**, Madison, v.33, p.1186-1193, 1993.
- CORNELIUS, P.L.; CROSSA, J. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment trials. **Crop Science**, Madison, v.39, p.998-1009, 1999.
- CORNELIUS, P.L.; SEYEDSADR, M.; CROSSA, J. Using the shifted multiplicative model to search for "separability" in crop cultivar trials. **Theoretical and Applied Genetics**, New York, v.84, p.161-172, 1992.
- CROSSA, J. Statistical Analyses of multilocation trials. **Advances in Agronomy**, New York, v.44, p.55-85, 1990.
- CROSSA, J.; FOX, P. N.; PFEIFFER, W. H.; RAJARAM, S.; GAUCH, H. G. AMMI adjustment for statistical analysis of an international wheat yield trial. **Theoretical and Applied Genetics**, New York, v.81, p.27-37, 1991.
- DIAS, C.T.S.; KRZANOWSKI, W.J. Model selection and cross validation in additive main effect and multiplicative interaction models. **Crop Science**, Madison, v.43, p.865-873, 2003.
- DUARTE, J.B.; VENCOVSKY, R. **Interação genótipo × ambiente: uma introdução à análise "AMMI"**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 60p. (Série Monografia, 9).

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

FALCONER, D.S. **Introdução à genética quantitativa**. Trad. de Almeida, M. e Silva, J.C. Viçosa: UFV, 1987. 279p.

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

FISHER, R.A.; MACKENZIE, W.A. Studies in crop variation. II. The manurial response of different potato varieties. **Journal of Agricultural Science**, Cambridge, v.23, p.311-320, 1923.

FREEMAN, G.H. Statistical methods for the analysis of genotype-environment interactions. **Heredity**, London, v.31, p. 399-454, 1973.

GABRIEL, K.R. The biplot graphic display of matrices with application to principal component analysis. **Biometrika**, Cambridge, v.58, p.453-467, 1971.

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

GAUCH, H.G. Model selection and validation for yield trials with interactions. **Biometrics**, Washington, v.44, p.705-715, 1988.

GAUCH, H.G. **MATMODEL version 2.0**: AMMI and related analyses for two-way data matrices. Microcomputer Power, Ithaca, New York, 1990.

GAUCH, H.G.; FURNAS, R.E. Statistical analysis of yield trials with MATMODEL. **Agro-nomy Journal**, Madison, v.83, p.916-920, 1991.

GAUCH, H.G.; ZOBEL, R.W. Imputing missing yield trial data. **Theoretical and Applied Genetics**, New York, v.79, p.753-761, 1990.

GAUCH Jr, H.G. **Statistical analysis of regional yield trials**: AMMI analysis of factorial designs. Amsterdam: Elsevier, 1992, 277p.

GAUCH Jr, H.G.; ZOBEL, R.W. Predictive and postdictive success of statistical analysis of yield trials. **Theoretical and Applied Genetics**, New York, v.76, p.1-10, 1988.

GAUCH Jr, H.G.; ZOBEL, R.W. AMMI analysis of yield trials. IN: KANG, M.D.; GAUCH Jr, H.G. **Genotype-by-environment interaction**, Boca Raton: CRC Press, 1996. chap.4, p.85-112.

GOLLOB, H.F. A statistical model which combines features of factor analytic and analysis of variance techniques. **Psychometrika**, New York, v.33, p.73-115, 1968.

GONZALEZ, L.M.A.R. **Estudo da interação de genótipos × ambientes para a cultura do algodão (*Gossypium* sp) na região oriental do Paraguai**. 1988. 91p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1988.

GOODMAN, L.A; HABERMAN, S.J. The analysis of nonadditivity in two-way analysis of variance. **Journal of the American Statistical Association**, Boston, v.85, p.139-145, 1990.

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

HEGEMANN, V.; JOHNSON, D.E. On analyzing two-way AoV data with interaction. **Technometrics**, Alexandria, v.18, n.3, p.273-281, 1976.

HOGG, R.V.; CRAIG, A.T. **Introduction to mathematical statistics**. New York: Macmillan Publishing, 1978, 438p.

JOHNSON, D.E. Some new multiple comparison procedures for the two-way AoV model with interaction. **Biometrics**, Washington, v.32, p.929-934, 1976.

JOHNSON, D.E.; GRAYBILL, F.A. An analysis of a two-way model with interaction and no replication. **Journal of the American Statistical Association - Theory and Methods**, Boston, v.67, n.340, p.862-868, 1972.

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions**. New York: Wiley, 1995, v.2, 719p.

KANG, M.S. Using genotype-by-environment interaction for crop cultivar development. **Advances in Agronomy**, New York, v.62, p.199-252, 1998.

KEMPTON, R.A. The use of biplots in interpreting variety by environment interactions. **Journal of Agricultural Science**, Cambridge, v.103, p.123-135, 1984.

KRZANOWSKI, W.J. Some exact percentage points of a statistic useful in analysis of variance and principal component analysis. **Technometrics**, Alexandria, v.21, p.261-263, 1979.

LAVORANTI, O.J. **Estabilidade e adaptabilidade fenotípica através da reamostragem “bootstrap” no modelo AMMI**. 2003. 166p. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2003.

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

MAIA, M.C.C. **Adaptabilidade e estabilidade de linhagens experimentais de soja selecionadas para caracteres agronômicos e tolerância a insetos, através de método uni-multivariado com reamostragem**. 2004. 124p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2004.

MANDEL, J. Non-additivity in two-way analysis of variance. **Journal of the American Statistical Association**, Boston, v.56, p.878-888, 1961.

MANDEL, J. The partitioning of interactions in analysis of variance. **Journal of Research of the National Bureau of Standards Series B**, Washington, v.73, p.309-328, 1969.

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

MIRANDA, F.T.S. **Métodos de regressão e uni-multivariados para a redução do número de repetições em experimentos intermediários de um programa de melhoramento de soja**. Piracicaba, 2004. 110 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

MOSER, B.K. **Linear models: a mean model approach**. San Diego: Academic Press, 1996, 228p.

NUNES, G.H.S. **Interação genótipos $\times$ ambientes em eucalipto: implicações sobre a seleção e formas de atenuar seu efeito**. 2000. 160p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Universidade Federal de Lavras, Lavras, 2000.

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

PERKINS, J.M.; JINKS, J.L. Environmental and genotype $\times$ environmental components of variability. III. Multiple lines and crosses. **Heredity**, London, v.23, p.339-456, 1968.

PIEPHO, H.P. Best Linear Unbiased Prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interactions (AMMI) analysis. **Theoretical and Applied Genetics**, New York, v.89, p.647-654, 1994.

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

PIKE, E.W.; SILVERBERG, T.R. Designing mechanical computers. **Machine Desing**, Cleveland, v.24, p.131-137, 159-163, 1952.

PIMENTEL-GOMES, F. **Curso de estatística experimental**. Piracicaba: Nobel. 2000. 477p.

PIOVESAN, P. **Validação cruzada com correção de autovalores e regressão isotônica nos modelos AMMI**. 2007. 77p. Dissertação (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2007.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, 2007.

SAS Institute inc. **SAS 9.1** Cary, 2003.

SOUZA, J. de. **Análise em componentes principais e suas aplicações: métodos estatísticos nas ciências psicossociais**. Brasília: Thesaurus. 1988, v.2, 68p.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sbg, 1992. 486p.

WEBER, W.E.; WRICKE, G.; WESTERMANN, T. Selection of genotype and prediction of performance by analysis genotype-by-environment interactions. In: KANG, M.S; GAUCH, H.G. **Genotype by environment interactions**. Boca Raton: CRC Press, 1996. chap.13, p.353-371.

WILLIAMS, E.J. The interpretation of interactions in factorial experiments. **Biometrika**, Cambridge, v.39, p.65-81, 1952.

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

YAN, W.; KANG, M.S. **GGE biplot analysis: a graphical tool for breeders geneticists, and agronomists**. Boca Raton: CRC Press, 2002, 271p.

YOCHMOWITZ, M.G.; CORNELL, R.G. Stepwise tests for multiplicative components of interaction. **Technometrics**, Alexandria, v.20, p.79-84, 1978.

ZOBEL, R.W.; WRIGHT, M.J.; GAUCH Jr, H.G Statistical analysis of a yield trial. **Agro-nomy Journal**, Madison, v.80 , p.388-393, 1988.

ANEXOS

ANEXO A - Programa para simulação dos experimentos

```

data U;
u=7201.925; *média geral;

array G {20} (-34.05 -535.225 -1064.25 -142.175 -361.525 -458.075 -940.1
-620.5 0 73.6 -551.25 -1484.375 -956.1 -1532.1 -1142.875 -1610.5
-1529.275 -309.775 -1103.8 -1803.65);

array E {34} (-5230.9 -5885.675 -3679.425 -2489.425 -1698.175 -3521.7
-1475.675 -1536.925 0 -3903.175 -4101.925 -2797.425 -4821.925 -1752.7
-3179.175 -2020.175 -1513.8 -2164.425 -2522.925 -2762.75 -3956.925
-4043.175 -4896.65 -2659.425 -2794.425 -4390.675 -2115.675 -4086.925
-1788.175 -5501.925 -5350.675 -3262.5 -1994.425 -1864.85);

array GE {20,34} (742.875 -324.7 -324.2 -1032.2 -713.45 -791.85
1224.05 -434.7 0 390.3 60.3 -388.575 0.3 -632.175 -274.95 -231.7
-155.45 876.55 -786.95 792.275 118.2 67.8 -272.075 -1133.45 -739.7
-28.45 -322.2 1.55 -783.45 -238.45 96.55 132.35 -922.2 -688.975
1085.45 440.225 713.075 678.975 1223.975 1164.2 488.975 322.725 0
67.725 911.475 1788.15 658.975 573.625 198.35 749.35 292.85 556.475
667.225 1663.525 278.55 218.975 594.65 6.475 -541.025 486.475 112.725
1021.475 -37.275 632.725 728.975 545.925 843.975 -153.75 1018.225 983
924.475 506.75 309.25 1178.225 1016.75 988 0 1209.25 1388 949.15 935.5
1018.85 660.75 1259.25 843.125 931.75 1254.75 1301.05 927.6 1208 1097.825
288 336.75 1874.25 251.75 1078 -405.75 1014.25 946.75 1316.45 -322
492.725 428.65 205.925 311.5 108.425 623.425 -196.075 -90.325 288.425 0
72.175 534.675 180.45 -144.075 122 31.8 139.3 -431.325 612.175 274.175
365.4 255.525 478.425 348 -1054.075 -529.075 258.425 203.425 420.925
-1106.575 -200.325 -137.825 -108.975 533.425 -316.7 1133 -172.225 114.45
116.525 485.275 652.875 849.025 7.775 0 -179.725 14.025 13.675 747.775
-202.725 31.025 738.275 662.275 1059.025 157.025 1675.45 364.05 1079.025
416.7 119.025 207.775 407.775 486.525 501.525 376.525 236.525 500.275
1044.825 736.525 488.775 608.3 543.075 126.275 314.325 591.825 1247.025
270.575 491.825 0 167.875 549.325 654.925 385.575 791.85 386.075 973.7
1.7 309.325 777.825 769.575 -163.6 163.075 778.775 518.075 -439.425
344.325 -134.425 636.825 -1045.675 116.825 329.325 1125.175 1018.075
-82.925784.075 1470.1 1326.675 768.85 1928.85 1279.925 -137.4 133.85
0 1070.1 1177.6 1373.35 846.35 1802.075 1573.725 2287.1 1331.6 2243.85
1774.1 1010.525 1184.275 813.85 1790.45 488.85 -1.15 1981.35 136.35 775.1
-63.65 750.1 1097.6 1507.8 762.6 197 -285.525 1098 1357.3 860.5 1358
587.625 -335.75 313 0 438 816.75 124.175 609.25 746.3 477.375 995.375
379.5 1309.25 371.5 1096.6 223.025 73 106.675 -487 56.75 1000.5 426.75
394.25 -945.75 595.5 678 1467.425 745.5 -575.125 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 424.125 -386.1 -273.825 163.9
-279.85 -393.075 -547.35 172.65 0 -301.1 313.9 413.575 -453.6 559.55
-1124.475 1008.025 37.025 963.9 -236.1 797.9 393.9 302.65 620.425 -1378.6
-1098.6 1235.15 407.65 453.9 -1788.6 -319.85 510.15 154.175 48.9 -38.125
572.725 488.75 687.225 416.25 287.5 590 851.25 323.75 0 805 605 671.1
263.75 1060.875 65.875 1521.5 778.25 1055 866.25 996.05 86.225 228.75
1003.825 113.75 -716.25 1691.25 166.25 505 -1282.5 47.5 210 1206.3 1167.5
-217.85 1439.6 1296.875 1618 243.125 -861.875 1167.2 316.875 639.375 0
855.625 1324.375 1239.975 1375.625 789.45 1314.5 2029.5 1141 1168.125
630.875 1658.675 1501.875 2224.375 1580.625 504.375 -205.625 2204.375
1475.625 2058.125 936.875 1339.375 1621.875 1630.95 1273.125 -228.575

```

```

803.825 1098.6 974.45 1128.6 1422.35 501.675 1176.1 931.1 0 709.85
1496.1 1217.65 758.6 1562.15 648.975 1925.85 913.1 761.1 1070.1 880.775
580.275 811.1 1354.275 299.85 332.35 2126.1 1043.6 804.85 -667.65 577.35
1152.35 824.025 1597.35 -281.85 1672.325 1258.35 1367.475 2224.6 1715.85
3034.925 720.85 817.1 0 1185.85 1973.35 2068.775 1602.1 1387.425 1702.6
1402.6 1128.475 1064.6 659.1 955.4 1368.775 1799.6 1593.85 1850.85
2088.35 2339.6 788.35 2149.6 644.6 1532.1 1789.6 2486.725 1833.35
1235.8 1748.1 1036.625 1174.6 1622.875 1245.375 1822.4 564.125 714.125
0 761.625 766.625 1861.8 1121.625 1152.1 882.75 1129.75 1042.5
1017.875 289.375 1488.375 1040.375 1657.875 1512.3 -229.625 601.625
1189.125 796.625 1670.375 -65.875 806.625 1480.375 1644.425 1249.125
280.6 2175.725 1538 1554.725 1136.75 679.25 2432.4 368 683 0 1326.75
1450.5 2625.35 1288 1693.275 1962.75 2399.875 1654.125 1793 901.5 1488.6
2144.675 2386.75 2094.1 924.25 1516.75 2093 1850.5 2259.25 -358.25
1440.5 1975.5 2760.5 1877.625 1958.725 2672 1050.525 1475.5 2424.275
1514.275 1573.2 588.025 931.775 0 1038.025 1390.525 1489.275 1139.275
1412.975 905.525 1292.4 1140.275 229.275 131.775 768.975 1517.6 2294.275
1769.375 50.525 855.525 2343.025 1473.025 1946.775 113.025 1421.775
1820.525 1773.05 2049.275 1756.675 135 163.525 873.4 168.525 397.275
861.6 124.775 57.275 0 -508.975 507.275 130.775 824.775 215.6 673.275
-32.975 715.15 949.775 651.275 1644.275 33.1 -2.725 664.35 -260.225
-436.475 906.025 456.025 747.275 -31.475 342.275 641.025 1071.375 69.775
274.4 1197.775 606.3 1198.65 1781.3 898.8 2038.6 320.05 786.3 0 815.05
1262.55 2036.75 1221.3 1679.8 228.3 1231.8 474.45 78.8 463.55 179.875
1226.3 1858.8 1772.35 177.55 885.05 1618.8 790.05 1831.3 -706.2 737.55
1666.3 1621.95 2201.3 759.5 1087.625 1199.9 1264.75 -615.1 -43.85
204.35 1832.4 426.15 0 1943.65 2332.4 2218.05 1624.9 113.875 1334.4
1421.275 648.525 2633.65 1724.4 2536.125 1510.325 867.4 2520.15 -20.1
-545.1 2204.9 1957.4 2166.15 196.15 1151.15 1628.65 2022.025 -106.35
-1407.3);

```

```

do exp=1 to 10000;
do i=1 to 20; *gen;
do j=1 to 34; *site;
  Y=U+G[i]+E[j]+GE[i,j]+ 297.61*rannor(0);
  output;
end;
end;
end;

```

```

proc sort data=U;
by Exp i j;
run;

```

```

proc means data=U noprint;
by Exp i j;
var Y;
output out = medias mean=YM;
run;

```

```

proc iml;
use medias;
read all into YM;
Y_data=YM;
r=4;
g=max(y_data[,2]);

```

```

e=max(y_data[,3]);
nEXP=max(y_data[,1]);
Y=shape(y_data[,6],nEXP*g,e);

*média de genotipos;
Yi=j(nEXP*g,e,0);
do i=1 to nEXP*g;
Yi[i,]=j(1,e, sum(Y[i,])/e );
end;

*média de ambientes;
Yj=j(g*nEXP,e,0);
do l=1 to nEXP;
do j=1 to e;
Yj[(1+g*(l-1)):(g+g*(l-1)),j] = j (g, 1, sum(Y[(1+g*(l-1)):(g+g*(l-1)),j])/g);
end;
end;

*média geral;
Y_barra=j(g*nEXP,e,0);
do l=1 to nEXP;
Y_barra[(1+g*(l-1)):(g+g*(l-1)),1:e] = j(g,e,sum(Y[(1+g*(l-1)):(g+g*(l-1)),1:e])/(g*e));
end;

*Interação
GE = Y - Yi - Yj + Y_barra;

*Soma de quadrados de cada genotipo;
SQGi=j(g*nExp,1,0);
do i=1 to g*nExp;
SQGi[i,]=r*sum(GE[i,]##2);
end;
SQGi=t(shape(SQGi,nExp,g));

*Soma de quadrados de cada ambiente;
SQEj=j(e,nExp,0);
do l=1 to nEXP;
do j=1 to e;
SQEj[j,1]=r*sum(GE[(1+g*(l-1)):(g+g*(l-1)),j]##2);
end;
end;
SQEj=t(SQEj);

quit;

```

ANEXO B - Programa para aplicação do Teste F modificado

```

proc glm data=TRIAL1 outstat=anavaexp;
class REP SITE GEN;
model YIELD= REP(SITE) GEN SITE GEN*SITE/ss3;
run;
*Procurando QMRes na Anova;
data anavaexp2;
set anavaexp;
if _source_ = 'ERROR';
SQRes=ss;
gl_Res=df;
QMRes=SQRes/gl_Res;
keep gl_Res QMRes;

*Procurando QMInt na Anova;
data anavaexp3;
set anavaexp;
if _source_ = 'SITE*GEN';
SQInt=ss;
gl_Int=df;
QMInt=SQInt/gl_Int;
keep gl_Int QMInt;

*Obter a tabela de dupla entrada dos residual (GE);
proc sort data=TRIAL1;
by site gen;
run;

proc means data=TRIAL1 noprint;
by site gen;
var YIELD;
output out = medias mean=YIELDM;
run;

proc sort data=medias;
by gen site;
run;

proc transpose data=medias out = medias2;
by gen;
id site;
var YIELDM;
run;

proc iml;
use anavaexp2;
read var {QMRes} into QMRes;
read var {gl_Res} into gl_Res;
use anavaexp3;
read var {QMInt} into QMInt;
read var {gl_Int} into gl_Int;
use medias2;
read all into YIELDM;
Y=YIELDM[,2:ncol(YIELDM)];

```

```

g=nrow(Y);
e=ncol(Y);
r=4;

Yi=j(g,e,0);
do i=1 to g;
Yi[i,]=j(1,e,sum(Y[i,])/e);
end;

Yj=j(g,e,0);
do j=1 to e;
Yj[,j]=j(g,1,sum(Y[,j])/g);
end;

Y_barra = j(g,e,sum(Y)/(g*e));
GE = Y - Yi - Yj + Y_barra;

*Soma de quadrados de cada gen e cada amb;
SQGi=j(g,1,0);
do i=1 to g;
SQGi[i,]=r*sum(GE[i,]##2);
end;

SQEj=j(e,1,0);
do j=1 to e;
SQEj[j,]=r*sum(GE[,j]##2);
end;

*Graus de liberdade
glGi=(g-1)*(e-1)/g;
glEj=(g-1)*(e-1)/e;
QMGi=SQGi/glGi;
QMEj=SQEj/glEj;

* Testando a significancia com QMRes;
FGi_1=QMGi/QMRes;
FEj_1=QMEj/QMRes;
pvalorGi_1=1-probf(FGi_1,glGi,gl_Res);
pvalorEj_1=1-probf(FEj_1,glEj,gl_Res);
glGi=j(g,1,glGi);
glEj=j(e,1,glEj);

print glGi SQGi QMGi FGi_1 pvalorGi_1, glEj SQEj QMEj FEj_1 pvalorEj_1;
run;

```