

ANE GABRIELLE DE OLIVEIRA

**USO DE MODELOS MULTIPLICATIVOS NO ESTUDO DA INTERAÇÃO  
GENÓTIPO X AMBIENTE**

Dissertação apresentada à Universidade Federal de Viçosa *Campus* Rio Paranaíba, como parte das exigências do Programa de Pós-Graduação em Agronomia – Produção Vegetal, para obtenção do título de *Magister Scientiae*.

RIO PARANAÍBA  
MINAS GERAIS – BRASIL  
2015

ANE GABRIELLE DE OLIVEIRA

**USO DE MODELOS MULTIPLICATIVOS NO ESTUDO DA INTERAÇÃO  
GENÓTIPO X AMBIENTE**

Dissertação apresentada à Universidade Federal de Viçosa *Campus* Rio Paranaíba, como parte das exigências do Programa de Pós-Graduação em Agronomia – Produção Vegetal, para obtenção do título de *Magister Scientiae*.

APROVADA: 27 de novembro de 2015.

---

Prof. Dr. José Marcelo Soriano Viana

---

Prof. Dr. Vinícius Ribeiro Faria  
(Coorientador)

---

Prof. Dr. Pedro Ivo Vieira Good God  
(Orientador)

A minha irmã Carla Luciana (*in memoriam*),  
meu amor e inspiração.

**Dedico**

## AGRADECIMENTOS

A Deus, que com a sua bondade me deu sabedoria e força para lutar em busca dos meus objetivos.

Aos meus pais, pelo esforço que fizeram para que eu concluísse mais esta etapa da minha vida. Por sempre acreditarem em mim, me amparar nos momentos de dificuldade e me ensinar a trilhar bons caminhos.

A minha irmã Carla grande incentivadora, companheira e minha inspiração. Fez do meu sonho o dela e esteve sempre comigo, me ajudando e aconselhando. Infelizmente não está comigo para compartilhar mais essa vitória, mas certamente se eu tive forças para concluir este trabalho, foi por ela.

Aos meus familiares pela motivação, confiança e por sempre torcerem por mim.

Ao meu orientador Professor Dr. Pedro Ivo Vieira Good God, por ter acreditado e depositado sua confiança em mim. Agradeço pela paciência, dedicação, palavras de incentivo e por compartilhar seus conhecimentos.

Aos meus coorientadores Prof. Dr. Vinícius Ribeiro e Prof. Dr. Fabyano Fonseca, pela ajuda, orientações e atenção.

Ao Prof. Dr. José Marcelo, pelas contribuições e disponibilidade em participar da banca de defesa.

Aos colegas do Laboratório de Genética e Melhoramento de Plantas UFV-CRP, pela ajuda e companheirismo.

Aos meus amigos pelo incentivo, compreensão, por tornarem mais leve os momentos difíceis e por sempre trazerem alegria para os meus dias.

A Universidade Federal de Viçosa – Campus Rio Paranaíba (UFV-CRP) e ao Programa de Pós-Graduação em Agronomia – Produção Vegetal, pela oportunidade de realização do curso de Mestrado.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudo.

## Índice

Resumo.....	vi
Abstract .....	vii
Introdução geral.....	1
Referências bibliográficas .....	3
CAPÍTULO I - Aplicação do modelo AMMI na caracterização de linhagens experimentais de soja.....	5
Resumo.....	5
Abstract .....	6
1. Introdução.....	7
2. Material e métodos.....	8
2.1. Experimentos .....	8
2.2. Análises estatísticas .....	9
2.3. Análises gráficas .....	11
3. Resultados e Discussão .....	12
4. Conclusões.....	16
5. Referências Bibliográficas.....	17
CAPÍTULO II - Modelo fator analítico multiplicativo misto na caracterização de linhagens experimentais de soja para adaptabilidade e estabilidade .....	20
Resumo.....	20
Abstract .....	21
1. Introdução.....	22
2. Material e métodos.....	24
2.1. Experimentos .....	24
2.2. Análises estatísticas .....	25
3. Resultados e Discussão .....	27

4. Conclusões.....	33
5. Referências Bibliográficas.....	33
Conclusões gerais .....	37
Anexos .....	38

## RESUMO

OLIVEIRA, Ane Gabrielle, M.Sc., Universidade Federal de Viçosa, novembro de 2015. **Uso de modelos multiplicativos no estudo da interação genótipo x ambiente.** Orientador: Pedro Ivo Vieira Good God. Coorientadores: Vinicius Ribeiro Faria e Fabyano Fonseca e Silva.

A escolha do método estatístico para o estudo da adaptabilidade e estabilidade é uma tarefa que exige atenção do melhorista. Dentre os métodos mais utilizados, destaca-se o modelo de efeitos principais aditivos e interação multiplicativa (AMMI). Entretanto, a análise AMMI possui limitações quando se trata de um conjunto de dados desbalanceado e com heterogeneidade de variâncias. O modelo fator analítico multiplicativo misto (FAMM) é uma alternativa para analisar dados com estas características. O objetivo deste estudo foi aplicar os métodos AMMI e FAMM em um conjunto de dados desbalanceado para avaliar a adaptabilidade e estabilidade de linhagens de soja. As produtividades de 48 genótipos de soja em sete ambientes foram avaliadas, com diferentes números de repetições, replicados no delineamento em blocos casualizados. A classificação dos genótipos quanto aos padrões de adaptabilidade e estabilidade utilizando os modelos AMMI e FAMM em geral não se manteve, indicando que as duas técnicas podem levar a conclusões diferentes no que se refere à recomendação de genótipos. Esta diferença na classificação pode estar associada às condições de desbalanceamento e heterogeneidade de variâncias. O modelo FAMM apresenta a vantagem de prever valores genotípicos e considerar o uso da informação de parentesco, o que pode tornar mais confiável o estudo dos padrões de adaptabilidade e estabilidade.

## ABSTRACT

OLIVEIRA, Ane Gabrielle, M.Sc., Universidade Federal de Viçosa, november, 2015. **Multiplicative models in the study of genotype x environment interaction.** Advisor: Pedro Ivo Vieira Good God. Co-advisors: Vinicius Ribeiro Faria and Fabyano Fonseca e Silva.

The choice of statistical method to study the adaptability and stability is a task that requires attention from the breeder. Among the most used methods, stands out Additive Main Effects And Multiplicative Interaction (AMMI) model. However, AMMI analysis has limitations when it comes to a set of unbalanced data and with heterogeneity of variances. The factor analytic multiplicative mixed (FAMM) model is an alternative to analyze data with these characteristics. The objective of this study was to apply the AMMI and FAMM methods in an unbalanced set of data to evaluate the adaptability and stability of soybean lines. Yields of 48 soybean genotypes in seven environments were evaluated, with different numbers of repeats, replicated in randomized block design. The classification of genotypes regarding the standards of adaptability and stability using the AMMI and FAMM models generally not maintained, indicating that the two techniques can lead to different conclusions as regards the recommendation of genotypes. This difference in classification may be associated with conditions of unbalanced and heterogeneity of variances. The FAMM model has the advantage of predicting genotypic values and consider the use of pedigree information, which may make more reliable the study of adaptability and stability standards.



## INTRODUÇÃO GERAL

Compreender os fatores que levam os genótipos a apresentarem performance superior é essencial para o sucesso de um programa de melhoramento de plantas (Malosetti et al., 2013). Dentre esses fatores está a interação genótipo x ambiente (GA), a qual possui papel fundamental, explicando o comportamento diferenciado dos genótipos dependendo do ambiente no qual é cultivado.

Para que um genótipo possa ser considerado superior ele deve apresentar, além performance superior, uma previsibilidade de comportamento e uma adaptação ao maior número de ambientes possíveis (Yan e Kang, 2003). Para a avaliação desses aspectos, são realizados nas fases finais de um programa de melhoramento o estudo da adaptabilidade e estabilidade dos genótipos, onde estes são testados em diversos ambientes que variam quanto ao local e/ou ano de cultivo.

Para a análise dos dados de experimentos multi-ambientais existem vários métodos. A maioria deles são realizados por meio de análise de variância conjunta (ANOVA) e/ou por meio do ajuste de regressão linear (Resende e Thompson, 2004). De maneira geral estes métodos podem ser classificados em não paramétricos (Lin e Binns, 1988), com base em análise de variância (Wricke, 1965), regressões lineares (Eberhart e Russell, 1966), análises multivariadas (Zobel et al., 1988) e também modelos mistos (Resende, 2002). Porém, os métodos considerados tradicionais (baseadas na ANOVA e regressão linear), apresentam limitações para a análise de dados oriundos de situações comuns da experimentação de campo.

O desbalanceamento é característica comum de experimentos multi-ambientais e também um dos principais complicadores quanto ao uso dos métodos tradicionais. Durante um programa de melhoramento, é comum a perda e inclusão de parcelas durante o período de avaliação dos genótipos, assim como experimentos com números diferentes de repetições e genótipos que não são avaliados em todos ambientes. Todas essas são situações que contribuem para o desbalanceamento do conjunto de dados (Figueiredo et al., 2014). Os métodos tradicionais mencionados não permitem a análise de experimentos desbalanceados. Geralmente são necessárias a exclusão de genótipos e/ou repetições e também a imputação dos dados, a fim de promover o balanceamento destes.

A análise AMMI (*Additive Main Effects And Multiplicative Interaction Model*) tem sido amplamente usada para avaliar a adaptabilidade e estabilidade de

genótipos. Nela em um único modelo são combinados, componentes aditivos para os efeitos principais (genótipos e ambientes) e componentes multiplicativos para os efeitos da interação (Zobel et al., 1988). A popularidade do método está associada à sua capacidade de explicar a natureza das interações, que pode ser feita observando os escores dos componentes principais (PC) e também pela sua fácil interpretação gráfica em um diagrama de dispersão multivariado, chamado biplot (Yokomizo et al., 2013). No entanto a análise AMMI, assim como os métodos tradicionais, é limitada para o uso em dados desbalanceados, fazendo necessária a exclusão ou imputação dos dados, que pode ser feita através, por exemplo, da técnica de médias ajustadas.

Um modelo alternativo para a análise de experimentos com desbalanceamento, sem a necessidade de ajuste do conjunto de dados, é o modelo fator analítico multiplicativo misto (FAMM), proposto por Piepho (1998). O método é análogo ao AMMI, pois nele ocorre a multiplicação de coeficientes genotípicos e ambientais (conhecido como cargas e escores fatoriais, respectivamente). Porém, nesse modelo é possível tratar os efeitos de genótipos como aleatórios e predizê-los com base no Melhor Preditor Linear Não Viesado (BLUP), considerando os efeitos de genótipos correlacionados entre ensaios. Outra vantagem deste método é a capacidade de analisar dados com heterogeneidade de variâncias residuais e experimentos não ortogonais.

O método AMMI é uma técnica amplamente difundida entre os melhoristas enquanto que o método FAMM, apesar de se adequar aos problemas da experimentação de campo, ainda é pouco utilizado. Segundo Figueiredo et al. (2015) provavelmente este fato está associado a complexidade de execução das análises pelo método FAMM. Assim, o objetivo deste trabalho foi estudar a adaptabilidade e estabilidade de genótipos de soja utilizando os modelos FAMM e AMMI nas fases finais de um programa de melhoramento e verificar a viabilidade de uso desses métodos a dados desbalanceados.

## Referências bibliográficas

- EBERHART, S.A.; RUSSELL, G. Stability parameter for comparing varieties. **Crop Science**. v. 6, p.36-40, 1966.
- FIGUEIREDO, A.G., RENZO PINHO, G.V., SILVA, H.D, BALESTRE, M. Application of mixed models for evaluating stability and adaptability of maize using unbalanced data. **Euphytica**. v.202, p. 393-409, 2015.
- KELLY, A.M., SMITH, A.B., ECCLESTON, J.A., CULLIS, B.R. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. **Crop Science**. v.47, p.1063–1070, 2007.
- LIN, C.S.; BINNS, M.R. A superiority measure of cultivar performance for cultivar, location, data. **Canadian Journal Plant Science**. v.68, p.193-198, 1988.
- MALOSETTI, M.; RIBAUT, J.M.; EEUWIJK, F.A.V. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. **Frontiers in Physiology**. v.4, p. 1-17, 2013.
- PIEPHO, H. P. Empirical best linear unbiased prediction in cultivar trials using factor analytic variance-covariance structures. **Theoretical and Applied Genetics**. v.97, p.195-201, 1998.
- RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília, DF: EMBRAPA Informação Tecnológica, 2002. 975 p.
- RESENDE, M.D.V.; THOMPSON, R. Factor analytic multiplicative mixed models in the analysis of multiple experiments. **Revista Matemática e Estatística**. v.22, p. 31-52, 2004.
- WRICKE, G. Zur Berechnung der Ökovalenz bei Sommerweizen und Hafer. **Zeitschrift für Pflanzenzüchtung**, v.52, p.127-138, 1965.
- YAN, W; KANG, M.S. **GGE Biplot analysis. A graphical tool for breeders**. Geneticists and Agronomists. CRC Press, 2003.
- YOKOMIZO, G.K.I; DUARTE, J.B; VELLO, N.A; UNFRIEND, J.R. Análise AMMI da produtividade de grãos em linhagens de soja selecionadas para resistência à ferrugem asiática. **Pesquisa Agropecuária Brasileira**. v.48, p. 1376-1384, 2013.

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

## CAPÍTULO I

### **Aplicação do modelo AMMI na caracterização de linhagens experimentais de soja**

**RESUMO** - A produtividade é um carácter complexo resultante, dentre outros fatores, da interação entre genótipos e ambientes (GA). Os ensaios de adaptabilidade e estabilidade são necessários para o entendimento da interação GA e recomendação de cultivares. O modelo de efeitos principais aditivos e interação multiplicativa (AMMI) destaca-se como uma das principais metodologias para o estudo do padrão de estabilidade e adaptabilidade de genótipos. O objetivo deste trabalho foi determinar os padrões de adaptabilidade e estabilidade de um conjunto de dados originalmente desbalanceados utilizando o método AMMI. As produtividades de 48 genótipos de soja conduzidos em sete ambientes foram mensuradas, com diferentes números de repetições e conduzidas no delineamento em blocos casualizados. Foi necessária a imputação dos dados perdidos e o descarte de algumas repetições para o balanceamento dos dados. A análise AMMI capturou aproximadamente 60% da variância da interação GA. Dentre os genótipos analisados, a maior parte apresentou adaptabilidade geral e estabilidade de comportamento, sendo assim, de recomendação geral. O método AMMI mesmo com restrições em analisar dados desbalanceados, possibilitou a identificação de genótipos promissores com alta produtividade, estáveis e com adaptabilidade geral, que poderão ser indicados para cultivo comercial.

## **Aplication of the AMMI model in characterization of soybeans experimental lines**

**ABSTRACT** - Productivity is a complex character resulting, among other factors, the interaction between genotypes and environments (GE). Adaptability and stability tests are necessary for understanding the GE interaction and recommendation of cultivars. The Additive Main Effects And Multiplicative Interaction Model (AMMI) stands out as one of the main methods for the study of the pattern of stability and adaptability of genotypes. The objective of this work was to determine the adaptability and stability standards of a set of data originally unbalanced using the AMMI method. Yields of 48 soybean genotypes conducted in seven environments were measured with different numbers of repetitions and conducted in a randomized block design. The imputation of missing data and the disposal of some repetition for balancing the data and so further analysis was required. The AMMI analysis captured approximately 60% of GE interaction. Among the genotypes studied, the majority presented general adaptability and stability of behavior, therefore, of general recommendation. The genotypes that stood out were 14, 17 and 37 present high yields and desirable behavior regarding adaptability and stability standards. The AMMI method same with restrictions in analyzing unbalanced data, enabled the identification of promising genotypes with high productivity, stable and with general adaptability, which may be indicated for commercial cultivation.

## 1. Introdução

Nos programas de melhoramento da soja um dos principais objetivos é a obtenção de genótipos mais produtivos. A produtividade é um carácter complexo resultante, dentre outros fatores, da interação entre genótipos e ambientes (GA). Experimentos que estudam a interação GA são fundamentais para o entendimento do comportamento dos genótipos nos diferentes ambientes, sendo essenciais nas fases finais de um programa de melhoramento e para a recomendação de cultivares.

Mesmo sendo de grande importância, as estimativas da interação GA não são suficientemente informativas para explicar o comportamento de cada genótipo diante das variações ambientais. Para tal objetivo, realizam-se análises de adaptabilidade e estabilidade, que priorizam genótipos com alta produtividade de grãos, com estabilidade quanto ao desempenho e adaptabilidade ampla aos diferentes ambientes de cultivo (Cruz e Regazzi, 2001).

Existe atualmente uma grande quantidade de metodologias para o estudo da adaptabilidade e estabilidades dos genótipos. Todas elas se assemelham quanto ao tipo de estudo realizado, sendo o que as diferenciam são os procedimentos estatísticos associados a obtenção das estimativas (Yan e Tinker, 2006). De maneira geral estes métodos podem ser classificados em não paramétricos (Lin e Binns, 1988), os baseados em análise de variância (Wricke, 1965), regressões lineares (Eberhart e Russell, 1966), análises multivariadas (Zobel et al., 1988) e também modelos mistos (Resende, 2002).

A escolha do melhor método pode ser uma tarefa difícil. Para isso deve ser levada em consideração, o conjunto de dados a serem analisados e a capacidade do método em apresentar respostas mais próximas o possível da realidade. Algumas técnicas tradicionais de análise da interação GA, baseados em análise de variância e regressão linear, possuem limitações quanto a habilidade em explicar grande parte da interação GA e a não informação sobre interações específicas positivas ou negativas com ambientes (Duarte e Vencovsky, 1999).

Uma alternativa para contornar a limitações dos métodos tradicionais é a análise AMMI (*Additive Main Effects And Multiplicative Interaction Model*). Em um único modelo são combinados, componentes aditivos para os efeitos principais (genótipos e ambientes) e componentes multiplicativos para os efeitos da interação. O método consiste em uma prévia análise de variância comum à matriz de médias

onde são estimados os efeitos principais, a parte aditiva (média geral, efeitos de genótipos e ambiental) e o desdobramento da interação, constituindo a parte multiplicativa do modelo que pode ser analisada por análise de componentes principais (ACP) (Gauch, 1988; Zobel et al. 1988). A ACP é uma técnica de análise exploratória de dados, através dela é possível descobrir a estruturas de dados que representam uma grande porcentagem da variância total, e criar novas construções hipotéticas (componentes principais) que podem ser utilizados para prever ou classificar observações em grupos (Timm, 2002).

Dentre as vantagens do método AMMI está a capacidade de explicar a natureza das interações, que pode ser feita observando os escores dos componentes principais (PC), ou de maneira mais simples por meio da representação gráfica nos chamados biplot, onde são plotados os escores dos PC (Oliveira et al., 2003). Na literatura é relatado que a análise AMMI tem se mostrado capaz de avaliar os efeitos da interação GA e vários autores tem utilizado este método na cultura da soja (Rocha et al., 2004; Oliveira et al., 2006; Yokomizo et al., 2013).

No entanto o modelo AMMI possui restrições quando se trata de dados desbalanceados, algo que é comum em experimentos de avaliação de cultivares. Uma alternativa para utilizar modelo AMMI em dados desbalanceados é inicialmente imputar as observações ausentes e, posteriormente, na matriz de dados completa (observados + imputados), usar o modelo AMMI para compreender a interação (Arciniegas-Alarcón e Dias, 2009). O objetivo deste trabalho foi estudar os padrões de adaptabilidade e estabilidade de linhagens experimentais de soja empregando dados desbalanceados, utilizando o método AMMI.

## **2. Material e métodos**

### **2.1. Experimentos**

As 48 linhagens usadas neste trabalho são provenientes do Programa de Melhoramento da Qualidade da Soja (PMQS) da Universidade Federal de Viçosa, onde as populações foram obtidas segundo Pinheiro et al. (2013). Os genótipos foram avaliados em sete ambientes que combinam diferentes locais e safras. Na safra 2010/11 e 2011/12, os genótipos foram avaliados em Rio Paranaíba-MG; no ano



agrícola 2012/13 os genótipos foram testados em Rio Paranaíba-MG, Capinópolis-MG e Florestal-MG. Finalmente, na safra 2013/14 os locais de avaliação foram Rio Paranaíba-MG e Florestal-MG. O delineamento adotado foi blocos casualizados, com número de repetições que variaram de dois a três.

As parcelas foram compostas de quatro fileiras de 5,0 m de comprimento com espaçamento de 0,5 m entre fileiras. Como parcela útil foram considerados as duas fileiras centrais. Durante a condução dos experimentos as progênies passaram por uma seleção visual a fim de eliminar plantas/progênies que apresentaram comportamento agrônômico indesejável, como por exemplo alta incidência de doenças. As sementes foram colhidas no estágio reprodutivo R8, onde 90% das vagens encontraram-se maduras. Após isso as sementes foram pesadas e determinada a produtividade em kg.ha<sup>-1</sup>.

## 2.2. Análises estatísticas

Preliminarmente, para que o número de repetições entre ambientes fosse balanceado, foram descartadas as repetições que apresentaram o maior número de parcelas perdidas. Em seguida, os valores de parcelas perdidas foram imputados por meio do cálculo das médias ajustadas proposto por Gomes (1990).

Em seguida ajustou-se o modelo de blocos casualizados, processado para cada ambiente. Após a adequação do conjunto de dados, foi realizada a análise de variância conjunta, conforme descrito a seguir:

$$y_{ijk} = u + b/a_{jk} + g_i + a_j + ga_{ij} + e_{ijk} \quad (1)$$

Onde:  $y_{ijk}$  é o valor fenotípico do genótipo  $i$  no  $j$ -ésimo ambiente e  $k$ -ésimo bloco;  $u$ : média geral;  $g_i$ : efeito do  $i$ -ésimo genótipo;  $b/a_{jk}$ : efeito do  $k$ -ésimo bloco dentro do  $j$ -ésimo ambiente;  $a_j$ : efeito do  $j$ -ésimo ambiente;  $ga_{ij}$ : efeito da interação do  $i$ -ésimo genótipo com o  $j$ -ésimo ambiente;  $e_{ijk}$ : erro aleatório.

O modelo AMMI para o estudo da adaptabilidade e estabilidade foi realizado segundo Duarte e Vencovsky (1999) o qual é descrito por:

$$Y_{ij} = \mu + g_i + a_j + \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \bar{e}_{ij} \quad (2)$$

em que:  $Y_{ij}$  é a resposta média do genótipo  $i$  no ambiente  $j$ ;  $\mu$  é a média geral dos ensaios;  $g_i$  é o efeito do genótipo  $i$ ;  $a_j$  é o efeito do ambiente  $j$ ;  $\lambda_k$  é o  $k$ -ésimo valor singular (escalar) da matriz de interações original (denotada por  $GA$ );  $\gamma_{ik}$  é o elemento correspondente ao  $i$ -ésimo genótipo no  $k$ -ésimo vetor singular coluna da matriz  $GA$ ;  $\alpha_{jk}$  é o elemento correspondente ao  $j$ -ésimo ambiente no  $k$ -ésimo vetor singular linha da matriz  $GA$ ;  $\rho_{ij}$  é o ruído associado ao termo  $(ga)_{ij}$  da interação clássica do genótipo  $i$  com o ambiente  $j$ ;  $e_{ij}$  e  $\bar{e}_{ij}$  é o erro experimental médio.

Na equação (1) o termo que representa a  $ga_{ij}$  é dado por:

$$ga_{ij} = \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \bar{e}_{ij} \quad (3)$$

Onde:

$$\rho_{ij} = \sum_{k=n+1}^p \lambda_k \gamma_{ik} \alpha_{jk} \quad (4)$$

A ACP (Análise de Componentes Principais) decompõe o efeito da interação  $ga_{ij}$  de maneira que esta passe a ser descrita como uma soma de  $p$  componentes, cada um resultante da multiplicação de  $\lambda_k$ , expresso na mesma unidade de  $Y_{ij}$ , por um efeito genotípico ( $\gamma_{ik}$ ) e um efeito ambiental ( $\alpha_{jk}$ ), ambos adimensionais. O termo traz uma informação relativa à variação devida à interação  $GA$ , no  $k$ -ésimo componente principal. De forma que a soma  $\lambda_k$  dos  $p$  componentes recompõem toda a variação ( $SQ_{GA} = \sum_{k=1}^p \lambda_k^2$ ). A recuperação da variação considera somente aqueles componentes mais fortemente determinados pela interação (linhas e colunas da matriz  $GA$ ), ou seja: o padrão (parte determinística ou sistemática). Assim, a interação do genótipo  $i$  com o ambiente  $j$  é descrita pela equação abaixo, descartando-se o resíduo adicional  $\rho_{ij}$  descrito anteriormente na equação 4. Por isso, o método AMMI é visto como um procedimento capaz de separar padrão e ruído na análise.

$$ga_{ij} = \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} \quad (5)$$

Na ACP os eixos captam, sucessivamente, porções cada vez menores da variação presente na matriz GA ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ) e assim é possível que os primeiros componentes sejam suficientes para descrever a  $SQ_{GA}$  (Duarte e Vencovsky, 1999). Neste trabalho para a escolha do número de eixos principais que explicaram e posteriormente foram usados para representação gráfica, foi levado em consideração a proporção da soma de quadrados da interação GA retida até o n-ésimo eixo principal que é dada por:

$$\sum_{k=1}^n \lambda_k^2 / SQ_{GA} \quad (6)$$

Onde:

$$SQ_{GA} = \sum_{ij} (\hat{g}\hat{a})_{ij}^2 \quad (7)$$

Para determinar qual família de modelos AMMI (AMMI0, AMMI1, ..., AMMI<sub>n</sub>) a ser usada foi testada a significância dos componentes pelo teste F. Os modelos significativos foram considerados capazes de explicar a interação GA.

As análises estatísticas foram realizadas utilizando programa R 3.0.2. (R Core Team, 2013). Para a análise AMMI foi utilizado o pacote Agricolae (Mendiburu, 2014).

### 2.3. Análises gráficas

A fim de facilitar a interpretação, os escores de genótipos e de ambientes foram representadas em um gráfico denominado biplot (Gabriel, 1971), que permite descrever a estabilidade e adaptabilidade dos genótipos aos ambientes de teste. A interpretação é feita observando-se a distância dos escores que representam a interação GA da origem do gráfico. Assim, o genótipo com o menor escore é dito como mais estável. Para determinar a adaptabilidade dos genótipos foi observada a proximidade dos genótipos e ambientes no gráfico, já que os genótipos e ambientes com o mesmo sinal interagem positivamente, indicando em qual ambiente o genótipo deve ser preferencialmente cultivado.

Para a construção dos gráficos foi usado o programa estatístico R 3.0.2. (R Core Team, 2013) e o pacote utilizado foi o Agricolae (Mendiburu, 2014).

### 3.Resultados e Discussão

Da soma de quadrados total, a porcentagem de cada efeito foi de 70,54% para ambientes, 3,44% para genótipos e 13,83% para a interação GA (Tabela 1). A análise conjunta da interação entre genótipos e ambientes foi significativa ( $p < 0,001$ ), indicando uma resposta diferenciada dos genótipos aos ambientes usados no estudo. O comportamento diferenciado dos genótipos, é em grande parte atribuído a interação GA, já que as magnitudes da soma de quadrados foram quatro vezes maiores que a dos genótipos. Em experimentos que avaliam a interação GA, a produtividade é influenciada, em maior parte, pelo ambiente (Meotti et al., 2012). Esta variabilidade dos efeitos justifica a necessidade de um estudo detalhado da interação para a recomendação dos novos genótipos.

**Tabela 1 .** ANOVA (modelo AMMI) para produtividade de 48 genótipos de soja testados em sete ambientes

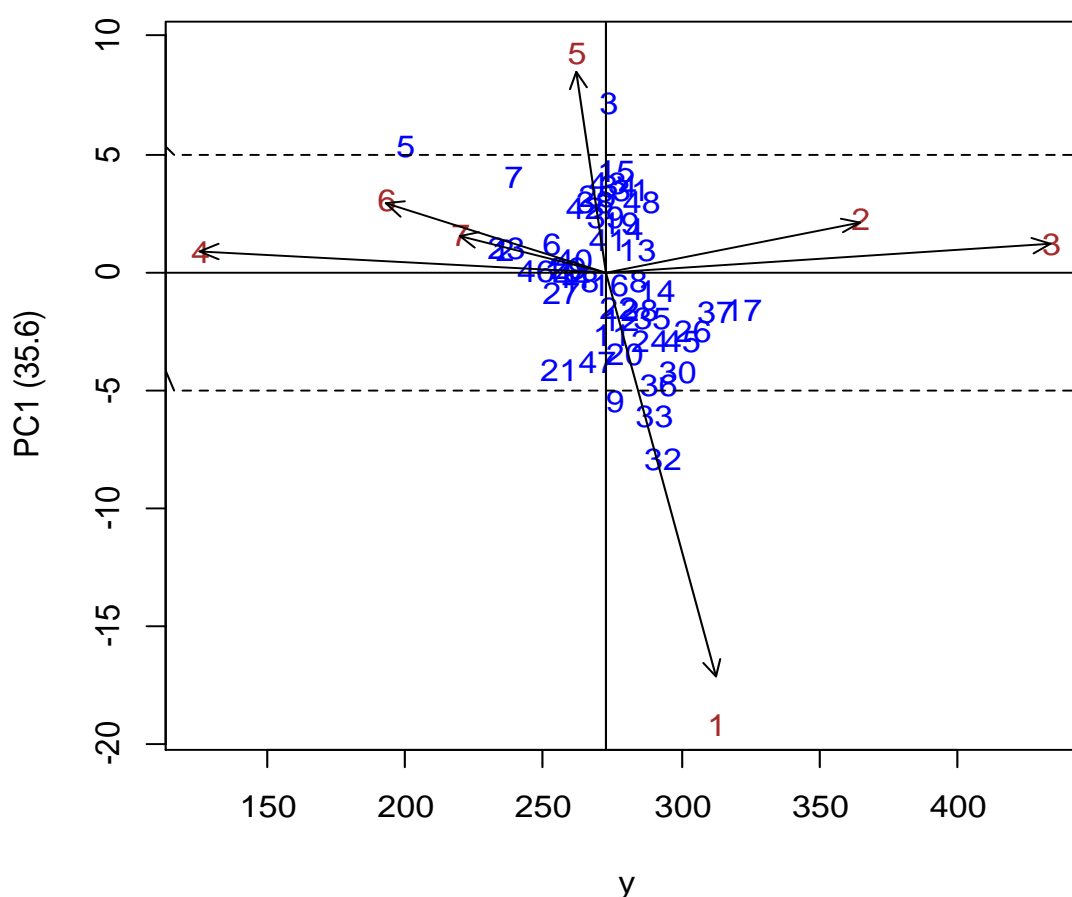
Fontes de variação	GL	SQ	QM	F	Pr > F	Explicação (%)
Repetições (Ambientes)	7	147239,00	21034,00	7,21	4,9200E-08	1,62
Genótipos	47	312110,00	6641,00	2,28	1,5510E-05	3,44
Ambientes	6	6405434,00	1067572,00	50,75	2,0320E-05	70,54
Interação GA	282	1255939,00	4454,00	1,53	1,1320E-04	13,83
PC1	52	447385,98	8603,58	2,95	0,0000	35,62
PC2	50	302467,12	6049,34	2,07	0,0001	24,08
PC3	48	221800,44	4620,84	1,58	0,0118	17,66
PC4	46	163833,24	3561,59	1,22	0,1660	13,04
PC5	44	80756,31	1835,37	0,63	0,9684	6,40
PC6	42	40153,44	956,03	0,33	1,0000	3,20
Resíduo	329	960350,00	2919,00			
Total	671	9081072,00				
Média(kg.ha <sup>-1</sup> ): 2729,50			CV(%):19,79			

GL: graus de liberdade; SQ: soma de quadrados; QM: quadrado médio, CV: coeficiente de variação.

Para o modelo AMMI a interação GA foi decomposta em seis componentes principais, mas apenas os três primeiros componentes principais foram significativos ( $p < 0,01$ ). O modelo AMMI3 (modelo considerando os 3 primeiros componentes principais) explicou 77,36% da interação GA. O PC1 foi altamente significativo

( $p < 0,001$ ) e capturou 35,62% da variação, enquanto os PC2 e PC3 explicaram 24,08% e 17,66% respectivamente. Estes resultados demonstram a confiabilidade do modelo AMMI em explicar e particionar as variações da interação GA, facilitando o estudo gráfico dos genótipos aos diferentes ambientes.

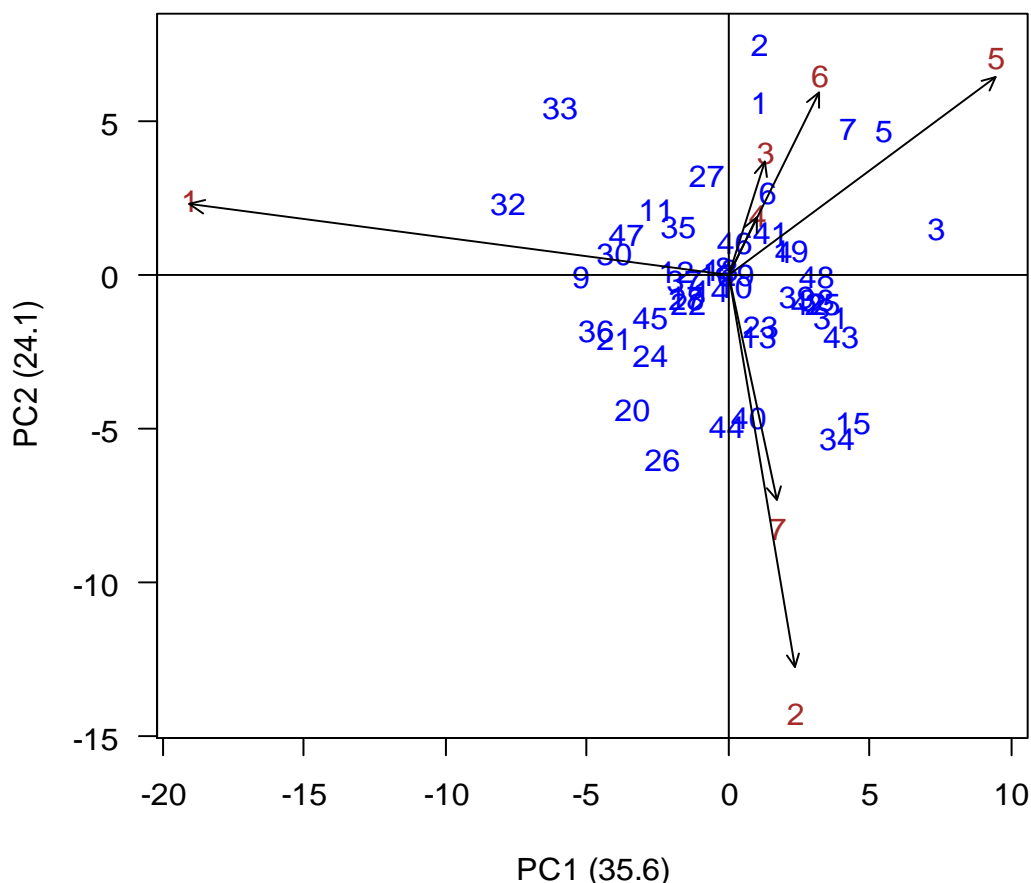
Na representação gráfica do modelo AMMI é possível observar o efeito da interação GA e dos ambientes (Figura 1). A distribuição dos genótipos, de uma maneira geral, em torno do eixo das abscissas, caracterizou uma tendência dos genótipos a uma menor interação GA para a produtividade e uma maior estabilidade. Os genótipos que apresentaram maiores interações foram 3, 5, 9, 32 e 33, uma vez que possuem maiores escores para o PC1 e assim são classificados como instáveis. O estudo da estabilidade dos genótipos permite a previsibilidade de comportamento desses aos ambientes de cultivo. Genótipos estáveis possuem recomendação ampla, enquanto os considerados instáveis, são mais influenciados pela interação GA.



**Figura 1:** Representação gráfica do modelo AMMI1 (com um componente principal). Na abscissa estão as produtividades ( $\text{g.m}^{-2}$ ) e na ordenada estão representados os escores do PC1 de 48 genótipos de soja testados em 7 ambientes. Os ambientes estão representados em vermelho.

A maior produtividade média relacionou-se aos ambientes 3 e 2, 4282,54 kg.ha<sup>-1</sup> e 3644,87 kg.ha<sup>-1</sup>, respectivamente. Ambos são classificados com ambientes estáveis, porque estão mais próximo de zero no eixo PC1. O ambiente 1 apresentou alta produtividade média (3084,06 kg.ha<sup>-1</sup>) e foi considerado instável assim como o ambiente 5 (2404,99 kg.ha<sup>-1</sup>), pois estão posicionados mais distantes da origem do PC1, sendo, portanto, os que mais contribuíram para a interação GA. Os demais ambientes, A4, A6 e A7 são estáveis e tiveram produtividades baixas, de 1282,19 kg.ha<sup>-1</sup>, 1946,97 kg.ha<sup>-1</sup> e 2206,46 kg.ha<sup>-1</sup>, respectivamente. O objetivo de testar a estabilidade dos ambientes é informar sobre a confiabilidade dos genótipos no ambiente testado (Oliveira et al., 2003). Associar as médias de produtividade de cada ambiente com a sua estabilidade permite inferir sobre suas características e a entender o comportamento dos genótipos.

Apesar do modelo AMMI3 ter sido validado pelo teste F, optou-se para a representação gráfica o modelo AMMI2 (Figura 2). A representação com dois eixos é de mais fácil interpretação, já que o acúmulo de dimensões torna o modelo menos preciso. Isso porque é observado um incremento nos ruídos e uma menor explicação do padrão da interação GA quando são adicionadas novas dimensões no modelo (Crossa et al., 1991; Lavoranti et al., 2004). Os componentes PC1 e PC2 corresponderam juntos a aproximadamente 60% da interação GA. Alguns autores indicam o uso de dois componentes (Yan and Rajcan, 2002), três componentes (Tukamuhabwa et al, 2012) e até quatro deles (Sivapalan et al., 2000), mas a escolha do melhor modelo não deve ser feita previamente. Segundo Timm (2002), a proporção ideal deve estar entre 70% e 90%, mas cabe ao pesquisador determinar quanto que o modelo reduzido deve explicar da variação total. Portanto, foi decidido seguir a interpretação, mesmo com a porcentagem de explicação estando abaixo da recomendada.



**Figura 2:** Representação gráfica do modelo AMMI2 (com dois componentes principais). Na abscissa estão representados os escores do PC1 com e nas ordenadas estão representados os escores do PC2 de 48 genótipos de soja testados em 7 ambientes. Os ambientes estão representados em vermelho.

No presente trabalho os ambientes com médias de produtividade acima da média geral foram classificados em favoráveis e os com médias inferiores a média geral foram dados como desfavoráveis. Para condições de adaptabilidade específicas de ambientes os genótipos 1, 2 e 7 mostraram adaptabilidade específica ao A6 que é considerado desfavorável, portanto esses genótipos podem ser indicados para ambientes com baixo nível tecnológico. Os demais genótipos encontraram-se próximo a origem do gráfico, o que indica adaptabilidade geral. Esses resultados condizem com um estudo feito anteriormente para este mesmo conjunto de dados utilizando o método de Eberhart e Russel (Mavungo Jr., 2014). Genótipos com adaptabilidade geral expressam comportamento semelhante tanto em ambientes considerados desfavoráveis quanto favoráveis.

Na Tabela 3, os genótipos foram ranqueados para um estudo mais detalhado da estabilidade. Dentre os genótipos destacam-se os 14, 17 e 37 que apresentaram

altas produtividades (acima da média) e ainda são tidos como estáveis. Já o genótipos 2, 5 e 7 tiveram baixas produtividade e são instáveis, ou seja, possuem maior interação GA. Para fins de recomendação de cultivares, para os genótipos considerados estáveis é também desejável que apresentarem altas produtividades (Pacheco et al., 2005). No entanto, genótipos considerados como estáveis e que apresentem baixas produtividades podem ser indicados para ambientes que otimizem a sua produção, como por exemplo, ambientes com alto grau tecnológico.

**Tabela 3.** Rank de estabilidade e produtividades dos genótipos pelo método AMMI. Os genótipos estão ordenados do mais estável para o menos estável.

				continuação...			
Gen	RE	RP	Média de produtividade (kg.ha <sup>-1</sup> )	Gen	RE	RP	Média de produtividade (kg.ha <sup>-1</sup> )
10	1	39	2574,16	48	25	12	2856,23
29	2	38	2591,01	38	26	33	2686,97
18	3	35	2632,76	25	27	31	2699,00
8	4	14	2845,09	24	28	11	2889,43
16	5	26	2743,58	31	29	17	2816,09
14	6	8	2908,65	40	30	36	2611,57
46	7	43	2471,26	47	31	32	2692,87
37	8	2	3122,59	44	32	37	2602,58
17	9	1	3220,68	30	33	5	2985,89
22	10	22	2770,41	43	34	29	2731,70
28	11	13	2846,45	21	35	41	2557,40
23	12	45	2369,67	1	36	47	2346,16
12	13	19	2783,51	36	37	7	2916,01
13	14	15	2844,91	20	38	18	2796,18
41	15	27	2735,94	9	39	24	2755,70
4	16	16	2826,10	26	40	3	3039,77
35	17	10	2896,61	34	41	21	2773,01
19	18	20	2777,41	7	42	44	2392,06
39	19	30	2725,16	15	43	23	2765,06
6	20	42	2527,79	2	44	46	2360,77
27	21	40	2564,87	5	45	48	2000,53
45	22	4	2999,64	3	46	28	2731,99
42	23	34	2648,49	33	47	9	2904,97
11	24	25	2745,79	32	48	6	2932,64

Gen: genótipo; RE: rank estabilidade; RP: rank produtividade

#### 4. Conclusões

O método AMMI, mesmo com restrições em analisar dados desbalanceados, possibilitou a identificação de genótipos promissores com alta produtividade,



estáveis e, a maioria deles, com adaptabilidade geral, que poderão ser indicados para cultivo comercial

Os padrões de adaptabilidade e estabilidade obtidos corroboraram com estudos anteriores utilizando o método de Eberhart e Russel.

## 5.Referências Bibliográficas

- ARCINIEGAS-ALARCON, S.; DIAS, C.T.S. Análise AMMI com dados imputados em experimentos de interação genótipo x ambiente de algodão. **Pesquisa Agropecuária Brasileira**. v. 44, p. 1391-1397, 2009.
- 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**. v. 81,p. 27-37, 1991.
- CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. 2 ed. rev. Viçosa: UFV, 2001.
- DUARTE, J. B.; VENCOVSKY, R. **Interação genótipo x ambiente: uma introdução à análise "AMMI"**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 60 p. (Série Monografias, 9).
- EBERHART, S.A.; RUSSELL, G. Stability parameter for comparing varieties. **Crop Science**. v. 6, p.36-40, 1966.
- GABRIEL, K. R. The biplot graphic of matrices with application to principal component analysis. **Biometrika**. v. 58, n. 3, p. 453-467, 1971.
- GAUCH, H. G. Model selection and validation for yield trials with interaction. **Biometrics**, Washington, v. 44, p. 705-715, 1988.
- GOMES, F. P. **Curso de Estatística Experimental**. 13.ed. Piracicaba: Nobel, 1990. 64p.
- LAVORANTI, O.J.; LINS, A.C.; FERREIRA, C.A. **Modelagem AMMI para estudos de interação em modelos estatísticos de efeitos fixos**. Comunicado Técnico 124, 2004.
- LIN, C.S.; BINNS, M.R. A superiority measure of cultivar performance for cultivar, location, data. **Canadian Journal Plant Science**. v.68, p.193-198, 1988.

- MAVUNGO JR, F. **Adaptabilidade e estabilidade de produtividade em progênies de soja**. 2014. Dissertação (Mestrado em Agronomia - Produção Vegetal) Universidade Federal de Viçosa, Minas Gerais.
- MENDIBURU, F. *Agricolae: Statistical Procedures for Agricultural Research*. R package version 1.2-1. 2014
- MEOTTI, G.V.; BENIN, G.; SILVA, R.R.; BECHE, E.; MUNARO, L.B. Épocas de semeadura e desempenho agrônômico de cultivares de soja. **Pesquisa Agropecuária Brasileira**. v.47, p.14-21, 2012.
- OLIVEIRA, A.B.; DUARTE, J.B.; PINHEIRO, J.B. Emprego da análise AMMI na avaliação da estabilidade produtiva em soja. **Pesquisa Agropecuária Brasileira**. v.38, p.357-364, 2003.
- OLIVEIRA, M.A.R.; PÍPOLO, V.C.; SCHUSTER, I.; VICENTE, D.; DELLAGOSTIN, M.; OLIVEIRA, E.F. Soybean stability and adaptability in Southern and Central Brazil. **Crop Breeding and Applied Biotechnology**. v.6, p.55-64, 2006.
- PACHECO, R.M.; DUARTE, J.B.; VENCOVSKY, R.; PINHEIRO, J.B.; OLIVEIRA, A.B. Use of supplementary genotypes in AMMI analysis. **Theoretical and Applied Genetics**. v.110, p.812-818, 2005.
- PINHEIRO, L.C.M.; GOD, P.I.V.G.; FARIA, V.R.; OLIVEIRA, A.G.; HASUI, A.A.; PINTO, E.H.G.; ARRUDA, K.M.A.; PIOVESAN, N.D.; MOREIRA, M.A. Parentesco na seleção para produtividade e teores de óleo e proteína em soja via modelos mistos. **Pesquisa Brasileira Agropecuária**. v.48, p.1246-1253, 2013.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.
- RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília, DF: EMBRAPA Informação Tecnológica, 2002. 975 p.
- ROCHA, M.M.; VELLO, N.A.; LOPES, A.C.A.; MAIA, M.C.C. Yield stability of soybean lines using additive main effects and multiplicative interaction analysis – AMMI. **Crop Breeding and Applied Biotechnology**, v.4, p.391-398, 2004.
- SIVAPALAN, S.; BRIEN, L.O.; ORTIZ-FERRARA, G.; HOLLAMBY, G.L.; BARCLAY, I.; MARTIN, P.J. An adaptation analysis of Australian and

- CIMMYT/ICARDA wheat germplasm in Australian production environments. **Australian Journal of Agricultural Research**. v.51, p. 903-915, 2000.
- TIMM, N. H. **Applied multivariate analysis**. New York Springer Verlag, 2002. 693p.
- TUKAMUHABWA, P.; ASIIMWE, M.; NABASIRYE, M.; KABAYI, P.; MAPHOSA, M. Genotype by environment interaction of advanced generation soybean lines for grain yield in Uganda. **African Crop Science Journal**. v. 20, p.107- 115, 2012.
- WRICKE, G. Zur Berechnung der Ökovalenz bei Sommerweizen und Hafer. **Zeitschrift für Pflanzenzüchtung**, v.52, p.127-138, 1965.
- YAN, W.; I. RAJCAN. Biplot analysis of test sites and trait relations of soybeans in Ontario. **Crop Science**. v.42, p.11-20, 2002.
- YAN, W.; TINKER, N.A. Biplot Analysis of Multi-Environment Trial Data: Principles and Applications. **Canadian Journal of Plant Science**, v.86, p.623-645, 2006.
- YOKOMIZO, G.K.I; DUARTE, J.B; VELLO, N.A; UNFRIEND, J.R. Análise AMMI da produtividade de grãos em linhagens de soja selecionadas para resistência à ferrugem asiática. **Pesquisa Agropecuária Brasileira**. v. 48, p. 1376-1384, 2013.
- ZOBEL, R.W.; WRIGHT, M.J.; GAUCH, H.G. Statistical analysis of a yield trail. **Agronomy Journal**. v.80, p.388-393, 1988.

## CAPÍTULO II

### **Modelo fator analítico multiplicativo misto na caracterização de linhagens experimentais de soja para adaptabilidade e estabilidade**

**RESUMO** - A escolha da melhor metodologia para o estudo dos padrões de adaptabilidade e estabilidade de genótipos é uma tarefa importante e complexa para um melhorista. É necessário usar modelos que sejam adequados ao conjunto de dados analisados e capazes de lidar com as adversidades da experimentação de campo. O modelo fator analítico multiplicativo misto (FAMM) tem se mostrado o mais apropriado para análises deste tipo, pois possui a capacidade de analisar dados desbalanceados, com heterogeneidade de variância, experimentos não ortogonais, além de apresentar os resultados em valores genotípicos. Neste trabalho objetivou-se aplicar o modelo FAMM como ferramenta para análise de um conjunto de dados desbalanceados de linhagens experimentais de soja. As produtividades de 48 genótipos de soja em sete ambientes foram avaliadas, com diferentes números de repetições em um delineamento em blocos casualizados. O conjunto de dados era desbalanceado e apresentou heterogeneidade de variâncias residuais. O modelo FAMM que melhor se ajustou para esse conjunto de dados foi o com dois fatores (FA2). O modelo com dois fatores explicou 81,47% da variância total. A coincidência de agrupamento dos genótipos obtidos pelo biplot e o dendrograma indicam a viabilidade de uso do método FAMM para o estudo da adaptabilidade e estabilidade em conjuntos de dados desbalanceados e com heterogeneidade de variâncias.

**Factor analytic multiplicative mixed model in characterization of soybeans  
experimental lines for adaptability and stability**

**ABSTRACT** - The choice of the best methodology for the study of adaptability standards stability of genotypes is an important and complex task for a breeder. It is necessary to use models that are appropriate for the data analyzed and able of handle with adversity field trial. The factor analytic multiplicative mixed (FAMM) model has proven most suitable for analyzes of this kind, because it has the ability to analyze unbalanced data with heterogeneity of variance, and non-orthogonal experiments and also presents the results in genotypic values. This study aimed to apply the FAMM as a tool for analyzing a set of unbalanced data soybean experimental lines. The yields of 48 soybean genotypes in seven environments were evaluated, with different numbers of repetitions in a randomized block design. The data set was unbalanced and showed heterogeneity of residual variances. The FAMM model best fit for this dataset was the two-factor (FA2). The model with two factors explain 81.47% of the total variance. The coincidence of cluster of the genotypes obtained by biplot and dendrogram indicating the feasibility of the use of FAMM method for the study of adaptability and stability for unbalanced data sets and heterogeneity of variances.

## 1.Introdução

A interação genótipo x ambiente (GA) é um dos maiores complicadores da seleção e indicação de cultivares (Garbuglio et al., 2007). A necessidade de métodos estatísticos capazes de interpretar a magnitude desta interação é fundamental para proporcionar respostas mais concisas que auxiliem na capitalização dessa interação (Yokozimo et al., 2013). Na literatura há uma variedade de metodologias para análises de experimento multi-ambientais, que desempenham um papel importante ao avaliar genótipos em vários estágios de programas de melhoramento de plantas.

Tradicionalmente a análises de experimentos em vários ambientes tem sido feita por meio da análise de variância conjunta (ANOVA) e/ou por ajuste de regressão linear (Resende e Thompson, 2004). Estes métodos apresentam limitações tais como a falta de habilidade para explicar grande parte dessa interação, a não informação sobre interações específicas positivas ou negativas com ambientes e a falta de linearidade nas relações GA em alguns casos. A incapacidade dos modelos de componentes de variância em lidar com esta estrutura de dados já foi discutida anteriormente por Piepho e Van Eeuwijk (2002).

Uma alternativa muito utilizada para contornar a limitações dos métodos tradicionais é a análise AMMI (Efeitos Principais Aditivos e Interação Multiplicativa). O método consiste em uma prévia análise de variância comum à matriz de médias, onde são estimados os efeitos principais, a parte aditiva (média geral, efeitos de genótipos e ambientais) e o desdobramento da interação, constituindo a parte multiplicativa do modelo, que pode ser analisada por decomposição dos valores singulares ou análise de componentes principais (ACP) (Zobel et al., 1988).

Em um programa de melhoramento, é comum a ocorrência de desbalanceamento dos dados, devido, por exemplo, à condução de experimentos com número diferentes de repetição e também pela perda ou descarte de parcelas. Outro fato comum em experimentos desse tipo é a existência de heterogeneidade de variâncias (Splike et al., 2005; Hu et al., 2013). Ao realizar qualquer análise estatística é importante avaliar como o modelo se ajusta aos dados e se os dados atendem aos pressupostos do modelo. Recentemente, as abordagens via modelos mistos tornaram-se populares, uma vez que proporcionam uma estrutura flexível em que dados desbalanceados são facilmente analisados (Kelly et al, 2007).

Estudos mostram a eficiência dos modelos mistos como uma alternativa para análises de experimentos em vários ambientes (Smith et al., 2005; Piepho e Mohring, 2010; Stefanova e Burchell, 2010; Hu e Splike, 2011). Entretanto, com um grande número de ambientes, o modelo misto multivariado é praticamente impossível de ser ajustado pois pode exceder os limites computacionais (Resende e Thompson, 2004). Outra dificuldade é que a maioria dos algoritmos de computação não acomodam situações em que tanto a matriz de variância verdadeira ou sua estimativa, é singular (Kelly et. al., 2007). Uma opção de modelo parcimonioso para o modelo misto multivariado é o modelo fator analítico multiplicativo misto (FAMM) (Resende e Thompson, 2004).

O FAMM foi proposto inicialmente por Piepho (1998) e tem como princípio a análise da fatores (AF). A AF aborda o problema de analisar a estrutura das interrelações (correlações) entre um grande número de variáveis, definindo um conjunto de dimensões latentes comuns, chamado de fatores (Figueiredo et al., 2015). Cada fator irá agrupar um conjunto de ambientes, altamente correlacionados entre si e fracamente correlacionados com outros ambientes, agrupando-os nos demais fatores.

Assim como o método AMMI, no FAMM ocorre a multiplicação de coeficientes genotípicos e ambientais (conhecido como cargas e escores fatoriais, respectivamente). Porém, nesse modelo é possível tratar os efeitos de genótipos como aleatórios e predizê-los com base nos Melhor Preditor Linear Não Viesado (BLUP). Além disso, é possível considerar a correlação entre genótipos nos diferentes ensaios. Segundo Smith et al. (2005), se o objetivo é ter estimativas dos efeitos genotípicos tão próximas possível dos verdadeiros, o ideal é que eles sejam considerados como aleatórios.

O método FAMM ainda é pouco utilizado para estudo de adaptabilidade e estabilidade, e segundo Figueiredo et al. (2015) provavelmente este fato está associado a complexidade de execução das análises. Segundo Kelly et al, (2007) o FAMM é eficaz nas fases precoces de seleção, e estes mesmos autores apontam para necessidade de verificar tal eficiência nos estágios finais do programa de melhoramento. Sendo assim, objetivou-se com este trabalho aplicar o modelo fator analítico multiplicativo misto (FAMM) para análise de um conjunto de dados desbalanceados de linhagens experimentais de soja em estágios finais de um programa de melhoramento.

## 2. Material e métodos

### 2.1. Experimentos

As 48 linhagens usadas neste trabalho são provenientes do Programa de Melhoramento da Qualidade da Soja (PMQS) da Universidade Federal de Viçosa, onde as populações foram obtidas segundo descrito por Pinheiro et al. (2013).

Os genótipos foram avaliados em sete ambientes que combinam diferentes locais e safras. Na safra 2010/11 e 2011/12, os genótipos foram avaliados em Rio Paranaíba-MG; no ano agrícola 2012/13 os genótipos foram testados em Rio Paranaíba-MG, Capinópolis-MG e Florestal-MG. Finalmente, na safra 2013/14 os locais de avaliação foram Rio Paranaíba-MG e Florestal-MG. O delineamento adotado foi blocos casualizados.

As parcelas eram compostas por quatro fileiras de 5,0 m de comprimento com espaçamento de 0,5 m entre fileiras. Como parcela útil, foram considerados as duas fileiras centrais. Durante a condução dos experimentos as progênies passaram por uma seleção visual a fim de eliminar plantas/progênies que apresentaram comportamento agrônômico indesejável, como por exemplo alta incidência de doenças. As sementes foram colhidas no estágio reprodutivo R8, onde 90% das vagens estavam maduras. Após isso as sementes foram pesadas e determinada a produtividade em  $\text{kg} \cdot \text{ha}^{-1}$ .

Durante a condução dos experimentos ocorreram perdas de parcelas, o que contribuiu para o desbalanceamento do conjunto de dados (Tabela 1). Para cada experimento foi utilizado números diferentes de testemunhas e mesmo em experimentos com números iguais de testemunhas, estas se diferenciavam quando ao cultivar utilizado.

**Tabela 1.** Número de repetições, número de testemunhas e porcentagem de desbalanceamento nos sete ambientes conduzidos.

Ambiente	Nº repetições	Nº testemunhas	Nº parcelas perdidas	% desbalanceamento
E1	2	5	9	18,75
E2	2	8	1	2,08
E3	3	8	0	0
E4	3	8	2	4,16
E5	3	8	3	6,25
E6	3	3	0	0
E7	3	2	5	10,42



Foram utilizadas as informações de parentesco das linhagens. A genealogia foi construída a partir do banco de dados do Programa de Melhoramento da Qualidade da Soja da Universidade Federal de Viçosa.

## 2.2. Análises estatísticas

A análise conjunta dos genótipos nos sete ambientes foi realizada segundo a metodologia de modelos mistos (Henderson, 1975):

$$y = Xb + Zu + \varepsilon \quad (1)$$

em que  $y$  é o vetor de observações da característica avaliadas,  $b$  é o vetor dos efeitos de fixos com matriz de incidência  $X$ ,  $u$  é o vetor dos efeitos aleatórios com matriz de incidência  $Z$  e  $\varepsilon$  é o vetor de erros ou resíduos aleatórios. Para este trabalho os efeitos de ambiente foram considerados fixos e os efeitos de genótipos e da interação GA foram considerados aleatórios.

O modelo FAMM foi realizado segundo o proposto por Smith et al., (2001) e Resende e Thompson (2004). Para o FAMM, o vetor de efeitos aleatórios é postulado em termos dos efeitos dos genótipos não observáveis em diferentes ambientes. Portanto a construção da extensão do modelo misto para o FAMM é dada por:

$$y = Xb + Z[\lambda f + \delta] + \varepsilon \quad (2)$$

$$\text{Sendo } u = [\lambda f + \delta] \quad (3)$$

em que  $\lambda$  é o vetor dos carregamentos do fatoriais,  $f$  é o vetor dos escores fatoriais para os genótipos nos ambientes e  $\delta_{ij}$  é do vetor de erro representando a falta de

ajuste do modelo. A distribuição conjunta de  $f$  e  $\delta$  é dada por:  $\begin{pmatrix} f \\ \delta \end{pmatrix} \sim N$

$$\left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I_k \otimes I_g & 0 \\ 0 & \Psi \otimes I_g \end{pmatrix} \right] \text{ onde } \Psi \text{ é a diagonal de variâncias específicas onde } \Psi_i$$

representa a variância específica para o  $i$ -ésimo ensaio. Isto implica que todas as covariâncias entre os níveis de  $u$  são devidos aos fatores comuns, enquanto que os

fatores específicos explicam a variação adicional de elementos individuais de  $u$  (Meyer et al., 2009).

Os fatores podem ser não ou correlacionados. As variáveis são agrupadas por meio de suas correlações, ou seja, aquelas pertencentes a um mesmo grupo serão fortemente correlacionadas entre si, mas pouco correlacionadas com as variáveis de outro grupo. Cada grupo de variáveis representará um fator (Johanson e Wichern, 2007).

O modelo apresenta distribuição com estrutura de médias e variâncias para os efeitos aleatórios como  $u \sim N(0, A \otimes \Sigma)$ , onde  $A$  é a matriz de parentesco e  $\Sigma$  é a matriz de variância-covariância genética, que é dada por:

$$\Sigma = \Lambda \Lambda' + \Psi \quad (4)$$

sendo:

$$\Lambda \Lambda' = V D V \quad (5)$$

Onde  $V$  refere-se a matriz de autovetores e  $D$  é a matriz diagonal dos autovalores.

Para análise de múltiplos experimentos, a abordagem da análise de fatores pode ser usada para propiciar uma classe de estruturas para a matriz de variância e covariância  $\Sigma$ , associadas aos efeitos de  $u$ . O modelo para os efeitos de genótipos em cada ambiente conduz a um modelo para  $\Sigma$ , no qual a variância genotípica no ambiente  $j$  (Equação 6), a covariância genotípica entre os ambientes  $j$  e  $j'$  (Equação 7) e a correlação genotípica entre os ambientes  $j$  e  $j'$  (Equação 8) são dadas por:

$$\sigma_{gjj} = \sum_{r=1}^k \lambda_{jr}^2 + \Psi_j \quad (6)$$

$$\sigma_{gjj'} = \sum_{r=1}^k \lambda_{jr} \lambda_{j'r} \quad (7)$$

$$\rho_{gjj'} = \sum_{r=1}^k \lambda_{jr} \lambda_{j'r} / \left[ \left( \sum_{r=1}^k \lambda_{jr}^2 + \Psi_j \right) \left( \sum_{r=1}^k \lambda_{j'r}^2 + \Psi_{j'} \right) \right]^{1/2} \quad (8)$$

Os BLUPs dos escores fatoriais dos genótipos em cada ambiente ( $f_s$ ) podem ser obtidos por:

$$f_s = \left[ \hat{\Lambda}' (\hat{\Lambda} \hat{\Lambda}' + \hat{\Psi})^{-1} \otimes I_g \right] u_s \quad (9)$$

Onde  $\hat{\Lambda}$  são cargas fatoriais estimadas;  $\hat{\Psi}$  é matriz diagonal das variâncias específicas estimadas e  $u_s$  é o efeito de genótipo em cada ambiente. Também os BLUPs dos resíduos das interações GA podem ser obtidos por:

$$\delta = \left[ \tilde{\Psi} \left( \hat{\Lambda} \hat{\Lambda}' + \hat{\Psi} \right)^{-1} \otimes I_g \right] u_s \quad (10)$$

Pode ser visto que o modelo fator analítico requer estimativa dos parâmetros  $\Lambda$  e  $\Psi$ , os quais compõem a matriz de variância-covariância, e podem ser estimados por REML através do algoritmo de informação média (AI) (Thompson et al., 2003).

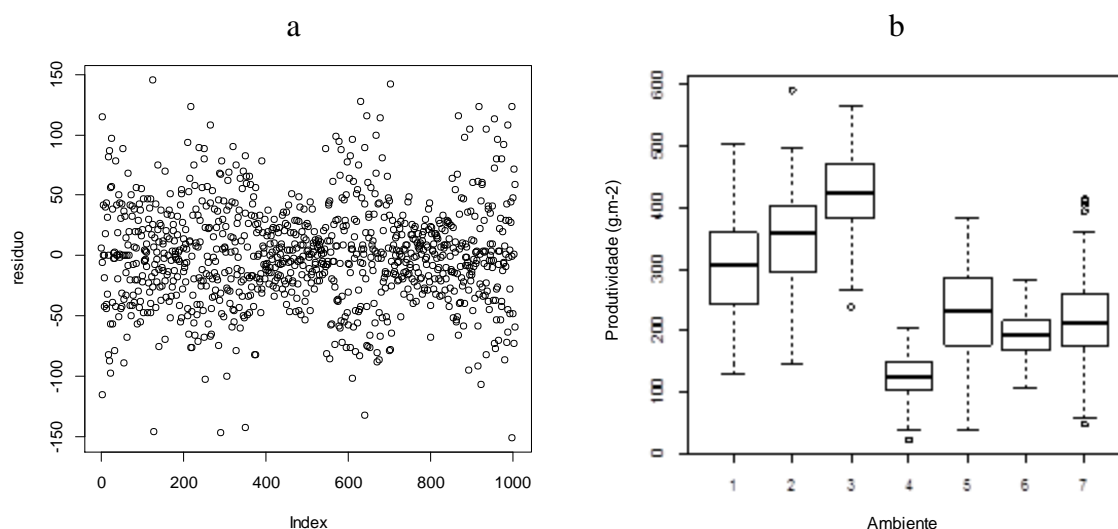
Todos os modelos deste trabalho foram construídos utilizando o programa estatístico ASReml (Gilmour et al., 2009). No programa foi acionado o qualificador !AISINGULARITIES para forçar a continuidade da análise dos dados, mesmo em condições em que ocorreram singularidade na matriz de informação média (AI). Os gráficos foram construídos plotando os escores fatoriais obtidos através da análise de fatores juntamente com as cargas fatoriais dos ambientes. Para isso foi usado o software estatístico R 3.0.2. (R Core Team, 2013).

Com a finalidade de comparar e validar as informações obtidas pelo FAMM, foi feita a análise de cluster e construído o dendrograma. A distância utilizada foi a euclidiana com ligação completa. Para isso foi usado o software estatístico R 3.0.2. (R Core Team, 2013) e o pacotes utilizados foram o Cluster (Maechler et al., 2013) e o NbClust (Charrad et al. 2014).

### 3. Resultados e Discussão

O gráfico de dispersão residual (Figura 1a) indicou a ocorrência de heterogeneidade e não independência dos resíduos. A heterogeneidade dos resíduos pode ser atribuída, neste caso, ao desbalanceamento do conjunto de dados e a falta de independência dos resíduos indicam uma correlação entre os resíduos. A violação da pressuposição de homogeneidade, inviabiliza utilização dos métodos usuais de análise de adaptabilidade e estabilidade baseados na ANOVA. Uma opção tomada por muitos pesquisadores é o relaxamento desta pressuposição, e assim prosseguir com as análises usando os modelos aditivos ou multiplicativos. No entanto ao ignorar esta condição podem ocorrer uma grande taxa de erro Tipo I para comparações

(Teste F) do efeito principal de genótipos (Hu et al., 2013). Resende e Thompson (2004) e Smith (2001) notaram que modelos ajustados com a suposição de variâncias heterogêneas são superiores do que os que admitem homogeneidade de variâncias residual. Na figura 1b, os ambientes 4 e 6, que apresentaram menores produtividades, também apresentaram uma menor amplitude de variação comparados aos mais produtivos. Geralmente, a variação tende a ser maior em ambientes melhores do que em ambientes com pior desempenho, embora o oposto pode ser observado (Malosetti et al., 2013).



**Figura1.** (a) – Gráfico de dispersão residual. (b) - Boxplot para produtividade de genótipos de soja em sete ambientes.

A análise conjunta mostrou efeitos significativos ( $p < 0,01$ ), para genótipos, ambiente e a interação GA. Demonstrou também que o comportamento relativo dos genótipos foi influenciado pelas condições ambientais, tornando viável os estudos de adaptabilidade e estabilidade para esse conjunto de dados. A média geral de produtividade dos genótipos testados foi de  $2693 \text{ kg.ha}^{-1}$  e foi superior a das testemunhas de  $2411 \text{ kg.ha}^{-1}$ .

Em relação às correlações fenotípicas e genotípicas, verificou-se pouca similaridade entre os pares quanto se trata da magnitude das correlações, no entanto com relação ao sentido, as correlações foram muito semelhantes (Tabela 2). De modo geral as correlações genotípicas usando a estrutura FA2 foram baixas. Mas, apesar disso, o modelo FMM é mais confiável para estimação das correlações entre ambientes que os métodos tradicionais, pois ele usa informação de todos os

ambientes simultaneamente para estimar a correlação entre os pares de ambientes, enquanto que os métodos tradicionais usam apenas as informações dos pares de ambientes para qual esta sendo calculada a correlação (Resende e Thompson, 2004).

**Tabela 2:** Matriz de correlação fenotípica (diagonal inferior), herdabilidade (diagonal) e correlação genotípica (diagonal superior) associada ao modelo FA2

Ambiente	1	2	3	4	5	6	7
1	<b>0,570</b>	0,266	0,417	0,704	-0,590	0,214	0,279
2	0,139	<b>0,480</b>	0,078	0,383	-0,198	0,326	0,339
3	0,120	0,025	<b>0,218</b>	0,210	-0,229	-0,026	0,003
4	0,071	0,048	0,018	<b>0,033</b>	-0,521	0,845	0,879
5	-0,059	-0,025	-0,019	-0,017	<b>0,032</b>	-0,271	-0,307
6	0,012	0,022	-0,001	0,015	-0,005	<b>0,010</b>	0,998
7	0,110	0,138	0,001	0,094	-0,032	0,058	<b>0,345</b>

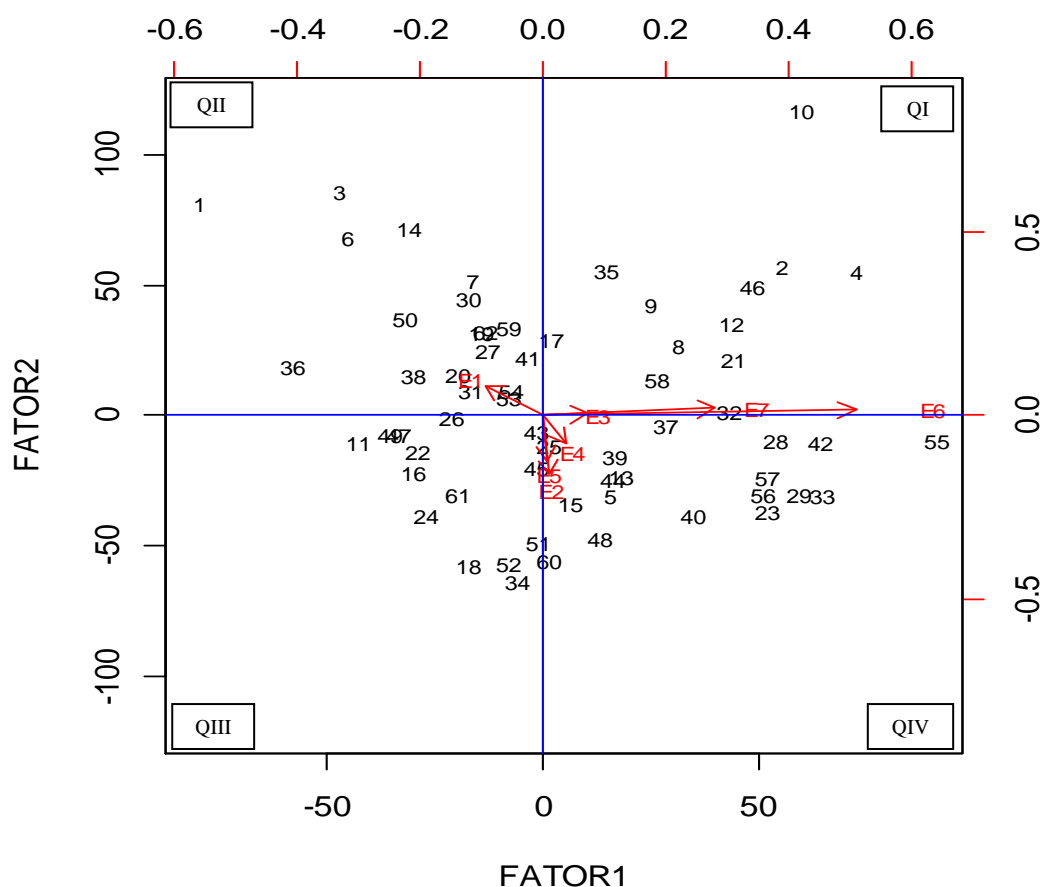
As correlações genéticas entre os ambientes variaram de -0,590 a 0,998, sendo que os ambientes 6 e 7, que apresentaram a maior correlação, foram conduzidos no mesmo ano agrícola. O E4 foi o que apresentou maiores valores de correlações com os ambientes 1, 6 e 7, mas para os demais ambientes os valores foram baixos. As baixas correlações fenotípicas e genotípicas, indicam que os genótipos não apresentaram comportamento semelhante nos ambientes, pois há uma mudança na classificação deles de um ambiente para o outro fazendo com que as magnitudes das correlações sejam menores. Este resultado pode estar associado, não só a influência dos ambientes, mas também à qualidade dos experimentos conduzidos e a perda de parcelas. Como o modelo FAMM usa informação de todos ensaios para estimar as correlações, um mau desempenho em um ambiente pode influenciar negativamente a estimação das correlações para demais.

Os ambientes apresentaram baixas herdabilidade (Tabela 2) evidenciando que os efeitos de ambientes são mais pronunciados que os genotípicos, diminuindo a confiabilidade em usar estes ambientes para o estudo da adaptabilidade e estabilidade dos genótipos.

O modelo FAMM que melhor se ajustou para esse conjunto de dados foi o com dois fatores (FA2). Os modelos foram avaliados pelo testes da razão de verossimilhança, pela comparação do Log L. O modelo FA(1) foi inferior ao FA(2). Já os demais modelos não convergiram, o que pode ter ocorrido devido a problemas com o algoritmo de ajuste do modelo. No entanto as análises não foram comprometidas, pois os dois fatores utilizados no estudo foram suficientes para

explicarem juntos 81,47% da variância total. O primeiro fator explicou aproximadamente 50% e o segundo 30% da variância total.

Os genótipos e ambientes com valores baixos de coordenadas para os eixos são os que relativamente menos contribuíram para interação GA, sendo portanto, fenotipicamente estáveis (Figura 2). Assim os genótipos considerados mais estáveis foram 25, 43, 45 e as testemunhas 53 e 54. Estes genótipos são de ampla recomendação pois também apresentarem produtividades médias desejáveis. Já os genótipos mais distantes da origem são os que mais contribuem para a interação (1, 3, 4, 6, 10, 14, 34 e 36), pois esses apresentam resposta específica na interação e, portanto, são considerados instáveis.



**Figura 2.** Biplot do modelo FAMM com dois fatores para a produtividade de 48 linhagens de soja mais testemunhas. Os ambientes estão representados de vermelho.

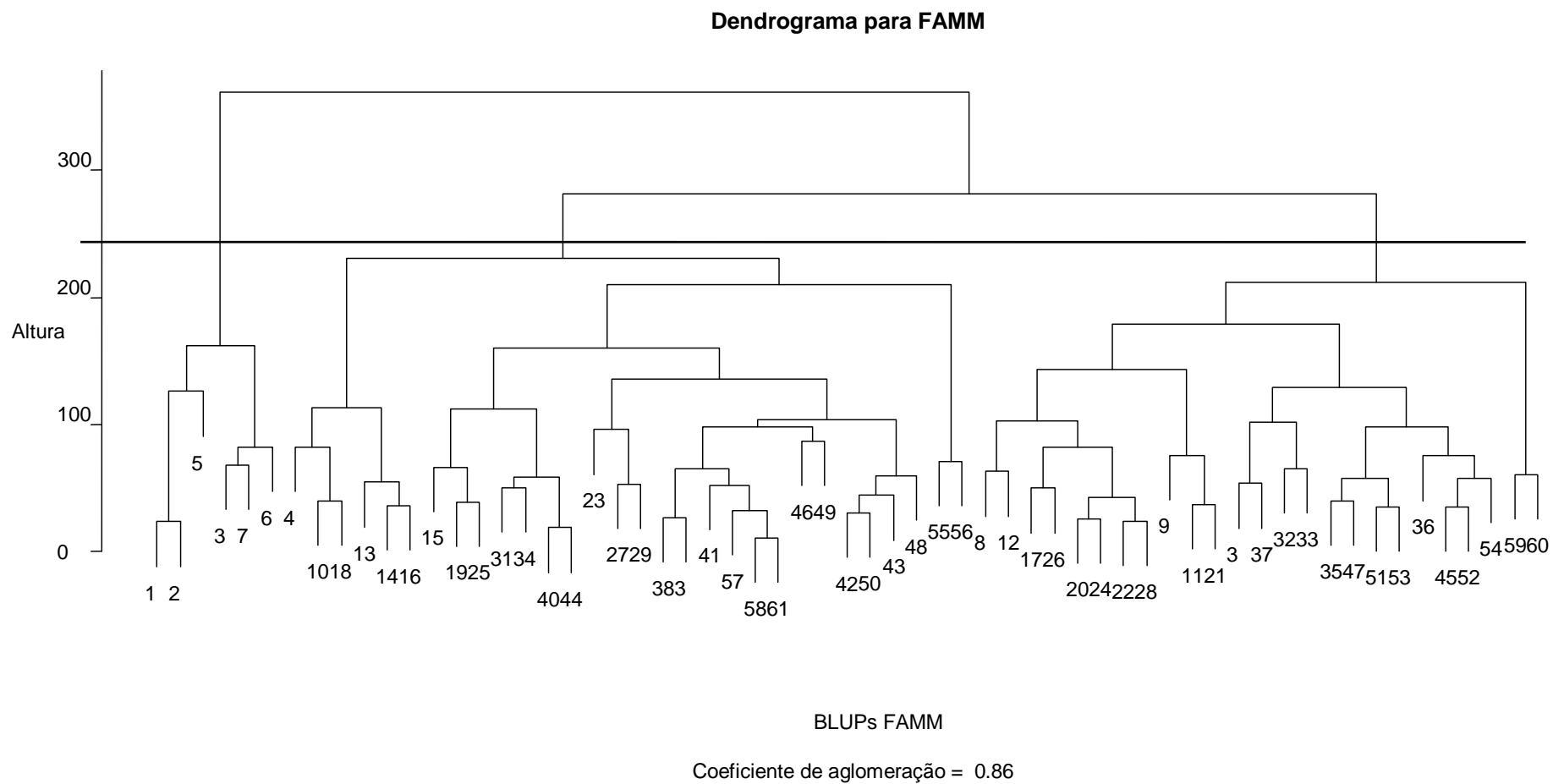
Os ambientes 6 e 7 são os que contribuíram mais para a interação, pois foram os que apresentaram escores mais altos. Neste estudo, um mesmo local apresentou padrão de comportamento diferente ao longo dos anos, indicando a importância de realizar ensaios repetidos durante os anos a fim de interpretar melhor a interação GA

(Figueiredo et al, 2015). Pelo biplot fica evidente que os ambientes conduzidos no mesmo ano, como por exemplo os ambientes 6 e 7 e também os ambientes 5 e 4, tiveram uma tendência de comportamento semelhantes apesar de terem sido conduzidos em localidades diferentes.

O estudo da adaptabilidade é feito observando a disposição dos genótipos nos quadrantes do gráfico e sua proximidade aos eixos. Quanto maior a proximidade ao eixo do fator (1 ou 2), mais adaptado ao ambiente que o fator explica o genótipo será. Nos quadrantes QII e QIV encontram-se os genótipos com adaptabilidade específica aos ambientes agrupados nos fatores 2 e 1 respectivamente. No QI ficam os genótipos que apresentam adaptabilidade aos dois grupos de ambientes e no QIII estão os genótipos não adaptados a nenhum dos grupos de ambientes estudados. Os ambientes 3, 6 e 7, que estão próximos do eixo das abscissas, correspondem aos ambientes que são explicados pelo fator 1. Os ambientes 2, 4 e 5 estão mais próximos do eixo das ordenadas são explicados pelo fator 2. Já o ambiente 1 não apresentou um padrão de explicação sendo, portanto, influenciado pelos dois fatores.

O mais importante neste tipo de estudo é o entendimento de como os genótipos variam em todo espectro ambiental. No entanto, apenas o biplot não é capaz de oferecer uma definição clara dos subconjuntos de genótipos. Portanto o uso da análise de cluster auxilia na construção de subgrupos de genótipos em resposta aos diferentes ambientes. Burguenõ et al. (2008) observou que a análise de agrupamento e a análise fatorial em conjunto são uma eficiente estratégia para estudar a performance de genótipos e ambientes.

A análise de cluster de ambientes e genótipos foi realizada com base na distância euclidiana e o método de ligação completa. O dendrograma (Figura 3) mostra os padrões dos genótipos descritos no biplot na Figura 2. Nele os genótipos foram agrupados em 3 conjuntos, o que concorda com a representação no biplot, o qual apresentou uma divisão dos ambientes em três grupos distintos. O primeiro conjunto contém poucos genótipos, os quais a maioria se encontram no segundo quadrante do biplot e apresentam adaptação específica aos genótipos agrupados no fator 2. Os outros dois conjuntos não se concentraram em quadrantes específicos do gráfico, mas mesmo assim é possível ver conjuntos de genótipos que coincidem tanto no biplot quanto no agrupamento do dendrograma. Por exemplo, os genótipos 23, 29, 56, 57 que estão no mesmo grupo e se encontram no quadrante IV do biplot, estes possuem adaptabilidade específica aos ambientes representados pelo fator 1.



**Figura 3:** Dendrograma dos BLUPs obtidos pelo FAMM dos genótipos e testemunhas e ligação completa pelo método de cluster



Os genótipos 8, 9, 12, 21 e 35, claramente no mesmo agrupamento, estão contidos no quadrante I. Segundo Garbuglio et al. (2007) os genótipos que se encontram no quadrante I tendem a ter as melhores médias de produtividade e no quadrante III as menores. No entanto neste trabalho as melhores produtividades estiveram bem distribuídas nos quadrantes I, II e IV. O fato de serem genótipos em fases finais do programa de melhoramento e já terem passado por várias seleções para produtividade é esperado que eles apresentem maiores produtividades, explicando a homogeneidade de distribuição das melhores médias. A aplicabilidade do modelo FAMM na literatura tem sido feita para conjunto de dados com um grande número de ambientes (Kelly et al., 2007; Smith et al., 2015), com a finalidade de usar fatores que expliquem um conjunto de ambientes altamente correlacionados.

#### **4. Conclusões**

A diferença no agrupamento dos genótipos pelo método FAMM e o dendrograma pode ter sido consequência do número reduzido de ambientes, já que o modelo FAMM é indicado para o caso de grande número de ambientes. No entanto os grupos que apresentaram concordância de agrupamento entre o biplot do método FAMM e o dendrograma indicam a viabilidade de uso do método para os estágios finais do programa de melhoramento.

#### **5. Referências Bibliográficas**

- BURGUEÑO, J.; CROSSA, J.; CORNELIUS, P.L.; YANG, R.C. Using factor analytic models for joining environments and genotypes without crossover genotype  $\times$  environment interaction. **Crop Science**. v. 48, p. 1291-1305, 2008.
- CHARRAD, M.; GHAZZALI, N.; BOITEAU, V.; NIKNAFS, A. (2014). NbClust: NbClust package for determining the best number of clusters. R package version 2.0.
- FIGUEIREDO, A.G., RENZO PINHO, G.V., SILVA, H.D, BALESTRE, M. Application of mixed models for evaluating stability and adaptability of maize using unbalanced data. **Euphytica**. v.202, p. 393-409, 2015.
- GARBUGLIO, D.D.; GERAGE, A.C.; ARAÚJO, P.M.; JUNIOR, N.S.F.; SHIOGA, P.S. Análise de fatores e regressão bissegmentada em estudos de

- estratificação ambiental e adaptabilidade em milho. **Pesquisa Agropecuária Brasileira**. v.42, p.183-191, 2007.
- GILMOUR, A.R.; GOGEL, B.J.; CULLIS, B.R. THOMPSON, R. **ASReml-R reference manual, release 3.0** Technical report. NSW Department of Primary Industries. 2009.
- HENDERSON, C.R. Best Linear Unbiased Estimation and Prediction under a Selection Model. **Biometrics**. v.31, p.423-447, 1975.
- HU, X.; YAN, S.; SHEN, K. Heterogeneity of error variance and its influence on genotype comparison in multi-location trials. **Field Crops Research**. v.149, p.322–328, 2013.
- HU, X.Y.; SPILKE, J. Variance–covariance structure and its influence on variety assessment in regional crop trials. **Field Crops Research**. v. 120, p. 1–8, 2011.
- JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. Englewood Cliffs: Prentice-Hall, 2007. 767 p
- KELLY, A.M., SMITH, A.B., ECCLESTON, J.A., CULLIS, B.R. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. **Crop Science**. v.47, p.1063–1070, 2007.
- MAECHLER, M.; ROUSSEEUW, P.; STRUYF, A.; HUBERT, M.; HORNIK, K.(2013). cluster: Cluster Analysis Basics and Extensions. R package version 1.14.4.
- MALOSETTI, M.; RIBAUT, J.M.; EEUWIJK, F.A.V. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. **Frontiers in Physiology**. v.4, p. 1-17, 2013.
- MEYER, K. Factor-analytic models for genotype  $\times$  environment type problems and structured covariance matrices. **Genetics Selection Evolution**. v. 41, p. 1-11, 2009.
- PIEPHO, H. P. Empirical best linear unbiased prediction in cultivar trials using factor analytic variance-covariance structures. **Theoretical and Applied Genetics**. v.97, p.195-201, 1998.
- PIEPHO, H.P., VAN EEUWIJK, F. Stability analysis of crop performance evaluation. In: Kang, M.S. **Crop improvement: Challenges in the 21st century**. Haworth Press, New York, 2002, p 315–351

- PIEPHO, H.P.; MÖHRING, J. Generation means analysis using mixed models. **Crop Science**. v.50, p. 1674–1680, 2010.
- PINHEIRO, L.C.M.; GOD, P.I.V.G.; FARIA, V.R.; OLIVEIRA, A.G.; HASUI, A.A.; PINTO, E.H.G.; ARRUDA, K.M.A.; PIOVESAN, N.D.; MOREIRA, M.A. Parentesco na seleção para produtividade e teores de óleo e proteína em soja via modelos mistos. **Pesquisa Brasileira Agropecuária**. v.48, p.1246-1253, 2013.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- RESENDE, M.D.V.; THOMPSON, R. Factor analytic multiplicative mixed models in the analysis of multiple experiments. **Revista Matemática e Estatística**. v.22, p. 31-52, 2004.
- SMITH, A. B., CULLINS, B. R.; THOMPSON, R. The Analysis of Crop Cultivar Breeding and Evaluation Trials: An Overview of Current Mixed Model Approaches. **Journal of Agricultural Science**, v. 143, 449-462, 2005.
- SMITH, A.B.; CULLIS, B.R.; THOMPSON, R. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. **Biometrics**, v. 57, p. 1138-1147, 2001.
- SMITH, A.B; GANESALINGAM, A.; KUCHEL, H.; CULLINS, B. R. Factor analytic mixed models for the provision of grower information from national crop variety testing programs. **Theoretical and Applied Genetics**. v.128, p.55-72, 2015.
- SPIPKE, J.; PIEPHO, H. P.; HU, X. Analysis of Unbalanced Data by Mixed Linear Models Using the mixed Procedure of the SAS System. **Journal Agronomy and Crop Science**. v. 191, p.47-54, 2005.
- STEFANOVA, K.T.; BUIRCHELL, B. Multiplicative mixed models for genetic gain assessment in lupin breeding. **Crop Science**. v. 50, p. 880–891, 2010.
- THOMPSON, R.; CULLIS, B. R.; SMITH, A. B.; GILMOUR, A. R. A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. **Australian and New Zealand Journal of Statistics**. v.45, p.445-59, 2003.
- YOKOMIZO, G.K.; DUARTE, J.B.; VELLO, N.A.; UNFRIED, J.R. Análise AMMI da produtividade de grãos em linhagens de soja selecionadas para

resistência à ferrugem asiática. **Pesquisa agropecuária brasileira**  
v.48, p. 1376-1384, 2013.

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

## CONCLUSÕES GERAIS

A classificação dos genótipos quanto as padrões de adaptabilidade e estabilidade usando os modelos AMMI e FAMM na maior parte não coincidiram.

O descarte de repetições e a imputação dos dados para realizar da análise pelo modelo AMMI, além de tornar as análises mais trabalhosas, podem também comprometer a recomendação dos genótipos aos ambientes de cultivo.

O AMMI é uma ferramenta mais utilizada pelos melhoristas, isso se dá ao fato de ser de mais fácil manipulação, no entanto o FAMM, apesar de ser uma análise mais complexa, apresenta resultados mais confiáveis, pois não é necessário o ajuste dos dados ao modelo e seus resultados são expressos em valores genéticos.

## ANEXOS

### ANEXO A: Rotina no R utilizada para a aplicação do método AMMI

```
#pacote usado para análise dos dados
library(agricolae)
#Data input
dadosA<-read.table("dadosAMMI.txt",header=T)
dadosA
str(dadosA)
dadosA$site<-as.factor(dadosA$site)
dadosA$rep<-as.factor(dadosA$rep)
dadosA$id<-as.factor(dadosA$id) #id: refere-se aos genótipos
attach(dadosA)
ls()
#modelo AMMI
mod.AMMI<-AMMI(ENV=site,GEN=id,REP=rep,Y=y,console=TRUE,PC=TRUE)
# biplot PC1 vs Produtividade
plot(mod.AMMI, first=0,second=1,number=TRUE)
# biplot PC1 vs PC2
plot(mod.AMMI,type=1,las=1,number=TRUE,length=0.08,lwd=1.0)
```

### ANEXO B: Rotina no ASReml para a aplicação do método FAMM

```
##modelo completo (sem fatores)
!WORKSPACE
test matrix 100x100 220311
id      !P      #indica a existência de pedigree
site    !A      #campo alfanumérico
bloco   !A      #campo alfanumérico
test    !A      #campo alfanumérico
y                #variável resposta
pedigree2.ped !REPEAT      #ignorar repetições de linhas no arquivo pedigree
dados.dat !SKIP1 !SECTION site #pular linha de cabeçalhos #especifica o fator de dados que
definem as seções de dados
!MAXIT 500      #define o número máximo de interações
!AISINGULARITIES #usada para continuar as análises mesmo que uma singularidade for
detectada na matriz de informação média (AI)
y ~ mu site site/bloco test !r id      #modelo sem fatores #!r: especifica a parte aleatória do modelo
```

```

##com 1 fator (FA1)
!WORKSPACE
test matrix 100x100 220311
id      !P      #indica a existência de pedigree
site    !A      #campo alfanumérico
bloco   !A      #campo alfanumérico
test    !A      #campo alfanumérico
y              #variável resposta
pedigree2.ped !REPEAT
dados.dat !SKIP1 !SECTION site
!MAXIT 500
!AISINGULARITIES
y ~ mu site site/bloco test !r id.xfa(site,1) #!r: especifica a parte aleatória do modelo
0 0 1
id.xfa(site,1) 2
id
xfa(site,1) 0 XFA1 !GP #indicando qual é modelos fator analítico (FA1)
7*0.5 # Psi, variância específica inicial
7*0.5 # loadings iniciais para o primeiro fator

```

```

##com 2 fatores (FA2)
!WORKSPACE
test matrix 100x100 220311
id      !P      #indica a existência de pedigree
site    !A      #campo alfanumérico
bloco   !A      #campo alfanumérico
test    !A      #campo alfanumérico
y              #variável resposta
pedigree2.ped !REPEAT
dados.dat !SKIP1 !SECTION site
!CONTINUE #recuperar os parâmetros de variância do arquivo .rsv
!MAXIT 500
!AISINGULARITIES
y ~ mu site site/bloco test !r id.xfa(site,2)
0 0 1
id.xfa(site,2) 2
id
xfa(site,2) 0 XFA2 !GP #indicando qual é modelos fator analítico (FA2)
7*1      # Psi, variância específica inicial
7*100    # loadings iniciais para o primeiro fator

```

7\*100 # loadings iniciais para o segundo fator

#### ANEXO C: Rotina no R para construção do biplot com os escores do modelo FAMM

```
# lendo PSI and loadings
XFAm <- matrix(scan("dados_graf.txt"),7,3)
dimnames(XFAm) <- list(paste('E',1:7,sep=""),c("Psi","Load1","Load2"))
ss <- svd(XFAm[,-1])
Lam <- XFAm[,-1] %*% ss$v
colnames(Lam) <- c("Load1","Load2")
Gvar <- Lam %*% t(Lam) + diag(XFAm[,1])
cLam <- diag(1/sqrt(diag(Gvar))) %*% Lam
# Escores
XFA5ln <- read.table("pedro6.sln")
XFA5 <- matrix(scan("pedro7.txt"),62,2)
scores <- XFA5 %*% ss$v
dimnames(scores) <- list(paste(' ',1:62,sep=""),paste('FATOR',1:2,sep=""))
#Biplot para FA2
biplot(scores[,1:2],Lam[,1:2],cex=0.7,ylim=c(-115,115))
abline(h=0,col=4) #linha horizontal
abline(v=0,col=4) #linha vertical
```

#### ANEXO D: Rotina no R para a análise de cluster e construção do dendograma

```
require(cluster)
Famm=read.table("famm.txt",header=T)
Famm
d.famm<-daisy(Famm, metric = "euclidean", stand = FALSE)
d.famm
d.famm1<-agnes(Famm, metric = "euclidean", method="complete", stand = F)
d.famm1
plot(d.famm1)
####Pacote 'NbClust'
install.packages("NbClust")
require(NbClust)
?NbClust
NbClust(Famm, distance = "euclidean", min.nc=2, max.nc=8, method = "complete", index = "ch")
```