

# ANÁLISE DA INTERAÇÃO GENÓTIPO POR AMBIENTE EM ENSAIOS MULTI-AMBIENTAIS: COMPARAÇÃO ENTRE OS MODELOS AMMI E GGE BILOT

Elizabete Cristina Röpk<sup>1</sup>  
Jéssica Maria Magalhães Moreno<sup>2</sup>  
Kuang Hongyu<sup>3</sup>.

**RESUMO:** O trigo, um dos grãos mais produzidos globalmente, foi o segundo mais cultivado em 2022, com 808,441 milhões de toneladas, seguindo o milho e precedendo o arroz. Sendo um alimento básico desde a antiguidade, é usado na produção de diversos produtos alimentícios. A interação genótipo-ambiente ( $G \times E$ ) é essencial na avaliação de novas variedades de plantas, incluindo o trigo, devido às respostas distintas que podem apresentar às variações ambientais. A seleção do método adequado para estudos de  $G \times E$  depende dos dados experimentais. Esses estudos têm sido essenciais no melhoramento de plantas, visando compreender e lidar com a interação genótipo-ambiente ( $G \times E$ ). Os ensaios multi-ambientes (METs) têm como principais objetivos investigar a  $G \times E$ , avaliar adaptabilidade e estabilidade genotípica, estabelecer relações entre genótipos e ambientes e fazer previsões de valor de reprodução dos genótipos. A compreensão e gestão da interação  $G \times E$  são fundamentais para o sucesso no melhoramento de plantas. As abordagens recentes incluem o modelo AMMI e o modelo GGE biplot, ambos representados graficamente. A combinação dos modelos AMMI e GGE Biplot proporciona uma abordagem mais eficiente para explorar a interação genótipo x ambiente, apoiando a seleção de genótipos e ambientes superiores. Os resultados obtidos para os dados MET mostraram uma semelhança significativa entre os modelos, com o AMMI explicando 89,8% da variação nos dois primeiros componentes principais, enquanto o GGE Biplot explicou 87,7%. O modelo AMMI é confiável para identificar locais de cultivo favoráveis e genótipos com bom desempenho médio, enquanto o GGE Biplot destaca-se na análise detalhada da interação genótipo-ambiente, facilitando a seleção de genótipos estáveis e adaptados.

**Palavras-chave:** Mega ambiente, ideótipo, estabilidade, componentes principais.

## ANALYSIS OF GENOTYPE BY ENVIRONMENT INTERACTION IN MULTI-ENVIRONMENT TRIALS: COMPARISON BETWEEN AMMI AND GGE BILOT MODELS

**ABSTRACT:** Wheat, one of the most produced grains globally, was the second most cultivated in 2022, with 808.441 million tonnes, following corn and preceding rice. Being a staple food since antiquity, it is used in the production of various food products. The genotype-environment interaction ( $G \times E$ ) is essential in the evaluation of new plant varieties, including wheat, due to the distinct responses they may present to environmental variations. The selection of the appropriate method for  $G \times E$  studies depends on the experimental data. These studies have been essential in plant breeding, aiming to understand and deal with the genotype-environment interaction ( $G \times E$ ). The main objectives of multi-environment assays (METs) are to investigate  $G \times E$ , to evaluate genotypic adaptability and stability, to establish relationships between genotypes and environments, and to make predictions of genotype reproductive value. Recent approaches include the AMMI model and the GGE biplot model, both of which are graphically represented. The combination of the AMMI and GGE Biplot models provides a more efficient approach to exploit the genotype x environment interaction, supporting the selection of genotypes and superior environments. The results obtained for the MET data showed a significant similarity between the models, with the AMMI explaining 89.8% of the variation in the first two principal components, while the GGE Biplot explained 87.7%. The AMMI model is reliable for identifying favorable growing sites and genotypes with good average performance, while the GGE Biplot excels in the detailed analysis of genotype-environment interaction, facilitating the selection of stable and adapted genotypes.

**Keywords:** Mega environment, ideotype, stability, principal components.

<sup>1, 2</sup> Graduanda no Departamento de Estatística, Instituto de Ciências Exatas e da Terra – ICET/UFMT (elizabeteropck@gmail.com; jessicamariamoreno@outlook.com)

<sup>3</sup>Professor Adjunto do Departamento de Estatística, ICET/UFMT (prof.kuang@gmail.com)

## 1. INTRODUÇÃO

O trigo é um dos três grãos mais produzidos globalmente, juntamente com o milho e o arroz. Em 2022, sua produção mundial foi de 808,441 milhões de toneladas, ficando em segundo lugar, atrás do milho (1,163.5 milhões de toneladas) e à frente do arroz (776,461 milhões de toneladas) (FAOSTAT, 2024). Este grão tem sido o alimento mais consumido pelas populações ocidentais desde a antiguidade. É empregado na produção de farinha, integral, semolina, cerveja e uma diversidade de produtos alimentícios (KUANG apud FABIANE, 2017).

Quando novas variedades de plantas são introduzidas, é crucial avaliar seu desempenho em diferentes ambientes. Isso ocorre porque essas variedades podem apresentar respostas distintas às variações ambientais, um fenômeno conhecido como interação genótipo-ambiente ( $G \times E$ ). O trigo, por exemplo, não está imune a esses efeitos, que podem complicar a seleção de genótipos com ampla adaptação ou estabilidade (YOKOMIZO et al., 2016).

Essa interação genótipo por ambiente ( $G \times E$ ) tem sido um tópico de interesse para biometristas e geneticistas quantitativos desde o início do século XX (YAN & KANG, 2003). A percepção de que a  $G \times E$  pode ser indesejável ou pode confundir a avaliação de genótipos levou ao esforço considerável para desenvolver índices de estabilidade que possam quantificar e selecionar contra a  $G \times E$ . Existem vários métodos estatísticos disponíveis para avaliar a  $G \times E$ , com o objetivo de compreender melhor este efeito.

A escolha do método mais adequado depende dos dados experimentais. Os estudos de  $G \times E$  têm recebido destaque e aplicabilidade ampla nas últimas duas décadas (Hongyu et al., 2015). No campo do melhoramento de plantas, a importância desses estudos de  $G \times E$  é evidente. Os principais objetivos dos ensaios multi-ambientes (METs) são: (i) investigar a  $G \times E$ , (ii) avaliar a adaptabilidade e estabilidade genotípica, (iii) estabelecer relações entre os testes de ambientes, entre os genótipos, e entre genótipos e ambientes (ou locais) simultaneamente, e (iv) fazer previsões do valor de reprodução dos genótipos que permitirão uma seleção precisa dos pais para o próximo ciclo de reprodução.

A presença da  $G \times E$  complica este processo e geralmente é expressa como respostas inconsistentes de alguns genótipos em relação a outros, devido à mudança de classificação genotípica, ou como mudanças nas diferenças absolutas entre genótipos sem mudança de classificação (GAUCH, 2013; HONGYU et al., 2015). Portanto, a compreensão e a gestão da  $G \times E$  são fundamentais para o sucesso do melhoramento de plantas.

Existem várias teorias que avaliam a adaptabilidade e estabilidade, cada uma com seus próprios conceitos e procedimentos de estimação biométrica (Camargo-Buitrago et al., 2011). Entre as abordagens mais recentes, temos o modelo AMMI, um método estatístico que busca entender a estrutura das interações entre genótipos e ambientes, e o modelo GGE biplot, que leva em conta tanto o efeito principal do genótipo quanto a interação entre genótipo e ambiente.

Ambas as análises, que se baseiam em gráficos biplot, representam uma matriz de dados. A principal diferença entre os modelos AMMI e GGE reside na fase inicial da análise: o modelo GGE analisa diretamente o efeito de  $G + G \times E$ , enquanto o AMMI separa  $G$  de  $G \times E$  e, na fase final da análise, forma gráficos biplot para ambas as metodologias. No entanto, essa separação não confere superioridade à análise AMMI (GAUCH, 2013; HONGYU et al., 2015).

Ao comparar as metodologias AMMI e GGE Biplot na formação de mega ambientes, foi relatada uma coincidência de 95,2% nos resultados entre as metodologias (Camargo-Buitrago et al. (2011). Portanto, o modelo GGE biplot é mais adequado para identificar mega ambientes, selecionar ambientes representativos e discriminativos e indicar cultivares mais adaptadas e estáveis para ambientes específicos (YAN, 2011).

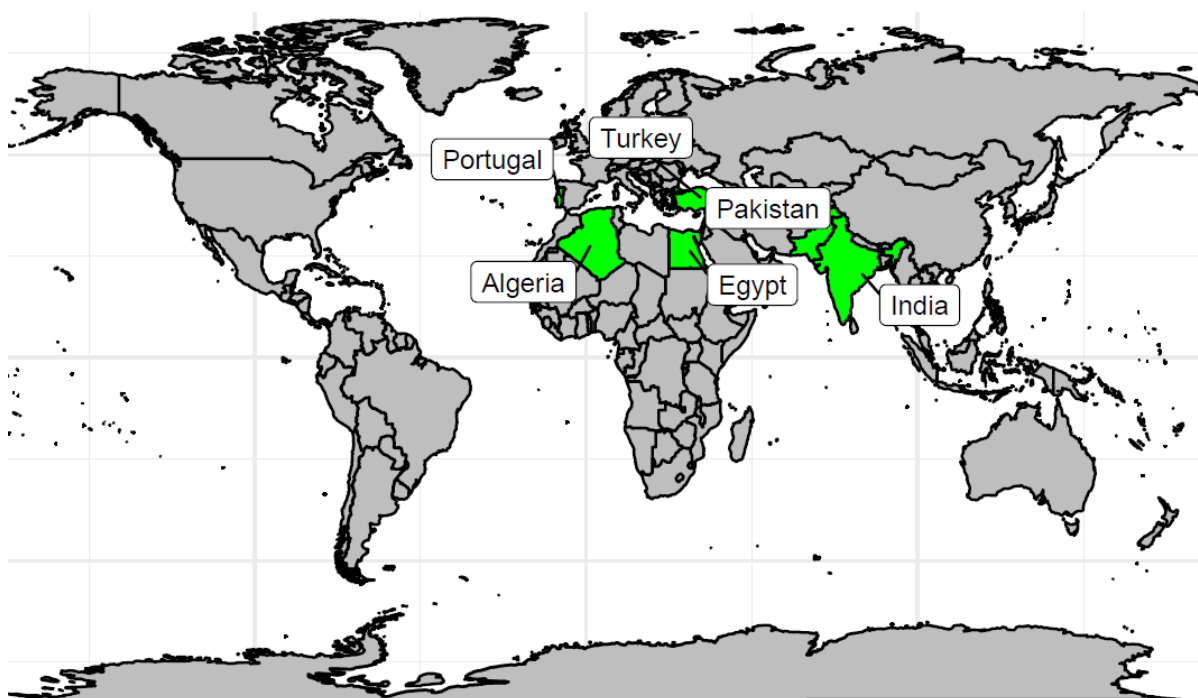
Por outro lado, a análise AMMI pode ser eficazmente utilizada para identificar condições ambientais superiores para exploração agrícola (seleção de locais de cultivo) e genótipos de desempenho médio superior (Yan, 2011; Hongyu et al., 2015). Este estudo se propõe a empregar a técnica de análise de GGE biplot para dados multi-ambientes (MET) com os seguintes objetivos: Primeiro, explorar mega ambientes para entender melhor o ambiente alvo; Segundo avaliar genótipos e ambientes individualmente dentro de cada mega ambiente; Terceiro, investigar as razões por trás da interação genótipo-ambiente; Por último, desenvolver uma nova abordagem para comparar os modelos AMMI e GGE biplot.

## 2. MATERIAIS E MÉTODOS

### 2.1. Biplot para tabela de dupla entrada

Foi utilizado neste trabalho, o conjunto de dados MET para avaliar o rendimento de trigo. Os dados foram fornecidos pelo CIMMYT (International Maize Wheat Improvement Center) em experimentos realizados em seis países, envolvendo 16 genótipos de trigo. Cada genótipo foi avaliado em 06 ambientes diferentes, com 02 repetições realizadas em 2020. O IBWSN (53rd International Bread Wheat Screening Nursery) é um projeto que tem como objetivo avaliar diversas linhas de trigo em diferentes condições ambientais.

Através deste projeto, os pesquisadores podem identificar as linhas que são mais resistentes a doenças e mais produtivas em diversas condições. Essas descobertas podem levar ao desenvolvimento de novas variedades de trigo que podem ser cultivadas em regiões onde as doenças do trigo são comuns, contribuindo assim para a melhoria da produção de alimentos nessas áreas. Além disso, essas informações podem ser usadas para informar estratégias de manejo de culturas, auxiliando os agricultores a tomarem decisões mais informadas sobre o plantio e a colheita de trigo.



**Figura 1 – Mapa mundial mostrando a distribuição geográfica dos ambientes.**

**Fonte: Elaboração própria.**

O biplot, uma técnica introduzida por Gabriel (1971), é uma representação gráfica dos resultados de uma análise de componentes principais ou de uma decomposição em valores singulares (DVS). Neste método, cada elemento de uma tabela bidimensional pode ser interpretado através do produto de vetores e do cosseno do ângulo entre dois vetores, conforme explicado por Yan & Kang (2003). Se duas matrizes tiverem o mesmo número de linhas e colunas, elas podem ser multiplicadas. A matriz resultante dessa multiplicação mantém o número de linhas e colunas das matrizes originais.

Considerando uma matriz de dados bidimensional  $\mathbf{Z}$  com elementos  $z_{ij}$ , onde  $i = 1, \dots, g$  representa as linhas (genótipos) e  $j = 1, \dots, e$  representa as colunas (ambientes), essa matriz pode ser decomposta por DVS em  $p$  componentes principais (PC).

$$Z_{ij} = \sum_{k=1}^p \lambda_k a_{ik} y_{jk} + e_{ij} \quad (1)$$

onde  $p \leq \min(e, g - 1)$ . Todo PC é composto pela matriz de escores genotípicos  $\alpha_{ik}$ , matriz de escores ambientais  $\gamma_{jk}$ , valor singular  $\lambda_k$  e  $\epsilon_{jk}$  resíduo para o genótipo  $i$  no ambiente  $j$  que não é explicada pelo modelo. O modelo com algumas restrições é dado por:

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$  e com ortonormalidade nas escores  $\alpha_{jk}$ , isto é,  $\sum_{k=1}^g \alpha_{ik} \alpha_{ik'} = 1$ , (se  $k = k'$ ) e  $\sum_{k=1}^g \alpha_{ik} \alpha_{ik'} = 0$ , (se  $k \neq k'$ ); com as restrições semelhantes para  $\gamma_{jk}$  (Yan, 2011).

Quando a matriz  $\mathbf{Z}$  de posto  $p$  pode ser suficientemente aproximada por uma matriz de posto 2 (os primeiros dois componentes são os mais importantes, pois explicam maiores variações dos dados), isto é,

$$Z_{ij} = \lambda_1 \alpha_{i1} \gamma_{j1} + \lambda_2 \alpha_{i2} \gamma_{j2} + \epsilon_{ij}, \quad (2)$$

Pode ser apresentada graficamente em um biplot de dimensão 2 após um particionamento em valor singular apropriado.

$$Z_{ij} = (\lambda_1^f \alpha_{i1}) (\lambda_1^{-f} \gamma_{j1}) + (\lambda_2^f \alpha_{i2}) (\lambda_2^{-f} \gamma_{j2}) + e_{ij}, \quad (3)$$

em que  $f = [0; 0,5; 1]$  é o fator de partição em valor singular (PVS).

O biplot é construído por meio da representação gráfica  $\lambda_1^f \alpha_{i1}$  como abcissa,  $\lambda_2^f \alpha_{i2}$  como ordenada para cada genótipo, e ao mesmo tempo traçando  $\lambda_1^{-f} \gamma_{j1}$  como abcissa  $\lambda_2^{-f} \gamma_{j2}$  como ordenada para cada ambiente (YAN & KANG 2003). O expoente  $f$  tem a função de ajustar os escores das linhas e colunas, otimizando a visualização do biplot para um objetivo específico.

No cenário dos dados MET, os valores singulares são totalmente atribuídos aos escores do genótipo (linha) quando  $f = 1$ , conhecido como “particionamento de valor singular centrado no genótipo” ou PVS = 1 (Yan, 2002). Se  $f = 0$ , os valores singulares são completamente alocados aos escores do ambiente (coluna), chamado de “particionamento de valor singular  $\lambda_k$  centrado no ambiente” ou PVS = 2. Quando  $f = 0,5$ , as raízes quadradas dos valores singulares são distribuídas igualmente entre os escores de genótipo e ambiente, denominado “particionamento de valor singular simétrico” ou PVS = 3.

Na análise GGE biplot, a partição de valor singular centrada no genótipo e a centrada no ambiente são empregadas para a avaliação de genótipos e ambientes de teste, respectivamente (Yan, 2011). Uma característica decisiva do biplot é que a aproximação de

qualquer elemento da matriz original  $\mathbf{Z}$  de rank 2 pode ser estimada visualmente pelo produto escalar dos vetores de genótipo e ambiente, respectivamente, e o cosseno do ângulo entre eles. Esta é conhecida como a propriedade do produto escalar do biplot (YAN & HOLLAND, 2010; YAN, 2011).

## 2.2. Formação do GGE biplot

O modelo GGE biplot (Yan & Kang, 2003; Yan, 2011), que leva em consideração tanto o efeito principal do genótipo quanto a interação genótipo x ambiente ( $G \times E$ ), é fundamentado em gráficos biplot e permite a representação gráfica de uma matriz de dados. Este Biplot é estruturado com base nos dois primeiros componentes principais de uma Análise de Componentes Principais (ACP), empregando Modelos de Regressões Locais (SREG).

O primeiro componente, que se correlaciona fortemente com o efeito principal do genótipo, expressa a proporção do rendimento atribuída exclusivamente às características do genótipo. O segundo componente representa a parcela do rendimento devida à interação  $G \times E$  (YAN & HOLLAND, 2010; YAN, 2011).

Quando diferentes cultivares se adaptam a diferentes grupos de ambientes e a variação entre os grupos supera a variação interna, ocorre a formação de um mega ambiente (YAN & KANG, 2003). A definição de mega ambientes e a relação entre os ambientes auxiliam os melhoristas de plantas na identificação de genótipos com adaptação ampla ou específica a determinados ambientes ou grupos de ambientes (HONGYU et al., 2015).

Na análise GGE biplot, ao estudar mega ambientes, a média no gráfico não está relacionada à média geral, mas sim à média do mega ambiente. Esta abordagem facilita a identificação de genótipos com adaptação ampla ou específica a determinados ambientes ou grupos de ambientes (Yan & Kang, 2003).

Para um conjunto de dados MET, cada valor na tabela é a produtividade média de um genótipo em um ambiente ( $y_{ij}$ ), que é a soma da média geral ( $\mu$ ), o efeito principal do genótipo ( $G_i$ ) para o ambiente particular ( $E_j$ ), e a interação específica ( $G \times E$ ) entre o genótipo e o ambiente ( $\phi_{ij}$ ), ignorando quaisquer erros aleatórios (YAN, 2011):

$$y_{ij} = \mu + G_i + E_j + \phi_{ij}, \quad (4)$$

O modelo GGE biplot não separa os efeitos do genótipo e da  $G \times E$ , mantendo-os juntos em dois termos multiplicativos, que podem ser visualizados na equação (4) da metodologia SREG.

Desde que apenas o  $G$  e  $G \times E$  são pertinentes à avaliação genótipo, avaliação ambiente de teste, e delineamento de mega ambiente; o efeito principal do ambiente  $E$  e a média geral devem ser removidas de cada elemento para apenas manter  $G$  e  $G \times E$  na tabela de dupla entrada (YAN, 2011):

$$Y_{ij} - \mu - E_j = G_i + \phi_{ij}, \quad (5)$$

Os dados MET ambiente-centrado, após dimensionamento apropriado dos dados, são submetidas à DVS e análise de biplot (Yan et al., 2000; Yan, 2011). Yan & Tinker (2006) propôs uma “relação de informação” (IR) para avaliar a adequação de um biplot em exibir os padrões de uma tabela de dupla entrada. Suponha que esta tabela de dupla entrada tem  $g$  genótipos e ambientes. O número máximo de PCs é requerido para representar completamente esta tabela é  $k = \min(e, g - 1)$ . Se não há correlação entre os ambientes, todos os  $k$  PCs

devem ser completamente independentes e a proporção da variação total explicada por cada PC deve ser exatamente  $\frac{1}{k}$ .

Quando existir alguma correlação entre os ambientes, a proporção da variação explicada pelos primeiros PCs deve ser maior do que  $\frac{1}{k}$ , e a variação explicada por outros PCs deve ser inferior ou igual a  $\frac{1}{k}$  (Yan & Tinker, 2006; Yan, 2011). A IR pode ser calculada para cada PC, que é a proporção da variação total explicada por cada PC multiplicado por  $k$ . A interpretação é a seguinte: um PC com  $IR \geq 1$  contém padrões (associações entre ambientes), e um PC com  $IR < 1$  não contém qualquer padrão ou informação.

O biplot de dimensão 2 representa adequadamente os padrões nos dados, se apenas os dois primeiros PCs têm um  $IR \geq 1$  (Yan & Tinker, 2006). Todas as análises deste trabalho foram feitas por meio de rotinas computacionais implementadas no software R (R Development Core Team, 2023).

### 2.3. Mega Ambiente

Quando diversas cultivares se adaptam a diferentes grupos de ambientes e a variação entre os grupos supera a variação interna, ocorre a formação de um “mega ambiente” (YAN & KANG, 2003; HONGYU et al., 2015). A definição de mega ambientes e a relação entre os ambientes auxiliam os melhoristas de plantas na identificação de genótipos com adaptação ampla ou específica a determinados ambientes ou grupos de ambientes (SILVA & BENIN, 2012).

Na análise GGE biplot, ao estudar mega ambientes, a média no gráfico não está relacionada à média geral, mas sim à média do mega ambiente. Esta abordagem facilita a identificação de genótipos com adaptação ampla ou específica a determinados ambientes ou grupos de ambientes. O propósito da análise de mega ambiente é tentar dividir uma região de corte alvo em sub-regiões significativas para que a interação genótipo x ambiente ( $G \times E$ ) possa ser explorada (YAN & KANG, 2003; YAN & TINKER, 2006; YAN, 2011).

Quando um GGE biplot de dimensão 2 é considerado uma aproximação suficiente dos dados (também chamado de “Which-won-where”), torna-se uma ferramenta eficaz para a análise de mega ambiente. Esta perspectiva é composta por um polígono irregular e um conjunto de linhas retas (com o mesmo número de lados do polígono) que irradiam a partir da origem do biplot e intersectam cada um dos lados do polígono perpendicularmente. Os vértices do polígono são os marcadores de genótipos localizados mais distantes da origem do biplot em todas as direções, de modo que todos os genótipos estão contidos dentro do polígono (YAN, 2011).

Uma linha que cruza perpendicularmente um lado do polígono representa ambientes hipotéticos, nos quais os dois genótipos que definem esse lado do polígono apresentaram boas produtividades; a posição relativa dos dois genótipos seria invertida em ambientes em lados opostos da linha ( $G \times E$  cruzado). Assim, as linhas que se irradiam dividem o biplot em setores, para cada setor, existe um genótipo no vértice, que apresentou melhor desempenho para ambientes que se enquadram nesse setor (YAN & KANG, 2003; YAN, 2011).

### 2.4. Modelo AMMI

No modelo AMMI dois métodos são combinados integrando-os em um único modelo: a análise de variância e a decomposição por valor singular. Esse modelo pressupõe componentes aditivos para os efeitos principais de genótipos ( $g_i$ ) e ambientes ( $e_j$ ) e considera

componentes multiplicativos para os efeitos da interação  $(ge)_{ij}$  para a resposta média do  $i$ -ésimo genótipo no  $j$ -ésimo ambiente em  $r$  blocos (repetições) é representada pelo modelo matemático (GAUCH, 2013; DIAS, 2014; HONGYU et al., 2015).

$$Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^n \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij} + \epsilon_{ij}, \quad (6)$$

em que:  $Y_{ij}$  é a média do  $i$ -ésimo genótipo no  $j$ -ésimo ambiente, com  $i = 1, 2, \dots, g$  e  $j = 1, 2, \dots, e$ ;  $\mu$  é a média geral;  $g_i$  e  $e_j$  são os efeitos do  $i$ -ésimo genótipo e  $j$ -ésimo ambiente, respectivamente;  $\lambda_k$  é o  $k$ -ésimo valor singular da matriz  $(GE)$  e com  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ;  $\gamma_{ik}$  e  $\alpha_{jk}$  são elementos dos  $k$ -ésimos valores singulares correspondentes ao  $i$ -ésimo genótipo e  $j$ -ésimo ambiente respectivamente;  $\rho_{ij}$  são o resíduo da interação  $G \times E$ ;  $n$  é o número de eixos ou componentes principais retidos pelo modelo;  $r$  é o número de blocos e  $\epsilon_{ij}$  o erro médio experimental associado ao  $i$ -ésimo genótipo no  $j$ -ésimo ambiente, assumidos independentes, e  $\epsilon_{ij} \sim N\left(0, \frac{\sigma^2}{r}\right)$ .

Na primeira fase os efeitos principais, na parte aditiva (média geral, efeitos de genótipos e ambientes), são ajustados por uma análise de variância comum aplicada à matriz de média  $(Y(g \times e))$  resultando em um resíduo de não aditividade, isto é, na interação  $G \times E$ , dada por  $(ge)_{ij}$ , essa interação constitui a parte multiplicativa do modelo. Na segunda fase a interação é analisada pela decomposição por valores singulares (DVS) ou por análise de componentes principais da matriz de interações  $(GE(gxe) = [(ge)_{ij}])$ .

A matriz  $GE$  é a matriz de interação entre os genótipos e os ambientes, (matriz de resíduo dos efeitos principais), em que cada elemento  $(ge)_{ij}$  de  $GE$  são dados por (Hongyu et al., 2014):

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

sendo que  $Y_{ij}$  é a média das repetições do genótipo  $i$  no ambiente  $j$ , com  $i = 1, 2, \dots, g$  e  $j = 1, 2, \dots, e$ ;  $\bar{Y}_{i.}$  é a média do genótipo  $i$ ;  $\bar{Y}_{.j}$  é a média do ambiente  $j$  e  $\bar{Y}_{..}$  é a média geral do experimento.

Existem várias técnicas para atribuir os graus de liberdade a um modelo AMMI, um dos procedimentos usuais consiste em determinar os graus de liberdade associados a cada parcela da  $SQ_{GXE}$ , ou seja, associada a  $\lambda_k^2$ , relacionada a cada membro da família de modelos AMMI, obtém-se o quadrado médio ( $QM$ ) correspondente a cada parcela (ou modelo), em seguida, é obtido um teste F avaliando-se a significância de cada componente em relação ao  $QM_{\text{ErroMédio}}$ . Isso resulta num quadro de análise de variância semelhante ao tradicional, com desdobramento para fonte de variação da interação  $G \times E$  (Gauch, 2013).

Dessa forma, a definição do número de eixos a serem retidos para explicar o padrão relacionado à interação leva em consideração a proporção da  $SQ_{G \times E}$  acumulada até o  $n$ -ésimo eixo  $\frac{\sum_{k=1}^n \lambda_k^2}{SQ_{GXE}}$ .

O ponto de parada que determina a seleção do modelo (AMMI0, AMMI1, ..., ou AMMIF) baseia-se na significância do teste F para os sucessivos termos da interação. O método de Gollob (1968) é um dos mais utilizados para atribuir graus de liberdades a um modelo AMMI, a expressão do método é:  $GL_{PC_k} = g + e - 1 - 2k$ , com  $k = 1, 2, \dots, p$ ,  $p = \min(g-1, e-1)$ , em que  $PC_k$ : o  $k$ -ésimo eixo de componente principal. Na análise AMMI, a

DVS é aplicada à matriz de interações GE (matriz de posto conhecido p) (HONGYU et al., 2014):

$$G\hat{E}_{(gxe)} = \sum_{k=1}^n \lambda_k \gamma_k \alpha'_k = \tilde{U} \tilde{S} \tilde{V}',$$

Em que  $\tilde{U}_{(g \times n)}$  tem em suas colunas apenas os  $n$  primeiros vetores  $\gamma_{(g \times 1)}$ ;  $\tilde{S}_{(n)}$  é a matriz diagonal com os primeiros valores singulares:  $\lambda_1, \lambda_2, \dots, \lambda_p$  (com  $n < p$ ) e  $\tilde{V}'_{(n \times e)} \tilde{U}'_{(n \times e)}$ .

### 3. RESULTADOS E DISCUSSÃO

A análise de variância para os dados é apresentada na Tabela 1. Verifica-se que os resultados da ANOVA conjunta relacionada à produtividade de trigo ( $t/ha$ ) mostraram que o efeito de genótipo e ambiente foram significativos ( $p < 0,05$ ) e o efeito da interação foi significativo ( $p < 0,01$ ). Os resultados indicam que os genótipos exibiram comportamentos distintos em ambientes, indicando a existência de variações genéticas, o que torna viável a seleção de genótipos superiores. Dessa forma, é imprescindível realizar uma análise minuciosa do comportamento dos genótipos a fim de determinar a intensidade da interação com os ambientes.

**Tabela 1. Análise de variância conjunta dos experimentos com 16 genótipos avaliados em 6 ambientes e a decomposição das somas de quadrados da interação GxE.**

Fonte Variação	GL	SQ	QM	Valor F	Pr(>F)
Ambientes (E)	5	$3,35 \times 10^{12}$	$6,70 \times 10^{11}$	7,98	0,0126*
Genótipo (G)	15	$8,59 \times 10^{11}$	$5,72 \times 10^{10}$	2,06	0,0190*
GxE	75	$4,52 \times 10^{12}$	$6,02 \times 10^{10}$	2,17	0,0002***
PC1	19	$3,24 \times 10^{12}$	170667221829	6,15	0,0000
PC2	17	$8,20 \times 10^{11}$	48255086581	1,74	0,0496
PC3	15	$351 \times 10^{11}$	23445263143	0,84	0,6314
PC4	13	$7,27 \times 10^{10}$	5595188583	0,20	0,9988
PC5	11	$3,45 \times 10^{10}$	3145240533	0,11	0,9998
Erro	90	$2,49 \times 10^{12}$	$2,77 \times 10^{10}$	-	-
Total	185	-	-	-	-
01 Média	413847,4				
CV (%)	40,27				

\*\*\* Significativo a 0,001 de probabilidade e \* Significativo a 0,05 de probabilidade.

Pela Tabela 1, o primeiro componente principal (PC1) foi significativo com ( $p < 0,05$ ) pelo teste F de Gollob (1968), explicou 71,7% da variação da  $SQG \times E$  e com 19 graus de liberdade. O PC2 foi significativo com ( $p < 0,05$ ), explicou 18,1% da variação da  $SQG \times E$  e com 17 graus de liberdade.

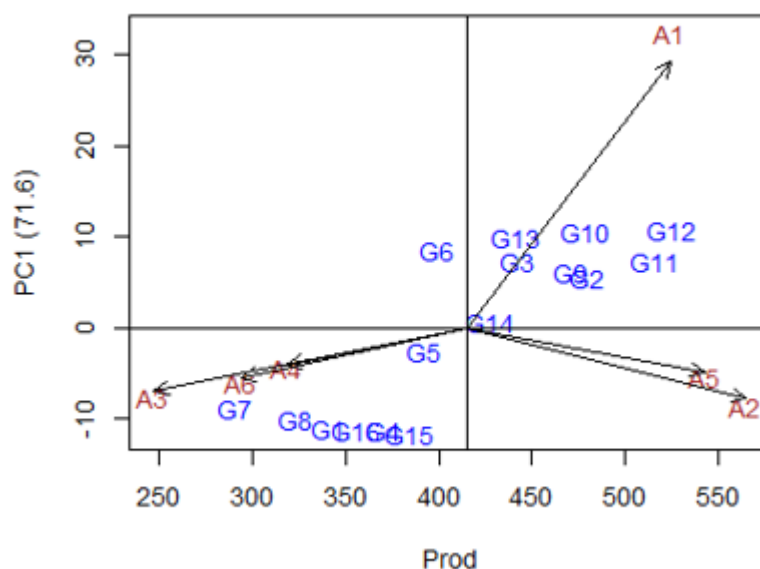
A falta de significância para o PC3 é insignificante e representa apenas ruído, ou seja, a variação aleatória não relacionada com o fenômeno da interação. Isso pode comprometer a



interpretação da estabilidade dos genótipos e ambientes na análise gráfica. Portanto, recomenda-se realizar a interpretação gráfica, considerando apenas as variações contidas nos dois eixos das componentes principais para avaliar a estabilidade dos genótipos e ambientes.

O biplot AMMI1 (Figura 2) é implementado para determinar tanto a estabilidade quando a produtividade dos genótipos. Neste contexto de biplot, os genótipos com escores do PC1 próximos de zero e localizados à direita da linha vertical são considerados os mais estáveis e com rendimento superior à média geral (OLIVOTO, 2021). Nessa análise, os genótipos mais estáveis e com rendimento superior à média geral foram: G14, G3, G13, G9, G2, G10, G11 e G12.

O genótipo G12 apresentou o melhor desempenho entre todos os genótipos, em relação de estabilidade e adaptabilidade (Figura 3), portanto, pode ser considerado como ideótipo (maior estabilidade e produtividade) e deve ser escolhido para recomendação pelos pesquisadores.



**Figura 2 - Biplot AMMI1 para o PC1 versus produção de trigo (t/ha) com 16 genótipos e 6 ambientes.**

O biplot AMMI2 (Figura 3) utiliza os dois primeiros PCs oriundos da decomposição por valor singular da matriz dos efeitos da interação e é empregado para inferir sobre os padrões da interação genótipo x ambiente (OLIVOTO, 2021). Na presente análise, os dois primeiros PCs explicam 89,8% da soma de quadrados da interação.

Verifica-se a formação de dois grupos de ambientes: Grupo 1 formado pelo local A1, A2 e A5 e Grupo 2 formado por A3, A4 e A6, portanto estes grupos apresentam características opostas (Figura 1). Quanto às adaptações específicas dos genótipos em ambientes, verifica-se que os genótipos G12, G10, G13, G16, G11, G3, G9 e G2 foram mais adaptados ao ambiente A1; o G5 ao ambiente A2; o G7 aos ambientes A3, A4 e A6 (Figura 2). Observa-se que os vetores relativos a esses genótipos apontam para uma direção comum, que pode ser orientada pelos ambientes (Gauch, 1992). Consta-se, ainda que os genótipos G7, G8, G4, G1, G15 e G16 não foram bem adaptados ao ambiente A1, e houve falta de adaptação de genótipos aos ambientes A2, A3, A4, A5 e A6 (Figura 3).

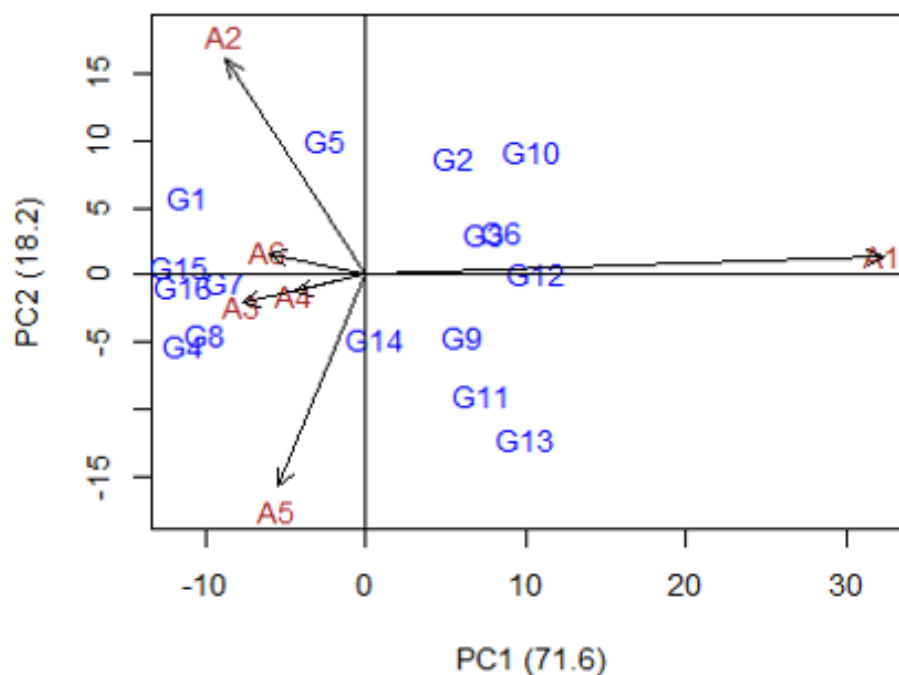


Figura 3 - Biplot AMMI2 para os dados de produtividade de trigo (*t/ha*) com 16 genótipos (G) e 6 ambientes (E)

### 3.1 GGE biplot para análise de dados MET (Ensaio Multi Ambientais)

Analisando a relação de informação (IR) dos seis componentes (Tabela 2), apenas o primeiro PC contém padrão ( $IR_1 = 4,34 > 1$ ) o segundo PC ( $IR_2 = 0,92 \cong 1$ ). Portanto, o biplot é considerado adequado para representar os padrões dos dados. O GGE biplot com base neste conjunto de dados é apresentado na Figura 4, na abscissa do biplot apresentam se os escores de PC1 e na ordenada os escores do PC2, dos genótipos e ambientes.

Tabela 2 – Valor singular, proporção explicada e relação da informação (IR) dos seis componentes principais (PCs)

PC	Valor Singular	Varição Explicado (%)	IR
1	1181,29	72,38	4,34
2	800,71	15,28	0,92
3	704,71	9,17	0,55
4	457,31	1,63	0,09
5	396,91	0,92	0,05
6	361,06	0,63	0,04

Os 16 genótipos foram nomeados como G1 a G16, enquanto os seis ambientes foram identificados A1 a A6. O biplot é dividido por uma linha que passa pela origem e divide as posições de um ambiente ou genótipo sendo denominada “vetor de ambiente” ou “vetor de

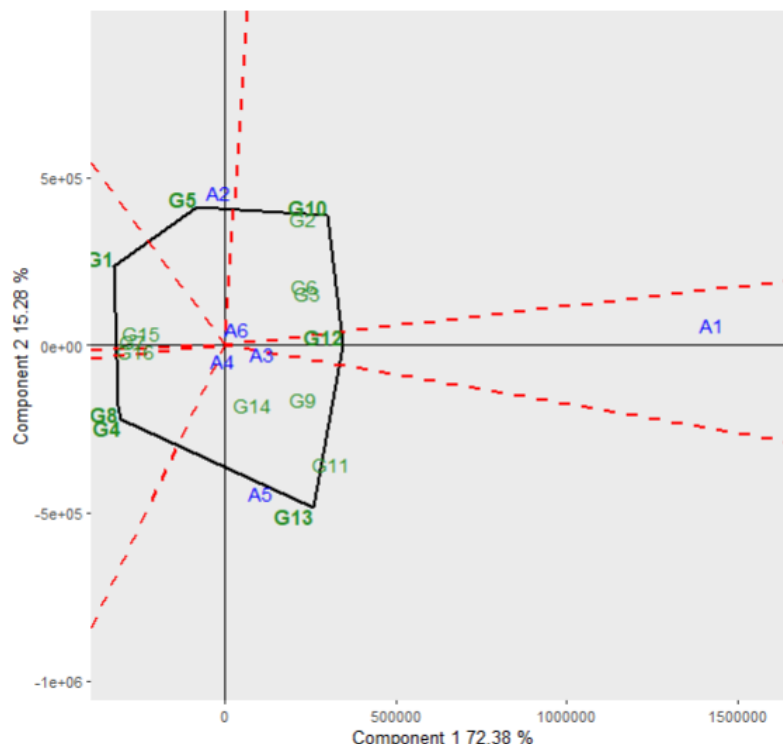
genótipo”. Esses vetores permitem a visualização das interações específicas entre um genótipo e um ambiente. A regra de interpretação é a seguinte: 1) quando o ângulo entre o vetor de um genótipo e o do ambiente é inferior a 90°, o desempenho desse genótipo em relação ao ambiente é superior à média; 2) se o ângulo for superior a 90°, o desempenho é inferior à média (pois o cosseno de um ângulo obtuso é menor do que 0); e quando o ângulo é aproximadamente 90°, indica que o desempenho está próximo da média (YAN, 2011; HONGYU et al., 2015).

A presença de interação significativa entre os genótipos e os ambientes (GxE) (Tabela 1) indica que o comportamento dos genótipos variou de um local para outro, o que influenciou na produtividade dos genótipos. Dessa forma, a análise GGE biplot é adequada para a seleção de genótipos mais adaptados e estáveis a determinados ambientes, bem como para a determinação de mega ambientes e seleção de ambientes de teste para os genótipos.

### 3.2 Análise mega ambiente

A Figura 4 permite visualizar o agrupamento dos ambientes de teste com base na interação  $G \times E$  entre os melhores genótipos. O biplot foi construído no ambiente centrado (centralização = 2), (escala = 0) e (PVS = 2). Os vértices do polígono são formados pelos genótipos: G1, G5, G10, G12, G13 e G4/G8, localizados mais distantes da origem fazendo com que os demais genótipos estejam contidos no polígono. Esses genótipos são os mais responsivos, enquanto os genótipos que estão no interior do polígono são menos responsivos em suas respectivas direções.

De acordo com Yan et al. (2007), o genótipo vértice, como regra geral, é o genótipo que apresenta o mais alto rendimento em todos os ambientes que dividem o mesmo setor. As linhas pontilhadas em vermelho, que saem da origem do biplot dividiu os seis ambientes em 4 grupos. Esses grupos são formados pelos ambientes (i) A1; (ii) A6; (iii) A2 e (iv) A3, A4 e A5.



**Figura 4 - O GGE Biplot (“Which-won-where”) (t/ha), para os dados de produtividade de trigo com informação de quais dos genótipos apresentaram melhor desempenho em quais ambientes.**

O genótipo G12 é o vértice do setor em que o ambiente A1 está situado; o mesmo acontece com o genótipo G5 no ambiente A2; o genótipo G10 forma o vértice no ambiente A6, em consequência é o genótipo que apresentou melhor desempenho neste ambiente quando comparado com os demais genótipos G6, G3 e G2 que compõem este ambiente; O genótipo G13 teve o melhor desempenho nos ambientes A3, A4 e A5 e é o genótipo mais adaptado (Figura 4). Os genótipos G1, G15, G7, G16, G8 e G4 não integram ambientes, significando que estes genótipos não foram produtivos em nenhum ambiente, ou seja, são os piores genótipos quando relacionados à produtividade em pelo menos alguns ou em todos os ambientes.

### 3.3 Avaliação de genótipos baseada em GGE biplot

O "ideótipo" é uma forma ideal de planta ou genótipo para um determinado ambiente e com finalidade de cultivo. Desta forma, o ideótipo é um genótipo que apresenta um alto desempenho médio e uma alta estabilidade em um mega ambiente. A visualização do GGE biplot "Média versus Estabilidade" trata-se de uma ferramenta eficaz para avaliar genótipos quanto ao desempenho e estabilidade (YAN et al., 2007; HONGYU et al., 2015; ATROCH, 2020).

As coordenadas médias de todos os ambientes de teste do biplot define o ambiente-média. Na figura 5, o eixo do ambiente-média (EAM) é a linha reta com uma seta que passa pela origem do biplot e do ambiente-média. Essa seta direciona para um maior desempenho médio para os genótipos. A linha sem setas perpendicular ao EAM, indica que quanto mais distante da origem maior a variabilidade de desempenho (menor estabilidade) nas duas direções. Nesse biplot, os valores singulares são integralmente compartimentados para os escores genotípicos, com base no  $PVS=1$  (YAN, 2011; HONGYU et al., 2015).

Pela Figura 5, o genótipo G12 foi o altamente estável e com maior produtividade média, enquanto o G11 foi o segundo mais produtivo, porém com alta instabilidade. O terceiro genótipo mais produtivo foram G13 e G10, no entanto apresentam alta instabilidade.

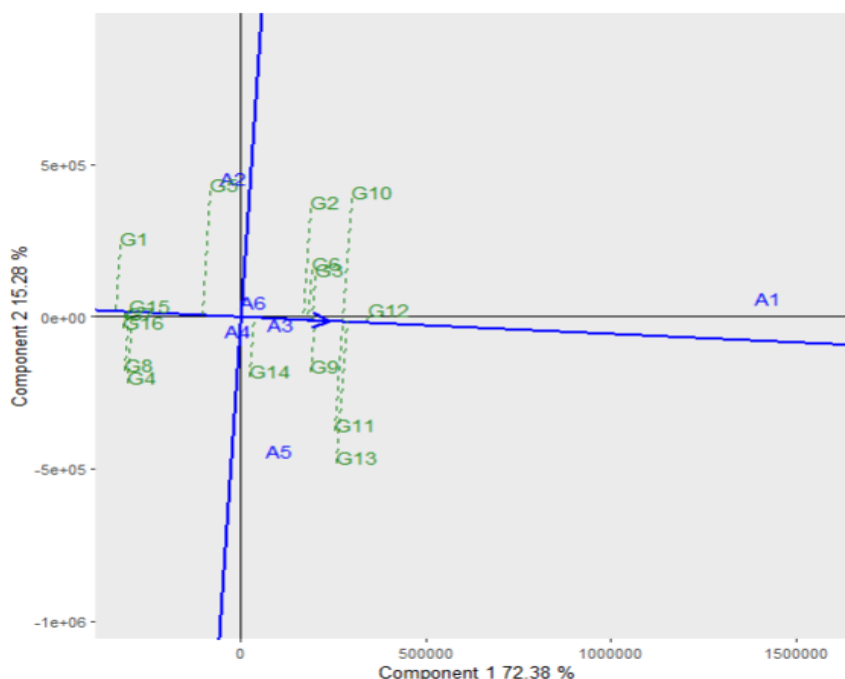
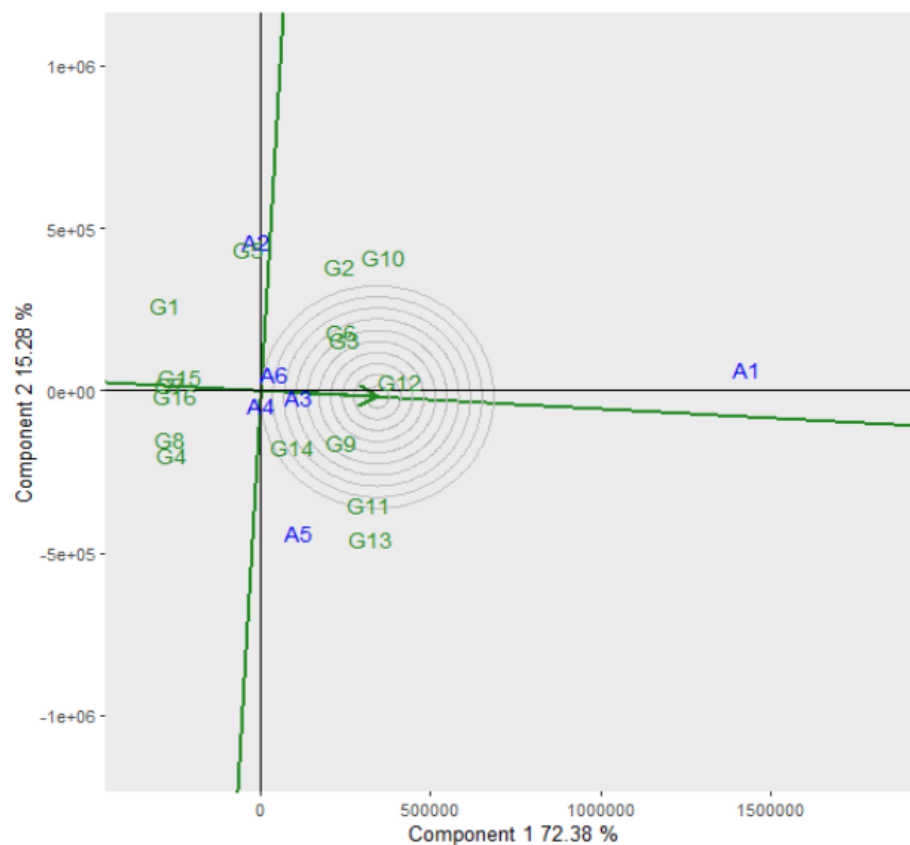


Figura 5 - O GGE Biplot ("Média versus Estabilidade") com eixo do ambiente-média (EAM) mostrando o desempenho médio e estabilidade dos genótipos

Um ideótipo pode ser um ponto localizado no centro dos círculos concêntricos sobre o EAM (Figura 6), na orientação positiva e com comprimento do vetor semelhante aos vetores mais longos dos genótipos sobre o lado positivo do EAM, ou seja, alto desempenho médio (HONGYU et al., 2015; ATROCH, 2020). Dessa forma, os genótipos localizados mais próximos do centro dos círculos concêntricos são mais desejáveis do que outros mais distantes. Assim, neste conjunto de dados o genótipo G12 foi o ideótipo. Na sequência, os melhores genótipos foram G9, G3, G6, G14 e G11.

Temos um conceito importante em relação a estabilidade apresentado na Figura 6, a expressão “estabilidade elevada” faz sentido somente quando está relacionada a um genótipo estável que resulta em alta média de desempenho (YAN & TINKER, 2006; YAN, 2011; HONGYU et al., 2015).

No biplot da figura 6, todos os genótipos são comparados com o genótipo “ideal”. O genótipo ideal é aquele representado pelo menor círculo com uma seta apontando para ele, e é definido como aquele que tem o maior rendimento em todos os ambientes, sendo absolutamente estável. Segundo, Yan et al. (2007), os genótipos são classificados de acordo com a sua distância do genótipo ideal. Neste estudo, o genótipo G12 é o ideal, seguido pelos genótipos G9, G3, G6, G14 e G11 que superaram os demais.



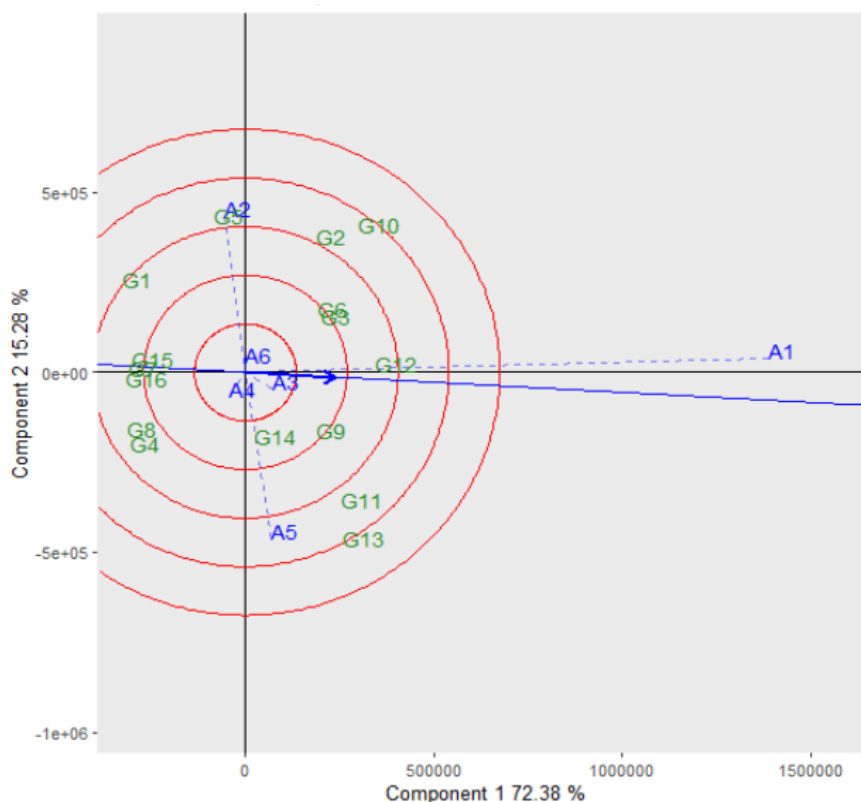
**Figura 6 - O GGE Biplot com eixo do ambiente-média (EAM) para classificar os genótipos em relação ao ideótipo (no centro dos círculos concêntricos).**

### 3.4 Avaliação do ambiente de teste

A avaliação do ambiente de teste tem por objetivo identificar ambientes que podem ser utilizados para selecionar os genótipos superiores de forma eficaz para um mega ambiente. O ambiente selecionado deve apresentar maior discriminação dos genótipos e representatividade. A finalidade da visualização no GGE biplot, Figura 7, é apresentar essa discriminação e representatividade.

Se o GGE biplot estiver baseado numa escala não padronizada, com dados centrados no ambiente (Escala = 0), o comprimento do vetor do ambiente será proporcional ao desvio padrão (SD) das médias dos genótipos e será igual a raiz quadrada da variância fenotípica ( $\sigma_p$ ) para um ambiente de teste, podendo ser utilizado como medida para o poder de teste (YAN, 2015; HONGYU et al., 2015).

Ambientes de teste com vetores mais longos demonstram uma maior capacidade de distinção entre os diferentes genótipos. Por outro lado, ambientes com vetores curtos têm uma capacidade discriminatória reduzida, resultando em uma maior semelhança e pouca ou nenhuma informação entre os genótipos. Portanto, não são recomendados para uso como ambientes de teste (HONGYU et al., 2015; YOKOMIZO et al., 2017). Um vetor curto, pode significar que o ambiente não esteja bem representado pelos PC1 e PC2, caso o biplot não represente adequadamente o  $G + GE$  dos dados. O ambiente A1 apresentou o vetor mais longo, significando que é mais discriminante em relação aos genótipos. Os ambientes A3, A4 e A6 apresentaram vetores curtos.



**Figura 7 - O GGE Biplot “discriminação e representatividade” para mostrar a capacidade de discriminação e de representatividade dos ambientes de teste.**

A Figura 7, tem uma segunda utilidade, a qual é indicar ambientes de representatividade do ambiente teste. Ambientes de teste com menores ângulos com o EAM, nesse caso, A1 é o mais representativo do que aqueles que apresentam ângulos maiores (A2 e A5); O ambiente de teste A1 é discriminante e representativo, logo é o ambiente de teste ideal para a seleção de genótipos adaptados. Os ambientes A2 e A5, são discriminantes, porém não representativos, assim são indicados para a seleção de genótipos adaptados especificamente para um mega ambiente, ou então para escolha de genótipos instáveis, caso o ambiente teste seja um único mega ambiente.

#### 4. CONCLUSÃO

Os modelos AMMI e GGE Biplot utilizados de maneira conjunta resultam na exploração mais eficaz da interação genótipo x ambiente com suporte a seleção de genótipos e ambientes superiores. Comparando os resultados obtidos nos modelos para os dados MET em estudo, os resultados apresentados são muito semelhantes. O modelo AMMI resultou numa maior explicação da variação 89,8% para os dois primeiros componentes principais, enquanto o GGE Biplot explicou 87,7%.

O modelo AMMI é uma escolha segura para identificar locais de cultivo mais favoráveis e genótipos com desempenho médio superior. Por outro lado, o modelo GGE Biplot é especialmente eficaz na análise da interação entre genótipo e ambiente, permitindo uma identificação mais precisa de mega ambientes e facilitando a seleção de genótipos estáveis e bem adaptados a ambientes específicos e mega ambientes.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- ATROCH, A. L.; FILHO, F. J. N.; SANTOS, J. A.; ATROCH, E. M. A. C; OLIVEIRA, I. J. Avaliação de guaranazeiro em ensaios multi-ambientes pelo modelo GGE biplot. *Biodiversidade*, v.19, n.4, p.44-56, 2020.
- CAMARGO-BUITRAGO, I.; INTIRE, E. Q. M.; GORDÓN-MENDOZA, R. Identificación de mega ambientes para potenciar el uso de genótipos superiores de arroz em Panamá. *Pesquisa Agropecuária Brasileira*, v.46, n.9, p.1601-1069, 2011.
- DIAS, C.T.S.; KRZANOWSKI, W.J. Choosing components in the additive main effect and multiplicative interaction (AMMI) models. *Scientia Agricola*. v.63, n.2, p.169-175, 2006.
- DIAS, C. T. S. ET AL. Metodologia AMMI: com aplicação ao melhoramento genético. Piracicaba: ESALQ, p. 169, 2014. Disponível em: <https://www.livrosabertos.abcd.usp.br/portaldelivrosUSP/catalog/download/50/43/199?inline=1>. Acesso em 15 mar. 2024.
- FAOSTAT, <https://fenix.fao.org/faostat/internal/en/#home>, Acesso em 29 mar. 2024.
- GABRIEL, K.R. The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, v.58, n.3, p.453-467, 1971.
- GAUCH, H. G.; A Simple Protocol for AMMI Analysis of Yield Trials. *Crop Science*, v.53, n.5, p.1860-1869, 2013.
- HONGYU, K.; GARCIA-PEÑA, M.; ARAÚJO, L. B.; DIAS, C. T. S. Statistical analysis of yield trials by AMMI analysis of genotype x environment interaction. *Biometrical Letters*, Poznan, v.51, p.89-102, 2014.
- HONGYU, K. et al. Comparação entre os modelos AMMI e GGE biplot para os dados de ensaios multi-ambientais. *Revista Brasileira de Biometria*, v. 33, n. 2, p. 139-155, 2015.
- HONGYU, K.; SILVA, F. Modelo de interação genótipo  $\times$  ambiente para dados de ensaios multiambientais, 2017.
- OLIVOTO, T.; Software R para avaliação de dados experimentais: Um foco em experimentos agrônomicos, 2021. Disponível em <https://tiagoolivoto.github.io/e-bookr/interaction.html>. Acesso em 15 mar. 2024.
- SILVA, R. R.; BENIN, G. Análises Biplot: conceitos, interpretações e aplicações. *Ciência Rural*, Santa Maria, v.42, n.8, p.1404-1412, 2012.
- YAN, W. Singular-value partition for biplot analysis of multi-environment trial data. *Agronomy Journal*, v.94, n.5, p.990-996, 2002.
- YAN, W.; KANG, M. S. GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists. CRC Press, Boca Raton, FL, USA, 2003.



YAN, W; TINKER, A. Biplot analysis of multi environment trial data: principles and applications. Canadian Journal of Plant Science, v.86, n.3, p.623-645, 2006.

YAN, W.; KANG, M.S.; MA, B.; WOODS, S.; CORNELIUS, P.L. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Science, v.47, n.2, p.643-653, 2007.

YAN, W; HOLLAND, J. B. A heritability-adjusted GGE biplot for test environment evaluation. Euphytica. v.171, n.3, p.355-369, 2010.

YAN, W. GGE Biplot vs. AMMI Graphs for Genotype-by-Environment Data Analysis. Journal of the Índia Society of Agricultural Statistics, v.65, n.2, p.181-193, 2011.

YOKOMIZO, G.K.; DIAS, J. S.A.; DIAS, C.T.S.; HONGYU, K. Análise AMMI em caracteres vegetativos de genótipos de bananeiras no Amapá. Rev. Cienc. Agrar., v. 59, n. 1, p. 1-8, 2016.

YOKOMIZO, G. K.; FARIAS NETO, J.T.; OLIVEIRA, M. S. P.; HONGYU, K. Análise GGE biplot na avaliação de características de cachos em açaizeiros da região amazônica. Mundo Amazônico, 2017.