Nativa, Sinop, v. 12, n. 4, p. 597-604, 2024.
Pesquisas Agrárias e Ambientais
DOI: https://doi.org/10.31413/nat.v12i4.17484
ISSN: 2318-7670

Bayesian approach: an alternative to the additive main effect and multiplicative interaction models for genotypes through environmental interactions

Zakia BATOOL [1], Waqas MUNIR [1], Mohsin ABBAS *[2], Umar Waqas KHAN [3], Mahwish RABIA [4]

[1] Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
[2] Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan.
[3] Department of Mathematics and Statistics, University of Agriculture, Faisalabad, Pakistan.
[4] Department of Statistics, Government College Women University, Sialkot, Pakistan.
*E-mail: abbasmohsin202@gmail.com

Submitted on: 04/20/2024; Accepted on: 09/24/2024; Published on: 11/12/2024.

ABSTRACT: Genotypes of different genetic structures behave differently in various environmental conditions. Genotype-by-environment interaction (GEI) is referred to as differential responses of different genotypes across different environments; GEI is of great importance because of the higher performance of genotypes that GEI assesses. However, the presence of GEI makes analysis more complicated. To up-root these assessment complications, several methods have been proposed, such as Principal Component Analysis (PCA), Cluster Analysis, Additive Main effects, Multiplicative Interaction (AMMI) models, and Genotype plus Genotype by Environment interaction (GGE). These methods neither overcome the problem of over-parameterization nor use the prior information. This study aims to use a technique to address these problems; for this purpose, wheat crop data comprised of 30 genotypes tested across 13 different locations of Punjab, Pakistan, for two consecutive years was used. The layout of the experiment was a Randomized Complete Block Design (RCBD). In this study, a comparison was made between Classical methods AMMI, GGE biplot, and Bayesian approach using Von-Mises Fisher distribution as prior. Classical methods showed that genotype V-11098 was the most desirable based on stability and high-yield performance. The Bayesian approach was used for GEI because it simplifies statistical interpretation by relaxing some constraints. It uses the prior information and provides solutions using the Markov Chain Monte Carlo (MCMC) algorithm. Bayesian strategy for analysis of GEI was used to assess the general, specific performance of genotypes and risk related to genotype. Analysis revealed that bi-linear terms $\mu_{25,1}$ for genotype NS-10 genotype and $\nu_{13,1}$ for environment S13 (Piplan-14) were found significant, indicating that these affect interaction. It was observed that the Bayesian approach could nicely explore GE interaction.

Keywords: Genotype-by-environment interaction (GEI); Von Mises-Fisher Distribution; Markov Chain Monte Carlo (MCMC).

Abordagem bayesiana: uma alternativa aos modelos de efeito principal aditivo e interação multiplicativa para genótipos por meio de interações ambientais

RESUMO: Genótipos de diferentes estruturas genéticas apresentam comportamentos distintos em várias condições ambientais. A interação genótipo-por-ambiente (GEI) é referida como respostas diferenciais de diferentes genótipos em diferentes ambientes; a GEI é de grande importância devido para definição do maior desempenho dos genótipos que a GEI avalia. No entanto, a presença da GEI torna a análise mais complicada. Para erradicar essas complicações de avaliação, vários métodos foram propostos, como Análise de Componentes Principais (PCA), Análise de Cluster, Efeitos Principais Aditivos, modelos de Interação Multiplicativa (AMMI) e Genótipo mais interação Genótipo por Ambiente (GGE). Esses métodos não superam o problema da superparametrização nem usam as informações anteriores. Este estudo visa usar uma técnica para abordar esses problemas; para esse propósito, foram usados dados de safra de trigo compostos por 30 genótipos testados em 13 locais diferentes de Punjab, Paquistão, por dois anos consecutivos. O layout do experimento foi um Delineamento de Blocos Completos Randomizados (RCBD). Neste estudo, foi feita uma comparação entre os métodos clássicos AMMI, GGE biplot e abordagem Bayesiana usando a distribuição de Von-Mises Fisher como prior. Os métodos clássicos mostraram que o genótipo V-11098 foi o mais desejável com base na estabilidade e no desempenho de alto rendimento. A abordagem Bayesiana foi usada para GEI porque simplifica a interpretação estatística ao relaxar algumas restrições. Ela usa as informações anteriores e fornece soluções usando o algoritmo Markov Chain Monte Carlo (MCMC). A estratégia Bayesiana para análise de GEI foi usada para avaliar o desempenho geral e específico dos genótipos

e o risco relacionado ao genótipo. A análise revelou que os termos bilineares μ_25,1 para o genótipo NS-10 e ν_13,1 para o ambiente S13 (Piplan-14) foram considerados significativos, indicando que afetam a interação. Foi observado que a abordagem Bayesiana poderia explorar bem a interação GE.
Palavras-chave: interação genótipo-por-ambiente (GEI); distribuição de Von Mises-Fisher; cadeia de Markov Monte Carlo (MCMC).

1. INTRODUCTION

Multi-environment experiments are commonly conducted to identify genotypes with a high yield and less sensitivity to adverse environmental changes. Genotypes do not perform similarly in different environments, so genotype-environment interaction (GEI) is crucial. Analyzing GEI to assess how well genotypes perform in different environments is necessary. Characteristics such as the quality of wheat and yield varied differently in different environments. They were studied by Uhlen et al. (1998), which considers the environmental consequences of these characteristics. In multi-environment trials (METs), the GE effects are more evident and have three main aims: (a) Estimation and forecasting of experimental data yield levels can be done more precisely; (b) yield stability and genotype adaptations to different environments can be determined; and (c) selection of ideal genotypes can be made, which can guide the future for suitable sites see, Crossa et al. (1990) for more details.

There are many methods to analyze GEI, such as principal component analysis (PCA), cluster analysis, genotype, and genotype by environment interaction (GGE) bi-plot analysis Yan and Tinker (2005). However, these methods have some drawbacks, such as the fact that PCA fails to separate and identify significant genotypes and the main effects of the environment.

Cluster analysis provides the graphical grouping of genotypes or environments, so AMMI can be used to overcome these problems because it accomplishes the GE much more effectively by having maximum variation explained by the interaction sum of squares. The AMMI model can effectively analyze stability analysis because it explores the GEI; that is why AMMI models are frequently used for GEI Crossa et al. (1990). Further, the AMMI model can be visualized by using bi-plots. The AMMI model can be defined as follows:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \sum_{k=1}^{t} \lambda_k u_{ik} v_{jk} + \varepsilon_{ij}$$

where: Y_{ij} is the response, λ_k is the singular value for the k-th principal component axis, u_{ik} is the element of the left singular vector, and v_{jk} is the element of the right singular vector.

Although AMMI models are very useful for GEI, they have some problems, such as the fact that it is difficult to explain the G×E structure of the AMMI model when the data set is incomplete or has missing values, see Perez-Elizalde et al. (2011) and the references cited therein. The problem of over-parameterization arises, which makes the model much more complex, and also because the model does not use prior information Josse et al. (2014). Outliers or extremes can lead to misleading interpretations using classical AMMI as it uses the ordinary least squares (OLS) method for estimation, which those outliers can significantly affect; see Rodrigues et al. (2015) and the references cited therein for more details.

It was suggested that these problems could be tackled using Bayesian methodology. The Bayesian approach is very useful because it makes statistical interpretation easy using the standard MCMC algorithm; see Gelman et al. (2004) for more details. Viele and Srinivasan (2000) applied uniform before the first column of UI and on th$va_{jk}$and obtained useful results. Gibbs sampling with embedded Metropolis-Hasting random walks is another tool of the Bayesian approach to obtain estimates of interaction. The Bayesian paradigm for principal component analysis was presented by Hoffman Hoff (2007). Although it provides a similar computational environment, it differs in the linear part, which contains the grand mean and column effect. The Bayesian approach is widely used to analyze problems like GEI because it relaxes model constraints and uses prior knowledge of the data. Da Silva et al. (2015) suggested a Bayesian shrinkage AMMI instead of a classical Bayesian; However, it gave a high shrinkage value and incorporated credible intervals for biplots compared to Bayesian AMMI.

Perez-Elizalde et al. (2011) proposed a Bayesian approach for treating problems with the AMMI model using the Von-Mises Fisher distribution as a prior, aiming for two advantages: i) The Bayesian approach can be applied to GEI data sets because it facilitates analysis using prior information related to the experiment under study. ii) The distribution related to any interest can be obtained by posterior distribution. Further, the obtained credible intervals can be used for good interpretation and provide biplots similar to classical methods.

2. MATERIAL AND METHODS

Multi-environmental data from two consecutive years, consisting of 30 genotypes of wheat crops evaluated in 13 different environments, was used for analysis. The experiment layout was RCBD, with three replications for all genotypes over the years. In the first section, combined analysis of variance (ANOVA) and AMMI analysis were performed, while in the second section, Bayesian methodology was applied. First-year data was used to elicit prior information, and the Von-MISES Fisher distribution was applied per the proper prior suggestion by Perez-Elizalde et al. (2011).

3. RESULTS

A combined ANOVA of thirty genotypes in 13 environments for two years was performed, and the output is provided in Table 1. Analysis showed that genotype and environment were highly significant, meaning these effects have substantial differences. The significance of G×E interaction ($p \leq 0.01$) indicated that genotypes have different responses in different environments. Environment, genotypes, and G×E explained 84.3%, 3.5%, and 8.6% variation of the sum of squares. From the combined ANOVA provided in Table 1, it is evident that G×E has a significant effect; hence, the identification of superior genotypes across environments is not possible based on just

the performance of the mean yield. Furthermore, ANOVA cannot partition the interaction into various components, so other stability measures are required to explore G×E.

3.1. AMMI analysis and bi-plot representation

As Table 1 indicated, the main and interaction effects were significant, so G×E is partitioned into bilinear terms. The Interactive Principal Component Axis (IPCA) are arranged in ascending order based on their importance, as given in Table 2. G×E interaction component significance was measured using Gollob's F-test at 0.01 probability. The first four principal components explained the 59.2% variation of the GEI sum of squares. Hence, the AMMI-4 model can be regarded as the best-fit model for the wheat crop data. Although the first four components were found to be significant, to avoid complexity, the first two principal components can be used to analyze the results using the AMMI bi-plot.

Two principal interaction components were adequate for the predictive model, while other components mostly contained non-predictive variation. Therefore, the first two components assessed the interaction of thirty wheat genotypes with 26 environments. The first IPC axis explained 23.3% variation, while the second principal component for the interaction sum of the squares captured 14.0%. The first two PCAs cumulatively explain 37.3% variation for the interaction sum of squares with 104 degrees of freedom out of 725, indicating that a sufficient percentage of G×E was explained. So, further diagnosis can be made by graphing a biplot using the first two IPCAs.

Table 1. Combined ANOVA.
Tabela 1. ANOVA combinada.

Sources	df	SS	MS	p-value	% SS
Environment (E)	25	2226401709	89056068	0.00	84.3
Rep(ENV)	52	3653099	70252	0.19	
Genotype (G)	29	92402025	3186277	0.00	3.5
G×E	725	227172934	313342	0.00	8.6
Residual	1508	90129189	59767		

Table 2. Combined ANOVA according to the AMMI model and Gollob's test of interaction PC's.
Tabela 2. ANOVA combinada segundo o modelo AMMI e teste de interação de Gollob PC's

Sources	df	SS	MS	p-value	G×E (%)	Cumulative(%)
Environment (E)	25	2226401709	89056068			
Rep(ENV)	52	3653099	70252			
Genotype (G)	29	92402025	3186277			
G×E	725	227172934	313342			
IPCA 1	53	53024110	1000455	0.00	23.3	23.3
IPCA 2	51	31803670	623601	0.00	14.0	37.3
IPCA 3	49	27472720	560667	0.00	12.1	49.4
IPCA 4	47	22299260	474452	0.00	9.8	59.2
IPCA Residual	525	92573174	176329	0.00		
Residual	1508	90129189	59767			

Table 3. Posterior Mean, Standard Deviation (SD), Quartiles, at 0.90 and 0.95 (Highest Posterior Density) HPD intervals.
Tabela 3. Média posterior, desvio padrão (DP), quartis, nos intervalos HPD de 0,90 e 0,95 (maior densidade posterior).

						0.90 HPD Interval		0.95 HPD Interval	
Parameter	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	lower	upper	lower	upper
λ_1	2077.26	423.20	790.88	2084.27	2369.18	1369.07	2748.10	1240.48	2888.35
λ_2	1292.73	381.89	1036.29	1303.84	1553.56	675.14	1920.03	533.34	2018.84
λ_3	903.23	347.96	664.77	909.93	1148.50	316.96	1470.83	216.29	1579.14
λ_4	597.50	306.67	371.38	588.14	814.40	24.49	1022.25	0.14	1110.15
λ_5	379.46	254.98	173.69	348.85	556.13	0.00	734.44	0.00	836.66
λ_6	236.08	199.88	66.70	191.74	359.06	0.00	528.47	0.00	621.69
λ_7	140.88	150.84	22.27	87.94	213.78	0.00	365.08	0.00	458.43
λ_8	79.15	106.10	6.91	34.94	109.76	0.00	225.29	0.00	314.97
λ_9	43.05	70.21	2.27	13.20	51.13	0.00	128.98	0.00	190.38
λ_{10}	22.06	42.92	0.71	4.81	22.37	0.00	65.14	0.00	108.12
λ_{11}	11.51	26.51	0.22	1.68	9.66	0.00	31.60	0.00	58.90
λ_{12}	5.88	15.83	0.07	0.62	3.94	0.00	15.14	0.00	30.46
$\mu_{25,1}$	0.23	0.14	0.14	0.24	0.33	0.01	0.47	-0.04	0.50
$\nu_{13,1}$	-0.63	0.11	-0.71	-0.64	-0.55	-0.81	-0.45	-0.83	-0.39

3.1.1 AMMI bi-plot

Figure 1 displays the biplot of PCA-1 versus PCA-2, which contains a 37.3% variation. From the plot, it can be observed that 11BT004 has a positive interaction with S10 (Vehari-14) and L10 (Vehari-13) and a negative interaction with S13 (Piplan-14) and S13 (Piplan-13). Genotypes NR-411, 11BT004, and V-12304, which have high PC1 and PC2 scores, significantly affect interaction.

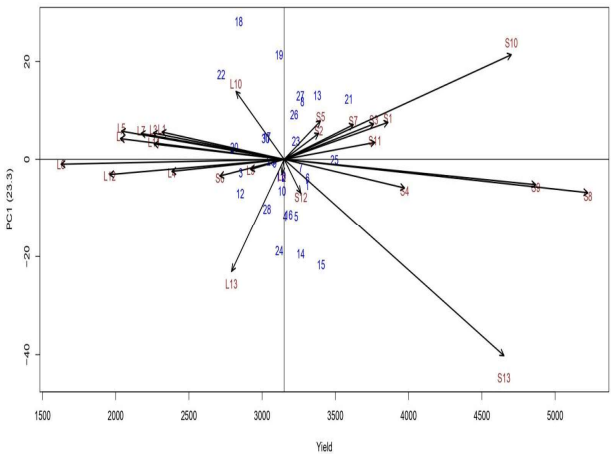

Figure 1. Biplot of interaction PCA Axis-1 versus Axis-2 of wheat data for 30 genotypes in 26 environments.
Figura 1. Biplot da interação PCA Eixo-1 versus Eixo-2 de dados de trigo para 30 genótipos em 26 ambientes.

The second biplot of IPCA Axis-1 versus the mean yield of environments and genotypes can be seen in Figure 2. The low-yielding genotypes were located in the upper left, and environments are shown in the lower left quadrant. Among the environments, S4 (BWN-14), S9 (Multan-14), and S10 (Vehari-14) were the highest-yielding environments, whereas S13 (Piplan-14) and S8 (Khanewal-14) were the most favorable environments. Among 26 environments, L6 (Gujranwala-13) was low-yielding and the most unfavorable.

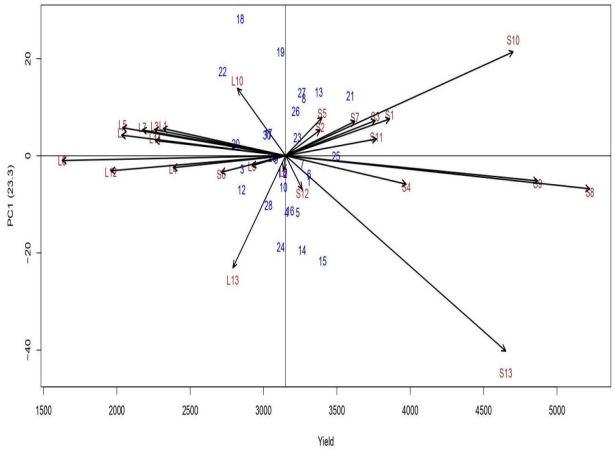

Figure 2: Biplot of interaction PCA Axis-1 versus mean yield of wheat data for 30 genotypes in 26 environments.

3.2. Bayesian analysis

The Bayesian AMMI estimate for λ_1 was 2077.26 and for λ_2 was 1292.73, and the standard deviation for λ_1 and λ_2 was 423.2 and 381.9, respectively. The posterior densities of the remaining singular values tend towards zero; complete posterior densities for the singular values are given in Table 3. Figure 2 shows a similar pattern: the first 2 and 3 posterior singular values have a slightly positive bell shape, while the later values tend to be directed towards zero. The cumulative proportion of variation for posterior densities by eigenvalues explored shows that about 90% of interaction variance can be explained by four components. It can be evaluated that the first component explains about 60% of the interaction variation, and the second component explains approximately 17%. It shows that two bilinear terms can be retained in the model, as the first two singular values can explain 80% of the cumulative variation. These results of two significant bilinear components are evidence of those normally present in the analyses of genotype by environment trials, where the requirement is to have more than one bilinear term in the model to handle the complexity of the G×E.

Wheat crop data for 30 genotypes in 13 environments is given in Figures 3 and 4. Figure 3 contains posterior densities and 0.95 HPD regions of the singular values, $\lambda_1, \lambda_2, \ldots, \lambda_{12}$ (C.1-C.12), while Figure 3 contains posterior densities and 0.95 HPD regions of the cumulative proportion of variance, $\phi_t = \sum_{k=1}^{t} \lambda^2{}_k / (\sum_{k=1}^{min\,(r,c-1)} \lambda^2{}_k)$, where t=1,2,… min(r, c-2).

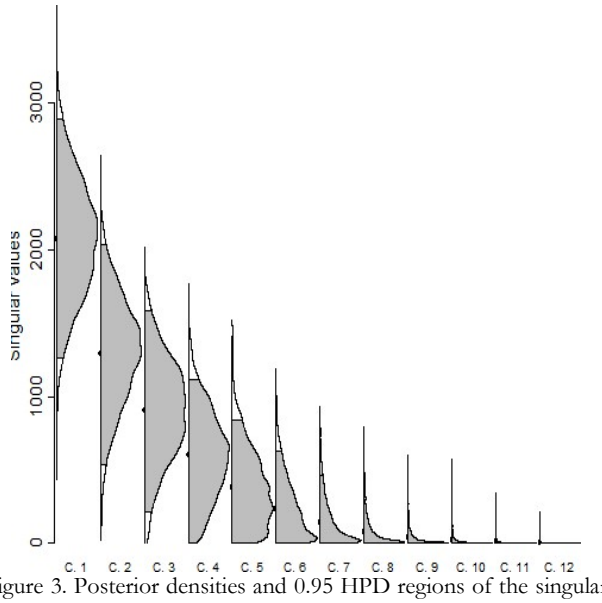

Figure 3. Posterior densities and 0.95 HPD regions of the singular values.
Figura 3. Densidades posteriores e regiões de 0,95 HPD dos valores singulares.

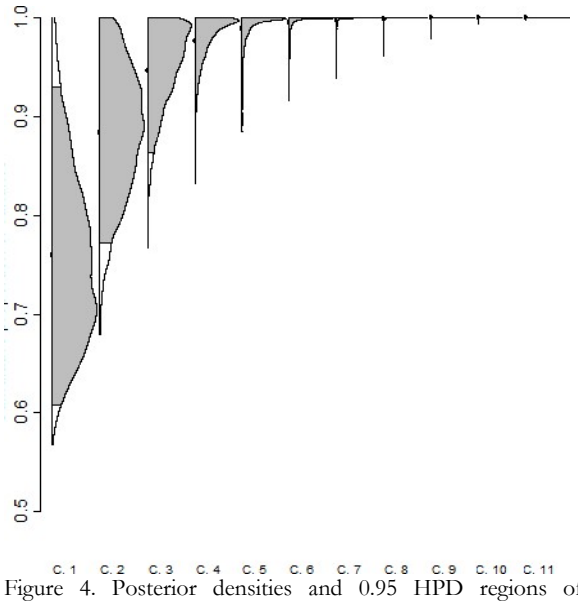

Figure 4. Posterior densities and 0.95 HPD regions of the cumulative proportion of variance.
Figura 4. Densidades posteriores e regiões de 0,95 HPD da proporção cumulativa de variância.

Table 4. Posterior mean, SD, quartiles, 0.90, and 0.95 HPD intervals were computed with 20,000 approximately independent samples simulated from the posterior distribution of 30 linear genotype effects and 13 environments for wheat crop data.
Tabela 4. Média posterior, DP, quartis, intervalos HPD de 0,90 e 0,95 foram calculados com 20.000 amostras aproximadamente independentes simuladas a partir da distribuição posterior de 30 efeitos de genótipos lineares e 13 ambientes para dados de cultura de trigo.

						0.90 HPD interval		0.95 HPD interval	
Parameters	Mean	SD	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	lower	upper	lower	upper
μ	3149.47	19.11	3136.41	3149.33	3162.22	3117.90	3180.29	3112.91	3187.78
α_1	144.47	102.01	76.88	144.77	213.02	-24.99	308.58	-48.61	350.92
α_2	-299.67	102.29	-368.71	-300.44	-230.39	-468.43	-132.72	-491.40	-93.06
α_3	80.48	100.34	12.61	80.22	147.79	-87.77	241.36	-113.71	280.58
α_4	343.76	101.80	274.76	343.74	411.53	184.22	519.40	145.95	543.93
α_5	68.06	102.55	-0.988	68.10	138.96	-93.04	240.88	133.03	264.83
α_6	111.16	101.75	42.36	111.66	179.85	-55.01	278.38	-91.35	305.81
α_7	-116.49	102.68	-185.74	-117.63	-47.77	-278.46	61.121	-311.26	91.21
α_8	-89.62	102.16	-159.11	-90.23	-21.23	-254.75	79.31	-285.64	113.89
α_9	-127.20	101.35	-196.58	-127.14	-57.99	-292.96	37.66	-332.64	63.68
α_{10}	-2.16	102.30	-69.51	-2.22	65.91	-169.02	164.47	-211.16	190.67
α_{11}	-296.26	101.48	-365.17	-296.33	-228.84	-468.71	-136.58	-490.76	-97.13
α_{12}	6.65	101.60	-61.12	7.11	75.31	-163.21	171.12	-195.24	204.97
α_{13}	82.29	102.6	13.19	81.11	151.35	-87.60	250.91	-117.20	282.40
α_{14}	161.84	103.42	93.04	162.78	230.93	-3.26	336.82	-40.87	360.99
α_{15}	114.64	101.61	46.99	115.30	182.10	-52.28	285.2	-87.42	311.87
α_{16}	125.49	101.25	58.68	124.24	193.01	-32.91	300.62	-65.19	330.20
α_{17}	-64.89	101.89	-134.56	-65.52	4.5839	-224.59	108.64	-262.62	132.10
α_{18}	-12.88	103.60	-83.14	-12.94	56.77	-181.72	154.50	-216.43	190.61
α_{19}	-29.45	102.42	-99.18	-29.20	40.05	-200.18	134.87	-244.40	156.10
α_{20}	227.33	100.42	160.09	227.98	295.82	69.60	396.76	31.49	424.60
α_{21}	112.95	102.55	43.99	112.25	181.99	50.03	286.92	-86.49	312.11
α_{22}	252.68	102.25	182.49	253.66	322.34	93.08	426.27	49.43	447.91
α_{23}	32.70	102.92	-36.10	31.65	102.47	-136.018	199.39	-167.03	233.40
α_{24}	-118.83	102.53	-187.85	-119.65	-49.46	-285.57	52.50	-318.81	81.03
α_{25}	-311.46	102.30	-380.53	-310.35	-243.37	-482.56	-145.76	-507.53	-109.09
α_{26}	-35.89	102.61	-105.44	-36.89	33.58	-197.72	138.63	-245.10	156.20
α_{27}	-340.15	101.36	-408.19	-339.38	-270.92	-502.16	-168.85	-536.61	-141.80
α_{28}	441.91	102.22	374.73	442.90	511.35	276.88	612.36	243.23	648.14
α_{29}	-429.16	103.26	-498.31	-428.42	-359.38	-602.33	-262.07	-628.11	-222.16
α_{30}	-32.3	101.5	-100.46	-31.99	36.55	-201.45	131.06	-227.45	171.49
β_1	-62.10	66.36	-106.74	-62.45	-17.49	-171.05	45.249	-188.46	69.61
β_2	-440.83	65.57	-485.00	-440.82	-395.33	-548.69	-335.02	-567.72	-312.09
β_3	-137.94	65.52	-182.41	-138.00	-93.69	-242.82	-28.64	-275.31	-18.64
β_4	28.50	65.90	-16.00	28.62	73.100	-77.38	137.87	-101.63	155.95
β_5	-429.95	65.85	-474.58	-430.60	-384.73	-533.56	-318.72	-558.36	-303.55
β_6	-980.62	65.59	-1024.24	-981.34	-936.84	-1084.69	-868.05	-1104.70	-845.19
β_7	-250.62	66.00	-294.87	-251.14	-205.93	-356.828	-140.71	-383.77	-126.10
β_8	1028.13	65.89	983.95	1028.66	1071.78	918.98	1135.92	899.01	1154.80
β_9	745.15	65.72	700.67	745.42	789.45	634.47	849.65	616.27	871.52
β_{10}	600.14	66.03	556.00	600.40	644.77	494.81	711.38	465.49	725.05
β_{11}	-131.97	65.28	-176.06	-131.55	-88.76	-236.91	-22.04	-261.44	-4.74
β_{12}	-539.01	65.08	-581.73	-539.37	-496.44	-650.78	-434.14	-662.55	-406.02
β_{13}	571.13	64.80	527.11	570.95	614.77	464.81	677.83	444.81	696.26

3.3. Credible regions of the first two bilinear terms of the linear bilinear model

The bilinear terms that do not include the null point (0,0) for the bivariate 0.90 and 0.95 HPD regions are referred to as significant terms, which means they significantly contribute to the explanation of interaction variation. In Table 3, the posterior mean, standard deviation (SD), quartiles ($q_{0.25}$, $q_{0.50}$, and $q_{0.75}$), 0.90 HPD, and 0.95 HPD intervals computed with 20,000 approximately independent samples simulated from the posterior distribution of the

residual variance (σ), all singular values ($\lambda_1, \lambda_2, \ldots, \lambda_{12}$) and the right and left singular vector elements of genotypes and environments, respectively, whose 0.90 HPD and 0.95 HPD intervals do not contain the null value (0, 0).

Data for wheat yield of 12 singular values and their 0.90 and 0.95 HPD are given. It can be seen that the first two eigenvalues have a high value, while the remaining values tend to move towards zero. Along with singular values, eigenvector values for those genotypes and environments are given, which do not have null points for 0.90 HPD and 0.95 HPD. The score of environment S13 ($\nu_{13,1}$) and genotype 25 ($u_{25,1}$) does not have a null point for both HPD intervals, which indicates that they have a significant contribution in G × E.

Table 5. List of locations included in this study (Location codes).
Tabela 5. Lista de locais incluídos neste estudo (Códigos de localização).

No.	Locations	Year-2013	Year-2014
1	Okara	L1	S1
2	MMRI	L2	S2
3	Dhakkar	L3	S3
4	Bahawal nagar (BWN)	L4	S4
5	KSK	L5	S5
6	Gujranwala (GRW)	L6	S6
7	Kot Naina	L7	S7
8	Khanewal	L8	S8
9	Multan	L9	S9
10	Vehari	L10	S10
11	Karor	L11	S11
12	Sargodha	L12	S12
13	Piplan	L13	S13

Table 6. List of Wheat genotypes included in this study.
Tabela 6. Lista de genótipos de trigo incluídos neste estudo.

No.	Genotype Name	Genotype Code	No.	Genotype Name	Genotype Code
1	V-12266	1	16	NR-411	16
2	V-11047	2	17	V-11365	17
3	V-12284	3	18	V-12265	18
4	V-11098	4	19	Millat-11	19
5	V-11061	5	20	Lasani-08	20
6	V-11022	6	21	Nr-409	21
7	V-11092	7	22	V-11143	22
8	NW.10.1111-7	8	23	V-11041	23
9	TW11510	9	24	V-11032	24
10	V-12275	10	25	NS-10	25
11	GH	11	26	11BT004	26
12	V-11137	12	27	9459-1	27
13	NN-GAN-3	13	28	V-12304	28
14	V-11138	14	29	11B2074	39
15	V-11046	15	30	11B2049	30

Biplots for the bivariate 0.90 HPD and 0.95 HPD regions of column scores are shown in Figure 5. Visual analysis shows that S10 forms a distinct group of environments with a positive value, while S13 shows a negative correlation with S10. Some degree of similarity was also demonstrated between S13 and S9; S10 overlaps with S12, S8, and S4, which means they form a homogeneous group. S2, S3, S1, S5, S7, and S11 form a less significant group. In Figure 6, a bi-plot for row score is shown. Among all genotypes, 25 ($u_{25,1}$) was far away, showing a positive value, followed by genotypes 20, 29, and 26, overlapping each other and showing some similarities. These genotypes form a distinct group. Genotype 25 is the only one that does not contain a null point (0,0) and significantly contributes to the interaction. The remaining genotypes from another group in the first bilinear components are not different from zero.

Biplots based on the 0.90 and 0.95 HPD intervals overlap, so some significant environments may not be portrayed. Hierarchical clustering was performed based on the posterior mean of Euclidean distance between the rows interacting with the complete linkage method (Figure 7). It can be seen that S10 forms a single group and then merges into another group; similarly, S9 and S13 form a cluster, indicating that there is some similarity between them.

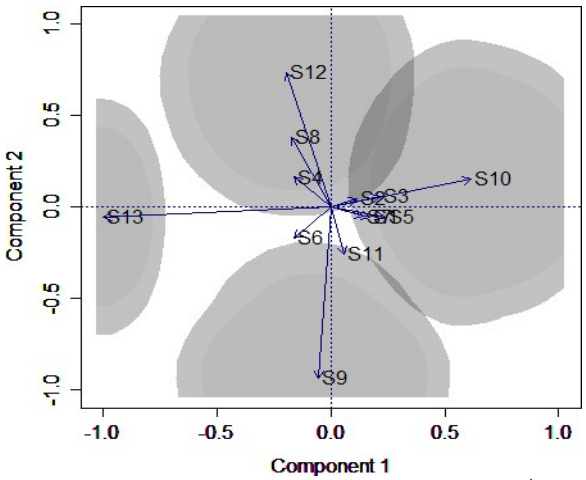

Figure 5. Plot of the bivariate column scores $V'D^{1/2}$ and the bivariate 0.95 (gray external contour) and 0.90 (gray internal contour) HPD regions.
Figura 5. Gráfico das pontuações da coluna bivariada V^' D^(1/2) e das regiões HPD bivariadas 0,95 (contorno externo cinza) e 0,90 (contorno interno cinza).

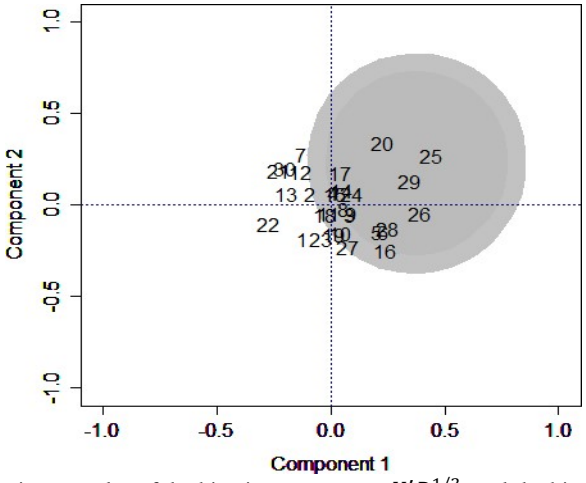

Figure 6. plot of the bivariate row scores $U'D^{1/2}$ and the bivariate 0.95 (gray external contour) and 0.90 (gray internal contour) HPD regions.
Figura 6. Gráfico das pontuações das linhas bivariadas U^' D^(1/2) e das regiões HPD bivariadas 0,95 (contorno externo cinza) e 0,90 (contorno interno cinza).

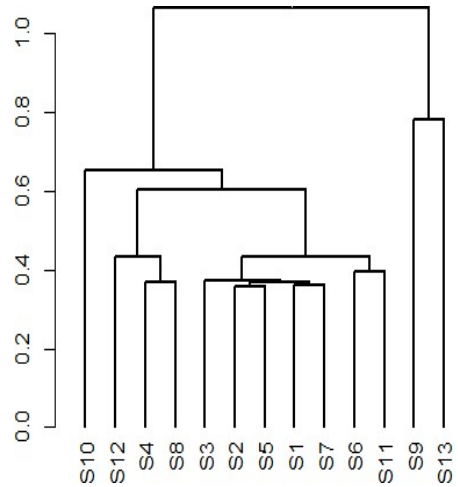

Figure 7. Dendrogram of the 13 environments using the first two right singular vectors.
Figura 7. Dendrograma dos 13 ambientes usando os dois primeiros vetores singulares direitos.

Similarly, a cluster analysis was performed using the posterior mean for genotypes using the first two left singular vectors (see Figure 8). Genotypes merge into four major clusters: Gen-20, Gen-29, and Gen-25, forming a cluster with similar characteristics. This technique for identifying homogeneous genotypes and environmental subsets is analogous to that provided by Burguenoa et al. (2008).

4. CONCLUSIONS

The study's objectives were to examine the genotype-environment interaction using a Bayesian approach, create a graphical representation of bi-plots using prior information, and elaborate on the fact that Bayesian models can be easily adapted to GEI.

A Bayesian approach using the von Mises-Fisher distribution was applied to elicit prior first-year data. Credible regions were obtained, and bilinear terms that do not contain null points (0, 0) were considered significant. Biplots using first and second bilinear terms were drawn, and it was observed that bilinear terms $u_{25,1}$ for genotype NS-10 genotype and $v_{13,1}$ for environment S13 (Piplan-14) have a significant effect on interaction.

The result concludes that the Bayesian bilinear model can be applied to yield trial data because it can deal with unbalanced data. Prior information related to genotypes, environments, and interactions can be incorporated using a Bayesian approach for means and variance. This approach can also be adopted for cells with unequal sizes.

It is suggested that this methodology can be carried out in other bilinear models by minimizing some restrictions and setting some parameters equal to zero. This approach will prove a good alternative to classical methods for analyzing genotypes through environmental interaction, which is often a major concern for yield experimenters.

6. REFERENCES

BURGUENOA, J.; CROSSA, J.; CORNELIUSB, P. L.; YANG, R.-C. Using factor analytic models for joining environments and genotypes without crossover genotype environment interaction. **Crop Science**, v. 48, n. 4, p. 1291-1305, 2008. https://doi.org/10.2135/cropsci2007.11.0632

CROSSA, J.; GAUCH, H. G.; ZOBEL, R. W. Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. **Crop Science**, v. 30, n. 3, p. 493-500, 1990. https://doi.org/10.2135/cropsci1990.0011183X003000030003x

DA SILVA, C. P.; DE OLIVEIRA, L. A.; NUVUNGA, J. J.; PAMPLONA, A. K. A.; BALESTRE, M. A bayesian shrinkage approach for ammi models. **Plos One**, v. 10, n. 7, e0131414, 2015. https://doi.org/10.1371/journal.pone.0131414

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian Data Analysis**. Chapman & Hall/CRC, 2004. 675p. https://doi.org/10.1201/b16018

HOFF, P. D. Model averaging and dimension selection for the singular value decomposition. **Journal of the American Statistical Association**, v. 102, p. 675-684, 2007. https://doi.org/10.48550/arXiv.math/0609042

JOSSE, J.; van EEUWIJK, F.; PIEPHO, H.-P.; DENIS, J.-B. Another look at bayesian analysis of ammi models for genotype-environment data. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 19, p. 240-257, 2014. https://doi.org/10.1007/s13253-014-0168-z

PEREZ-ELIZALDE, S.; JARQUIN, D.; CROSSA, J. A general Bayesian estimation method of linearbilinear models applied to plant breeding trialswith genotype × environment interaction. **Journal of Agricultural, Biological, and Environmental Statistics**, v. 17, p. 15-37, 2011. https://doi.org/10.1007/s13253-011-0063-9

RODRIGUES, P. C.; MONTEIRO, A.; LOURENCO, V. M. A robust ammi model for the analysis of genotype-by-environment data. **Bioinformatics**, v. 32, p. 58-66, 2015. https://doi.org/10.1093/bioinformatics/btv533

UHLEN, A. K.; HAFSKJOLD, R.; KALHOVD, A.-H.; SAHLSTROM, S.; LONGVA, A.; MAGNUS, E. M. Effects of cultivar and temperature during grain filling on wheat protein content, composition, and dough mixing properties. **Cereal Chemistry**, v. 75, p. 460-465, 1998. https://doi.org/10.1094/cchem.1998.75.4.460

VIELE, K.; SRINIVASAN, C. Parsimonious estimation of multiplicative interaction in analysis of variance using kullback-leibler information. **Journal of Statistical Planning and Inference**, v. 84, p. 201-219, 2000. https://doi.org/10.1016/S0378-3758(99)00151-2

YAN, W.; TINKER, N. A. An integrated biplot analysis system for displaying, interpreting and exploring genotype × environment interaction. **Crop Science**, v. 45, p. 1004-1016. 2005. https://doi.org/10.2135/cropsci2004.0076

Acknowledgments: The authors would like to thank the reviewers for their constructive comments on improving the quality of the paper.

Author Contributions: Z.B.: conceptualization, methodology, formal analysis, investigation, writing – original draft; W. M.: supervision, writing – review & editing; M.A.: supervision, formal analysis, investigation, writing – review & editing; U.W.K.: resources, writing – review & editing; M.R.: formal analysis, writing – review & editing.

Funding: This research did not receive a specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Review by institutional committee: Not applicable.

Ethics Committee: The present study compares classical methods AMMI, GGE biplot, and a Bayesian approach using Von-Mises Fisher distribution as a prior; all analyses were based on existing data.

Data availability: The data from this research can be emailed to the corresponding author upon request.

Conflict of Interest: The authors have no conflicts of interest to declare.