Identification of Stable Faba Bean (*Vicia faba* L.) Genotypes for Seed Yield in Ethiopia Using GGE Model

Gebeyaw Achenef Haile, Gizachew Yilma Kebede

Ethiopian Institute of Agricultural Research, Kulumsa Agricultural Research Center, Assela, Ethiopia

Email address:
gebesday08@gmail.com (G. A. Haile)

Abstract: Among various pulse crops, Faba bean is widely produced in Ethiopia. The crop is usually adaptable in mid and high altitude area (1800-3000 m.a.s.l.). The applications of GGE biplot ease the graphic comparison and identification of higher genotypes for supporting decision on variety selection and recommendation in different locations. Twelve advanced faba bean genotypes were conducted in 2019 across seven locations in Ethiopia using randomized complete block design with four replications. The aim of this study were to evaluate faba bean genotypes for high mean yield and identify stable varieties across locations, select ideal environment in order to design effective breeding strategy through clustering mega environments. The IPC_1 and IPC_2 together explained 58.14% of the total G X E interaction. The line transient through the biplot origin and vertical to the E1 axis splits genotypes that yielded below the mean or in the left hand side (G7, G9, G10, G2, and G1) and genotypes that yielded above the mean were all other genotypes found in the right hand side. The AEC vertical axis designated yield stability measure of genotypes. The smaller the length of the line perpendicular to the horizontal AEC axis at E1 (Assasa) indicated the more stable the genotype and vice versa. G2 and G4 were the best stable genotypes, whereas G12 and G7 were the most unstable genotypes relative to other genotypes. The test locations with longest vectors from biplot origin are more selective of the genotypes hence, E3, E1 and E2 considered more discriminating environments for the testing genotypes and least representative due to large deviation from AEC. According to the center of the concentric circles G8, represents the position of perfect genotype. The polygon view of GGE biplot identified two mega environments E1 (Assasa) and E2 (Kulumsa) as one mega environment and G11 (EH 09046-3) was the vertex genotype. The second mega environment comprises E3 (Bekoji), E4 (Kofele), E5 (Adet), E6 (Debark) and E7 (Holetta) and G12 (Tumsa) was the winning genotype for these environments. This designated there is no genotypes showed superior performance across all environments.

Keywords: Average Environment Coordinate (AEC), GGE Biplot, G x E interaction, Mega-environment, Stability

1. Introduction

Faba bean (*Vicia faba* L.) is cool season adaptable legume among various pulse crops, Faba bean is widely produced in Ethiopia. The crop is usually adaptable in mid and high altitude area (1800-3000 m.a.s.l.) [10]. Faba bean is usually produced under rain-fed conditions. However, faba bean production and productivity is becoming incapable due to inadequate improved cultivar that can be tolerant to disease and abiotic stress and high yielder varieties. In sufficient soil nutrient and high G x E interaction also another yield limiting factor in Ethiopia [7].

Selection of various crop varieties including faba bean for better production in suitable environments is usually challenging by lack of stable varieties due to existence of Genotype (G) and Environment (E) interaction effect [9]. Therefore, to evaluate stable variety Multi-environment trials (MET) are essential and these have to be done each year for grain yield and other yield component traits [2]. However, this target is made more complicated by genotype × environment interaction (GEI). Therefore, it is difficult to understand the general pattern of the data without some kind of graphical presentation [2]. The Biplot technique provides a powerful solution to this problem [5]). A biplot that displays the GGE of a MET data, referred to as a GGE Biplot it is an ideal tool for MET data analysis [11]. Therefore, the GGE biplot model is described as strong tool for effective analysis and comment of multi-environment data structure in breeding

programs [14, 8].

GGE biplot is a data visualization tool, which graphically displays G x E interaction in a two-way table [14, 8]. The applications of GGE biplot facilitate the visual comparison and identification of superior genotypes for supporting decision on variety selection and recommendation in different environments. Moreover, GGE biplot is an effective tool for mega environment analysis "which-won-where" pattern, where by particular genotypes can be recommended to specific mega-environments, genotype evaluation, the mean performance, stability and environmental evaluation [14, 17]. The GGE biplot has frequently been used for explaining G x E interaction and to determine high yielding and wide adaptability cultivars [18]. This graphical model also has broader relevance for agricultural researchers because they pertain to any two-way data matrices, and such data emerge from many kinds of experiments [4, 2]. Yield is the combined effect of G, E, and GE. Hence, GGE biplot technique separates two principal components, PC1 and PC2, which are also referred to as primary and secondary effects, respectively. The principal components are derived from subjecting environment-centered yield data (the yield variation due to GGE i.e. removing environment) to singular value decomposition. Then the pattern of genotypic response across environments can be graphically determined in a GGE biplot [15, 2]. Genotype and G x E interaction which are the two factors, important to elite cultivar selection, these factors are graphically shown through GGE biplot, which is important for visual evaluation of both genotypes and environments [14].

Plant breeders should take G x E interaction in to consideration and have measured its magnitude comparative to genotype and environment effects which affect grain yield and other yield related traits. Moreover, identification of varieties that provide best yield at a specific environment would be important to crop breeders and producers by using multi-environment trials data (many years and locations). This finding describes the usefulness of GGE biplot to evaluate faba bean genotypes for high mean yield and stability across locations, select ideal environment in order to design effective breeding strategy through delegating mega environments.

2. Materials and Methods

2.1. Descriptions of Experimental Area

The experiment was conducted at seven different locations from June to December, 2018 in the main cropping season under rain fed condition. These locations represent the varying agro ecologies of the major central faba bean growing areas of Ethiopia. The description of the test locations in terms of geographical position, altitude and climatic conditions and soil properties is given in Table 1.

Table 1. Summary of Experimental Locations.

| Locations | Geographical position | | Altitude m.a.s.l | averagerainfall | Temperature | | agro-ecology | Soil Type |
	Latitude	Longitude			Min.	Max.		
Asassa	07°06′12″N	39°11′32E	2300	620	5.8	23.6	THMH	Clay
Kulumsa	08°01′00″N	39°09′32E	2200	820	10.5	22.8	TSMMH	Clay
Bekoji	07°31′22″N	39°14′46E	2780	1010	7.9	16.6	CHMH	Clay
Holeta	09°04′12″N	38°29′45E	2400	1044	6.05	22.4	TMMH	Nitosol
Kofele	07°04′27″N	38°46′45E	2660	1211	7.1	18.0	CHMH	Nitosol
Debark	130 7' N	37053'E	2900	1044	8.6	19.8	CHMH	Nitosol
Adet	110 16' N	372 29'E	2240	1119.1	11.8	25.8	THMH	Nitosol

THMH: Tepid Humid Mid-Highland; TSMMH: Tepid Sub Moist Mid-Highland; CHMH: Cool Humid Mid-Highland; TMMH: Tepid Moist Mid-Highland.

Table 2. Descriptions of Experimental Materials.

Code	Genotypes	Pedigree
G1	Gora (standard check)	EH91020-8-2 X BPL44-1
G2	EH 010002-1-1	EH00126-1 X ILB938
G3	EH 010008-5	EKLS/CSR02017-1-4 X ILB938
G4	EH 010051-1	EKLS/CSR02018-1-1 X ATOM
G5	EH 010058-1	EKCSR/01004-2-1 X ATOM
G6	EH 010058-2	EKCSR/01004-2-1 X ATOM
G7	EH 09012-1	EH95132-1 X ILB938
G8	EH 09017-5	EH00014-1 X ILB4726
G9	EH 09021-1	EH01012-1 X ILB4726
G10	EH 09028-3	Wolki X ILB4726
G11	EH 09046-3	Wolki X ILB1563
G12	Tumsa (standard check)	Tesfa X ILB4726

2.2. Experimental Materials

A total of twelve faba bean genotypes that comprise ten advanced breeding lines and two recently released varieties (standard checks) were used for field experiment. The list of genotypes, pedigree information and their code were described in Table 2.

2.3. Experimental Design and Procedure

The experiment was conducted using a randomized complete block design (RCBD) with four replications. For each experimental unit a plot size of 4m length by 1.6m width (6.4 m^2) was used with inter row spacing of 40cm and the spacing between plants was 10cm. The distance between plots and replications were 0.6m and 1.5m respectively. All the agronomic practices were applied uniformly to the experimental units according to the recommendations. Fertilizer was applied to each plot at the rate of 121 kg NPS ha^{-1} at the time of sowing.

2.4. Data Collected

The data were recorded in plot and single plant basis according to descriptors of [6] developed for faba bean. All yield and yield related traits data were recorded on the two middle rows of each experimental unit (net plot size 3.2 m^2). The plot-based data was collected from the entire rows. For individual plant based data was recorded from a total of five randomly taken plants from each plot and averaged for data analysis.

Genotype plus genotype by environment interaction (GGE) analysis partition the G + GE effects into principal components through singular value decomposition of environmentally centered yield data. The GGE biplot model of t principal components is given as follows:

$$\overline{Y_{ij}} - \mu i - \beta j = \sum_{k=1}^{t} \lambda \kappa\, \sigma i\kappa\, \gamma j\kappa + \varepsilon ij$$

Where, $\overline{Y_{ij}}$= the performance of genotype i in environment j, [13] μi= the grand mean, βj = the main effect of environment j, k = the number of principal components (PC); λk = singular value of the kth PC; and αik and γjk = the scores of ith genotype and jth environment, respectively for PC k; εij = the residual associated with genotype i in the environment j. Usually only the first two PCs are used especially if they account for the major portion of the G x E interaction [13]. Therefore, the basic model for GGE biplot is:

$$Y_{ij} - \mu - \beta_j = \lambda_1 \xi_{i1} \eta_{j1} + \lambda_2 \xi_{i2} \eta_{j2} + \epsilon_{ij}$$

Where, Y_{ij} is the mean for the ith genotype in the jth environment, μ is the grand mean β_j is the main effect of environment j, λ_1 and λ_2 are the singular values of the 1st and 2nd principal components (PC1 and PC2), ξ_{i1} and ξ_{i2} are the PC1 and PC2 scores, respectively, for genotypei^{th}, η_{j1} and η_{j2} are the eigenvectors for the jth environment for PC1 and PC2 and ϵ_{ij} is the residual error term [11].

3. Results and Discussion

3.1. Evaluations of Genotype Mean Yield Performance at Specific Environment

The ranking of 12 faba bean genotypes based on their mean yield and stability is described in (Figure 1). The PC$_1$ and PC$_2$ together explained 58.14% of the total G X E interaction. As [14] pointed out PC1 approximated the genotype main effect or mean performance and PC2 approximate the GGE interaction effect, which is used to measure genotypes instability. The line passing through biplot origin and the environment E1 axis is indicated by arrow head solid striate line called E1 axis.

The projection of genotype markers onto this axis approximates the mean yield of the genotypes. Accordingly, the biplot organized the orders of genotypes based on mean yield performance relative to environment E1 hence, G11>G3>G8>G6>G4>G12>G2>G10>G9>G5, > G1>G7. The line passing through the biplot origin and perpendicular to the E1 axis separates genotypes that yielded below the mean or in the left hand side (G7, G9, G10, G2, and G1) and genotypes that yielded above the mean were all other genotypes found in the right hand side. Despite inconsistencies in yield rank, the agreement was found between GGE biplot (Figure 1) and simple arithmetic mean method, which took into accounts both genotype effect and G x E interaction effect [16, 20].

The AEC vertical axis indicated yield stability measure of genotypes. The smaller the length of the line perpendicular to the horizontal AEC axis indicated the more stable the genotype and vice versa [1]. Accordingly, G2 and G4 were the best stable genotypes, whereas G12 and G7 were the most unstable genotypes relative to other genotypes. In general, G11 and G3 were not only showed higher yield but also the best stable genotypes. Conversely, genotype G7 was low yielding and less stable (Figure 1).
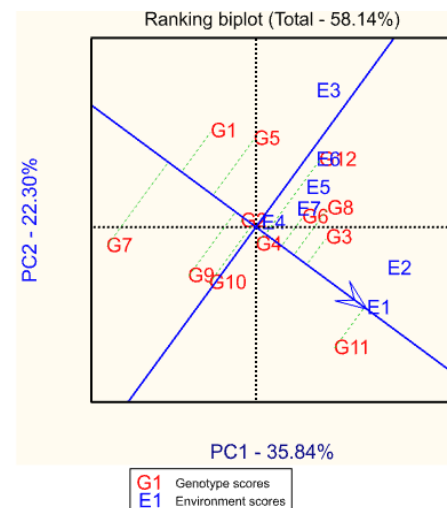

Figure 1. Ranking of the genotypes based on mean performance at E1 (Assasa) Environment.

3.2. Discriminating Versus Representativeness

The other application of GGE biplot is measuring representativeness to define an average environment and use it as a reference or benchmark. Therefore, the GGE graphical display result shows us an average environment is indicated by solid red line with arrow headed (Figure 2). The discriminating power vs. representative view of GGE biplot identifies test environments that effectively identify superior

genotypes for mega-environments. An ideal test environment is a virtual environment that has ability to discriminate the genotype and represent the mega-environment. Test environment with longest vectors from biplot origin are more discriminating of the genotypes hence, E3, E1 and E2 considered more discriminating environments for the testing genotypes and least representative due to large deviation from AEC. However, these environments are important for selecting specifically adapted genotypes.

If the test environment has a very short vector or is close to biplot origin, these environments are non-discriminating hence, less useful therefore, E4 and E7 were exhibited these characters this may be due to unfavorable rainfall conditions. Similarly, the test environment that has small angle with AEC (the average coordinates of all test environments) is more representative of the mega-environment than those that have larger angles with it for instance the angle between E3 and E1 revealed almost right angle hence, these environments were not good representative of an ideal environment as supported by [3, 15].
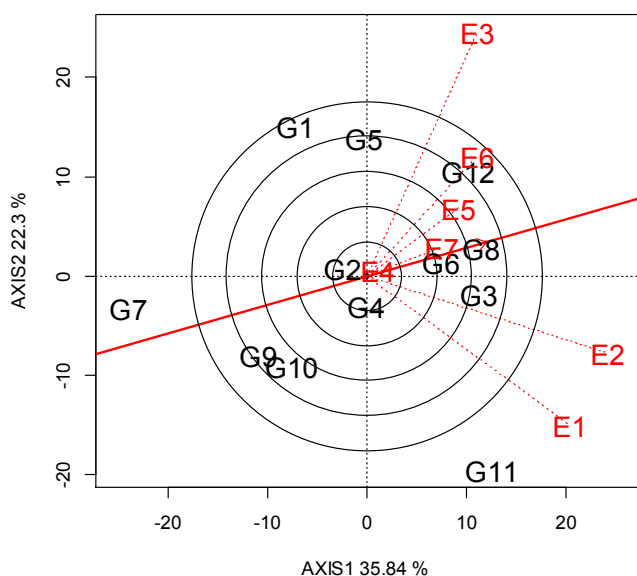

Figure 2. *Discriminations Vs. representativeness view of GGE biplot.*

The purpose of Multi location trial is, to evaluate the cultivar performance and to measure its stability [19]. Accordingly, environments found on to the center of the concentric circles represents the ideal environment, E4 and E7 were representative environments despite E7 was not discriminating cultivars, it is close to the mean environment axis which indicate representative samples for target environment. (Figure 5). Therefore, E2 found to be ideal environment since it is close to the mean environment axis i.e. representative and has ability to discriminating of genotypes. An environment is more desirable if it is closer to the „ideal" environment. Therefore, E5, and E7 were relatively highly desirable or representative test environments, whereas E1, E2 and E3 were relatively more discriminate environments but not representative test environments for mean environment.

3.3. Visual Identification of the Best Genotype(s) in Each Environment

The polygon view of the GGE biplot showed G12 was gave the best yield across five environments such as; E5 (Adet), E7 (Holetta), E3 (Bekoji), E6 (Debark) and E4 (Kofele) (Figure 3). The vertical lines to the polygon sides separate the biplot into six sectors. Environment/s found in a specific sector along with genotypes revealed that genotype was gave better yield and vice versa. In the following figure 3 indicated that most of the environments found in two sectors in quadrant one. Genotypes like G1, G11, and G6 were the vertex or the winning genotypes for the respective environments.
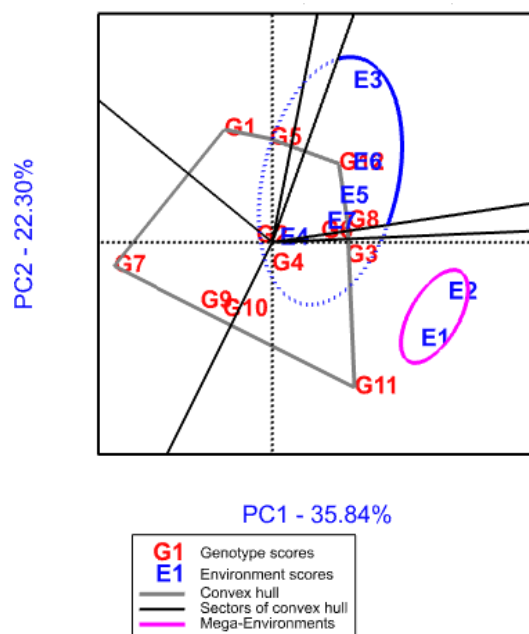

Figure 3. *Polygon view of the GGE biplot show the "which-won-where" pattern.*

According to [3] separating target environment into mega environment is recommended if crossover patterns are repeatable across year. Thus, similar results found from this study where the locations found within one sector considered as one mega-environment such as E1 and E2 fall in one mega-environment (Figure 3). Moreover, the other importance of the GGE biplot identifies the genotype which won in a particular testing location [12, 13]. Therefore, the apex genotypes (G1, G11, G12 and G7) were the most responsive genotypes to the environments; they were either the highest or the lowest genotypes in some or all of the environments. Vertical lines to the sides of the vertex hull were drawn, starting from the biplot origin, to divide the biplot into different sectors or quadrants, each having a vertex genotype (Figure 3). The vertex genotype for each quadrant was the one that gave the highest yield for the environments that found within that quadrant. Thus, genotype G11 was gave the uppermost yield in environments E1 and E2; similarly G12 was gave the highest yield in

environments E4, E7, E5, E3 and E6. The other vertex genotypes found in the lower left quadrant and far from the origin was G7 and not gave the highest yield in any of the tested environments. Hence, this genotype is the poorest genotypes in some or all of the test environments.

Genotype G12 yielded higher than G11 in all environments except E4 since all environments were on the sector of G12. The longer the line of the genotype, regardless of direction, the greater the Genotype by environment interaction associated with the genotype, which is a measure of variability or instability of the genotype across environments. Thus, the performance of genotypes G11, G7, G1, G9 and G10 were highly variable (less stable), whereas genotypes G2, G6 and G4 were highly stable (Figure 4) however stability per se is not necessarily a positive factor.
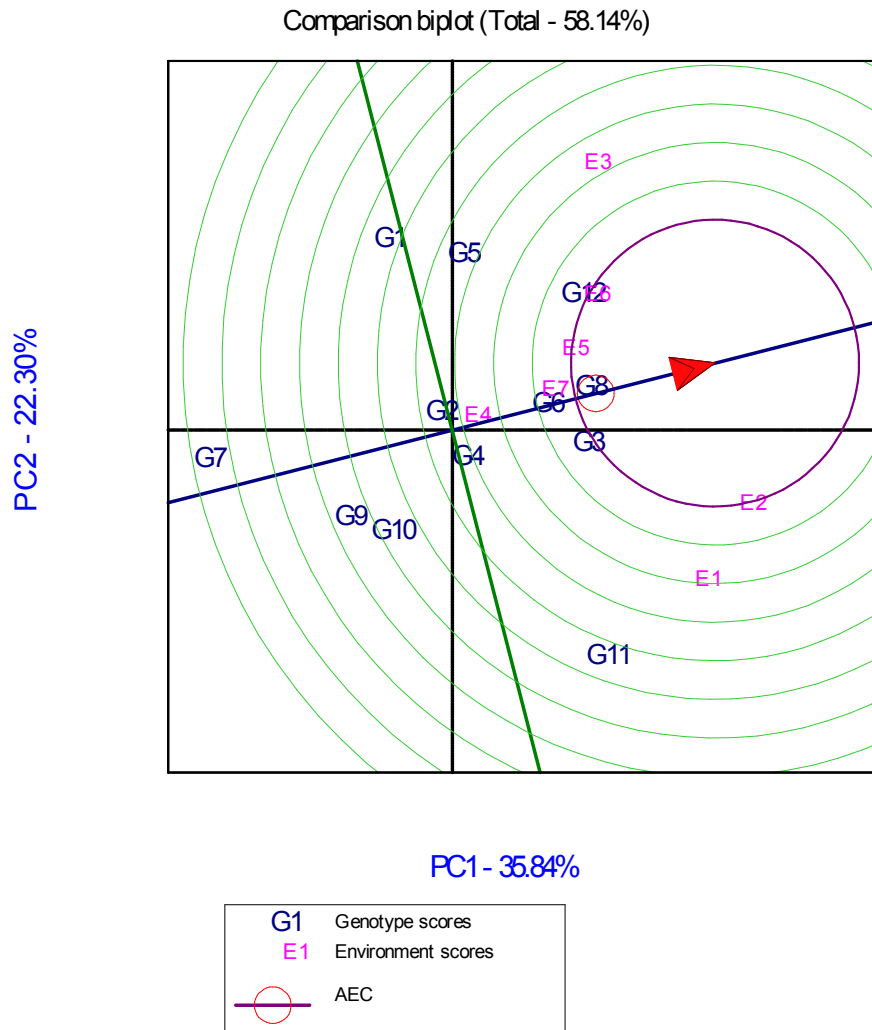

Figure 4. *Ranking of genotypes with respect to ideal genotype.*

The center of the concentric circles in Figure 4 represents the position of an 'ideal' genotype according to this G8 was considered as an ideal genotype which is defined by a projection on to the mean-environment axis that equals the longest vector of the genotypes that had above average mean yield and by a zero projection on to the perpendicular line (zero variability across environments). A genotype is more desirable if it is closer to the 'ideal' genotype. Thus, genotypes G6, G12 and G3 were more desirable than genotypes that were far from the concentric circle or ideal genotype. The low-yielding genotypes G7, G9, G1 and G10, were of course, undesirable because they were far away from the 'ideal' genotype (Figure 4).

3.4. Examine Environments

Average environmental axis (AEA) is a line passing through the origin and pointing to the positive direction with its distance equal to the longest vector. Besides, an ideal environment is a point on the AEA in the positive direction of the biplot origin and is equal to the longest vector of all environments [15, 2]. According to this E7 is identified as ideal environment despite far from the concentric circle, it is close to the mean environment axis. However, E5, E6, were identified as desirable representative environments. Though there was no ideal environment, E2 was better environment relative to other environments; E1 and E3 had longest

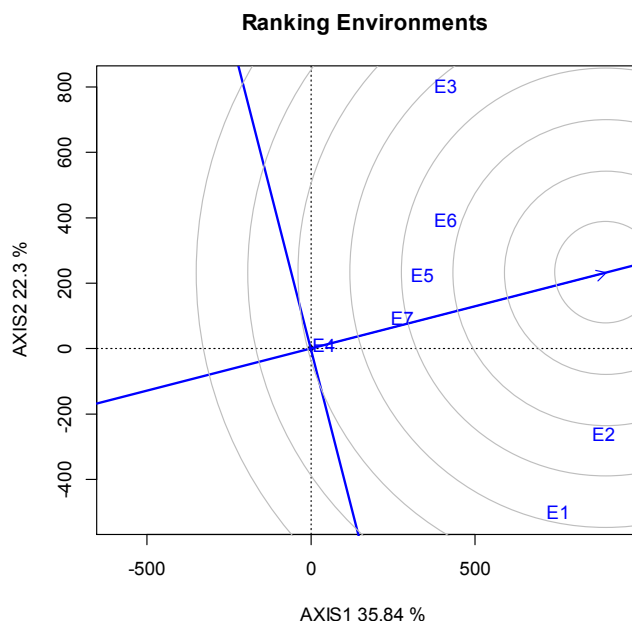projection form biplot origin and had better performance (Figure 5).

Ranking Environments

Figure 5. Ranking of genotypes with respect to ideal environments.

4. Conclusion and Recommendation

Genotype plus Genotype and Genotype x Environment interaction (GGE) model pattern analysis was applied to identify stable faba bean breeding lines or genotypes for grain yield to examine genotype and environment performance and to visualize the relationships of environment and genotypes. Despite there are various statistical models were applied to measure the stability of cultivar performance, recently many researchers used the GGE biplot technique to evaluate both mean yield and stability as well as easily identify which cultivar was superior at where environment. In this study, our result indicated that the genotype main effect or mean performance (PC1) and approximate the GGE interaction effect (PC2) explained together 58.14% to the total variation. In visual identification of the best genotypes, G12 was the winning genotypes in all environments. Two mega-environments were identified through this biplot technique.

References

[1] Ahmadi, J., Mohammadi, A., & Najafi Mirak, T. (2012). Targeting promising bread wheat (Triticum aestivum L.) lines for cold climate growing environments using AMMI and SREG GGE Biplot analyses. *Journal of Agricultural Science and Technology*, *14* (3), 645-657.

[2] Belay, F., Meresa, H., & Gebreslasie, A. (2018). Evaluation of the performance of some white seeded sesame (Sesamum indicum L.) genotypes using GGE biplot analysis in Northern Ethiopia. *Journal of Cereals and Oilseeds*, *9* (1), 1-9.

[3] Choudhary, M., Kumar, B., Kumar, P., Guleria, S. K., Singh,

N. K., Khulbe, R., ... & Rakshit, S. (2019). GGE biplot analysis of genotype× environment interaction and identification of mega-environment for baby corn hybrids evaluation in India. *Indian J. Genet*, *79* (4), 658-669.

[4] Dia M., Wehner T. C., and Arellano C. (2016). Analysis of Genotype x Environment Interaction Using SAS Programming. North Calorina State University, USA, 108 (5): 1838-1852.

[5] Erdemci, I. (2018). Investigation of genotype× environment interaction in chickpea genotypes using AMMI and GGE biplot analysis. *Turkish Journal Of Field Crops*, *23* (1), 20-26.

[6] Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, *58* (3), 453-467.

[7] IBPGR and ICARDA (International Board of Plant Genetic Resources and International Crop research in dray area (1985). descriptors for faba bean (*Vicia faba* L.) Moench. IBPGR, Rome.

[8] Nagdeve, S. S. M., & Deshmukh, D. T. (2018). GGE Bi-plot analysis in castor (Riccinus communis L.) for vidarbha region of Maharashtra state. *Electronic Journal of Plant Breeding*, *9* (2), 768-772.

[9] Papastylianou, P., Vlachostergios, D. N., Dordas, C., Tigka, E., Papakaloudis, P., Kargiotidou, A., Kostoula, S. (2021). Genotype X Environment Interaction Analysis of Faba Bean (Vicia faba L.) for Biomass and Seed Yield across Different Environments. *Sustainability*, *13* (5), 2586.

[10] Samonte, S. O. P., Wilson, L. T., McClung, A. M., & Medley, J. C. (2005). Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analyses. *Crop science*, *45* (6), 2414-2424.

[11] Schoeninger, V., Coelho, S. R. M., Bassinello, P. Z., Prado, N. V. D., Soares, T. Z. B., & Siqueira, V. C. (2020). Adaptability of brazilian beans cultivars to industrial canning. *Ciência Rural*, *50*.

[12] Tadele, M. (2020). Impact Assessment of Abiotic Production Constrains of Faba Bean (Vicia faba L.) and Breeding Mechanisms for Acid Soil, Drought and Waterlogging Affected Environments.

[13] Tolessa, T. T., Keneni, G., & Mohammad, H. (2015). Genetic progresses from over three decades of faba bean ('Vicia faba'L.) breeding in Ethiopia. *Australian Journal of Crop Science*, *9* (1), 41-48.

[14] Utpal, R., Paloti, M. C., Patil, R. S., & Katageri, I. S. (2018). Combining ability analysis for yield and yield attributing traits in interspecific (G. hirsutum L.× G. barabdense L.) hybrids of cotton. *Electronic Journal of Plant Breeding*, *9* (2), 458-464.

[15] Yan W., Cornelius PL., Crossa J. and Hunt LA. (2001). Comparison of two types of GGE biplots in genotype by environment interaction analysis. Crop Science, 41: 656- 663.

[16] Yan, W. (2015). Mega environment analysis and test location evaluation based on unbalanced multiyear data. *Crop Science*, *55* (1), 113-122.

[17] Yan, W., & Wu, H. X. (2008). Application of GGE biplot analysis to evaluate genotype (G), environment (E), and G× E interaction on Pinus radiata: A case study. *New Zealand Journal of Forestry Science*, *38* (1), 132-142.

[18] Yan W., Hunt L., Sheng Q. and Szlavnics Z. (2000). Cultivar evaluation and mega environment investigation based on GGE biplot. Crop Sci. 40: 597–605.

[19] Yan, W., & Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Canadian journal of plant science*, *86* (3), 623-645.

[20] Zhang Pan-pan, Song Hui, KE Xi-wang, Jin Xi-jun, Yin Li-. hual, Liu Yang, QU Yang, SU Wang, Feng Nai-jie, Zheng Dian-feng, Feng Bai-li (2016). GGE biplot analysis of yield stability and test location representativeness in Proso millet (Panicum miliaceum L.) genotypes. Journal of Integrative Agriculture 15 (6): 1218–1227.