INT. J. BIOL. BIOTECH., 17 (3): 551-560, 2020.

ASSESSMENT OF GENOTYPE BY ENVIRONMENT INTERACTION (GEI) TO DETERMINE MEGA-ENVIRONMENT IN HIGHLY ADAPTABLE/STABLE CULTIVARS OF COTTON (*GOSSYPIUM HIRSUTUM* L.)

Muhammad Zubair*[1], Lal Hussain Akhtar[1], Sabir Hussain[2], Imtiaz Ali[3], Muhammad Tariq Mahmood[4], Mashal Rehman[3], Rashid Minhas[1], Muhammad Shahjahan Bukhari[1], Muhammad Imran Akram[1] and Rahmat Ullah[1]

[1]*Agricultural Research Station, Bahawalpur, Pakistan*
[2]*Cotton Research Station, Bahawalpur, Pakistan*
[3]*Regional Agriculture Research Institute (RARI), Bahawalpur, Pakistan.*
[4]*Gram Breeding Research Station, Kullorkot, Pakistan*

Muhammad Zubair*[1],*Correspondence Author email:, mzf1483@gmail.com
Lal Hussain Akhtar[1], Email: grsbwp@gmail.com
Sabir Hussain[2], Email: sabirhussain.h@gmail.com
Imtiaz Ali[3], Email: imtiaz.malghani@gmail.com
Muhammad TariqMahmood[4], Email: taqaisrani@gmail.com
Mashal Rehman Email: mashal.rehman2210@gmail.com
Rashid Minhas[1], Email: rashidpbg@gmail.com
Muhammad Shahjahan Bukhari[1], Email: shahjhanbukhari@gmail.com
Muhammad Imran Akram[1], Email: aridian181@gmail.com
Rahmat Ullah[1], Email: rahmatullah2934@gmail.com

ABSTRACT

Cotton is a cash crop which contributes a high value addition in the economy of Pakistan. Sixty six cotton genotypes divided into three sets (Set-I, II, III) were tested in 14 unique environment's (areas). A statistically graphical approach GGE biplot was applied to check the yield stability of genotypes. The Set-I had 23 genotypes which were tested in 14 different environments, out of which 4 genotypes (G12, G15, G18 and G19) showed best performance. While, Set-II had 21 genotypes from which 2 genotypes (G4 and G14) were most stable genotypes. While in Set-III 22 genotypes were tested from which 2 genotypes (G12 and G17) were the best and stable performing genotypes.

Keywords: *Gossypium hirsutum*, Multivariate analysis, Average Environment Axis, Polygon, Vector

INTRODUCTION

Cotton is considered as white gold for Pakistani farmers. Cotton production was 9.861 million bales which were 17.5% decrease than last year production (11.946 million bales). Cotton contributed 0.8% share in GDP and 4.5% in agriculture (Economic Survey of Pakistan, 2018-19). This decrease in yield was largely due to shifting of cultivated area of cotton into sugar cane because of less net profit to the farmers. The yield was also affected due to harsh weather conditions, stunting of cotton growth and attack of insect pest (Economic Survey of Pakistan, 2018-19).

The environment plays one of the most important roles in the performance of genotypes. Some genotypes perform well in the favorable environment, but don't perform well another environment. The stability of genotypes in different environments and locations is the pre-requisite for the approval of variety, (Yan and Rajcan, 2002). Therefore, the candidate strains of Cotton of all Cotton Research Institutes/Stations of Pakistan are tested for their stability in yield at distinctive locations and environments. So that, the plant breeders are able to select the best performing genotypes by the utilization of multi-environment yield trial. This approach is troublesome due to the presence of Genotype by environment interaction (GEI).

Many statistical strategies have been developed to check the yield stability of genotypes at distinctive environments, but one of the foremost utilized strategies is the multiplicative analysis (GGE Biplot) (Gauch 1992; Ebdon and Gauch 2002; Moreno, *et al.*, 2003). The GGE Biplot has extraordinary significance for breeders and agriculture researchers because they contain two-way data and graphical representation. The GGE Biplot analysis decided the interaction between the genotype and the environment (Zobel *et al.,* 1988). This investigation moreover provides the information about the yield stability of genotype (Purchase *et al.*, 2000).

The main objective of this study was to determine the leading genotypes that perform best in various environments (locations) of Pakistan so that farmers don't suffer from yield losses and disease damage due to the varieties developed by plant breeders.

MATERIALS AND METHODS

The plant breeders and other scientists, developed genotypes that were tested in different locations and this experiment conducted every year. Different cottons research stations/institutes and Private seed companies send their best performing cotton genotype/genotypes to the Pakistan Central Cotton Committee (PCCC), Multan. Sixty six genotypes, including two checks (CIM-602, FH-142) contributed by different Cotton Institutes and seed companies were tested by PCCC, Multan under coded numbers. The varieties were divided into three Sets-I, II, III. The Set-I, II, III contained 23, 21 and 22 genotypes, including two checks, respectively. These 3 sets were tested at 14 different locations (environments) of Pakistan (Table 1).

The cultural and agronomic operations were done according to the requirement of the crop. Yield data of cotton were recorded during October-December by each institute and were compiled. The yield data of coded varieties were sent back to the Pakistan Central Cotton Committee (PCCC), Multan. The PCCC, Multan compiled the field data of each location. The compiled decoded data of all locations were sent to the all the contributors.

The yield data were subjected to the statistical analysis GGE Biplot with the help of statistical software, Plant Breeding Tools (PB Tools) version 1.3. These multi-environment analyses were done in order to check the Genotype by Environment Interaction (GEI) through graphical methods.

RESULTS

SET-I (23 Genotypes at 14 Locations)

Mega-environment Analysis for Set-I

The GGE biplot was made by plotting the scores of genotypes and environments of PC1 against their respective scores of second principal component (PC2) (Yan *et al.,* 2007). The GGE biplot consisted of irregular polygon that connected the genotypes with each other and lines that drew from the origin of biplot and intersect with polygon at right angles. The Polygon marks the genotypes which were located away from the biplot origin in different directions (Yan *et al.,* 2000; Yan *et al.,* 2007).

The 10 genotypes formed a polygon which was connected to each other. These genotypes were located furthest away from the origin and showed maximum variations as compared to other genotypes. The genotypes G21, G11, G22, G6, G9, G12, G18, G2, G16, and G4 are shown in Figure.1.

The lines started from the origin of biplot and intersected the polygon perpendicularly representing the winning environment with genotypes. Therefore, these lines divide the polygon into different sectors and each sector, having winning genotypes (Yan and Rajcan, 2002). If all environments fell in single sector, it indicated that the one genotype had higher yield in all environments. If environments fell into different sectors, it should have indicated that different genotypes had the highest yield in different environments. For example, the G9 genotype has higher yield in E11 and E2 environments. Similarly, G18, G2 and G12 had higher yield in E1, E4, E5, E6, E7, E8, E10, E12 and E14 (Fig. 1).

Genotype Evaluation for Set-I:

It is very important to evaluate the genotypes in a specific mega-environment. The axis of AEC abscissa (“Average Environment Axis”) is a single line with an arrow that passes through the origin of biplot (Fig. 2). The double arrow line passes from the axis of the AEC which is perpendicular to the AEC abscissa. The genotype marked/placed on the “average environment axis” represented the main effects of the genotypes. The arrow to the point of axis of abscissa represented the direction of best and highest performing genotypes. (Yan and Hunt, 2001; Yan *et, al.,* 2007). So, the best performing genotypes ranked with respect to G. G18>G19=G12>G15>G2>G20>G5=G9>G8>G1>mean G3>G13>G16>G7>G17>G10>G23=G6=G4>G22>G21>G11. Similar results were reported by other scientists (Yan *et al.,* 2007; Yan *et. al.,* 2000). Thus, the G18 was more stable genotype because of closet place near to zero point. This indicated that the genotype was performing well in mean mega-environments. So, G2 and G20 were less stable genotypes as compared to G15 and so on (Fig. 2).

Environments Evaluation for Set-I:

The main purpose of environment evaluation was to identify the test environment which ultimately diagnosed the superior genotypes for the mega-environment. The vector length of the environment was directly proportional to the standard deviation of genotypes means in the given environment which measured the selective power to the environment. Three types of test environments were used for selecting/rejecting the genotypes.

Table 3. Three types of test environments based on test environment evaluation.

	Discriminating	Non-discriminating
Representative	Type 2: Ideal for selecting superior genotypes	Type 1: Useless
Not representative	Type 3: Useful for culling inferior genotypes	

On the basis of vector length, the environments were divided into three types (Table-3). In the Type: 1, environments, the length of the vector is short, it means that these environments have little or no useful information about the genotype's performance (i.e. E1, E3, E4, E13). While, Type: 2, environment had vectors with long length as compared to Type-1 and formed a small angle with the AEC abscissa (i.e. E12, E8, E14) which mean that these environments were ideal for the selection of genotypes. The Type-3, are those environments that have long vector and formed a large angle with AEC abscissa (i.e. E11), it is not useful for selecting the genotypes (Figure-3). The findings of this study are similar to the results of Yan and Hunt (2001)**.**

SET-II (21 Genotypes at 14 Locations)

Mega-environment Analysis for Set-II

The genotypes formed irregular polygon and were away from the origin of the graph while 4 lines originated from the origin and intersected the polygon at different angles. The genotypes which formed a polygon indicated that these genotypes showed maximum variations as compared to other genotypes (i.e.G1, G17, G18, G21, G2, G4, G14, G7, G20, G6 (Figure 4). The G2 and G21 genotypes had higher yield in E2, E11 and E12 environments while G7 and G14 had higher yield in 7 environments (E6, E7, E8, E9, E10, and E14). Similarly, the G20 and G6 also had higher yield in E5 environment.

Genotype Evaluation for Set-II

The AEC depends on genotypes partitioning, on the AEC abscissa which is called as "mean vs stability" because it represents the average performance of genotypes across the stable environments within mega-environment. The double arrow line passes through abscissa and represents the mean performance of genotype. So, the genotypes were ranked according to Yan, (2000) review paper.

G14>G4>G7>G5>G15>G12>G2=G21>G16 mean G3>G18>G10=G11>G19>G17>G13>G1>G6=G20.

While, GGE indicated the G+GE means Genotype (G) + Genotype and Environment (GE) interaction and AEC abscissa represents the contribution of genotypes and AEC ordinate measure the stability and instability of a genotype/genotypes. The G14 and G4 were the most stable genotypes across the mega-environments as compared to the other genotypes because these two genotypes are closer to the small circle at AEC abscissa.

Environments Evaluation for Set-II

The main purpose was to test and evaluate the main environment in which the performance of genotypes was best (Yan and Rajcan, 2002).The AEC represents the "Discriminating power vs. Representativeness on GGE biplot with following questions?

1. Is environment which was tested have ability to discriminate between the genotypes i.e. it provides useful information about variations among the genotypes?
2. Is the test environment really mega environment?
3. Does the test environment give information about genotypes?

The E11 formed long vector and large angle with origin, so the discriminate the genotypes and E2, E4 environment while very short vectors denotes that these environments have little or no useful information about the best performance of genotypes. The length of vectors of following environments (E1, E6, E14) is medium and formed a short angle with origin having the useful information about the selection of superior genotypes and called mega-environments. Similar results were calculated by the other scientists like Dehghani *et al.* (2006).

Table1. Punjab Province locations.

S.No	**Name of location**	**Environment code used in the graph**
1	Pakistan Central Cotton Research Institute (CCRI) Multan, Cotton Research Institute (CRI) Multan,	E1
2	Cotton Research Station, Bahawalpur,	E7
3	Cotton Research Station Sahiwal,	E8
4	Cotton Research Institute Khanpur,	E9
5	Cotton Research Station Vehari	E10
6	NIAB, FSD Cotton Research Station, Faisalabad	E11
	Sindh Province Locations	
7	Central Cotton Research Institute, Sakrand	E12
8	Cotton Research Station, Mir Pur Khas.	E13
9	Cotton Research Station, Ghotki.	E14
10	Agriculture Research Institute, SAU, Tandojam.	E2
	Balochistan Province Locations	
11	Agriculture Research Institute, Sariab	E3
12	Cotton Research Station, Lasbella	E4
13	Cotton Research Station, Sibbi	E5
	KPK Province Locations	
14	Cotton Research Station, Dera Ismail Khan	E6

Table 2. Set-I Codes used in graph for genotypes.

S.No	**Genotype**	**Code**	**S.No**	**Genotype**	**Code**	**S.No**	**Genotype**	**Code**
1	AGC-Nazeer-1	G1	9	CIM-625	G9	17	FH-152	G17
2	B-2	G2	10	CIM-632	G10	18	GH-Deebal	G18
3	Bahar-07	G3	11	CRIS-600	G11	19	GH-Mubarak	G19
4	Bakhtawar-1	G4	12	Crystal-12	G12	20	IUB-65	G20
5	BH-201	G5	13	Cyto-313	G13	21	NS-181	G21
6	CEMB-55(DG)	G6	14	D-19	G14	22	SAU-1	G22
7	CEMB-88(DG)	G7	15	Eagle-2	G15	23	Sitara-15	G23
8	CIM-602 (Std-1)	G8	16	FH-142 (Std-2)	G16			
Set-II: Codes used in graph for genotypes								
1	AA-933	G1	8	CIM-602 (Std-1)	G8	15	RH-668	G15
2	Badar-1(CEMB 2)	G2	9	FH-142 (Std-2)	G9	16	Shaheen-1	G16
3	Bahar-2017	G3	10	MNH-1016	G10	17	Tarzan-5	G17
4	BH-221	G4	11	NIAB-1048	G11	18	Thakkar-808	G18
5	BS-18	G5	12	NIAB-545	G12	19	Tipu-1	G19
6	BS-80	G6	13	NIAB-Bt-2	G13	20	VH-Gulzar	G20
7	CEMB-100(DG)	G7	14	RH-662	G14	21	Weal-Ag-1606	G21
Set-III: Codes used in graph for genotypes								
1	BZU-05	G1	8	Evyol-148	G8	15	KZ-125	G15
2	CEMB-101(DG)	G2	9	FH-142 (Std-2)	G9	16	MNH-1020	G16
3	CIM-343	G3	10	FH-444	G10	17	MNH-1026	G17
4	CIM-602 (Std-1)	G4	11	FH-490	G11	18	NIA-85	G18
5	CIM-663	G5	12	GH-Haadi	G12	19	NIAB-898	G19
6	Cyto-515	G6	13	ICI-2121	G13	20	NS-191	G20
7	D-12	G7	14	IUB-69	G14	21	NU-21 (CEMB-2)	G21
						22	RH-Manthar	G22

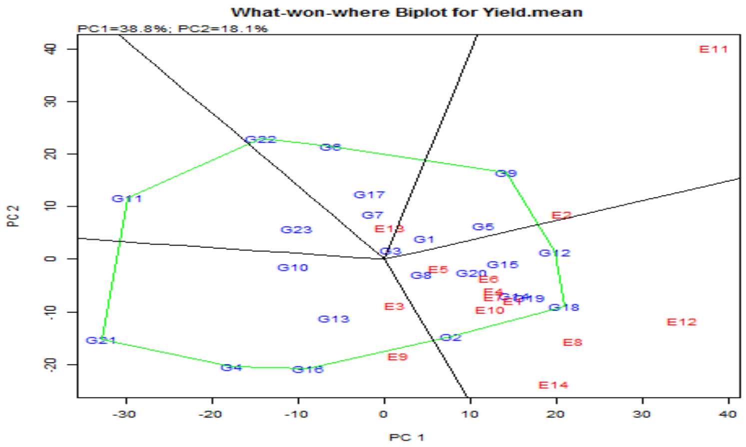

Fig. 1. (Set-I): GGE biplot “What-won-where biplot yield mean.

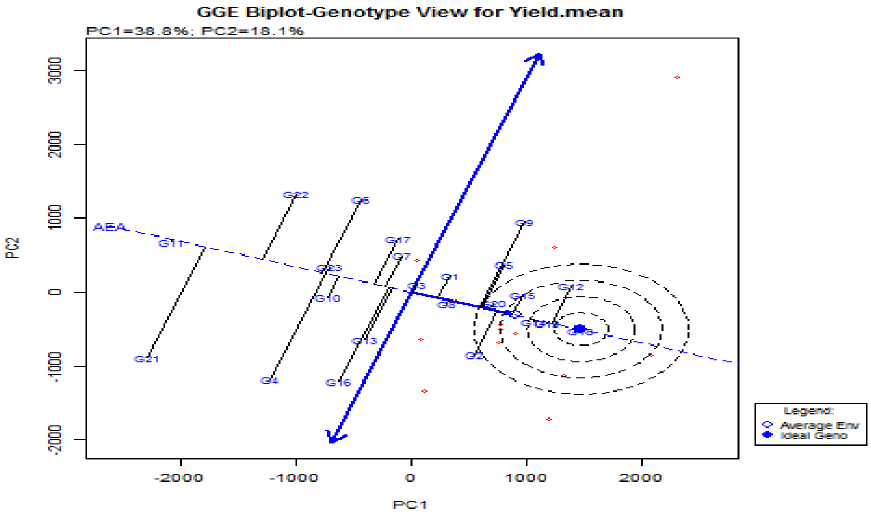

Fig. 2. (Set-I): GGE Biplot Genotype Evaluation Graph.

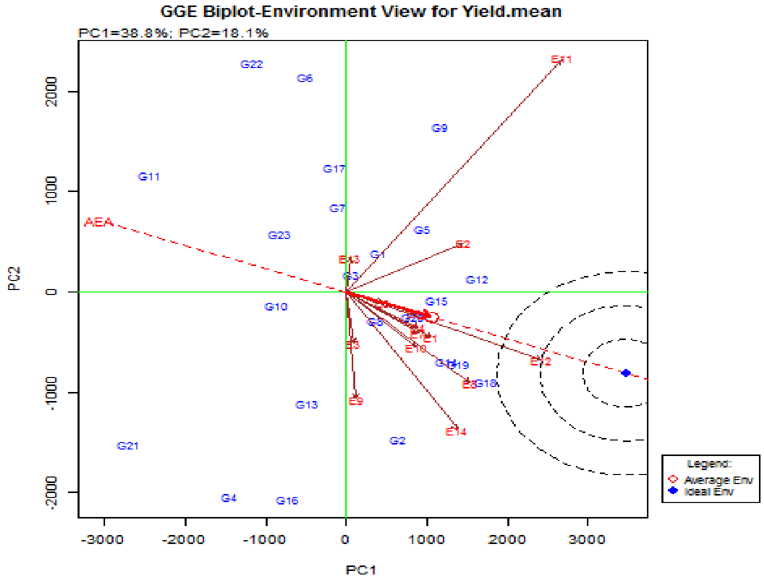

Fig. 3. (Set-I): GGE Biplot Environment view for yield mean.

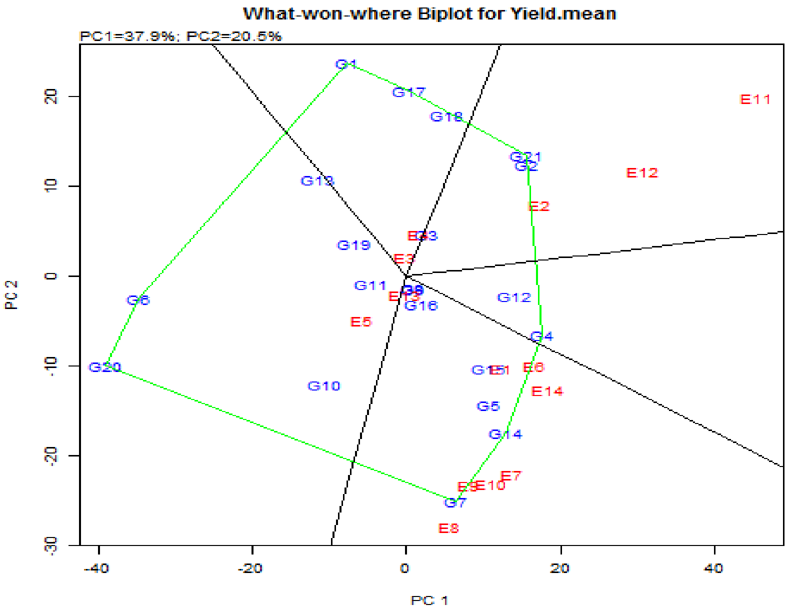

Fig. 4 (Set-II): GGE biplot "What-won-where biplot yield mean.

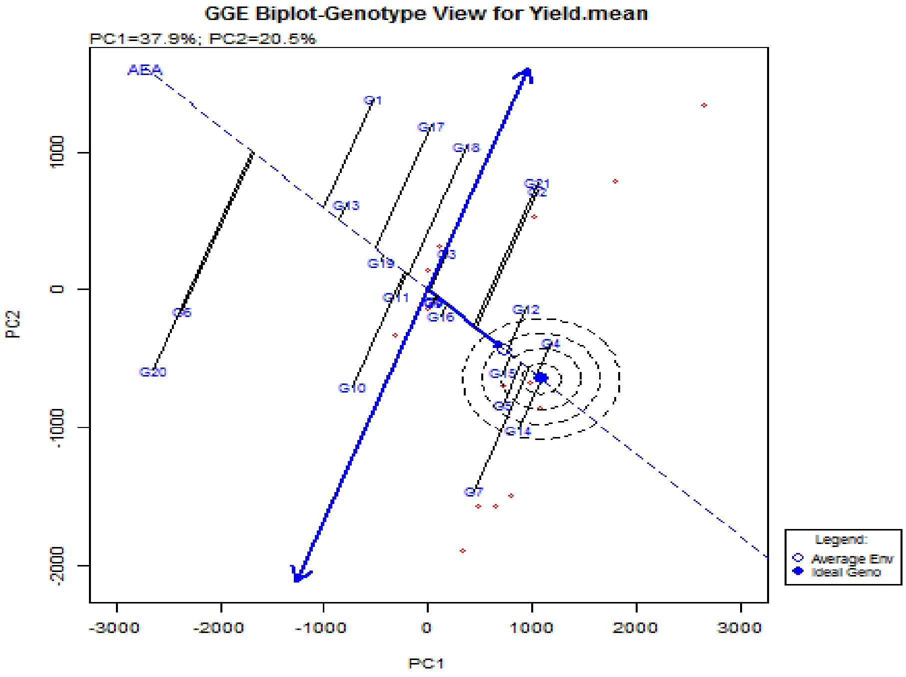

Fig. 5 (Ser-II: GGE Biplot Genotype Evaluation Graph).

SET-III (22 Genotypes at 14 Locations)

Mega-environment Analysis for Set-III

The genotypes G7, G18, G12, G17 and G16 were connected with irregular polygon and were away from the origin of the graph which showed most variable genotypes and 4 lines are originated and interconnect the polygon at different angles.

The G2 genotype had high yield in E2, E12 environments while the G17 genotype had higher yield in 6 environments (E1, E9, E10, E11, E13, and E14).

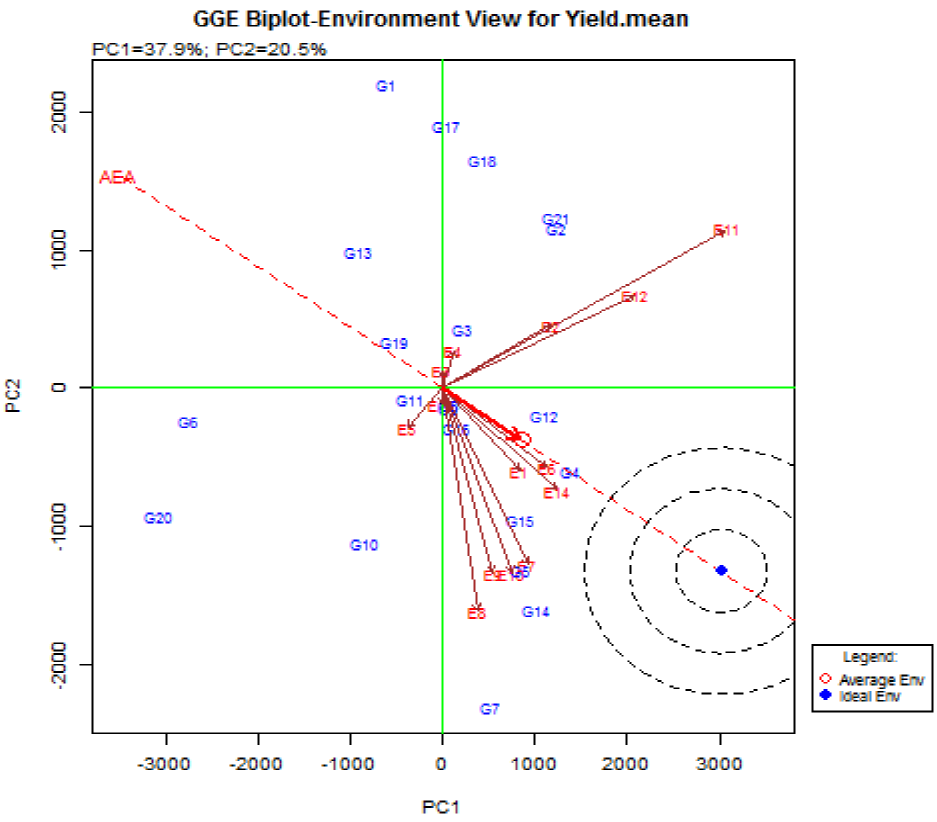

Fig. 6 (Set-II): GGE Biplot Environment view for yield mean.

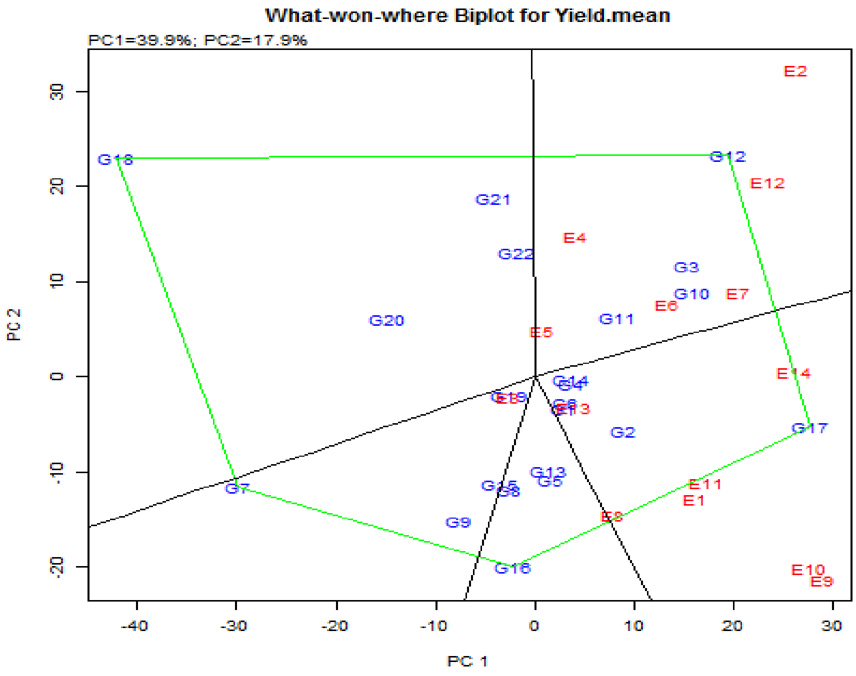

Fig. 7 (Set-III): GGE biplot "What-won-where biplot yield mean.

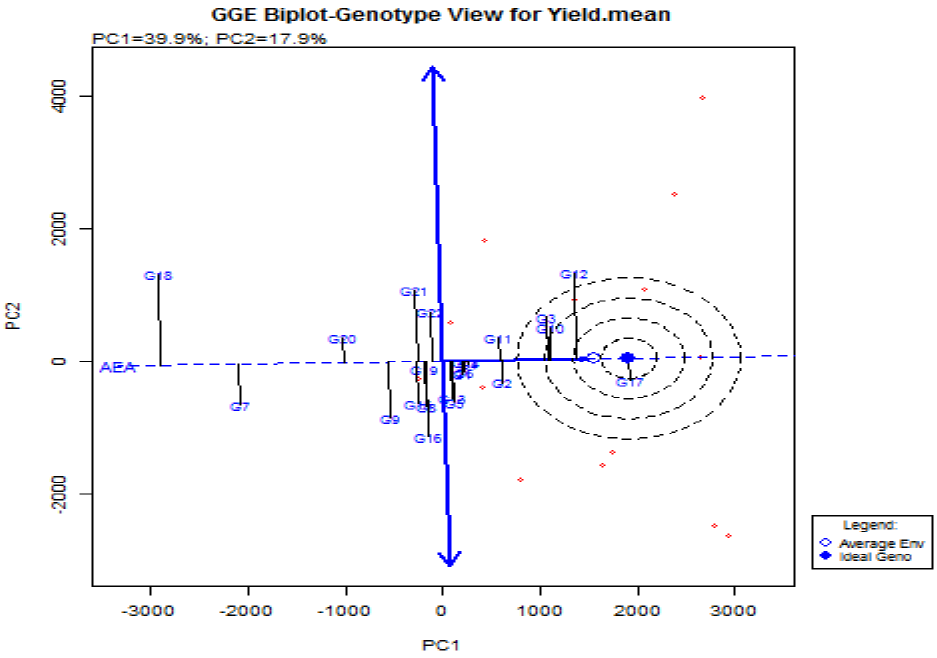

Fig. 8 (Set-III): GGE Biplot Genotype Evaluation Graph.

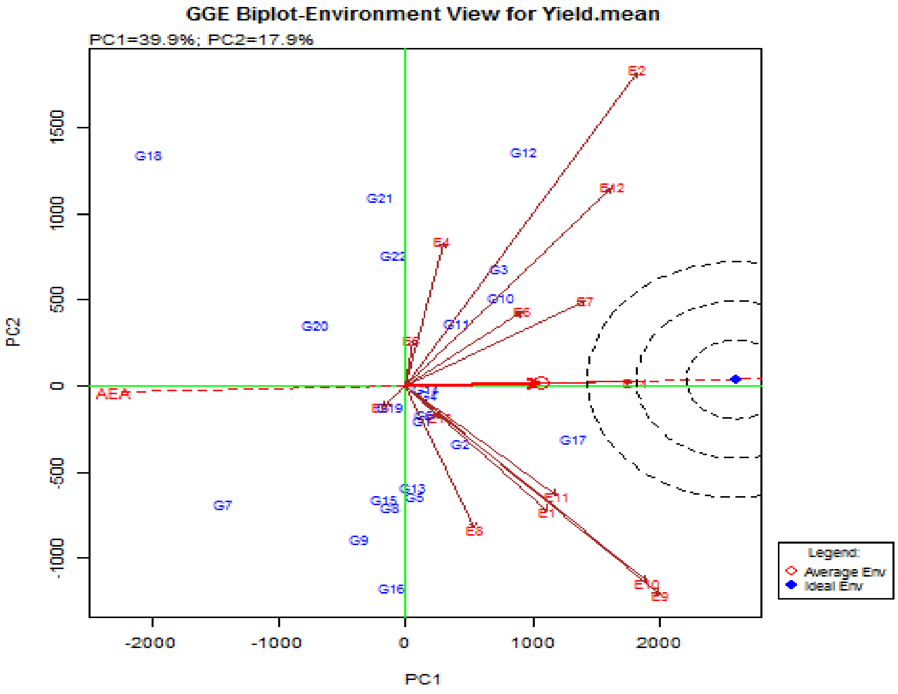

Fig. 9 (Set-III): GGE Biplot Environment view for yield mean.

Genotype Evaluation for Set-III

The placement of genotypes on the AEC abscissa indicates the genotypic performance in the stable environments. Therefore, the genotypes are getting score as their placement near to a small circle on AEC abscissa(Yan *et. al.*, 2000).

G17>G12>G10=G3>G11=G2>G4=G5=G15=G6=G1>G22>G19=G16>G21>G9>G20>G7>G18The genotypes G17 and G12 were the best and stable performing among the environments as compared to the other genotypes because these two genotypes were close to the small circle at AEC abscissa.

Environments Evaluation for Set-III

The E2 environment forms the long vector and large angle with origin, so this environment has less information about the genotypes and E5, E13 environments with very short vectors denotes that these environments have little or no useful information about the best performance of genotypes. The environments (E6 and E7) which length of vectors was medium and formed a short angle with origin, having the useful information about the selection of superior genotypes and called mega-environments.

DISCUSSION

SET-I

The results of some scientists were similar with our findings as yield of four genotypes was higher in different locations (environments) and one genotype CIM-625 had greater yield at two locations NIAB, CRS, Faisalabad and ARI, SAU, Tandojam (Yan *et al.*, 2000; Yan *et al.*, 2007) (Figure 1). The 10 genotypes form a polygon which was connected to one another. These genotypes are located furthest away from the origin and showed maximum variations as compared to other genotypes. These genotypes were G21, G11, G22, G6, G9, G12, G18, G2, G16 and G4 as shown in Figure 1. Similar results have been reported by other scientists like Yan *et al.* (2000) and Yan *et al.* (2007).

If G is too small that represents the genotypes on AEC abscissa is highly correlated with G and this correlation is 1.0 (Yan and Hunt, 2001; Yan *et. al.,* 2007). Thus, the G18 is more stable genotype because of closet place on near to zero point. This indicated that the genotype is highly performing in mean mega-environments. So, G2 and G20 were less stable genotypes as compared to G15 and so on (Fig. 2). The results were supported by the findings of Allard and Bradshaw (1964); Arshad *et al.* (2003), and Gasura *et al.* (2015)**.**

The Type: 2, environments had vectors with long length as compared to Type-1 and formed a small angle with the AEC abscissa (i.e. E12, E8, E14) which means that these environments were ideal for the selection of genotypes (Fig. 3). The findings of this study were similar to the results of Yan and Hunt (2001)**.**

SET-II

The genotypes which formed a polygon indicated that 10 genotypes showed maximum variations as compared to other genotypes and had higher yield in E2, E11 and E12) environments while CEMB-100(DG) and RH-662 had higher yield in 7 environments (E6, E7, E8, E9, E10 and E14). The G14 and G4 were the most stable genotypes across the mega-environments as compared to the other genotypes because these two genotypes were closer to the small circle at AEC abscissa. Similar findings were also reported by the Yan and Rajcan (2002).

The E11 formed long vector and large angle with origin, so the discriminate the genotypes and E2, E4 environment with very short vectors denotes that these environments have little or no useful information about the best performance of genotypes. The environments (E1, E6, E14) length of vectors was medium and formed a short angle with origin having the useful information about the selection of superior genotypes and called mega-environments. Similar results were reported by the other scientists like Dehghani *et al.* (2006) and Gasura *et al.* (2014).

SET-III

The G2 genotype had high yield in E2, E12, G7 environments while the G17 genotype had higher yield in 6 environments (E1, E9, E10, E11, E13, E14).

The placement of genotypes on the AEC abscissa indicates the genotypic performance in the stable environments. The G17 and G12 are the best and stable performing genotypes among the environments as compared to the other genotypes because these two genotypes are close to the small circle at AEC abscissa. Similar findings had been reported by Mukoy *et al.* (2018); Gauch *et al.* (2008).

The E2 environment formed the long vector and large angle with origin, so this environment has less information about the genotypes and E5, E13 environment with very short vectors denotes that these environments have little or no useful information about the best performance of genotypes. The environments (E6 and E7) where length of vectors is medium and formed a short angle with origin, having the useful information about the selection of superior genotypes and called mega-environments (Moore *et al.,* 2015; Sadabadi *et al.,* 2018).

CONCLUSION

The GGE biplot was made by plotting the scores of genotypes and environments of PC1 against their respective scores of second principal component (PC2). The GGE biplot of mega-environment analysis consisted of irregular polygon that connected the genotypes with each other and lines that drew from the origin of biplot and intersect with polygon at right angles.

The Set-I had 23 genotypes which were tested in 14 environments and these environments E12, E8, E14 having best performing and stable genotypes G18, G19, G12 and G15, While, Set-II had 21 genotypes were tested in 14 environments. The environments (E1, E6 and E14) where length of vectors is medium and formed a short angle with origin having the useful information about the selection of superior genotypes and called mega-environments. The G14 and G4 were the most stable genotypes across three environments (E1, E6 and E14) which are called the mega-environments. The Set-III had 22 genotypes tested at 14 locations. The G17 and G12 were the best and stable performing genotypes among the environments (E6 and E7) as compared to the other genotypes because these two genotypes are close to the small circle at AEC abscissa.

Acknowledgements

The authors are very thankful to breeding team and all staff of all cotton research stations who conducted the trial and compiled the data successfully.

REFERENCES

Allard, R.W. and A.D. Bradshaw (1964). Implication of genotype–environmental interaction in applied plant breeding. *Crop Science,* 5: 503–506.

Arshad, M., A. Bakhsh, A.M. Haggani and M. Bashir (2003). Genotype–environment interaction for grain yield in chickpea (*Cicer arietinum*). *Pakistan Journal of Botany,* 35: 181–186.

Dehghani, H., A. Ebadi and A. Yousefi (2006). Biplot analysis of genotype by environment interaction for barley yield in Iran. *Agronomy Journal,* 98: 388–393.

Ebdon, J.S. and H.G. Gauch (2002). Additive main effect and multiplicative interaction analysis of national turfgrass performance trials: I. Interpretation of genotype environment × interaction. *Crop Science*, 42: 489-496.

Gauch, H.G. (1992). *Statistical analysis of regional yield trials: AMMI analysis of factorial designs*. Elsevier Amsterdam, the Netherlands.

Gauch, H.G., H.P. Piepho and P. Annicchiarico (2008). Statistical analysis of yield trials by AMMI and GGE: further considerations. *Crop Science,* 48: 866–889.

Gasura, E., P.S. Setimela and C.M. Souta (2015). Evaluation of the performance of sorghum genotypes using GGE biplot. *Canadian Journal of Plant Science,* 95: 1205-1214.

Gasura, E., P.S. Setimela, A. Tarekegne, D. Icishahayo, R. Edema, P.T. Gibson and P. Okori, (2014). Variability of grain-filling traits in early maturing CIMMYT tropical maize inbred lines. *Crop Science,* 54(2): 530-536

Moreno, G.J., J. Crossa, and P.L. Cornelius (2003). Additive main effects and multiplicative interaction model-I. Theory on variance components for predicting cell means. *Crop Science*, 43: 1967-1975.

Moore, K.J. and P.M. Dixon (2015). Analysis of combined experiments revisited. *Agronomy Journal,* 107: 763-771.

Mukoyi, F., E. Gasura and G.S. Makunde (2018). Implications of correlations and genotype by environment interactions among cotton traits. *African Crop Science Journal*, 26 (2): 219 – 235.

Purchase, J.L., H. Hatting, and D. Van (2000). Genotype x environment interaction of winter wheat (*T.aestivum*) in South Africa: Stability analysis of yield performance. *South African Journal of Plant Soil*, 17(3): 101-107.

Sadabadi, M.F., G.A. Ranjbar, M.R. Zangi, S.K.K. Tabar and H.N. Zarini (2018). Analysis of stability and adaptation of cotton genotypes using gge biplot method. *Trakia Journal of Sciences*, 1: 51-61.

Yan, W., S.K. Manjit, M. Baoluo, W. Sheila and L.C. Paul (2007). GGE Biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science,* 47: 641–653.

Yan, W. and I. Rajcan (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Science*, 42: 11-20.

Yan, W. and L.A. Hunt (2001). Interpretation of genotype 3 environment interaction for winter wheat yield in Ontario. *Crop Science,* 41: 19–25.

Yan, W., L.A. Hunt, Q. Sheng and Z. Szlavnics (2000). Cultivar evaluation and mega-environment investigation based on GGE biplot. *Crop Science,* 40: 596–605.

Zobel, R.W., M.G. Wright and H.G. Gauch (1988). Statistical analysis of a yield trial. *Agronomy Journal,* 80: 388-393.

(*Accepted for publication May 2020*)