

African Crop Science Journal by African Crop Science Society is licensed under a Creative Commons Attribution 3.0 Uganda License. Based on a work at [www.ajol.info/](http://www.ajol.info/) and [www.bioline.org.br/cs](http://www.bioline.org.br/cs)  
DOI: <http://dx.doi.org/10.4314/acsj.v26i2.5>



## IMPLICATIONS OF CORRELATIONS AND GENOTYPE BY ENVIRONMENT INTERACTIONS AMONG COTTON TRAITS

F. MUKOYI<sup>1,2</sup>, E. GASURA<sup>1</sup> and G.S. MAKUNDE<sup>3</sup>

<sup>1</sup>Department of Crop Science, University of Zimbabwe, P. O. Box MP 167, Mt Pleasant, Harare, Zimbabwe

<sup>2</sup>Cotton Research Institute, Department of Research and Specialist Services, P. Bag 765, Kadoma, Zimbabwe

<sup>3</sup>Crop Breeding Institute, Department of Research and Specialist Services, P. O. Box CY550, Causeway, Harare, Zimbabwe

**Corresponding author:** [gasurae@yahoo.com](mailto:gasurae@yahoo.com), [egasura@agric.uz.ac.zw](mailto:egasura@agric.uz.ac.zw)

(Received 14 May, 2017; accepted 12 February, 2018)

### ABSTRACT

The existence of negative correlations, coupled with significant genotype x environment interaction (GE) among cotton traits poses both opportunities and challenges in breeding. This study aimed at determining the existence of GE on seed cotton (*Gossypium hirsutum* L.) yield, lint yield, earliness index, gin out turn ratio, 100 seed weight, split boll weight and fuzz grade and correlations, among these traits, as well as understanding the implications on variety evaluation and release. Ten genotypes from Cotton Research Institute, Zimbabwe, were evaluated across seven locations, during the 2012/13 and 2013/14 cropping years with three replications at each location. Seed cotton and lint yield, earliness index, gin out turn ratio, 100-seed weight, split boll weight and fuzz grade were measured on the genotypes at all locations. Analysis of variance for GE showed significant differences ( $P < 0.05$ ) on most traits, except earliness index. There were huge GE and error variance components on seed cotton yield and allied traits that reduced heritability and correlation among these traits. Seed cotton and lint yield were significantly correlated ( $P < 0.001$ ;  $r = 0.95$ ) to each other; while split boll weight had significant correlation ( $P < 0.004$ ;  $r = 0.82$ ) with 100-seed weight. The genotype plus GE biplot analysis showed that SZ 9523 had high and stable yield and desirable fuzz grade, which made it an ideal genotype for release. All locations belonged to one complex mega-environment with Chisumbanje Research Station more representative and most discriminating location which can be used in early generation testing of cotton.

**Key Words:** *Gossypium hirsutum*, lint, mega-environments

### RÉSUMÉ

L'existence des corrélations négatives, couplée avec l'interaction significative entre génotype et environnement (GE) parmi les traits de coton pose à la fois des opportunités et des défis en sélection. Cette étude a pour objectif de déterminer l'existence de GE sur le rendement des graines de coton (*Gossypium hirsutum* L.), le rendement en fibres, l'indice de précocité, le taux de gin obtenu, le poids de 100 graines, le poids des capsules fendues et la qualité du duvet et les corrélations, entre ces traits, de même que comprendre les implications sur l'évaluation de la variété et la délivrance. Dix génotypes de l'Institut de Recherche sur le Coton du Zimbabwe, ont été évalués dans sept localités, durant les années de production 2012/2013 et 2013/2014 avec trois réplifications dans chaque localité. Le rendement en graine et en fibre, indice de précocité, le taux de gin obtenu, le poids de 100 graines, le poids des capsules fendues et la qualité du duvet ont été mesurés sur les génotypes dans toutes les localités. L'analyse de variance pour GE a montré de différences significatives ( $P < 0,05$ ) sur la plupart des traits, sauf l'indice de précocité. Il y avait une grande composante de la variance de GE et l'erreur sur le rendement en graine

du coton et les traits liés qui ont réduit l'héritabilité et la corrélation parmi ces traits. Le rendement en graine et en fibre du coton ont été significativement corrélés ( $P < 0,001$ ;  $r = 0,95$ ) l'un à l'autre ; alors que le poids des capsules fissurées ont de corrélation significative ( $P < 0,004$ ;  $r = 0,82$ ) avec le poids des 100 graines. L'analyse de génotype plus GE biplot a montré que SZ9523 a eu de haute et stable rendement et désirable qualité du duvet, qui le rend un génotype idéal pour la délivrance. Toutes les localités ont formé un complexe méga-environnement avec la station de recherche de Chisumbanje plus représentative et plus discriminante qui peut être utilisée pour une évaluation du coton à une génération précoce.

*Mots Clés:* *Gossypium hirsutum*, fibre, méga-environnements

## INTRODUCTION

Cotton (*Gossypium hirsutum* L.) has global importance in the fiber industry (Campbell and Jones, 2005). Fiber quality and lint yield are complex traits and negatively affected by genotype x environment interaction (GE) (Percy, 2003; Karademir *et al.*, 2011). In Zimbabwe cultivation of cotton is mainly restricted to drier regions due to its deep rooting ability (Mapuranga *et al.*, 2015). Recently, cotton production in Zimbabwe has expanded in new areas which are now receiving less rainfall than before, due to climatic change (Mapuranga *et al.*, 2015). This poses new breeding challenges in selection of the best genotypes for all cotton production environments in Zimbabwe. The country is divided into five major-agro-ecological regions (Rukuni, 2006) with diverse physical conditions (Nyamapfene, 1991; Setimela *et al.*, 2005).

The temporal and spatial variations in biotic stresses, temperature, rainfall and soil characteristics influence lint (fibre) yield and quality (Campbell and Jones, 2005; Maleia *et al.*, 2010; Meredith Jr *et al.*, 2012). This results in differential performance of genotypes across locations; a phenomenon known as crossover GE (Finlay and Wilkinson, 1963). An appropriate stable variety is capable of utilising resources that are available in high potential environments; while maintaining above average yield in all other environments. This phenomenon is known as the dynamic concept of stability (Yan and Kang, 2002).

Quantification and visualisation of GE is important in variety development. The use of

biplots has allowed visualisation of this phenomenon in a graphical two-way data set (Gauch, 1993; Crossa *et al.*, 2002; Yan and Kang, 2002; Yan and Tinker, 2006). Approaches such as the additive main effect and multiplicative interaction biplot (Gauch, 1993, 2006, 2013); and the genotype main effect plus GE (GGE) biplot (Yan, 2001; Yan and Tinker, 2005, 2006) have been widely used. However, it is no exception that the two approaches have their computational and mathematical limitations (Mandel and Gauch, 1993; Laffont *et al.*, 2007; Yan *et al.*, 2007; Gauch *et al.*, 2008; Yang *et al.*, 2009). The GE two-way data can be subjected to different ways of singular value partitioning (SVP) (Yan and Tinker, 2006). The biplot model that is fitted to residuals after the removal of the environmental main effect (environment centered biplot) is called a GGE biplot or site regression (SREG) biplot (Crossa *et al.*, 2002; Yang *et al.*, 2009). A GGE biplot generated based on the SREG model has proved to be useful in grouping similar environments, identifying ideal testing sites, understanding the correlation of traits with either locations or genotypes and in identifying stable genotypes with high yield (Yan and Kang, 2002; Yan and Tinker, 2005, 2006). In Zimbabwe, applications of this technique have been reported in maize (Setimela *et al.*, 2007; Setimela *et al.*, 2010; Kamutando *et al.*, 2013), and recently in sorghum (Gasura *et al.*, 2015), but not yet extended to other crops including cotton. The objective of this study was to determine the importance and magnitude of GE and correlation among cotton traits, and

their implications in future cotton breeding and variety recommendation.

## MATERIALS AND METHODS

**Cotton germplasm and evaluation locations.** Five experimental (SZ 95-23, 280-94-10, 644-98-01, 648-98-11, and SZ 95-7) and five commercial (SZ 9314, CRI MS1, CRI MS2, LS 9219, and QM 301) cotton genotypes were evaluated at seven locations, during the 2012/13 and 2013/14 cropping years (Table 1). All the genotypes have medium staple except LS 9219 and 280-94-10.

**Experimental design and crop management.** The land was disc ploughed and harrowed to a fine tilth at all locations (Table 1). The cotton genotypes were hand planted in furrows at a rate of 3-5 seeds per station, following a randomised complete block design, with three replications at all locations (Table 1) for two years. Plots consisted of 5 rows of 6 m length, with an inter-row and in-row spacing of 1 m and 0.4 m, respectively. Compound L fertiliser (N: P: K: S ratio = 5: 18: 10: 8 and 0.25% boron) was manually banded in the planting furrows at a rate of 250 kg ha<sup>-1</sup>. At three weeks after planting, the crop was thinned to one plant per station, to achieve a recommended plant population of 33,333 plants ha<sup>-1</sup>. Top dressing with ammonium nitrate fertiliser (34.5% N) was done at a rate of 150 kg ha<sup>-1</sup>, at nine weeks after crop emergence. Pests were controlled using the recommended cotton pest scouting and control protocol (Mapuranga *et al.*, 2015). Weeding was done three times at all locations, using ox-drawn cultivators, followed by hand hoeing.

**Data collection.** Data were collected on split boll weight (g 100 bolls<sup>-1</sup>), seed cotton yield (t ha<sup>-1</sup>), earliness index, gin out-turn ratio, seed weight (g 100 seeds<sup>-1</sup>), lint yield (t ha<sup>-1</sup>) and fuzz grade. Split boll weight was obtained by weighing a total of 100 split bolls randomly picked from a plot. Total seed cotton yield was obtained by weighing and summing up all the

TABLE 1. Description of the cotton testing locations in Zimbabwe

*Location code	Location name	Natural region	Altitude (masl)	Average rainfall (mm)	Temperature range (°C)	Soil type
SMV	Shamva	2a	1 149	750-1000	15-28	Basaltic, loamy
WOZ	Wozhele	2b	1 245	750-1000	15-30	Alluvial
CRI	Kadoma	2b	1 156	750-800	15-35	Red clay loamy
KUW	Kuwirirana	3	996	650-800	18-35	Sand loamy
CHT	Chitekete	IV	914	650-800	20-40	Black vertisols
MUZ	Muzarabani	V	600	<500	20-40	Alluvial clay
CHS	Chisumbanje	V	406	<500	20-40	Black alluvial vertisols

\*In the subsequent text, environments 1-7 and 8-14 refers to evaluations done at these locations in descending order during the 2012/13 (first year) and 2013/14 (second year), respectively

picks (harvests) done at different times of crop phenology in kg plot<sup>-1</sup>, and then converted to t ha<sup>-1</sup>. Earliness index was determined by calculating the ratio of early harvest (the first pick) to the total yield harvested at the end of the picking period. Gin out-turn ratio (also called lint ratio or lint index) was calculated as the percentage of fiber produced from a seed cotton sample after ginning a total of 100 split bolls. Seed weight was determined by weighing a total of 100 seeds after ginning. Lint yield was obtained by multiplying the total cotton yield by the gin out turn ratio of the respective plot in kg and then converted to t ha<sup>-1</sup>. Fuzz grade was calculated on a scale of 1-10 where 1 is poor and 10 is very good fuzziness.

**Data analysis.** The genotype x location x year (G x L x Y) analysis of variance was performed on all measured parameters using a mixed model, with genotypes and locations fixed; while years were random in GenStat software version 17 (GenStat, 2014). The model followed was:

$$Y_{ijkl} = \mu + r_1(pt)_{jk} + g_i + p_j + t_k + (gp)_{ij} + (gt)_{ik} + (pt)_{jk} + (gpt)_{ijk} + e_{ijkl}$$

Where:

$Y_{ijkl}$  was the response of the  $i^{\text{th}}$  genotype in the  $j^{\text{th}}$  location and the  $k^{\text{th}}$  year in the  $l^{\text{th}}$  replication,  $\mu$  was the grand mean and  $r_1(pt)_{jk}$  was the effect of the  $l^{\text{th}}$  replication within locations and years;  $g_i$ ,  $p_j$  and  $t_k$  were the main effects of the genotype, locations and years,  $(gp)_{ij}$ ,  $(gt)_{ik}$ ,  $(pt)_{jk}$  were the first order interactions and  $(gpt)_{ijk}$  is the second order interaction, and  $e_{ijkl}$  was the micro-environmental deviation within locations and years (pooled error term). The terms  $i=1,2,3\dots 10$ ;  $j=1,2\dots 7$ ;  $k=1,2$ ; and  $l=1,2,3$ .

Variance components attributed to locations ( $\delta^2l$ ), years ( $\delta^2y$ ), locations x years ( $\delta^2ly$ ), genotypes ( $\delta^2g$ ), genotypes x location ( $\delta^2gl$ ), genotypes x years ( $\delta^2gy$ ), genotypes x locations x years ( $\delta^2gly$ ) and random error ( $\delta^2e$ ) were estimated as well (McIntosh, 1983;

Moore and Dixon, 2015). Variance components due to environments and genotype x environment were estimated by summing  $\delta^2l$ ,  $\delta^2y$  and  $\delta^2ly$ ; and  $\delta^2gl$ ,  $\delta^2gy$  and  $\delta^2gly$ , respectively. The broad sense heritability based on fixed genotypes on a single plot basis, single environment basis and across environments basis were estimated as:

$\delta^2g / (\delta^2g + \delta^2gl + \delta^2gy + \delta^2gly + \delta^2e)$ ;  $\delta^2g / (\delta^2g + \delta^2gl + \delta^2gy + \delta^2gly + \delta^2e/nr)$  and  $\delta^2g / (\delta^2g + \delta^2gl/nl + \delta^2gy/ny + \delta^2gly/nly + \delta^2e/nlry)$  Where  $nr$  = number of replications,  $nl$  = number of locations,  $ny$  = number of years,  $nly$  = number of location x year combinations and  $nlry$  was the number of years x location x replications. Pearson's phenotypic correlation analysis was done using GenStat software version 17 (GenStat, 2014) on all traits measured.

In order to determine the number of principal components to retain during GGE biplot analyses, a post-dictive evaluation was done for the model fitting using Gollob (1968) F-test (Dias *et al.*, 2003; Gauch, 2013). F-test (Gauch, 2013) showed that the two principal components of the biplot were significant, and thus could explain much of the variation (at least 64%) in the two-way data. Therefore, a GGE-2 (SREG-2) biplot analysis (Yan and Tinker, 2006) was done using Genstat Software version 17 (GenStat, 2014). The GGE biplot model was described by Yan *et al.* (2000), Yan and Hunt (2001) and Yan (2002) as:

$$Y_{ij} - \mu - \beta_j = \sum_{l=1}^k \lambda_l \xi_{il} \eta_{jl} + \varepsilon_{ij}$$

Where:

$Y_{ij}$  is the mean yield of the  $i^{\text{th}}$  genotype in the  $j^{\text{th}}$  environment,  $\mu$  is the grand mean,  $\beta_j$  is the main effect of the environment  $j$ ,  $\lambda_l$  is the singular value of the  $l^{\text{th}}$  principal component and  $k=2$  in this case,  $\xi_{il}$  is the eigen vector of the genotype  $i$  for PC  $l$ ,  $\eta_{jl}$  is the eigen vector of environment  $j$  for PC  $l$ ,  $\varepsilon_{ij}$  is the residual associated with genotype  $i$  in the environment

*j.* Based on this model, the results of all biplots presented in this work are mainly environment centered.

Grand means of all traits across environments were gathered to make a two-way table of genotypes x traits means. The genotypes means across environments data were standardised by dividing each trait mean value with the within trait standard deviation, as outlined by Yan and Tinker (2006). Standardisation removed the different units among different traits (Yan *et al.*, 2000; Yan and Hunt, 2001). The resultant data were subjected to the scatter biplot analyses using the symmetric focused SVP method, and was trait-centered. The sectors that grouped specific genotypes and traits were generated.

The existence of the crossover interactions was checked using the Gail-Simon test (Baker, 1988). The which-won-where scatter biplot, genotype comparison biplot, location comparison biplot and the location x trait biplot were generated using the appropriate SVP methods (Yan, 2002). Biplots for seed cotton yield and lint yield were generated separately. In the scatter biplot, the polygon view displaying the which-won-where pattern was formed by connecting the genotype markers furthest away from the biplot origin, such that the polygon contained all other genotypes (Yan, 2002). The polygon was then dissected by sectors running from the biplot origin such that each sector encloses certain environments and genotypes. Visualisation of the mean and stability of genotypes using a genotype comparison biplot was achieved by drawing an average environment coordinate that is represented by a small circle. A line that passes through the biplot origin and the average environment coordinate was drawn, followed by a perpendicular line. For the analyses of test location, an average environment coordinate was drawn as represented by a small circle. A line that passes through the biplot origin and the average environment coordinate was drawn, followed by a perpendicular line.

## RESULTS

### **Combined analysis of variance across years x locations and variance components.**

Locations had highly significant differences ( $P < 0.001$ ) for all traits measured, except fuzz grade (Table 2). Years were also significant ( $P < 0.05$ ) for all traits, except for seed cotton and lint yield (Table 2). Locations x years were significant ( $P < 0.01$ ) for all traits.

Genotypes were significantly different ( $P < 0.05$ ) for all traits except seed cotton yield and earliness index (Table 2). The GE was significant ( $P < 0.05$ ) for other traits either as genotypes x locations, genotypes x years or genotypes x locations x years interactions, but not significant for earliness index. Locations had highest contribution to the variation observed and GE also contributed to the variation observed on all the traits measured (Table 3). The GE to genotype variance component ratio was large for all traits studied. This ratio was much larger for lint yield (13 times) and seed cotton yield (22 times), than for the rest of the traits studied (Table 3).

### **Genotypes mean performance, heritability estimates and correlation analysis.**

The genotypic mean values for all traits are shown in Table 4. The heritability values for all traits were low based on individual plot analysis, but increased based on environment means across locations (Table 3). The Pearson phenotypic correlation analysis showed positive correlation between lint and seed cotton yield ( $r = 0.95$ ,  $P < 0.001$ ); while the split boll weight was positively correlated with 100 seed weight ( $r = 0.82$ ,  $P < 0.004$ ). The relationship between these traits was further supported in the genotype x trait biplot analysis (Fig. 1).

### **Mega-environment delineation and genotype performance evaluation.**

The Simon-Gail test (Baker, 1988) showed significant crossover GE interactions (not shown). The which-won-where pattern showed that different genotypes were winning

TABLE 2. Mean square values and their significance for cotton fiber yield and related traits

Source of variation	Degrees of freedom	Split boll weight (g)	Seed cotton yield (t ha <sup>-1</sup> )	Earliness index (%)	Gin out turn ratio (%)	Lint yield (t ha <sup>-1</sup> )	100 seed weight (g)	Fuzz grade
Locations	6	2.38347***	7614390***	7712.52***	162.127***	1257566***	26.6372***	0.16508 <sup>NS</sup>
Years	1	4192.323***	41241 <sup>NS</sup>	15302.79***	24.266*	808 <sup>NS</sup>	131.0409***	78.86667***
Locations x years	6	2.37822***	15643388***	9484.13***	56.541***	3035150***	46.2832***	0.42222**
Locations x years (replications)	28	0.08288*	315637***	37.44**	3.924**	58140***	0.1674 <sup>NS</sup>	0.10238 <sup>NS</sup>
Genotypes	9	0.67464***	233972 <sup>NS</sup>	33.39 <sup>NS</sup>	17.184***	57255*	3.6886***	0.13122*
Genotypes x locations	54	0.05137 <sup>NS</sup>	234603**	13.41 <sup>NS</sup>	3.014*	44747**	0.2954**	0.09541*
Genotypes x years	9	0.67378***	336857**	29.9 <sup>NS</sup>	3.051 <sup>NS</sup>	66770**	0.9318***	0.21587**
Genotypes x locations x years	54	0.05141 <sup>NS</sup>	213614**	14.33 <sup>NS</sup>	3.468**	39620*	0.2892*	0.11093**
Residual	252	0.05355	133918	19.8	2.017	25410	0.1853	0.06799

TABLE 3. Variance components and their percentage (in brackets) contribution to the total variance for cotton fiber yield and related traits

Source of variation	Split boll weight (g)	Seed cotton yield (t ha <sup>-1</sup> )	Earliness index (%)	Gin out turn ratio (%)	Lint yield (t ha <sup>-1</sup> )	100 seed weight (g)	Fuzz grade
Locations	0.04 (0.2)	121645.88 (14.5)	127.92 (23.8)	2.64 (34.0)	19990.43 (12.6)	0.44 (14.9)	0.001 (0.2)
Years	19.96 (98.9)	0 (0)	72.69 (13.5)	0.10 (1.2)	0 (0)	0.62 (21.1)	0.375 (77.2)
Locations x years	0.08 (0.4)	510925.03 (60.8)	314.89 (58.5)	1.75 (22.6)	99233.67 (62.6)	1.54 (52.0)	0.011 (2.2)
Genotypes	0.01 (0.1)	2382.24 (0.3)	0.32 (0.06)	0.36 (4.7)	758.21 (0.5)	0.08 (2.8)	0.002 (0.3)
Genotypes x locations	0 (0)	16780.83 (2.0)	0 (0)	0.17 (2.1)	3222.83 (2.0)	0.02 (0.6)	0.005 (0.9)
Genotypes x years	0.03 (0.1)	9663.76 (1.2)	0.48 (0.09)	0.05 (0.6)	1969.52 (1.2)	0.04 (1.2)	0.007 (1.5)
Locations x years x genotypes	0 (0)	26565.33 (3.2)	0	0.48 (6.2)	4736.67 (3.0)	0.03 (1.2)	0.014 (2.9)
Residual	0.05 (0.3)	133918.00 (15.9)	19.8 (3.7)	2.02 (26.0)	25410.00 (16.0)	0.19 (6.3)	0.068 (14.0)
Environments	20.08 (99.5)	632570.92 (75.3)	515.5 (95.8)	4.49 (57.9)	119224.10 (75.2)	2.60 (87.9)	0.387 (79.6)
Genotype x environment (GE)	0.03 (0.1)	53009.93 (6.3)	0.48 (0.09)	0.70 (9.0)	9929.02 (6.3)	0.09 (3.0)	0.026 (5.3)
GE/G	2	22.25	1.49	1.94	13.1	1.06	17.221
<b>Broad sense heritability</b>							
Based on a single plot	15.11	1.26	1.57	11.74	2.1	23.35	1.578
Based on single environments means	23.79	2.38	4.37	20.84	3.96	35.69	3.005
Based on across environments means	47.97	16.21	31.25	73.39	24.1	75.35	18.094

TABLE 4. Main effects of the genotypes and locations for cotton fiber yield and related traits

Genotype name	Split boll weight (g)	Seed cotton yield (t ha <sup>-1</sup> )	Earliness index (%)	Gin out turn ratio (%)	Lint yield (t ha <sup>-1</sup> )	100 seed weight (g)	Fuzz grade
644-98-11	3.014	1929	92.09	42.55	821.5	11.48	5.5
SZ-95-7	3.127	2038	91.45	42.86	876.8	11.68	5.571
280-94-10	3.075	1986	91.76	42.4	845.9	11.26	5.619
SZ-95-23	3.268	2042	90.05	42.08	860.1	11.88	5.571
CRI-MS-1	3.255	1962	91.56	41.9	827	12.07	5.571
CRI-MS-2	2.935	1980	90.39	42.54	841.9	11.27	5.429
SZ 9314	3.319	2015	90.32	42.42	854.1	11.72	5.476
QM 301	3.271	1899	89.67	42.99	815.9	11.87	5.524
LS 9219	3.238	1842	91.07	40.76	752.7	12.08	5.524
648-01-4	3.155	1837	89.68	42.71	785.4	11.58	5.5
Least significant difference	0.099	-	-	0.61	68.51	0.19	0.112
F-probability	<.001	0.079	0.093	<.001	0.019	<.001	0.048
Coefficient of variation	7.31	18.74	4.9	3.36	19.25	3.68	4.72
<b>Location name</b>							
Chisumbanje	3.023	1960	90.32	40.9	818.5	11.99	5.583
Chitekete	2.95	1942	93.06	41.25	799.7	11	5.6
Kadoma	3.394	1427	100	43.65	625.2	12.97	5.45
Kuwirirana	3.135	1954	100	45.27	885.2	11.03	5.533
Muzarabani	2.955	2460	82.48	40.88	1006.8	11.72	5.483
Svamva	3.301	1627	69.76	41.68	675.1	11.56	5.517
Wozhele	3.403	2302	100	42.62	986.5	11.55	5.533
Least significant difference	0.083	186.1	1.6	0.5106	81.06	0.219	0.094
F-probability	<.001	<.001	<.001	<.001	<.001	<.001	0.027
Coefficient of variation	7.31	18.74	4.9	3.36	19.25	3.68	4.72

Genotype x environment interaction analysis of cotton traits

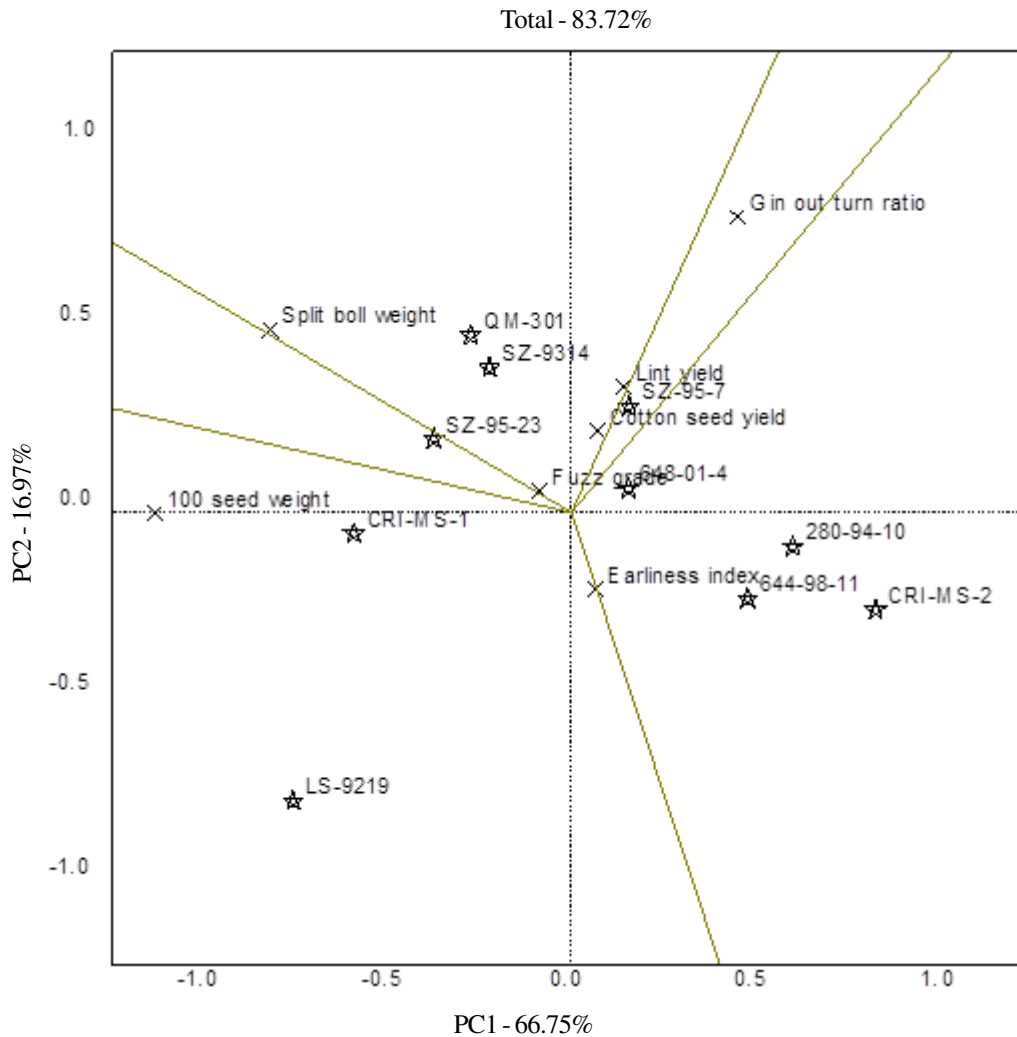


Figure 1. Genotype x trait biplot showing cotton genotypes with a given set of traits in a study in Zimbabwe. Data were standardised with the within-trait standard deviation. The biplot was produced based on trait focused SVP, no scaling, no transformation and the data were trait centered.

in different environments and the pattern was similar for both seed cotton yield and lint yield (Figs. 2 and 3). A polygon was formed by joining genotypes that were furthest from the biplot origin. The genotypes that were on the vertices of the polygon include 644-98-11, CRI-MS-2, SZ-9314, QM-301, LS-9219, and CRI-MS-1 (Figs. 2 and 3).

The genotype comparison biplots based on seed cotton yield and lint yield produced a similar pattern of genotype ranking, based on

mean yield and stability of these two traits (Fig. 4). Genotypes that include SZ 95-23, CRI MS1, CRI MS2 and SZ 9314 were above the average environment coordinate; while the rest were below the line (Fig. 4). However, the comparison biplots showed that genotype CRI-MS-2 was closer to the average environment coordinate; followed by SZ-95-23 and SZ-9314, for both seed cotton and lint yield (Fig. 4).



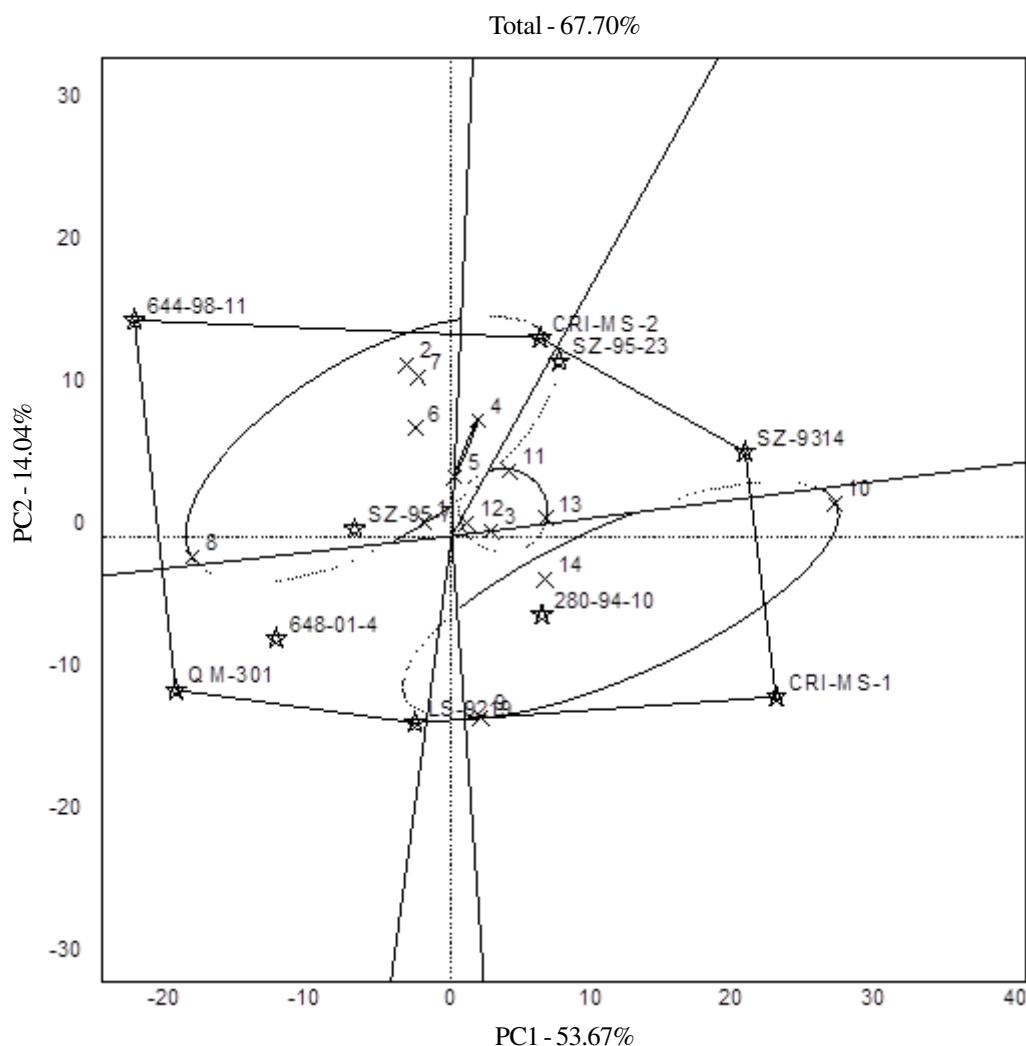


Figure 2. The which-won-where and mega-environment delineation biplot for the 10 cotton genotypes evaluated in five locations for two years for seed cotton yield in Zimbabwe. The biplot was produced based on symmetric focused SVP, no scaling, no transformation and the data were environment centered.

**Site performance evaluation.** The seed cotton yield biplot showed Chisumbanje Research Station in the south east lowveld was closest to the average environment coordinate for both seed cotton and lint yield (Figs. 5 and 6). Kadoma Research Station and Wozhele Farm, both in Mashonaland West came second for seed cotton and lint yield, respectively (Figs. 5 and 6). Kuwirirana, Wozhele and Chitekete were positively correlated (Fig. 7).

## DISCUSSION

**Variance components and heritability estimates.** In the current study, heritability improved significantly based on genotypic means across environments (years and location combinations). The heritability value could be increased by increasing the number of replications, years, locations and by using best experimental designs (Bernardo, 2002).

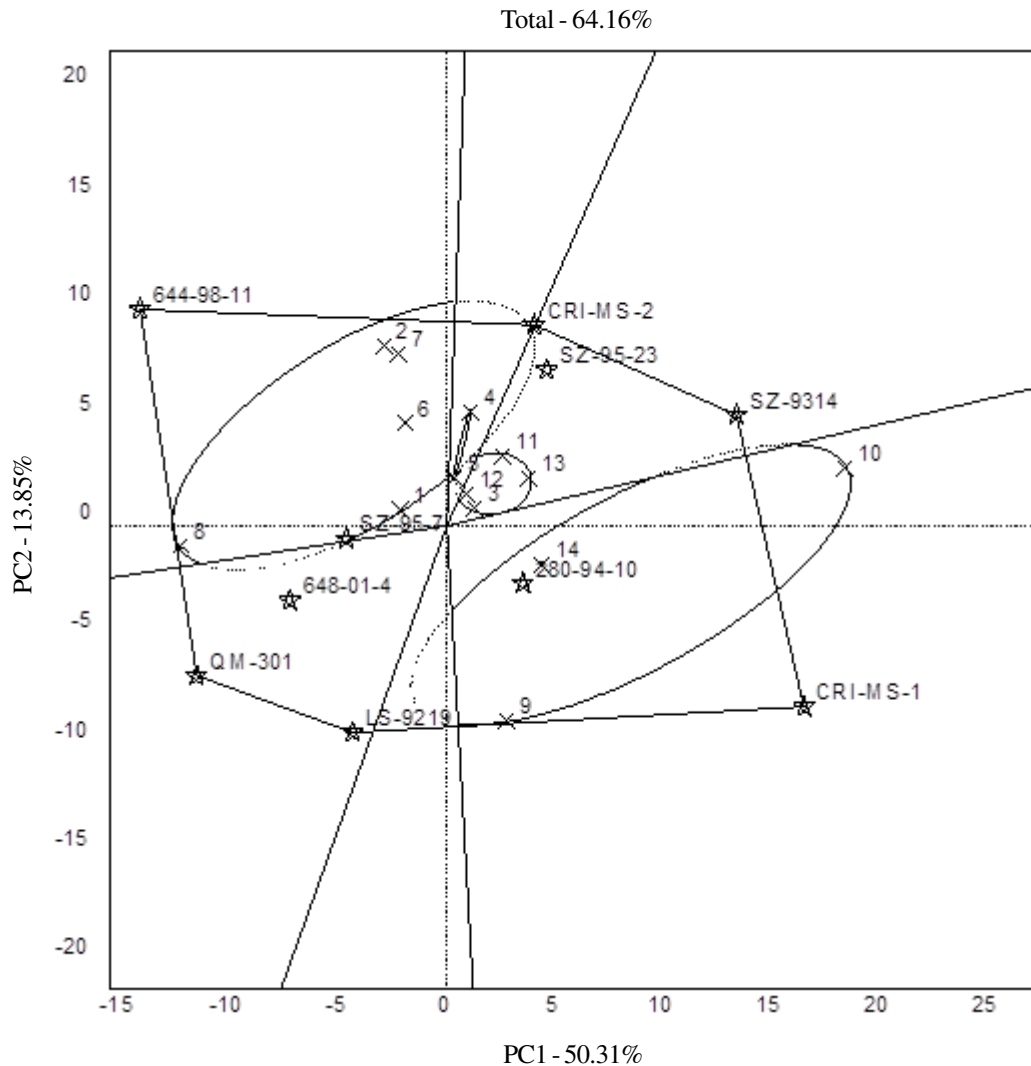


Figure 3. The which-won-where and mega-environment delineation biplot for the 10 cotton genotypes evaluated in five locations for two years for lint yield in Zimbabwe. The biplot was produced based on symmetric focused SVP, no scaling, no transformation and the data were environment centered.

Selection of suitable cotton varieties is hindered by the existence of the large error and GE variance components, that reduce heritability. This situation was observed on several studies in cotton (Campbell and Jones, 2005; Maleia *et al.*, 2010; Meredith Jr *et al.*, 2012). When the error and the GE variance components are huge, they reduce the repeatability (Annicchiarico *et al.*, 2000; Bernardo, 2002; Gasura *et al.*, 2013).

The highly significant correlations observed on individual plot basis (Fig. 7) were mainly due to non-genetic causes such as error and GE. On the other hand, correlations based on genotypic means across locations and years were accurate, since the heritability value was improved. This will be costly in variety testing by breeders but has an advantage of improving the selection efficiency, and hence, variety recommendation. When multiple traits are to

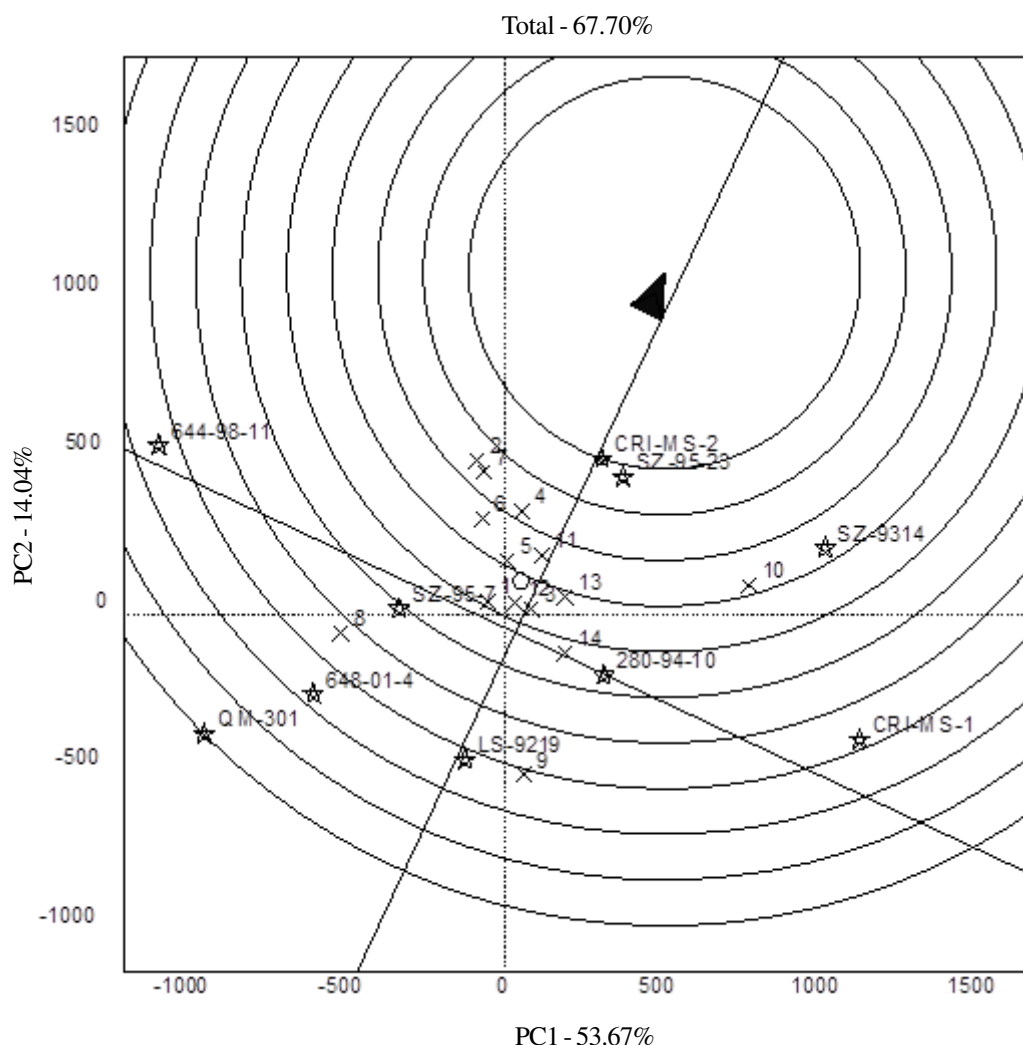


Figure 4. A genotype comparison biplot showing the best cotton genotypes based on seed cotton mean performance and stability across 14 environments (seven locations and two years) in Zimbabwe. To avoid congesting the graph, the environments are numbered 1-14 while genotypes are shown by their names. The biplot was produced based on genotype focused SVP, no scaling, no transformation and the data were environment centered.

be considered in variety evaluation and recommendation, a selection index must be developed (Gasura *et al.*, 2013) and used to improve the selection efficiency. Use of a selection index is ideal when the traits to use for indirect selection are highly correlated with the target trait and are simply inherited (Gasura *et al.*, 2013). A selection index has an

advantage of maximising genetic gains for the most important traits concerned.

**Significance of correlation analysis.** The highly significant positive correlations observed on split boll weight and seed yield reflected a true genetic relationship. Split boll weight is a function of the number of bolls

Total - 76.39%

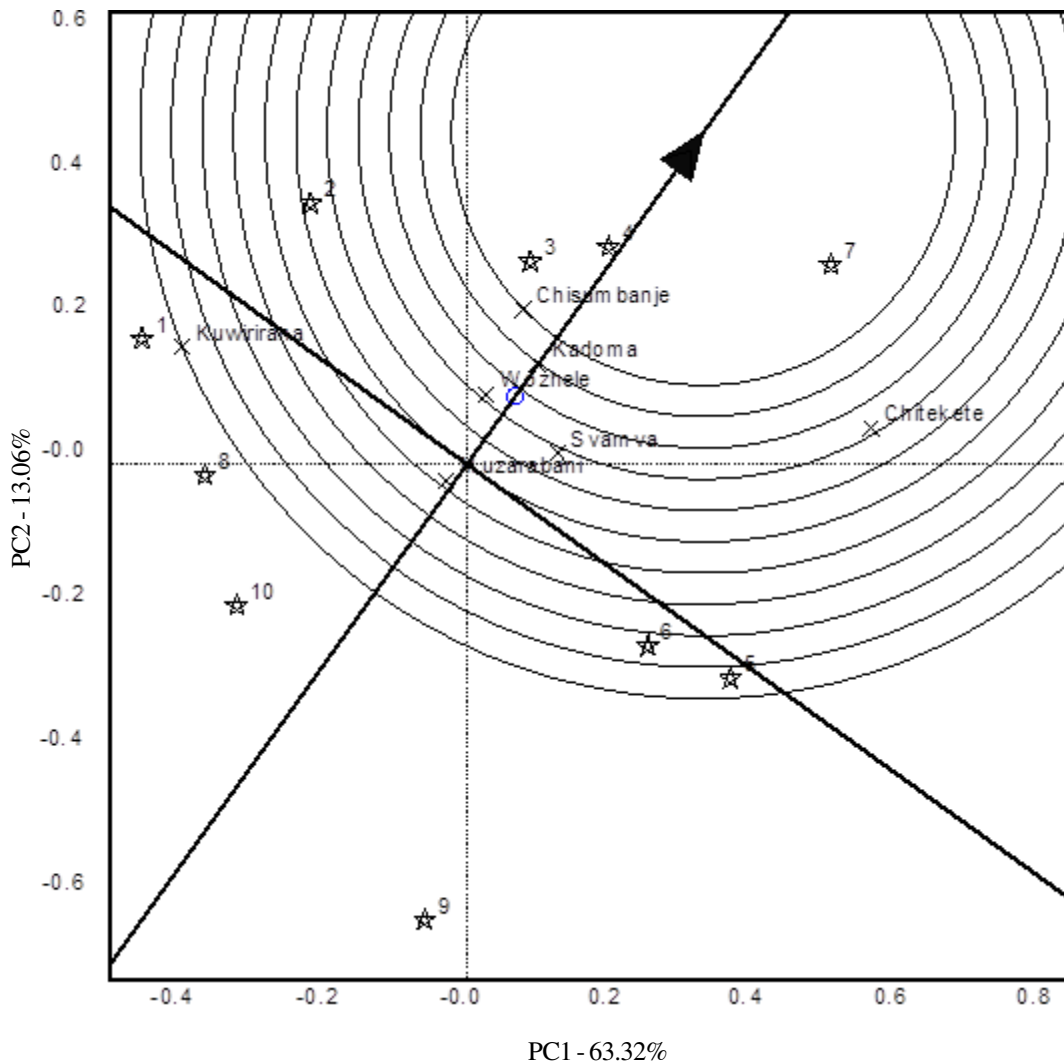


Figure 5. The ideal testing location for cotton seed yield among the locations used in evaluations in Zimbabwe. The biplot was produced based location focused SVP, no scaling, no transformation and the data were location centered. To avoid congesting the graph, the genotypes are numbered 1-10 while locations are shown by their codes.

sampled; while seed weight is a function of the total number seeds per boll (Killi *et al.*, 2005). The latter is highly heritable, and thus explains the positive correlation observed between split boll weight and 100-seed weight (Fig. 7). When complex traits such as yield are correlated with simply inherited traits, then simply inherited traits can be used to proxy the complex trait since such traits are not

affected much by GE (Meredith Jr *et al.*, 2012; Gasura *et al.*, 2014).

The high positive correlation observed for seed cotton and lint yield (Fig. 7) may reflect the relationship between samples that were ginned than the genetic cause. This is because the gin out-turn ratio was not related to lint yield. Furthermore the genotype mean square for seed cotton was non-significant. In general,

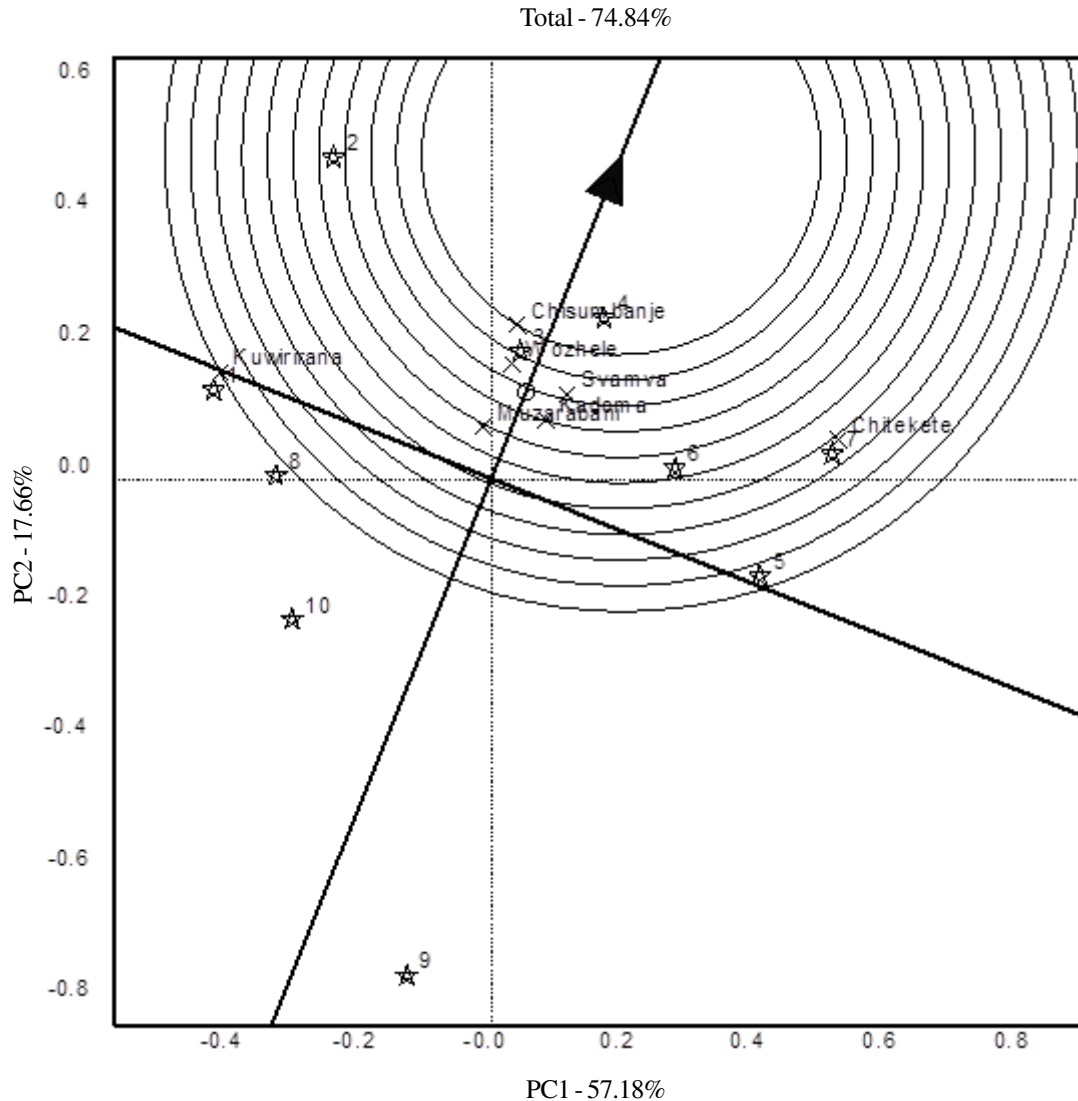


Figure 6. The ideal testing location for lint yield among the locations used in evaluations in Zimbabwe. The biplot was produced based location focused SVP, no scaling, no transformation and the data were location centered. To avoid congesting the graph, the genotypes are numbered 1-10 while locations are shown by their codes.

high earliness index, high seed cotton yield, high lint yield, high split boll weight, high gin out turn ratio, good fuzz grade and small seed size make the ideal genotype for commercialisation. The lack of correlation between lint or seed cotton yield with other traits such as earliness index, fuzz grade and seed size could be exploited to benefit the farmers. For example, short maturing genotypes could be selected as a drought

escaping mechanism for farmers in dry areas, without major penalties on yield traits. Lack of correlation between total cotton yield (seed cotton yield before ginning) and 100-seed weight suggests that it is possible to get genotypes with high lint yield from genotypes that possess high seed cotton yield.

**Genotype evaluation based on mean yield and stability.** The huge crossover type GE

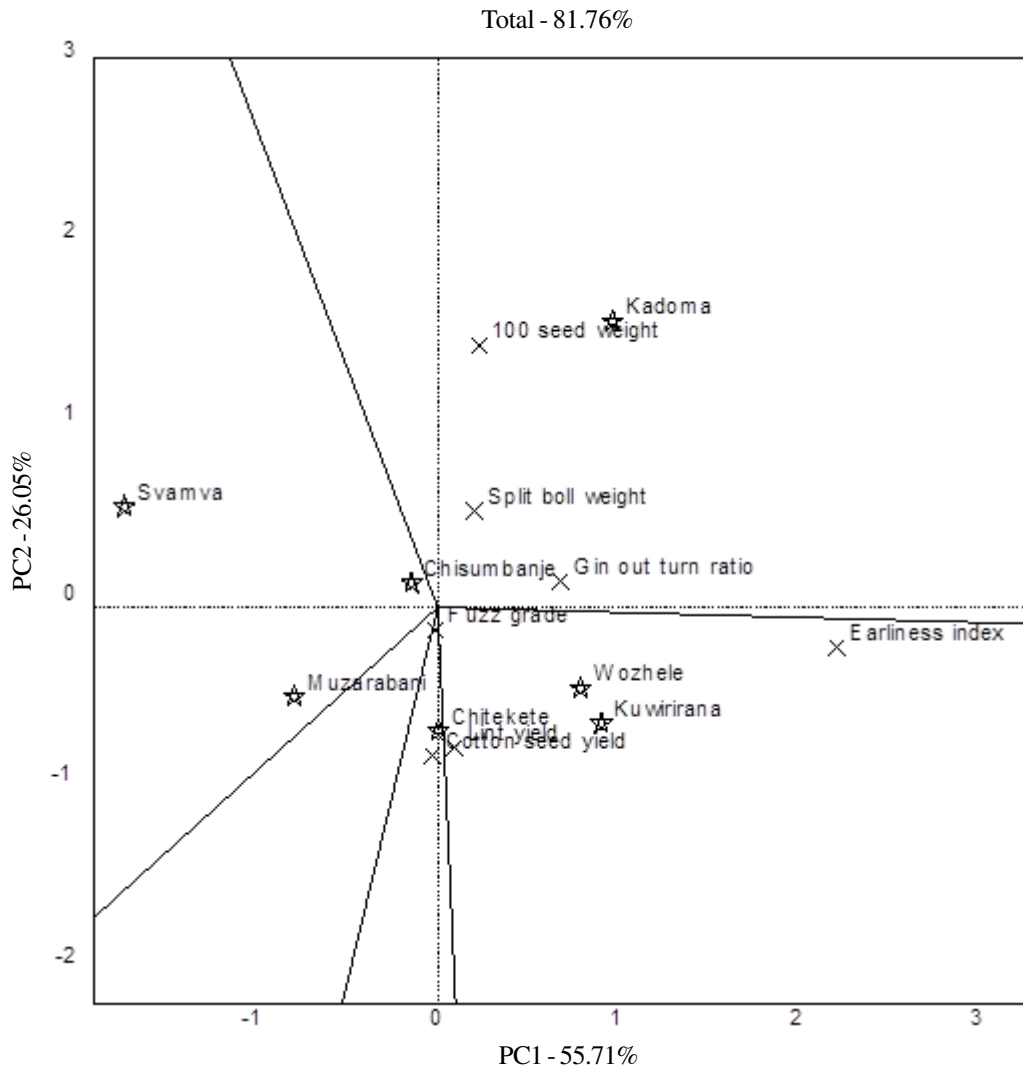


Figure 7. Location x trait biplot showing cotton locations with a given set of traits. Data were standardised with the within-trait standard deviation. The biplot was produced based on trait focused SVP, no scaling, no transformation and the data were trait centered.

observed (Figs. 1 and 2) posed challenges in recommending varieties for wider production. Crossover type GE interaction has also been reported to cause difficulties in crop breeding (Baker, 1988; Bernardo, 2002; Yan and Kang, 2002). The high ranking genotypes in most environments were 644-98-11, CRI-MS-2, SZ-9314 and CRI-MS-1. However, the which-won-where pattern based on seed cotton and lint yield was non-repeatable in individual years,

suggesting that the locations studied belong to a single and complex mega-environment (Yan and Kang, 2002; Yan and Tinker, 2005; 2006). In line with the existence of a complex environment, the sites studied showed some variation in rainfall, soil types, altitude and management conditions (Table 1). Where, a single mega-environment exists, a single breeding programme focusing the entire mega-environment is therefore recommended.

Furthermore, when a single mega-environment exists, Yan and Kang (2002) and Gauch (2013) highlighted that variety evaluation should be based on mean yield and stability performance.

In cotton, farmers are interested in the total cotton yield; while processors are much more interested in the lint yield (Killi *et al.*, 2005). Therefore, variety evaluation for cotton should consider these two traits with equal importance; while considering other related traits such as fiber quality (Campbell and Jones, 2005). In this regard, use of a selection index is helpful in order to balance the genetic gains obtained for each trait. Of interest, in this study the order of winning genotypes based on mean yield and stability for both seed cotton yield and lint yield did not change. This order remained as CRI-MS-2 > SZ-95-23 > SZ-9314 and was supported by a large positive correlation ( $r = 0.95$ ;  $P < 0.001$ ) that existed between these two traits. The best genotype was CRI-MS-2, a commercial variety (Fig. 4); although the across environment mean yield of this genotype was not significantly different from genotype SZ-95-23. Therefore, based on mean yield and stability, the next candidate genotype that can be recommended for release is SZ-95-23 (Fig. 4). Genotype SZ-9314 is the third best and could be considered for release based on other quality traits. The genotype x trait analysis shows the other added advantage of genotype SZ-95-23 and SZ-9314 to be of good fuzz grade compared to CRI-MS-2 and this warrants the value for cultivation and use of these promising genotypes.

**Test location evaluation.** Chisumbanje Research Station with long PC1 and small PC2 scores for seed cotton yield and lint yield (Figs. 5 and 6) was the ideal testing location for these two traits. The identification of Chisumbanje as the ideal testing location for lint yield and seed cotton yield was also confirmed by the location x trait biplot analysis that showed Chisumbanje to be closer to the biplot origin (Fig. 7). Furthermore, the large variance components of the location make the location

x trait biplot meaningful since the effect of the location variance component was large for most traits (Table 3). Therefore, location Chisumbanje can be used in early generation testing in reducing the number of genotypes for taking for multi-location trials. The most interesting phenomenon is that Chisumbanje is well suited for testing seed cotton yield (Fig. 5) and also lint yield (Fig. 6). Genotypes selected at this location will have high chance of being selected again, following multi-location testing. It is proposed that early generations observational trials be conducted at this location. Yan and Kang (2002) and Yan and Tinker (2005) demonstrated the use of GGE biplot in test location evaluation. The ideal testing site must be both discriminating (large PC1 scores) and representative of other test locations (small PC2 scores) (Yan and Kang, 2002; Yan and Tinker, 2005, 2006).

## CONCLUSION

The existence of huge GE and correlation among traits raises the need to include selection index in cotton breeding, selection and recommendation. Genotype SZ-95-23 can be recommended for commercialisation in Zimbabwe since it was stable across locations during the two years of testing. In addition, the genotype has high lint yield, seed cotton yield and desirable fuzz grade. One complex mega-environment was identified for cotton production in Zimbabwe and could be well served by one breeding programme. Chisumbanje Research Station represented the most discriminating and representative location for evaluating cotton trials in Zimbabwe.

## ACKNOWLEDGEMENT

We are grateful to the Cotton Research Station under the Crop Breeding Institute of the Ministry of Agriculture, Mechanisation and Irrigation Development, Zimbabwe for financial support.

## REFERENCES

- Annicchiarico, P., Pecetti, L., Boggini, G. and Doust, M.A. 2000. Repeatability of large-scale germplasm evaluation results in durum wheat. *Crop Science* 40:1810-1814.
- Baker, R. 1988. Tests for crossover genotype-environmental interactions. *Canadian Journal of Plant Science* 68:405-410.
- Bernardo, R. 2002. Breeding for quantitative traits in plants. Stemma Press Woodbury, Minnesota, USA. 369pp.
- Campbell, B. and Jones, M. 2005. Assessment of genotype x environment interactions for yield and fiber quality in cotton performance trials. *Euphytica* 144:69-78.
- Crossa, J., Cornelius, P.L. and Yan, W. 2002. Biplots of linear-bilinear models for studying crossover genotype x environment interaction. *Crop Science* 42: 619-633.
- Dias dos, S., Carlos, T. and Krzanowski, W.J. 2003. Model selection and cross validation in additive main effect and multiplicative interaction models. *Crop Science* 43:865-873.
- Finlay, K. and Wilkinson, G. 1963. The analysis of adaptation in a plant-breeding programme. *Crop and Pasture Science* 14: 742-754.
- Gasura, E., Setimela, P., Edema, R., Gibson, P.T., Okori, P. and Tarekegne, A. 2013. Exploiting grain-filling rate and effective grain-filling duration to improve grain yield of early-maturing maize. *Crop Science* 53: 2295-2303.
- Gasura, E., Setimela, P.S. and Souta, C.M. 2015. Evaluation of the performance of sorghum genotypes using GGE biplot. *Canadian Journal of Plant Science* 95: 1205-1214.
- Gasura, E., Setimela, P.S., Tarekegne, A., Icishahayo, D., Edema, R., Gibson, P.T. and Okori, P. 2014. Variability of grain-filling traits in early maturing CIMMYT tropical maize inbred lines. *Crop Science* 54:530-536.
- Gauch, H.G. 1993. Prediction, parsimony and noise. *American Scientist* 468-478.
- Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Science* 46: 1488-1500.
- Gauch, H.G. 2013. A simple protocol for AMMI analysis of yield trials. *Crop Science* 53:1860-1869.
- Gauch, H.G., Piepho, H.P. and Annicchiarico, P. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. *Crop Science* 48: 866-889.
- GenStat, 2014. GenStat for Windows 17<sup>th</sup> Edition. VSN International Ltd, Hemel Hempstead, United Kingdom.
- Gollob, H.F. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika* 33:73-115.
- Kamutando, C.N., Muungani, D., Masvodza, D.R. and Gasura, E. 2013. Exploiting genotype x environment interaction in maize breeding in Zimbabwe. *African Journal of Agricultural Research* 8:4058-4066.
- Karademir, C., Karademir, E. and Gencer, O. 2011. Yield and fiber quality of *F. Bulgarian Journal of Agricultural Science* 17:795-805.
- Killi, F., Efe, L. and Mustafayev, S. 2005. Genetic and environmental variability in yield, yield components and lint quality traits of cotton. *International Journal of Agricultural Biotechnology* 7:1007-1010.
- Laffont, J.-L., Hanafi, M. and Wright, K. 2007. Numerical and graphical measures to facilitate the interpretation of GGE biplots. *Crop Science* 47:990-996.
- Maleia, M.P., Vidigal Filho, P.S., Kvitschal, M.V. and Vidigal, M. 2010. Stability and adaptability of commercial cotton cultivars (*Gossypium hirsutum* L. race latifolium H.) in Mozambique. *African Journal of Agricultural Research* 5:539-550.



- Mandel, J. and Gauch, H.G. 1993. Non-Additivity and Biplots. *Biometrics* 1:952-954.
- Mapuranga, R., Chapepa, B. and Mudada, N. 2015. Strategies for integrated management of cotton bollworm complex in Zimbabwe: A review. *International Journal of Agronomy and Agricultural Research* 7(1): 23-35.
- McIntosh, M. 1983. Analysis of combined experiments. *Agronomy Journal* 75:153-155.
- Meredith Jr, W., Boykin, D., Bourland, F., Caldwell, W., Campbell, B., Gannaway, J., Glass, K., Phillips-Jones, A., May, L. and Smith, C. 2012. Genotype x environment interactions over seven years for yield, yield components, fiber quality, and gossypol traits in the regional high quality tests. *Journal of Cotton Science* 16:160-169.
- Moore, K.J. and Dixon, P.M. 2015. Analysis of combined experiments revisited. *Agronomy Journal* 107:763-771.
- Nyamapfene, K.W. 1991. The soils of Zimbabwe. Second Edition. Nehanda Publishers, Harare, Zimbabwe. 179pp.
- Percy, R.G. 2003. Comparison of bulk F<sub>2</sub> performance testing and pedigree selection in thirty Pima cotton populations. *Journal of Cotton Science* 7:170-178.
- Rukuni, M. 2006. Revisiting Zimbabwe's agricultural revolution. Zimbabwe's Agricultural Revolution Revisited, Harare, Zimbabwe: University of Zimbabwe Publications, Zimbabwe. 728pp.
- Setimela, P., Chitalu, Z., Jonazi, J., Mambo, A., Hodson, D. and Bänziger, M. 2005. Environmental classification of maize-testing sites in the SADC region and its implication for collaborative maize breeding strategies in the subcontinent. *Euphytica* 145: 123-132.
- Setimela, P., Crossa, J. and Bänziger, M. 2010. Targeting of early to intermediate maize hybrids for yield performance and yield stability using SREG model. *South African Journal of Plant and Soil* 27:125-132.
- Setimela, P., Vivek, B., Bänziger, M., Crossa, J. and Maiden, F. 2007. Evaluation of early to medium maturing open pollinated maize varieties in SADC region using GGE biplot based on the SREG model. *Field Crops Research* 103:161-169.
- Yan, W. 2001. GGEbiplot. A Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agronomy Journal* 93: 1111-1118.
- Yan, W. 2002. Singular-value partitioning in biplot analysis of multi-environment trial data. *Agronomy Journal* 94:990-996.
- Yan, W. and Hunt, L. 2001. Interpretation of genotype x environment interaction for winter wheat yield in Ontario. *Crop Science* 41:19-25.
- Yan, W., Hunt, L., Sheng, Q. and Szlavnic, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science* 40:597-605.
- Yan, W. and Kang, M.S. 2002. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC press, Boca Raton, London, United Kingdom.
- Yan, W., Kang, M.S., Ma, B., Woods, S. and Cornelius, P.L. 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science* 47:643-653.
- Yan, W. and Tinker, N.A. 2005. An integrated biplot analysis system for displaying, interpreting, and exploring genotype x environment interaction. *Crop Science* 45: 1004-1016.
- Yan, W. and Tinker, N.A. 2006. Biplot analysis of multi-environment trial data: Principles and applications. *Canadian Journal of Plant Science* 86:623-645.
- Yang, R.-C., Crossa, J., Cornelius, P.L. and Burgueño, J. 2009. Biplot analysis of genotype x environment interaction: Proceed with caution. *Crop Science* 49: 1564-1576.