GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data

**Sujay Rakshit · K. N. Ganapathy · S. S. Gomashe · A. Rathore · R. B. Ghorade ·
M. V. Nagesh Kumar · K. Ganesmurthy · S. K. Jain · M. Y. Kamtar ·
J. S. Sachan · S. S. Ambekar · B. R. Ranwa · D. G. Kanawade · M. Balusamy ·
D. Kadam · A. Sarkar · V. A. Tonapi · J. V. Patil**

Abstract Sorghum [*Sorghum bicolor* (L.) Moench] is a very important crop in the arid and semi-arid tropics of India and African subcontinent. In the process of release of new cultivars using multi-location data major emphasis is being given on the superiority of the new cultivars over the ruling cultivars, while very less importance is being given on the genotype × environment interaction (GEI). In the present study, performance of ten Indian hybrids over 12 locations across the rainy seasons of 2008 and 2009 was investigated using GGE biplot analysis. Location attributed higher proportion of the variation in the data (59.3–89.9%), while genotype contributed only 3.9–16.8% of total variation. Genotype × location interaction contributed 5.8–25.7% of total variation. We could identify superior hybrids for grain yield, fodder yield and for harvest index using biplot graphical approach effectively. Majority of the testing locations were highly correlated. 'Which-won-where' study partitioned the testing locations into three mega-environments: first with eight locations with SPH 1606/1609 as the winning genotypes; second mega-environment encompassed three locations with SPH 1596 as the winning genotype, and last mega-environment represented by only one location with SPH 1603 as

S. Rakshit (✉) · K. N. Ganapathy · S. S. Gomashe ·
V. A. Tonapi · J. V. Patil
Directorate of Sorghum Research, Rajendranagar,
Hyderabad, Andhra Pradesh 500 030, India
e-mail: rakshit@sorghum.res.in

A. Rathore
International Crops Research Institute for the Semi-Arid
Tropics, 502324 Andhra Pradesh, India

R. B. Ghorade
Dr. Panjabrao Deshmukh Krishi Vidyapeeth,
Akola, Maharashtra 444104, India

M. V. N. Kumar
ANGRAU Regional Agricultural Research Station,
Palem, Andhra Pradesh 509 215, India

K. Ganesmurthy
Department of Genetics and Plant Breeding,
Tamil Nadu Agricultural University,
Coimbatore, Tamil Nadu 641 003, India

S. K. Jain
Sorghum Research Station, Sardarkrushinagar Dantiwada
Agricultural University, Banaskantha, Deesa,
Gujarat 385 535, India

M. Y. Kamtar
Main Sorghum Research Station, University of
Agricultural Sciences, Dharwad, Karnataka 580005, India

J. S. Sachan
Crop Research Station (AICRP), CS Azad University of
Agriculture and Technology, Mauranipur, Jhansi,
Uttar Pradesh 284 204, India

S. S. Ambekar
Marathwada Agricultural University, Parbhani,
Maharashtra 431 402, India

B. R. Ranwa
Maharana Pratap University of Agriculture and
Technology, Udaipur, Rajasthan 313001, India

the winning genotype. This clearly indicates that though the testing is being conducted in many locations, similar conclusions can be drawn from one or two representatives of each mega-environment. We did not observe any correlation of these mega-environments to their geographical locations. Existence of extensive cross-over GEI clearly suggests that efforts are necessary to identify location-specific genotypes over multi-year and -location data for release of hybrids and varieties rather focusing on overall performance of the entries.

Keywords *Sorghum bicolor* · Multi-location data · GE interaction · GGE biplot · Stability · Mega-environment

Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is the fifth most important cereal crop across the world, which is mostly cultivated in the arid and semi-arid tropics (SAT) for its better adaptation to various stresses, including drought, heat, salinity and flooding (Harris et al. 2006; Ejeta and Knoll 2007). Globally sorghum is cultivated in 40.51 mha with maximum area in India (7.67 mha) followed by Sudan (5.61 mha), Nigeria (4.74 mha), Niger (3.32 mha) and other courtiers in SAT (http://faostat.fao.org). India ranks second to USA in terms of sorghum production. However, the productivity of sorghum in India and other SAT is much lower than that in USA. To ensure food and nutrition security to the large poor masses in this region productivity of sorghum needs to be augmented with breeding efforts. Multi-location testing of new

D. G. Kanawade
Agricultural Research Station (PDKV), Buldana, Maharashtra 443001, India

M. Balusamy
Agricultural Research Station-TNAU, Bhavanisagar, Tamil Nadu 638451, India

D. Kadam
Agricultural Research Station MPKV, 415, Karad, Maharashtra, India

A. Sarkar
National Academy of Agricultural Research Management, Rajendranagar, Hyderabad, Andhra Pradesh 500 030, India

cultivars plays an important role in breeding programme of any crop. However, in this process, major emphasis is given on the agronomic superiority of the new cultivars over the ruling cultivars in terms of grain and/or fodder yield. However, little or no emphasis is given on interaction of the cultivars with the target environments, which is mostly unpredictable. In this situation, a multi-location trial (MLT) can help to understand the performance of genotypes over diverse environments by studying the stability of the genotypes across environments (Scapim et al. 2000). Mostly the MLT data are rarely utilized to their full potential, though data are collected on many traits. In analyzing such data, mostly genotype evaluation is limited on genotype main effects (G), while genotype × environment interactions (GEI) are ignored as noise or confounding factor (Yan and Tinker 2006).

Since early twentieth century, GEI has been the prime focus area of research among plant biologists. Several authors have employed various statistical models to understand this complex GEI (reviewed by Yan and Kang 2003). Analysis of variance (ANOVA), principal component analysis (PCA), and linear regression (LR) are traditionally applied to treat complex MLT data (Zobel et al. 1988). ANOVA can only describe the main effects effectively as it is an additive model (Snedecor and Cochran 1980). On the contrary, PCA being a multiplicative model, does not describe the additive main effects (Zobel et al. 1988). LR models though combine both additive and multiplicative components, explaining both main effects and the interaction, the interaction gets confounded with the main effects compromising the power of general significance test (Wright 1971). Zobel et al. (1988) proposed additive main effects and multiplicative interaction (AMMI) model by integrating additive and multiplicative components into an integrated, powerful least squares analysis, which can explain GEI much effectively. Further advent and propagation of biplot methodology has greatly addressed the complex GEI in much simplistic graphical manner (Gabriel 1971). A biplot is a scatter plot that graphically summarizes two factors in such a way that relationships among the factors and underlying interactions between them can be visualized simultaneously. To understand GEI, two types of biplots, the AMMI biplot (Crossa 1990; Gauch 1992) and the GGE biplot (Yan et al. 2000; Yan and Kang 2003) are the most commonly used. In recent literature, utility of

AMMI analysis and GGE biplot analysis to visualize and interpret MLT data is being widely debated (Gauch 2006; Yan et al. 2007; Gauch et al. 2008; Yang et al. 2009). It may be kept in mind that the measured value of each cultivar in a test environment is a cumulative measure of genotype main effect (G), environment main effect (E) and the GE interaction (Yan and Kang 2003). For evaluation of cultivar, both G and GE must be considered simultaneously (Yan and Tinker 2006; Sabaghnia et al. 2008). The G + GE (GGE) biplot removes the E and integrates the G with GE interaction effect of a G × E dataset (Yan et al. 2000). Effectively it detects the GE interaction pattern in the data and can identify 'which-won-where' besides identifying different mega-environments (Yan et al. 2007). GGE biplot is almost close to the best AMMI models when different AMMI family models (AMMI0 to AMMIk) are compared (Dias et al. 2003; Ma et al. 2004). Moreover, AMMI is misleading in identifying 'which-won-where' (Yan et al. 2007). Thus GGE biplot is more logical and biological as compared to AMMI in explaining PC1 score, which represents genotypic effect rather than additive main effect (Yan 2002).

GGE biplot analysis has been carried out in understanding GEI in many crop species including soybean (Yan and Rajcan 2002), rice (Samonte et al. 2005), wheat (Kaya et al. 2006; Roozeboom et al. 2008), barley (Dehghani et al. 2006; Mohammadi et al. 2009), peanut (Putto et al. 2008), Lentil (Sabaghnia et al. 2008), oat (Yan et al. 2010), sorghum (Rao et al. 2011) and others. In spite of reports on utility of GE analysis in deciding superior genotypes and/or test environments in many crops, application of such techniques in grain sorghum MLTs is scanty. Recently, DeLacy et al. (2010a, b) have characterized the GEI across sorghum growing regions of India by analyzing ten years' multi-environment (MET) AIC-SIP data from 1986/87 to 1996/97. However, these studies have focused on post-rainy sorghum MLT data. Rao et al. (2011) have applied GGE biplot and AMMI analysis in identifying best performing sweet sorghum hybrids. In this the analysis is limited to sweet sorghum genotypes only. Till date to the best of our knowledge GEI in grain sorghum MLT data has not been analyzed effectively. In India new sorghum cultivars are tested across vast climatic and geographical conditions under the aegis of All India Coordinated Sorghum Improvement Project (AIC-SIP). These testing locations are distributed across

latitude and altitude with varied macro-climatic conditions, representing major sorghum growing situations in SAT. Hence, as a case study we analyzed the performance of ten rainy season grain sorghum hybrids across 12 locations for 2 years (rainy seasons of 2008 and 2009) using GGE biplot to demonstrate the utility of biplot graphical approach in analyzing and interpreting the complex GEI in MLT data.

Materials and methods

Plant materials and testing locations

Data used for this study was a sub-set of the AICSIP rainy season grain sorghum database, in which eight hybrids were evaluated along with two commercial checks, CSH 16 and CSH 23 in three replications across 12 locations (environments) during the rainy seasons of 2008 and 2009. Information on the hybrids used in the study are presented in Table 1. Detail features of the testing locations and dates of sowing are given in Fig. 1 and Table 2. The testing locations were distributed among seven states of the country, with four locations in Maharashtra, two each in Tamil Nadu and Karnataka, and one each in Andhra Pradesh, Uttar Pradesh, Rajasthan and Gujarat. During both the years, the crops were sown during June-July depending on the onset of monsoon at the particular location (Table 2). In each location, the experiment was conducted in randomized block design with six rows each of 6 m length with 45×15 cm^2 crop geometry. Crop management practices were standard across all locations. Data from internal four rows were considered for plot

Table 1 Information on the genotypes used in the study

Hybrid code	Pedigree/original name	Developing sector
SPH 1596	MDSH 297	Private
SPH 1603	GK 4032	Private
SPH 1604	GK 4033	Private
SPH 1606	KDSH 1179	Private
SPH 1609	HTJH 3201 (GTSH 06016)	Private
SPH 1611	Dhanarassi-909	Private
SPH 1615	KSH 6363	Private
SPH 1616	MLSH 60	Private
CSH 16	27A × C43	Public
CSH 23	MS 7A × RS 627	Public

Fig. 1 Geographical location of the testing environments
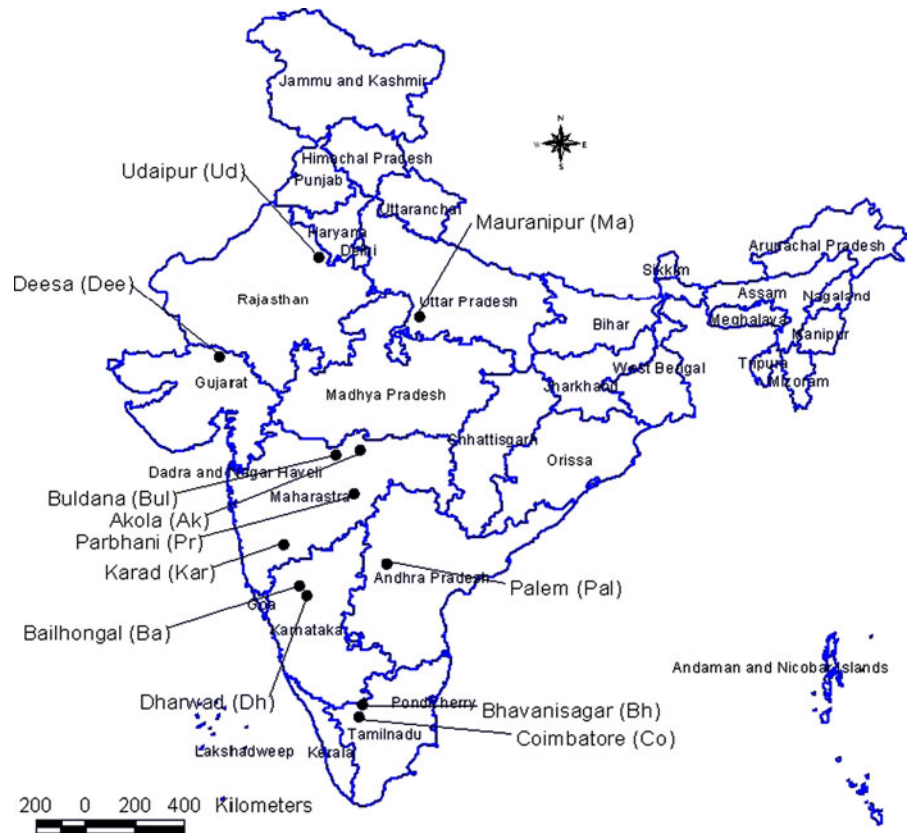
yield (grain and fodder) and days to 50% flowering calculation. Yield data were recorded at physiological maturity. Plot yield data were converted to kg ha^{-1} using the plot size as factor. Another statistic, harvest index (HI) was calculated as the ratio of grain mass to total above ground biomass and was used for analysis.

Data analysis

The statistical theory of GGE methodology has been explained in detail by Yan and Kang (2003). For environment centered matrix, the data were subjected to singular value decomposition (SVD) by estimating each element of the matrix using the following equation:

$$Y_{ij} = \mu + \beta_j + \sum_{k=1}^{k} \lambda_k \gamma_{ik} \delta_{jk}$$

where Y_{ij} is the performance of ith genotype in jth environment, μ is the general mean, β_i is the environment main effect, k is the number of principal components (PC) required for appropriate depiction of GGE, λ_k is the singular value of the kth PC (PC$_k$);

and γ_{ik} and δ_{jk} are the scores of ith genotype and jth environment, respectively for PC$_k$. The SVD was achieved through a scaling factor, f to derive alternate genotype ($n_{ik} = \lambda_k^f \lambda_{ik}$) and environment ($m_{jk} = \lambda_k^{f-1} \lambda_{ik}$) scores, respectively (Sabaghnia et al. 2008). Thus the G × E table was presented in a plot having m environment and n genotype points, respectively. We used the software GGEbiplot ver. 6.3 (Yan 2001) in the analysis. The MLT data was analyzed without scaling ('Scaling 0' option) to generate a tester centered (centering 2) GGE biplot as suggested by Yan and Tinker (2006). HI data was subjected to square root transformation before analysis. For genotype evaluation, genotype-focused singular value partitioning (SVP = 1) was used using the 'Mean versus stability' option of GGE biplot software, while for location evaluation, environment-focused singular value partitioning (SVP = 2) was employed (Yan 2001) using 'Relation among testers' option. 'Which-won-where' option was used to identify which genotype was the winner in a given set of environment and to identify mega-environments. Analysis of variance (ANOVA)

Table 2 Information on the trial environments

Site (code)	Elevation (msl)	Average rainfall (mm)	Date of sowing	
			2008	2009
Akola (Ak)	282	780	June 1	July 2
Bailhongal (Ba)	680	660	June 26	June 13
Bhavanisagar (Bh)	256	685	July 8	June 10
Buldana (Bul)	200	820	June 3	July 3
Coimbatore (Co)	412	730	June 16	June 10
Deesa (Dee)	136	350	July 7	July 14
Dharwad (Dh)	678	750	June 19	June 17
Karad (Kar)	597	703	June 30	July 6
Mauranipur (Ma)	271	954	July 8	July 23
Palem (Pal)	642	650	June 28	June 18
Parbhani (Pr)	357	750	June 25	June 29
Udaipur (Ud)	598	601	July 1	June 29

and heritability estimates generated in the biplot software were used for interpretation.

Results

Analysis of variance

Overall picture of relative magnitude of the G, L and GL is presented in Table 3 in the form of ANOVA. Analysis of variance clearly showed that genotype and location effects were significant for all the traits in both the seasons. GL effects were also significant for majority of cases except for HI in combined analysis. Table 3 also depicts the relative contribution of each source to the total variation (G + L + GL). It was observed that location was the most important source of variation for all the traits. In 2009, the influence of locations was relatively lower in comparison to 2008 for all the traits. For grain yield, location accounted for 82.8% of the variation in 2008 and 68.6% in 2009, while 76.3% of the variation was explained in combined analysis (Table 3). Contributions of genotype (G) were 3.9 and 5.8% in 2008 and 2009, respectively, while it explained 7.6% of the yield variation in combined analysis. Proportions of variation explained by GL were 13.3, 25.6 and 16.1% during 2008, 2009 and combined analysis, respectively. For fodder yield, location attributed to 81.5 and 68.3% of variation in 2008 and 2009, respectively, whereas genotype explained only 9.0 and 16.1% of variation during the seasons. Contribution of GL was near to

genotypic contribution for fodder yield. Among various traits studied, location contributed minimum of 59.3% for days to 50% flowering in 2009, where contribution of GL was high (25.7%). However, in 2008 location attributed more than 88% of the variation for flowering date. Similar trend was observed for HI as well.

Mean performance and stability of the genotypes across locations

Performance and stability of genotypes were visualized graphically through GGE biplot (Fig. 2). This can be evaluated by average environment coordination (AEC) method (Yan 2001, 2002). For this, environment centered (centering = 2) genotype-metric (SVP = 1) biplots for grain yield, fodder yield, days to 50% flowering and HI are presented in Fig. 2a–d, respectively. The first two PC explained 71.4% variation for grain yield, 87.0% for fodder yield, 78.4% of days for 50% flowering and 85.1% for HI. In Fig. 2a–d the line with single arrow head is the AEC abscissa. AEC abscissa passes through the biplot origin and marker for average environment and points towards higher mean values. The average environment has average PC1 and PC2 scores over all environments (Yan 2001). The perpendicular lines to the AEC passing through the biplot origin are referred to as AEC ordinate. These ordinates are depicted as double-arrowed lines in Fig. 2a–d. The greater the absolute length of the projection of a cultivar, the less stable it is. Furthermore, the average yield of genotypes is approximated by the projections of their markers to the AEC abscissa (Kaya et al. 2006). Accordingly, SPH 1606 and SPH 1609 were the best performing genotypes in terms of grain yield followed by SPH 1596 (Fig. 2a). On the other hand, SPH 1604 and CSH 23 were poorest yielders. It may be observed that SPH 1603, SPH 1604, SPH 1615 and SPH 1616 were least stable for grain yield with higher projection from the AEC abscissa. On the contrary, CSH 23 and CSH 16 were relatively stable though not higher grain yielders. SPH 1606 and SPH 1609 were highest grain yielder in both 2008 and 2009, however they were not stable (data not shown). For fodder yield, all the genotypes showed relative stability, with SPH 1603 and SPH 1604 as best fodder yielders (Fig. 2b). Similar trend was observed in individual years as well. CSH 23 was earliest among all the entries, while SPH 1609 took maximum duration to flower (Fig. 2c). Interestingly, though SPH 1606 was the best yielder along with SPH

Table 3 ANOVA and proportion of variation (G + L + GL) explained by genotype (G), location (L) and GL interaction for various traits

Trait/year		Source		
		G	L	GL
Grain yield 2008	MS	$2.5E \times 10^6**$	$4.3E \times 10^7**$	$7.7E \times 10^5**$
	Proportion of G + L + GL (%)	3.9	82.8	13.3
Grain yield 2009	MS	$3.6E \times 10^6**$	$3.5E \times 10^7**$	$1.4E \times 10^6**$
	Proportion of G + L + GL (%)	5.8	68.6	25.6
Grain yield combined	MS	$5.5E \times 10^6**$	$4.5E \times 10^7**$	$1.1E \times 10^6**$
	Proportion of G + L + GL (%)	7.6	76.3	16.1
Fodder yield 2008	MS	$8.5E \times 10^7**$	$6.3E \times 10^8**$	$8.2E \times 10^6**$
	Proportion of G + L + GL (%)	9.0	81.5	9.5
Fodder yield 2009	MS	$1.1E \times 10^8**$	$3.9E \times 10^8**$	$9.8E \times 10^6**$
	Proportion of G + L + GL (%)	16.1	68.3	15.6
Fodder yield combined	MS	$1.9E \times 10^8**$	$6.7E \times 10^8**$	$1.1E \times 10^7**$
	Proportion of G + L + GL (%)	16.8	72.8	10.4
Days to 50% flowering 2008	MS	126**	1554**	11*
	Proportion of G + L + GL (%)	5.9	88.3	5.9
Days to 50% flowering 2009	MS	148**	527**	25**
	Proportion of G + L + GL (%)	15.0	59.3	25.7
Days to 50% flowering combined	MS	266**	1352**	22**
	Proportion of G + L + GL (%)	12.3	76.6	11.1
Harvest index 2008	MS	0.02**	0.36**	0.00*
	Proportion of G + L + GL (%)	4.3	89.9	5.8
Harvest index 2009	MS	0.02**	0.09**	0.004**
	Proportion of G + L + GL (%)	12.3	63.0	24.7
Harvest index combined	MS	0.04**	0.29**	0.003
	Proportion of G + L + GL (%)	9.3	82.2	8.5

* $P < 0.05$, ** $P < 0.01$

1609, it was relatively early than others (Fig. 2c). This was the case in both the years. CSH 16, CSH 23, SPH 1606, SPH 1615 and SPH 1616 were relatively stable for days to 50% flowering. SPH 1596, though ranked third for grain yield it was best for HI, followed by CSH 16 and SPH 1606. The hybrids SPH 1604, SPH 1615 and SPH 1616 were not stable for HI.

Since all the entries are mainly for grain purpose, more focus was made on grain yield for further analysis. In addition Fig. 3 depicts the ranking of the genotypes for grain yield in terms of 'ideal genotype'. An 'ideal genotype' is high performer with high stability across environments (Yan and Tinker 2006). From our study it may be stated that SPH 1606 and 1609 followed by SPH 1596 were close to ideal genotypes. These genotypes had high grain yield performance among all genotypes (Table 4). To study specific adaptation of best grain

yielding genotype SPH 1606, test environments were ranked based on the relative grain yield of the genotype in given environment (Fig. 4). It may be observed that SPH 1606 had below average yield at Deesa, while near average yield at Mauranipur and Bhavanisagar. In the remaining locations it performed much above average with highest yielding at Buldhana. Same was the trend with SPH 1609. This is attributed to the fact that these two hybrids are highly correlated with near zero angle between their vectors.

Environment evaluation

Like in the previous section, the relationships among the test environments were studied by environment centered preserving of data (SPV = 2) without scaling. Combined analysis over 2 years for grain yield (Fig. 2e),
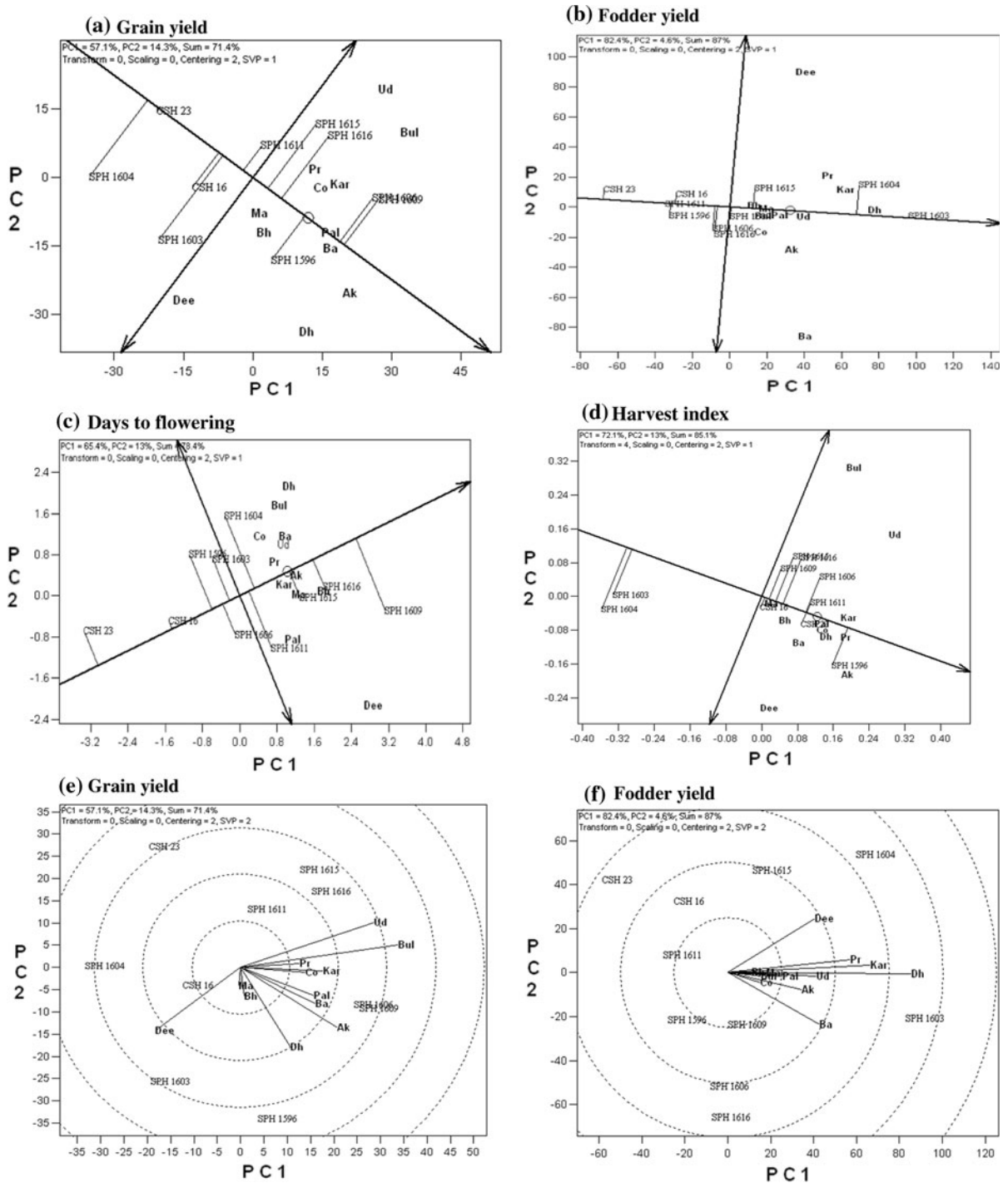
Fig. 2 GGE biplots of the combined analysis for grain yield, fodder yield, days to flowering and harvest index. **a–d** Mean versus stability of the genotypes. **e–h** Relation among the test locations. **i–l** Which-won-where analysis of the genotypes

fodder yield (Fig. 2f), days to 50% flowering (Fig. 2g) and HI (Fig. 2h) showed that the majority of the angles between their vectors are acute. Acute vector angles

are indicative of closer relationship among the environments (Yan and Tinker 2006). Thus majority of the locations were highly correlated with an exception
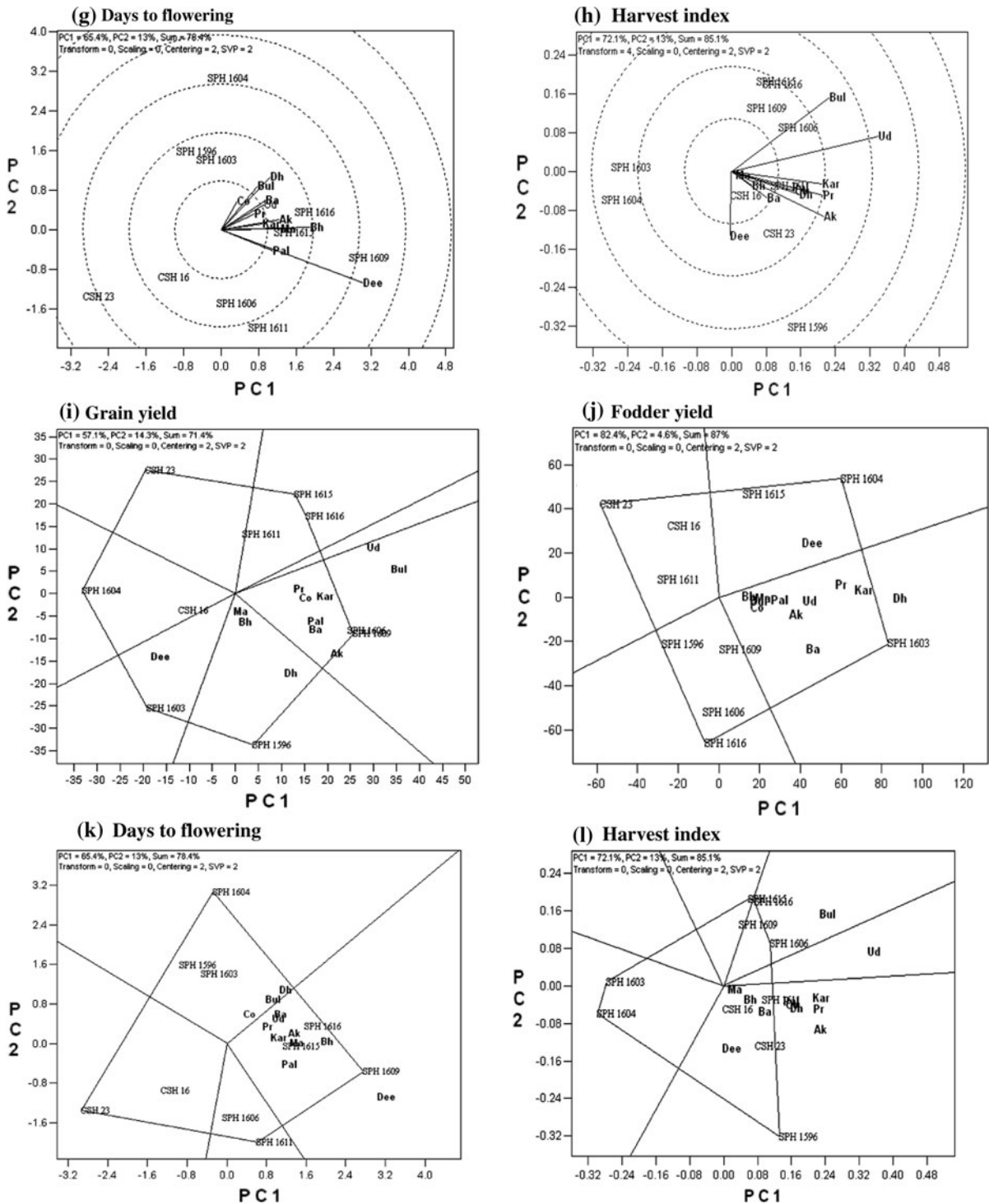
Fig. 2 continued

Fig. 3 Ranking of genotypes relative to an ideal genotype (the *small circle* on average environment coordinate, AEC)
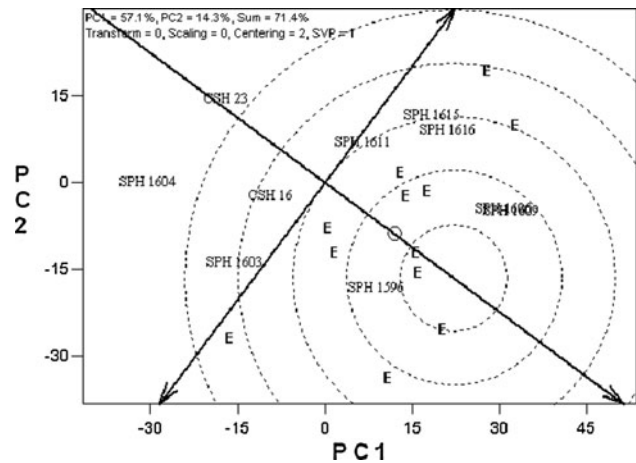

**Table 4** Year-wise character means of sorghum hybrids and locations under testing over 2 years

Location/ genotype	Grain yield			Fodder yield			Days to 50% flowering			Harvest index		
	2008	2009	Combined	2008	2009	Combined	2008	2009	Combined	2008	2009	Combined
Genotype[a]												
SPH 1596	4,238	4,133	4,186	11,635	12,342	11,989	68	65	67	0.29	0.26	0.27
SPH 1603	3,815	3,753	3,784	15,326	17,103	16,214	70	65	68	0.22	0.19	0.20
SPH 1604	3,533	3,442	3,487	14,722	16,471	15,597	70	66	69	0.22	0.18	0.20
SPH 1606	4,369	4,154	4,261	12,420	13,725	13,073	69	65	68	0.28	0.24	0.26
SPH 1609	4,185	4,226	4,205	12,592	13,834	13,213	73	69	71	0.27	0.23	0.25
SPH 1611	3,972	3,819	3,895	11,030	13,076	12,053	70	66	68	0.29	0.23	0.26
SPH 1615	4,134	3,931	4,033	13,020	13,752	13,386	71	67	70	0.27	0.23	0.25
SPH 1616	3,906	4,122	4,014	11,862	14,013	12,937	71	68	70	0.27	0.23	0.25
CSH 16	3,925	3,513	3,719	12,010	12,501	12,255	68	64	66	0.27	0.22	0.25
CSH 23	3,661	3,392	3,526	10,404	11,477	10,941	66	61	64	0.29	0.23	0.26
Location[b]												
AK	4,751	3,718	4,234	11,742	9,640	10,691	70	72	71	0.29	0.28	0.29
BA	5,061	2,572	3,816	7,508	14,630	11,069	77	NA	77	0.40	0.16	0.28
BH	5,821	3,119	4,470	10,490	10,357	10,423	65	63	64	0.36	0.23	0.29
BUL	3,934	4,720	4,327	11,036	10,009	10,523	74	62	68	0.26	0.32	0.29
CO	3,504	5,050	4,277	6,222	12,588	9,405	61	58	59	0.36	0.29	0.33
DEE	2,305	2,857	2,581	18,327	14,982	16,655	76	68	72	0.12	0.16	0.14
DH	4,367	5,473	4,920	18,270	18,345	18,308	70	67	68	0.20	0.23	0.22
KAR	5,349	4,101	4,725	13,453	18,514	15,983	64	70	67	0.29	0.19	0.24
MA	2,335	3,991	3,163	18,123	19,174	18,649	85	66	75	0.12	0.17	0.14
PAL	3,647	3,614	3,630	11,862	13,739	12,800	67	65	66	0.25	0.21	0.23
PR	2,360	1,999	2,180	16,852	9,110	12,981	65	69	67	0.13	0.19	0.16
UD	4,252	4,967	4,610	6,140	14,866	10,503	62	62	62	0.42	0.25	0.34
Grand mean	3,974	3,848	3,911	12,502	13,829	13,166	70	66	68	0.27	0.22	0.24
h_{bs}^2	0.69	0.60	0.81	0.90	0.91	0.94	0.91	0.82	0.92	0.88	0.82	0.92

NA not available

[a] Genotype means are based on 12 location data over 2 years

[b] Location means are based on ten genotype data over 2 years
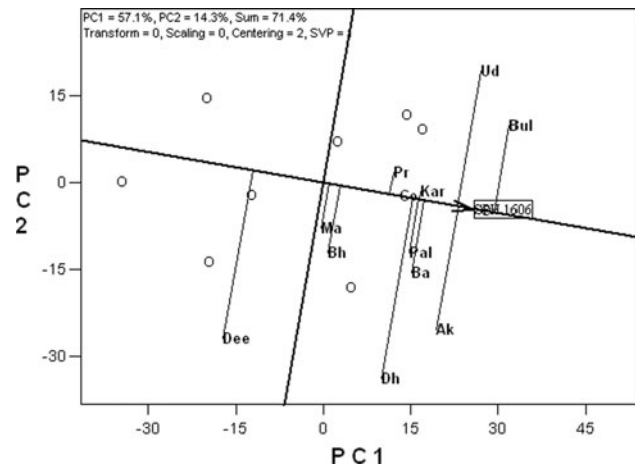
between Udaipur and Deesa for grain yield (Fig. 2e) or between Buldhana and Deesa for HI (Fig. 2h). The situation was much more apparent for fodder yield (Fig. 2f) and days to 50% flowering (Fig. 2g). Deesa consistently showed inverse relationship (negative correlation) to that of Udaipur or Buldhana for grain yield and HI as their vectors showed obtuse angles. Individual year data also supported the observation with some additional differences between the locations. For instance, in 2008 Akola showed opposite relation to Udaipur or Mauranipur with Dharwad and Bailhongal, while in 2009 Palem with Deesa and Parbhani, or Dharwad with Parbhani or Karad (data not shown). However, these relationships were not consistent over years, except that of Deesa with Udaipur and Buldhana as mentioned earlier. Mauranipur and Bhavanisagar were not correlated with Buldhana, Parbhani, Karad and Coimbatore with near right angles between them (Fig. 2e). Distance between two environments measures their ability in discriminating the genotypes. Thus 12 locations could be divided into four groups for grain yield; one with Udaipur, Buldhana, Parbhani, Karad and Coimbatore, while, second with Palem, Bailhongal and Akola. Third group was represented by Dharwad, Mauranipur and Bhavanisagar, while fourth was represented solely by Deesa.

Representativeness of the test environments is presented in Fig. 5 with projection of the environments to the Average environment axis (AEA). In the figure, the 'average environment' is represented by a small circle on the AEA. Environments with smaller angles with the AEA are most representative of the average test environments. Thus Palem was closest to the average environment followed by Bailhongal, Akola

and others. While ranking the genotypes in the near average environment, Palem showed that CSH 16, CSH 23, SPH 1603 and SPH 1604 had lower than average grain yield, while SPH 1611 and SPH 1615 performed near average yield. SPH 1606 and SPH 1609 yielded maximum at this location (Fig. 6).

Which-won-where and mega-environment identification

Which-won-where graph is constructed first by joining the farthest genotypes forming a polygon. Subsequently perpendicular lines are drawn from the origin of the biplot to each side of the polygon, separating the biplot into several sectors with one genotype at the vertex of the polygon. These lines are referred to as equality lines (Yan 2001). Genotypes at the vertices of the polygon are either the best or poorest in one or more environments. The genotype at the vertex of the polygon performs best in the environment falling within the sectors (Yan 2002; Yan and Tinker 2006). Which-won-where biplots for grain yield, fodder yield, days to 50% flowering and HI over 2 years are presented in Fig. 2i–l, respectively. The biplots indicated existence of crossover GE and existence of mega-environments, particularly for grain yield. Out of the four which-won-where biplots, it may be observed that grain yield biplot (Fig. 2i) is the most informative, as it could discriminate environments more effectively and the polygon (hexagon) is well distributed. The polygons for fodder yield (Fig. 2j) and days to flowering (Fig. 2k) had fewer vertices and the locations were not well separated. HI also could not segregate the locations much effectively (Fig. 2l). Thus being less informative they were not considered further.

Fig. 5 Ranking of environments based on discriminating ability and representativeness
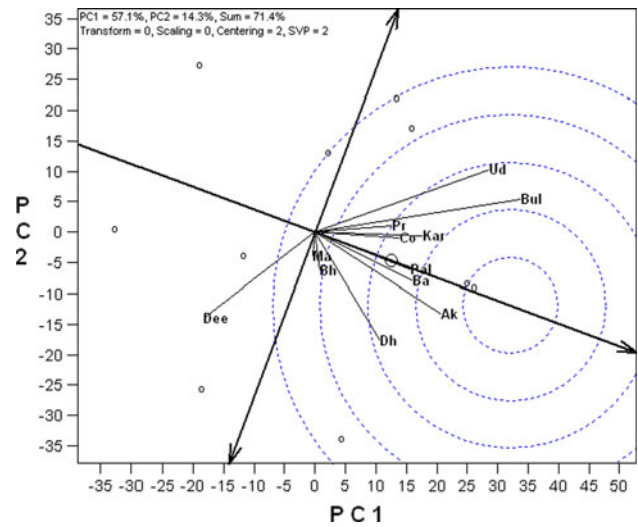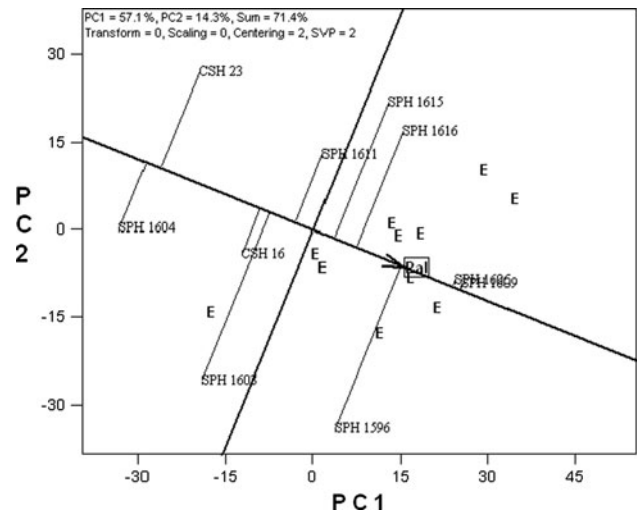

Fig. 6 Ranking of genotypes based on their performance in near ideal location, Palem

For grain yield, the hexagon has six genotypes, viz., SPH 1606/1609, SPH 1615, CSH 23, SPH 1604, SPH 1604 and SPH 1596 at the vertices (Fig. 2i). The hybrids SPH 1606/1609 performed best in Buldhana, Udaipur, Parbhani, Karad, Coimbatore, Palem, Bailhongal and Akola, while SPH 1595 being the best in Dharwad, Mauranipur and Bhavanisagar. SPH 1603 performed best in Deesa. The equality lines divided the biplot into six sectors effectively, of which three retained all the 12 locations. Thus the testing locations may be partitioned into three mega-environments: one with Udaipur, Buldhana, Parbhani, Karad, Coimbatore, Palem, Bailhongal and Akola with SPH 1606/1609 as the winning genotypes. Second mega-environment encompassed Dharwad, Mauranipur and Bhavanisagar with SPH

1596 as the winning genotype, while last megaenvironment was represented by Deesa with SPH 1603 as the winning genotype. We did not observe any correlation between the locations in one megaenvironment in terms of its geographical location, rainfall pattern and altitude (Table 2; Figs. 1, 2i).

Discussion

India is a vast country with diverse agro-climatic conditions. The growing conditions of the multi-location testing sites of AICSIP represent diverse sorghum production ecosystems across the world in terms of their latitude, altitude and macro-climatic conditions.

GGE biplot is a very potential tool to analyze MET data to interpret complex GEI interaction (Yan 2001; Yan and Tinker 2006). It can effectively detect the interaction pattern graphically besides identifying 'which-won-where' and delineation of mega-environments among the testing locations (Yan et al. 2007). However, this potential tool has not yet been employed to analyze the multi-location data of grain sorghum trials. We have studied the GEI among ten rainy season grain sorghum hybrids across 12 locations over 2 years using GGE biplot analysis. In our study, environment or location contributed 59.3–89.9% of the variation in the data, while contribution of genotype and their interaction with location was less (Table 3). Gauch and Zobel (1997) reported that normally in MET data, environment accounts for about 80% of the total variation. In bread wheat MET data, Kaya et al. (2006) reported as high as 81% of variation being explained by environment. Similar trend was reported by Dehghani et al. (2006) for barley yield trials in Iran. However, Sabaghnia et al. (2008) reported little lesser (52.1–68.8%) contribution of location to the total variation in lentil yield trial data from Iran. Putto et al. (2008) reported 50–80% of the total variation attributed due to location, while genotype main effect contributed 15–46% of the total variation. They observed only 4–5% contribution of GL interaction to the total variation in analyzing peanut yield of 17 diverse lines over 30 years, covering 130 locations. In our study, GL explained higher proportion of the variation than G alone. Higher proportion of GL as compared to G is indicative for possible existence of different mega-environments in testing locations (Yan and Hunt 2002; Mohammadi et al. 2009). This may not be true only under Indian situation but in other sorghum growing regions as well. Thus, the sorghum breeders need to take into consideration this point while breeding in their respective situations.

In GGE biplot analysis, the complex GEI are simplified in different PC and the data are presented graphically against various PC (Yan and Tinker 2006). If the first two PC explain more than 60% of the (G + GL) variability in the data, and the combined (G + GL) effect account for more than 10% of the total variability, then the biplot adequately approximates the variability in G × E data (Yang et al. 2009; Yan et al. 2010). In our study the first two PC explained more than 70% of the variability for all the four traits studied. In addition Table 3 indicates that G and GL together accounted for more than 10% of total variability. Thus the biplots may safely be interpreted as effective graphical representation of the variability in the MLT data. The graphical presentation of PC1 and PC2 (Fig. 2) has clearly brought out the complexity in the data set. The AEC ordinates point greater GE interaction effect (poor stability) in either direction (Yan and Tinker 2006), while the projections of markers of a genotype to the AEC abscissa approximates the average yield (Kaya et al. 2006). Thus it is evident from Fig. 1a that the highest grain yielders (SPH 1606 and SPH 1609) were not stable, while the most stable one, CSH 23 was among poorest yielders. However, this was not the trend for fodder yield. Wide variability in terms of stability was recorded for flowering date. Thus it was observed that a genotype showing stability for one trait may not necessarily be stable for other traits as well. This may be explained by the fact that each trait is governed by different set of genes and influence of environment on the cumulative expression of different set of genes will vary considerably, which is observed in variation in stability of genotypes for grain yield (Fig. 1a), fodder yield (Fig. 1b), flowering time (Fig. 1c) and HI (Fig 1d). Since HI is a derived factor of grain and fodder yield, it will have some resemblances with either of the traits individually. According to Lin and Binns (1988), soil and weather are the two main elements of an environment or location influencing the performance of a genotype. Out of these, soil element is generally persistent and may be regarded as fixed. On the other hand weather element has a predictable component represented by the general climatic zone, and unpredictable component contributed by year-to-year variation. Similarly, the GE interaction may also be subdivided (Allard and Bradshaw 1964). Lin and Binns (1988) extended this idea into estimating cultivar × predictable variation by averaging the cultivar × location mean over years, and cultivar × unpredictable variation by taking years within location. Thus Lin and Butler (1988) estimated fixed components by averaging a set of cultivar × location means over years with the assumption that GEI structure over years may be improved substantially if locations are grouped based on fixed component. So use of GGE biplot is justifiable since cultivar × predictable variation is controllable (Dehghani et al. 2006). Following similar strategies, several authors have identified high performing and stable genotypes in different crop species including barley (Dehghani et al. 2006), wheat

(Kaya et al. 2006), lentil (Sabaghnia et al. 2008), rapeseed (Dehghani et al. 2008) and others.

One advantage of graphical presentation of GEI is that genotypes closest to ideal genotype can be identified conveniently. Similar is the case with ideal environment as well. Ideal genotype (higher yielding and greater stability) is defined by having greatest vector length of highest yielding genotype with zero GEI as located at the center of the concentric circles in Fig. 3. Genotypes located closer to the 'ideal geno-type' are more desirable than others. Thus SPH 1606 and 1609 were closest to ideal genotype followed by SPH 1596 (Fig. 3). This would be difficult to conceive from mean table alone (Table 4). The (most ideal genotype), SPH 1606, performed best at Buldana, while near average yielded at Mauranipur and Bha-vanisagar, and lower than average yield at Deesa (Fig. 4). The above result suggests high crossover GE interaction, i.e. order of genotypes based on their performance varied depending on the testing environ-ment. Similar observation was made by other authors in different crops (Dehghani et al. 2006; Kaya et al. 2006; Sabaghnia et al. 2008; Dehghani et al. 2008). Saeed and Francis (1984) reported significant effect of cropping season rainfall and temperature on grain yield, contributing to the GEI. Dehghani et al. (2006) also suggested pre-seasonal and cropping season rainfall, temperature regime and relative humidity to contribute to GEI sum of squares. Soil types, light intensity etc. also influence GEI.

Using biplots, relationship between the testing environments can be understood easily considering the angle between their vectors. Vector of an environ-ment is the line connecting its marker to the origin of the biplot. Cosine of the angle between two vectors is indicative of their correlation (Yan and Tinker 2006). Thus our study clearly indicated that except Deesa with Udaipur or Buldana all testing locations were closely related (Fig. 2e–h) and most locations were close to the average environment, i.e. Palem (Fig. 5). Using graph-ical presentation, discrimination ability and represen-tativeness of the test environments can be detected conveniently. Projections of the environments with respect to the concentric circles are indicative of their discrimination ability (Yan 2001). Thus, Udaipur, Buldhana, Akola, Dharwad and Deesa with higher vector lengths were more discriminating than Maura-nipur and Bhavanisagar. Thus, near average locations, like Palem, Bailhongal and Akola are most representative

location and good test environments for selecting generally adapted genotypes. On the other hand, Udaipur, Buldhana and Deesa, being discriminating and non-representative are useful for selecting specif-ically adapted genotypes. Here comes the advantage of such graphical representation, where generally adapted environment and specific environment can be identi-fied conveniently. Closer relationships between the test environments indicated that same information could be obtained from fewer environments. Thus similar environments may be removed in future multi-location testing of sorghum hybrids. This point assumes much importance in order to optimally allocate the scarce resources while allocation MLTs.

Presence of wide obtuse angles between environ-ment vectors (Fig. 5), which indicates strong negative correlations among the test environments suggests existence of strong crossover GE across some loca-tions for grain yield (Yan and Tinker 2006). This indicated that genotypes performing better in one environment would be performing poor in another environment. At the same time, closer relationships among other locations are indicative of non-existence of crossover GE, suggesting ranking of genotype does not change from location to location. Mixture of crossover and non-crossover types of GEI in MET data is of very common occurrence (Kaya et al. 2006; Fan et al. 2007; Sabaghnia et al. 2008; Rao et al. 2011). This may be possible because some genotypes are highly responsive to change in the growing environ-ment, while others may be stable as response to environment is purely a combined properties of their gene combinations. 'Which-won-where' is the most attractive feature of GGE biplot, which graphically addresses crossover GE, mega-environment differen-tiation, specific adaptation etc. (Gauch and Zobel 1997; Yan et al. 2000; Yan and Tinker 2006; Putto et al. 2008, Rao et al. 2011). Such biplot is the succinct summary of the GEI pattern of a MET data. Based on this analysis, the testing locations were partitioned into three mega-environments. ME 1 was represented by Udaipur, Buldhana, Parbhani, Karad, Coimbatore, Palem, Bailhongal and Akola with SPH 1606/1609 as the winning genotypes. ME 2 consisted of Dharwad, Mauranipur and Bhavanisagar with SPH 1596 as the winning genotype, and ME 3 was represented by Deesa along with SPH 1603 as the winning genotype. This clearly suggested that though the testing is being conducted in many locations, almost similar

conclusion may be drawn from one or two representatives of each mega-environment. Thus the cost of testing may be reduced significantly. However, this mega-environment pattern needs to be verified through multi-year and -environment trials as conducted in wheat (Yan et al. 2000) and peanut (Casanoves et al. 2005; Putto et al. 2008). In the given situation smaller zonation of testing locations and focusing breeding efforts in a location-specific manner holds more importance, which is relevant to other crops as well.

Conclusion

The study has clearly and conveniently aided in identification of stable and superior hybrids in a graphical manner. It has also brought out that genotype showing stability for one trait not necessarily is stable for other. Thus breeders need to prioritize the trait they need to focus during breeding programme. Easy detection of mixed crossover effects using GGE biplot is added advantage of the procedure. Sorghum breeders across world need to consider this while breeding cultivars for varied geographical and agro-climatic regions. Location-specific adaptation of cultivars as detected in the present study clearly suggests that location-specific breeding needs more emphasis than focusing on wider adaptability. In this regard, participatory plant breeding assumes more importance than present research station oriented breeding programme. 'Which-won-where' analysis has demonstrated existence of mega-environments and many of the locations though geographically located far apart may generate similar information. Hence, to conduct the MET effectively with limited resources, discriminative locations encompassing representative locations may be included, rather than extending the trials extensively over related locations. Following similar analysis the sorghum breeders in other region need to identify mega-environments and then allocate testing sites accordingly. Another point that needs to be focused is that, in the existing procedure of varietal release, average of a given genotype over years and/or locations, and its superiority over the checks is only considered, while stability of genotypes is overlooked. Existence of extensive crossover GEI clearly suggests that the existing procedure does not realistically depict the actual situation. Rather, efforts are necessary to identify location-specific genotypes over multi-year and -location data to consider them for their release, since this will take into consideration the stability parameter of the genotypes. This is pertinent not only to sorghum alone but in other crops as well.

References

Allard RW, Bradshaw AD (1964) Implication of genotype-environmental interaction in applied plant breeding. Crop Sci 5:503–506

Casanoves F, Macchiavelli R, Balzarini M (2005) Error variation in multi-environment peanut trials. Crop Sci 45:1927–1933

Crossa J (1990) Statistical analyses of multilocation trials. Adv Agron 44:55–85

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

Dehghani H, Omidi H, Sabaghnia N (2008) Graphic analysis of trait relations of rapeseed using the biplot method. Agron J 100:1443–1449

DeLacy IH, Kaul S, Rana BS, Cooper M (2010a) Genotypic variation for grain and stover yield of dryland (rabi) sorghum in India 1. Magnitude of genotype × environment interactions. Field Crop Res 118:228–235

DeLacy IH, Kaul S, Rana BS, Cooper M (2010b) Genotypic variation for grain and stover yield of dryland (rabi) sorghum in India 2. A characterization of genotype × environment interactions. Field Crop Res 118:236–242

Dias CT, Dos S, Krzanowski WJ (2003) Model selection and cross validation in additive main effects and multiplicative interaction models. Crop Sci 43:865–873

Ejeta GJ, Knoll E (2007) Marker-assisted selection in sorghum pp. In: Vashney R, Tuberosa R (eds) Genomic assisted crop improvement: vol 2. Genomic applications in crops. Springer, Berlin, pp 187–205

Fan XM, Kang MS, Chen H, Zhang Y, Tan J, Xu C (2007) Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. Agron J 99:220–228

Gabriel KR (1971) The biplot graphic display of matrices with application to principal component analysis. Biometrika 58:453–467

Gauch HG (1992) AMMI analysis of yield trials. In: Kang MS, Gauch HG (eds) Genotype-by-environment interaction. CRC Press, Boca Raton, pp 1–40

Gauch HG (2006) Statistical analysis of yield trials by AMMI and GGE. Crop Sci 46:1488–1500

Gauch HG, Zobel RW (1997) Identifying mega-environment and targeting genotypes. Crop Sci 37:381–385

Gauch HG, Piepho HP, Annicchiarico P (2008) Statistical analysis of yield trials by AMMI and GGE: further considerations. Crop Sci 48:866–889

Harris K, Subodhi PK, Borrell A, Jordon D, Rosenow D, Nguyen H, Klein P, Klein R, Mullet J (2006) Sorghum staygreen QTL individually reduce post-flowering drought-induced leaf senescence. J Exp Bot 58:327–338

Kaya YM, Akcurra M, Taner S (2006) GGE-biplot analysis of multi-environment yield trials in bread wheat. Turk J Agric For 30:325–337

Lin CS, Binns MR (1988) A method of analyzing cultivar × location × year experiments: a new stability parameter. Theor Appl Genet 76:425–430

Lin CS, Butler G (1988) A data-based approach for selecting location for regional trial. Can J Plant Sci 68:651–659

Ma BL, Yan W, Dwyer LM, Fregeau-Reid J, Voldeng HD, Dion Y, Nass H (2004) Graphic analysis of genotype, environment, nitrogen fertilizer, and their interactions on spring wheat yield. Agron J 96:389–416

Mohammadi R, Aghaee M, Haghparast R, Pourdad SS, Rostaii M, Ansari Y, Abdolahi A, Amri A (2009) Association among non-parametric measures of phenotypic stability in four annual crops. Middle East Russ J Plant Sci Biotech 3(Special Issue I):20–24

Putto W, Patanothai A, Jogloy S, Hoogenboom (2008) Determination of mega-environments for peanut breeding using the CSM-CROPGRO-Peanut model. Crop Sci 48:973–982

Rao PS, Reddy PS, Ratore A, Reddy BVS, Panwar S (2011) Application GGE biplot and AMMI model to evaluate sweet sorghum (Sorghum bicolor) hybrids for genotype × environment interaction and seasonal adaptation. Indian J Agric Sci 81:438–444

Roozeboom KL, Schapugh T, Tuinstra MR, Vanderlip RL, Milliken GA (2008) Testing wheat in variable environments: genotype, environment, interaction effects, and grouping test locations. Crop Sci 48:317–330

Sabaghnia N, Dehghani H, Sabaghpour SH (2008) Graphic analysis of genotype by environment interaction for lentil yield in Iran. Agron J 100:760–764

Saeed M, Francis CA (1984) Association of weather variable with genotype × environment interactions in grain sorghum. Crop Sci 24:13–16

Samonte SOPB, Wilson LT, McClung AM, Medley JC (2005) Targeting Cultivars onto rice growing environments using AMMI and SREG GGE biplot analysis. Crop Sci 45:2414–2424

Scapim CA, Oliveira VR, Braccini AL, Cruz CD, Andrade CAB, Vidigal MCG (2000) Yield stability in maize (Zea mays L.) and correlations among the parameters of the Eberhart and Russell, Lin and Binns and Huehn models. Genet Mol Biol 23:387–393

Snedecor GW, Cochran WG (1980) Statistical methods, 7th edn. Iowa State University Press, Ames

Wright AJ (1971) The analysis and prediction of some two factor interactions in grass breeding. J Agric Sci Camb 76:301–306

Yan W (2001) GGEbiplot–a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. Agron J 93:1111–1118

Yan W (2002) Singular value partitioning for biplot analysis of multi-environment trial data. Agron J 4:990–996

Yan W, Hunt LA (2002) Biplot analysis of diallel data. Crop Sci 42:21–30

Yan W, Kang MS (2003) GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton

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

Yan W, Tinker NA (2006) Biplot analysis of multi-environment trial data: principles and applications. Can J Plant Sci 86:623–645

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

Yan W, Kang MS, Ma B-L, Woods S, Cornelius PL (2007) GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Sci 47:643–653

Yan W, Fregeau-Reid J, Pageau D, Martin R, Mitchell-Fetch J, Etieenne M, Rowsell J, Scott P, Price M, de Hann B, Cummiskey A, Lajeunesse J, Durand J, Sparry E (2010) Identifying essential test location for oat breeding in eastern Canada. Crop Sci 50:504–515

Yang RC, Crossa J, Cornelius PL, Burgueno J (2009) Biplot analysis of genotype × environment interaction: proceed with caution. Crop Sci 49:1564–1576

Zobel RW, Wright MJ, Gauch HG (1988) Statistical analysis of a yield trial. Agron J 80:388–393