REFEREED PAPER

INVESTIGATING SUGARCANE GENOTYPE X ENVIRONMENT INTERACTIONS UNDER RAINFED CONDITIONS IN SOUTH AFRICA USING VARIANCE COMPONENTS AND BIPLOT ANALYSIS

RAMBURAN S AND ZHOU M

South African Sugarcane Research Institute, Private Bag X02, Mount Edgecombe, 4300, South Africa
Sanesh.Ramburan@sugar.org.za *Marvellous.Zhou@sugar.org.za*

Abstract

The objective of this study was to evaluate the current regional sub-divisions, genotype stability, and magnitude of genotype x environment (GxE) interactions in the rainfed parts of the industry, to improve the efficiencies of selection and evaluation procedures. Cane yield (TCANE), estimated recoverable crystal (ERC) and tons ERC/ha (TERC) data of 15 genotypes from 153 environments (trial x ratoon combinations) harvested between 1999 and 2009 were analysed using variance components and genotype + genotype x environment (GGE) biplot analysis. Differences between trials accounted for the largest proportion of variation in TCANE (35%), ERC (44%), and TERC (47%), while genotype differences were smaller than the GxE interactions for TCANE and TERC, but not for ERC. Larger genotype x trial interactions compared to genotype x ratoon interactions suggested that more emphasis be placed on sampling more trial sites rather than testing more ratoons within trials. The GGE biplots for TERC revealed four mega-environments, each with environments from the coast, hinterland and midlands. Ratoons of the same trials and trials from the same location occasionally grouped into different mega-environments, suggesting large spatial and temporal (seasonal) variability in environmental conditions that can only be managed through extensive trial characterisation. Genotype N31 was identified as most stable and high yielding, suggesting the potential value of the variety as an alternative control to NCo376 for rainfed selection sites. These findings highlighted the need for increased GxE studies to enhance efficiencies of breeding and evaluation networks in the industry.

Keywords: biplot, genotype x environment, mega-environment, sugarcane, variance components

Introduction

Genotype x environment (GxE) interaction is a phenomenon that results in differential varietal responses in different environments. Such differential responses are considered a hindrance to breeding efforts, as superior varietal performance in one environment may not be observed in the target population of environments (TPE). To account for this, breeders conduct multi-environment trials (MET) to evaluate the performance of potential commercial varieties under a range of different conditions. Besides identifying superior varieties suited to particular environments, METs also provide information on the nature of the TPEs, and how they should be subdivided for further selection and evaluation. Such subdivisions are necessary for the

implementation of regional specific breeding strategies, which can ensure the greatest gains from selection, as environmental variance between selection sites is minimized. Subdivision of the TPE is also important for the development of recommendation domains. A mega-environment may be defined as a portion of a crop species' growing region with a homogeneous environment that causes some genotypes to perform similarly (Gauch and Zoebel, 1997), and is normally identified through analyses of MET data.

Investigations of mega-environments are a prerequisite for any meaningful variety evaluation and recommendation procedures (Yan and Hunt, 2001). This implies that variety improvement and evaluation networks established without preliminary mega-environment investigations could be at risk of reduced genetic gains from selection due to inaccurate regional subdivisions. The rainfed parts of the South African sugar industry are currently divided into the coastal (C), hinterland (H) and midlands (M) regions. Commercial varieties developed for these regions originate from five separate selection programmes (two each in the C and M regions, and one in the H region). The choice of representative farms for each programme was based on expert knowledge of geographic, soil and climatic variability between the different sugarcane producing regions and the representativeness of stations to general growing conditions within the respective regions (Nuss, 1998). The choice of farms proceeded without any formalized preliminary study of the nature of the TPE and their effects on variety performance. Since their establishment, the representativeness of the selection farms to the TPEs, and their effectiveness in this regard has not been evaluated.

In addition to selection trials, post-release variety evaluation trials are conducted throughout the rainfed regions of the industry, under a wider range of agro-climatic conditions, primarily for recommendation purposes. Trials within this network are conducted mainly on commercial farms with different combinations of soil types, harvest age, harvest season, and husbandry practices, and are more representative of industry conditions. More importantly, the extensive spatial distribution of the trial network suggests that it is a more representative sample of the TPE's compared to selection trials alone. Since the inception of this network, no attempts have been made to investigate the nature of the GxE interactions in the rainfed parts of the industry using contemporary GxE analysis techniques. Data from such networks may also be used to gain an understanding of the relative contributions of different components of variation on yield variability (Smith *et al*., 2005). For example, in the South African industry, information on the relative contributions of variety, site, ratoon, and their interactions have never been evaluated. Such information is essential to inform decisions on selection and evaluation, investigate contributions of repeatable and non-repeatable components of GxE interactions, and improve the efficiencies of trial networks (DeLacy *et al*., 2010).

Methods of analyzing MET data have included the conventional analysis of variance, linear regression (Finlay and Wilkinson, 1963), stability analysis (Shukla, 1972), additive main effects and mulitiplicative interaction (Gauch, 1992). The genotype + genotype x environment (GGE) biplot method developed by Yan *et al*. (2000) has recently gained popularity as a tool to investigate various aspects of GxE interactions. The method has been successfully applied to other crops such as wheat (Sharma *et al*., 2010), barley (Dehghani *et al*., 2006), and soybean (Yan and Rajcan, 2002). Studies with sugarcane in Florida, USA (Glaz and Kang, 2008) and Guatemala (Queme *et al*., 2010) have also been conducted; however, the methodology has not

been attempted in South Africa. Variance components analyses have also been used widely to evaluate the relative contributions of relevant terms to variability and to assess the precision of trial networks to optimise resources (Ceretta and van Eeuwijk, 2008). For example, De la Vega and Chapman (2010) used linear mixed models to estimate variance components attributed to trials, hybrids, replicates, and incomplete blocks, and tracked genetic progress for yield of commercial sunflower in Argentina.

Despite the availability of appropriate data, the many useful applications, and the available methods of analysis, no attempts have been made to improve sugarcane selection and evaluation procedures in South Africa using contemporary GxE analysis techniques. A systematic analysis of a series of METs has not been conducted for the rainfed production regions, and information on mega-environment constitution, industry-wide genotypic performance and stability, and components of variation are lacking. Recent concerns around the relevance of selection sites and differential performance of varieties in the TPE compared to performance during selection have prompted such an analysis. The objective of this study was to utilise post-release variety trial data to investigate the magnitude of GxE interactions, evaluate variance components, and identify mega-environments within the rainfed parts of the South African sugar industry.

Materials and Method

Trial datasets

The MET dataset used for this study was comprised of 43 trials conducted in 18 different locations, and harvested over one to six ratoons during the period 1999-2009. The majority of the trials (32) formed part of SASRI's post-release variety evaluation project; while other (11) late stage plant breeding selection trials were included as well. Most trials were established on commercial fields across the three regions and allowed to run for as many ratoons as the commercial plantings surrounding the trials. Plant breeding trials that were included in the dataset only ran for a maximum of three ratoons (plant plus two ratoons), and were conducted on SASRI selection stations. For the purpose of this study, an environment was defined as a trial x ratoon combination. As with most studies of this nature, the 'location' was a loose spatial reference identified by a town name, and trials at different sites in each location were numbered in the order of establishment. Table 1 describes the environments, which are abbreviated according to the region, location, site, and ratoon number, respectively. For example, CEM21 refers to a first ratoon crop, of the second trial conducted at Empangeni, in the coastal region. Variety (will be used interchangeably with the term genotype) numbers per trial varied from five to 12, and not all varieties were included in each trial. Consequently, only the 15 most commonly tested varieties were included in the analyses, with the majority of the trials containing the commercial control NCo376.

All trials were conducted as randomised complete block designs with four to six replicates. Trial plots consisted of five or six rows that were between 8 and 10 m long, spaced 1.0 to 1.2 m apart. Weed and fertilizer management proceeded as per commercial farm practice, while pests and diseases were not controlled in an attempt to evaluate variety responses. At harvest, three or four net rows were cut and bundled by hand and weighed using a hydraulic grab apparatus equipped with a load cell to determine cane yield in tons cane/ha (TCANE). A 12-stalk sucrose sample

was taken from each plot to determine estimated recoverable crystal per cent (ERC). For each trial plot, the tons ERC/ha (TERC) was calculated as the product of TCANE and ERC.

Table 1. Definitions of the environments making up the multi-environment dataset. Trials were conducted at various locations in the coastal (C), hinterland (H) and midlands (M) regions and were coded as environments defined by the respective site and ratoon number.

Ref	Region	Location	Trial (within location)	Ratoon	Environ code	Ref	Region	Location	Trial (within location)	Ratoon	Environ code
1	C	Amatikulu (AK)	1	0	CAK10	80	H	Doringkop (DK)	1	0	HDK10
2	C	Amatikulu (AK)	1	1	CAK11	81	H	Doringkop (DK)	1	1	HDK11
3	C	Amatikulu (AK)	2	0	CAK20	82	H	Doringkop (DK)	1	2	HDK12
4	C	Amatikulu (AK)	3	0	CAK30	83	H	Doringkop (DK)	1	3	HDK13
5	C	Amatikulu (AK)	3	1	CAK31	84	H	Doringkop (DK)	1	4	HDK14
6	C	Amatikulu (AK)	3	2	CAK32	85	H	Doringkop (DK)	2	1	HDK21
7	C	Amatikulu (AK)	3	3	CAK33	86	H	Doringkop (DK)	2	2	HDK22
8	C	Amatikulu (AK)	4	0	CAK40	87	H	Doringkop (DK)	2	3	HDK23
9	C	Amatikulu (AK)	4	1	CAK41	88	H	Doringkop (DK)	3	0	HDK30
10	C	Amatikulu (AK)	4	2	CAK42	89	H	Doringkop (DK)	3	1	HDK31
11	C	Amatikulu (AK)	4	3	CAK43	90	H	Doringkop (DK)	3	2	HDK32
12	C	Amatikulu (AK)	4	4	CAK44	91	H	Eshowe (ES)	1	0	HES10
13	C	Amatikulu (AK)	4	5	CAK45	92	H	Eshowe (ES)	1	1	HES11
14	C	Empangeni (EM)	1	0	CEM10	93	H	Eshowe (ES)	1	2	HES12
15	C	Empangeni (EM)	2	0	CEM20	94	H	Eshowe (ES)	1	3	HES13
16	C	Empangeni (EM)	2	1	CEM21	95	H	Kearsney (KS)	1	1	HKS11
17	C	Empangeni (EM)	2	2	CEM22	96	H	Kearsney (KS)	1	2	HKS12
18	C	Empangeni (EM)	2	3	CEM23	97	H	Kearsney (KS)	2	1	HKS21
19	C	Empangeni (EM)	2	4	CEM24	98	H	Kearsney (KS)	2	2	HKS22
20	C	Empangeni (EM)	2	5	CEM25	99	H	Kearsney (KS)	3	0	HKS30
21	C	Empangeni (EM)	3	0	CEM30	100	H	Kearsney (KS)	3	1	HKS31
22	C	Empangeni (EM)	3	1	CEM31	101	H	Kearsney (KS)	3	2	HKS32
23	C	Empangeni (EM)	3	2	CEM32	102	H	Kearsney (KS)	4	0	HKS40
24	C	Empangeni (EM)	3	3	CEM33	103	H	Kearsney (KS)	4	1	HKS41
25	C	Empangeni (EM)	3	4	CEM34	104	H	Kearsney (KS)	4	2	HKS42
26	C	Empangeni (EM)	3	5	CEM35	105	H	Kearsney (KS)	4	3	HKS43
27	C	Empangeni (EM)	3	6	CEM36	106	H	Kearsney (KS)	4	4	HKS44
28	C	Empangeni (EM)	4	0	CEM40	107	H	Kearsney (KS)	4	5	HKS45
29	C	Empangeni (EM)	4	1	CEM41	108	H	Kearsney (KS)	5	0	HKS50
30	C	Empangeni (EM)	4	2	CEM42	109	H	Kearsney (KS)	5	1	HKS51
31	C	Empangeni (EM)	4	3	CEM43	110	H	Kearsney (KS)	5	2	HKS52
32	C	Empangeni (EM)	4	4	CEM44	111	H	Kearsney (KS)	5	3	HKS53
33	C	Empangeni (EM)	4	5	CEM45	112	H	Kearsney (KS)	5	4	HKS54
34	C	Empangeni (EM)	4	6	CEM46	113	H	Melmoth (MM)	1	0	HMM10
35	C	Empangeni (EM)	5	0	CEM50	114	H	Melmoth (MM)	1	1	HMM11
36	C	Empangeni (EM)	6	0	CEM60	115	H	Melmoth (MM)	1	2	HMM12
37	C	Empangeni (EM)	6	1	CEM61	116	H	Paddock (PD)	1	0	HPD10
38	C	Empangeni (EM)	6	2	CEM62	117	H	Paddock (PD)	2	0	HPD20
39	C	Empangeni (EM)	7	0	CEM70	118	H	Paddock (PD)	2	2	HPD22
40	C	Empangeni (EM)	7	1	CEM71	119	H	Paddock (PD)	2	3	HPD23
41	C	Empangeni (EM)	7	2	CEM72	120	H	Paddock (PD)	2	4	HPD24
42	C	Hibberdene (HB)	1	0	CHB10	121	M	Bruyns Hill (BH)	1	0	MBH10
43	C	Hibberdene (HB)	1	1	CHB11	122	M	Bruyns Hill (BH)	1	1	MBH11
44	C	Hibberdene (HB)	1	2	CHB12	123	M	Bruyns Hill (BH)	1	2	MBH12

45	C	Hibberdene (HB)	1	3	CHB13	124	M	Bruyns Hill (BH)	2	0	MBH20
46	C	Scottburgh (SC)	1	0	CSC10	125	M	Bruyns Hill (BH)	2	1	MBH21
47	C	Scottburgh (SC)	1	1	CSC11	126	M	Bruyns Hill (BH)	2	2	MBH22
48	C	Scottburgh (SC)	1	2	CSC12	127	M	Bruyns Hill (BH)	3	0	MBH30
49	C	Scottburgh (SC)	1	3	CSC13	128	M	Bruyns Hill (BH)	3	1	MBH31
50	C	Scottburgh (SC)	1	4	CSC14	129	M	Bruyns Hill (BH)	3	2	MBH32
51	C	Scottburgh (SC)	1	5	CSC15	130	M	Bruyns Hill (BH)	4	0	MBH40
52	C	Scottburgh (SC)	1	6	CSC16	131	M	Bruyns Hill (BH)	4	1	MBH41
53	C	Scottburgh (SC)	2	1	CSC21	132	M	Bruyns Hill (BH)	4	2	MBH42
54	C	Scottburgh (SC)	2	2	CSC22	133	M	Dalton (DL)	1	1	MDL11
55	C	Scottburgh (SC)	2	3	CSC23	134	M	Dalton (DL)	1	2	MDL12
56	C	Scottburgh (SC)	2	4	CSC24	135	M	Eston (ES)	1	0	MES10
57	C	Umfolozi (UF)	1	0	CUF10	136	M	Eston (ES)	1	1	MES11
58	C	Umfolozi (UF)	1	1	CUF11	137	M	Glenside (GS)	1	1	MGS11
59	C	Umfolozi (UF)	1	2	CUF12	138	M	Glenside (GS)	1	2	MGS12
60	C	Umfolozi (UF)	1	3	CUF13	139	M	Glenside (GS)	1	3	MGS13
61	C	Umfolozi (UF)	1	4	CUF14	140	M	Glenside (GS)	2	0	MGS20
62	C	Umfolozi (UF)	1	5	CUF15	141	M	Glenside (GS)	2	1	MGS21
63	C	Umfolozi (UF)	2	0	CUF20	142	M	Glenside (GS)	3	0	MGS30
64	C	Umfolozi (UF)	3	0	CUF30	143	M	Glenside (GS)	3	1	MGS31
65	C	Umfolozi (UF)	3	1	CUF31	144	M	Glenside (GS)	3	2	MGS32
66	C	Umfolozi (UF)	3	2	CUF32	145	M	Mid Illovo (MI)	1	0	MMI10
67	C	Umfolozi (UF)	3	3	CUF33	146	M	Mid Illovo (MI)	1	1	MMI11
68	C	Umfolozi (UF)	4	0	CUF40	147	M	Mid Illovo (MI)	1	3	MMI13
69	C	Umfolozi (UF)	4	1	CUF41	148	M	York (YK)	1	0	MYK10
70	C	Umfolozi (UF)	4	2	CUF42	149	M	York (YK)	1	1	MYK11
71	C	Umfolozi (UF)	4	3	CUF43	150	M	York (YK)	1	2	MYK12
72	C	Verulam (VM)	1	0	CVM10	151	M	York (YK)	2	0	MYK20
73	C	Verulam (VM)	1	1	CVM11	152	M	York (YK)	2	1	MYK21
74	C	Verulam (VM)	1	2	CVM12	153	M	York (YK)	2	2	MYK22
75	C	Zinkwazi (ZK)	1	0	CZK10						
76	C	Zinkwazi (ZK)	1	1	CZK11						
77	C	Zinkwazi (ZK)	1	2	CZK12						
78	C	Zinkwazi (ZK)	1	3	CZK13						
79	C	Zinkwazi (ZK)	1	4	CZK14						

Variance components analysis

To evaluate the variance contributions of different components, the following random effect model was fitted to the plot data:

$$Y_{ijkl} = \mu + T_j + B(T)_{l(j)} + G_i + GT_{ij} + GB(T)_{il(j)} + R_k + RT_{jk} + RB(T)_{kl(j)} + RG_{ik} + RGT_{ijk} + e$$

where Y_{ijkl} is the phenotypic performance of observation l of genotype i in ratoon k of trial j, μ is the grand mean, T_j is the trial main effect, $B(T)_{l(j)}$ is the effect of block within trial, G_i is the genotype main effect, GT_{ij} is the effect of the genotype x trial interaction, $GB(T)_{il(j)}$ is the effect of genotype x block (within trial), R_k is the main effect of ratoon, RT_{jk} is the effect of the ratoon x trial interaction, $RB(T)_{kl(j)}$ is the effect of the ratoon x block (within trial) interaction, RG_{ik} is the effect of the ratoon x genotype interaction, RGT_{ijk} is the effect of the genotype x trial x ratoon interaction, and e is the error. Restricted maximum likelihood (Patterson and Thompson,

1971) using the sparse Average Information algorithm (Gilmour *et al*. 1995) was used to estimate the variance components, as implemented in GenStat 12.0 (Anon, 2009).

GGE biplot analysis

The two-way table of variety x environment data for TCANE, ERC and TERC were analysed using GGE Biplot software (Yan, 2001). GGE Biplot analysis produces biplots derived from principal components analysis of the environment centered data (data minus the environment means), which therefore represents the genotype main effect and the GxE interaction. The genotypes and environments are represented by points on a two-dimensional plot of principal component (PC) scores (conventionally PC1 and PC2). The distance of an environment from the biplot origin is a measure of its ability to discriminate between genotypes, the distance between two environments measures their dissimilarity in discriminating the genotypes, and the angle between them represents their correlation. Acute angles represent positive correlations, obtuse angles represent negative correlations, and right angles represent no correlations between environments (Yan and Tinker, 2005). In general, the proximity of environments on the biplots is an indication of their similarity, and the proximity of genotypes to environments is an indication of the degree of positive interaction with the environment. GGE Biplot provides a range of viewing options to investigate relationships between environments and genotypes, identify mega-environments, examine the representativeness of test environments as selection sites, rank genotypes based on performance in single environments, and much more. Extensive descriptions of these user options are given by Yan and Tinker (2005) and will be described later.

Results

Variance components

The model is defined to split all variance components so that they are separated as much as possible, thereby allowing for an evaluation of the variance components of interest. In this model, terms such as ‘ratoon x block (within trial)’ and ‘genotype x block (within trial)’ are not of interest, as there would be no need to interpret replicate effects in a trial. These terms are simply included to remove the variability in the trial associated with the design.

The estimates of variance components are presented in Table 2. The largest proportion of the total variation in TERC (47%) was accounted for by the main effect of trials followed by the environment (trial x ratoon combination) component (24%). In terms of the components related to the GxE interaction (genotype, genotype x trial, genotype x ratoon, and genotype x trial x ratoon), the genotype x trial and genotype x trial x ratoon accounted for the largest variation. This highlights the importance of trial site effects on the GxE interaction, as well as the influence of trial site effects on the ratooning ability of varieties. The main effect of genotype accounted for 1.2% of the total variation compared to the 7.2% accounted for by the genotype x environment interaction (as expressed by the sum of the genotype x trial, genotype x ratoon, and genotype x trial x ratoon interactions). This suggests that only a small proportion of the total variation was due to the mean differences between varieties, and that the genotype x environment interaction was more pronounced. The variation accounted for by the main effect of ratoon was less than the variation accounted for by the individual ratoon interaction components. The genotype x ratoon interaction accounted for a minor proportion (0.25%) of the total variation

relative to the trial x ratoon interaction (24.4%), suggesting that ratooning ability is influenced to a larger extent by site differences compared to variety differences.

The proportion of variation accounted for by the different terms for TCANE was similar to that of TERC. Most of the variation was accounted for by the main effect of trials (35%), followed by the trial x ratoon interaction component (33%). The genotype x environment interaction (as expressed by the sum of the genotype interaction variance components) accounted for more variation (6.7%) than the main effect of genotype (4.2%), suggesting that TCANE, together with TERC, is influenced to a larger extent by GxE interactions than by genotypic main effects. However, when only considering terms related to the GxE interaction (genotype, genotype x trial, genotype x ratoon, and genotype x trial x ratoon), the genotype main effect accounted for most of the variation, followed by the genotype x trial, genotype x trial x ratoon, and genotype x ratoon interactions, respectively. The small variation accounted for by the genotype x ratoon interactions relative to the genotype x trial interaction once again highlights the greater influence of site factors rather than variety on ratooning ability. The strong genotype x trial x ratoon interaction suggests that variety cane yields and their ratooning ability was trial specific.

Similarly, for ERC, the largest proportion of the variation was accounted for by the main effect of trials (44%) followed by the trial x ratoon interaction (i.e. environment main effect; 28.6%). In contrast to TCANE and TERC, the main effect of genotype accounted for a higher proportion (5.5%) of the variation than the combined GxE interaction components (5.1%). This suggests that ERC is a trait that is more strongly influenced by genetic composition than GxE interactions. Once again, the variation accounted for by the genotype x ratoon interaction was minor (0.1%) compared to the contribution of the trial x ratoon interaction (28%). When only considering terms related to the GxE interaction (genotype, genotype x trial, genotype x ratoon, and genotype x trial x ratoon), the genotype main effect was once again the largest. This was followed by the genotype x trial x ratoon effects, indicating that genotype variability for sucrose content across ratoons was largely controlled by trial location. The genotype x trial component was the third largest, followed by the genotype x ratoon interaction. This suggests that ERC among varieties was least influenced by the effect of ratoons.

Biplot analysis

Mega-environment analysis

Figure 1 shows the polygon view of GGE biplots for TCANE, ERC and TERC based on the two-way table of GxE means. The polygon is drawn on vertex genotypes that were the most responsive (i.e. the best or poorest performers at some or all of the environments). Perpendicular lines are drawn from each side of the polygon to the origin, such that the biplot is divided into sectors representing different mega-environments. Environments within a mega-environment sector have the same effects on variety performance and should be considered as a homogenous group.

For TERC, the biplot explained a total of 42% of the total variation of the two-way data matrix (Figure 1a), highlighting the complexity of the GxE structure in the dataset. The TERC biplot showed four distinct mega-environments represented by the N31, N12, N29, and N36 sectors, respectively. The majority of the midlands (M) environments were located within the N31 and N12 sectors, while the N29 and N36 sectors contained both coastal (C) and hinterland (H) environments. In general, the ratoons within trials clustered together strongly, followed by a

clear separation of the M environments from C and H environments. The mega-environments identified here are not in keeping with the geographical regional divisions utilised currently for the rainfed parts of the industry (midlands, coast, hinterland). For example, if the current regional separations were appropriate the biplot would consist of three mega-environments corresponding to the C, H and M regions. The mega-environments identified here may not necessarily group according to any geographical zoning, and attempts should be made to identify the main environmental factors characterising these mega-environments.

Table 2. Variance components estimates and percentage of total phenotypic variance for tons cane/ha (TCANE), estimated recoverable crystal per cent (ERC %), and tons estimated recoverable crystal (TERC). Standard errors of estimates are in brackets.

	TERC		TCANE		ERC%	
	Variance	%	Variance	%*	Variance	%
Trial	8.41 (2.22)	47.34	312 (92.4)	35.0	2.54 (0.69)	44.2
Block (within trial)	0.28 (0.06)	1.60	20.2 (4.4)	2.3	0.03 (0.009)	0.5
Genotype	0.22 (0.22)	1.24	37.5 (15.8)	4.2	0.31 (0.12)	5.5
Genotype x trial	0.62 (0.61)	3.48	33.5 (5.7)	3.8	0.11 (0.02)	1.9
Genotype x block (within trial)	0.79 (0.79)	4.47	53.4 (4)	6.0	0.02 (0.01)	0.3
Ratoon	0.08 (0.08)	0.47	0	0	0	0
Ratoon x trial	4.34 (0.64)	24.43	294 (42.6)	32.9	1.64 (0.23)	28.6
Ratoon x block (within trial)	0.23 (0.03)	1.30	14.2 (2.1)	1.6	0.04 (0.01)	0.7
Genotype x ratoon	0.05 (0.02)	0.25	2.6 (1.2)	0.3	0.01 (0.007)	0.1
Genotype x trial x ratoon	0.61 (0.06)	3.45	23.2 (2.7)	2.6	0.18 (0.02)	3.1
Error	2.13 (0.05)	11.98	104 (2.9)	11.7	0.87 (0.02)	15.2

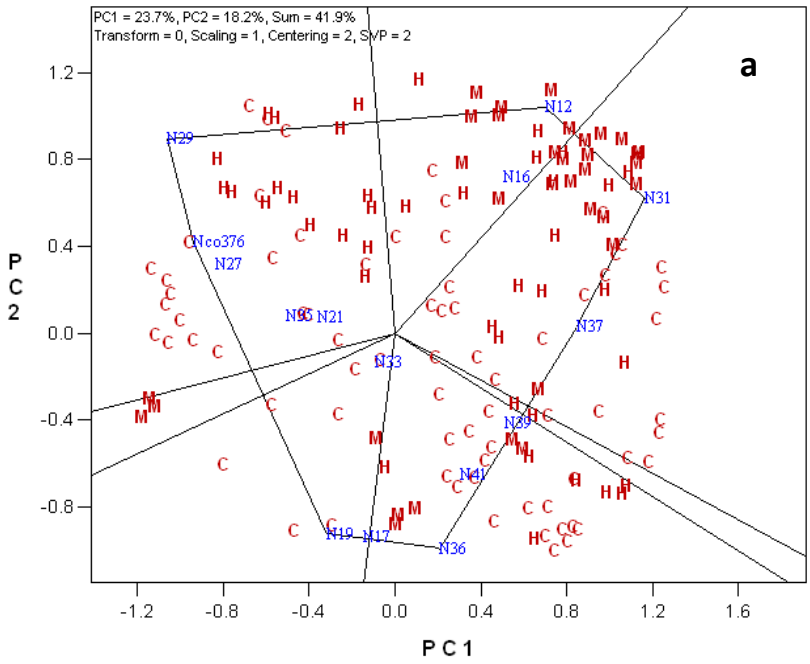
*Percentge of total phenotypic variance

The TCANE biplot (Figure 1b) also accounted for approximately 42% of the variation in the GxE table of means and was very similar to the biplot produced for TERC. However, this biplot exhibited three mega-environments instead of four, as the N12 mega-environment observed in the TERC biplot now fell within the N31 sector. The majority of the M environments fell within the N31 sector, which also contained some C and H environments. This implies that N31 was the highest yielding variety in the environments grouped into that mega-environment. The N36 sector contained both C and H environments, as well as four M environments that were different from the other M environments. The N29 sector contained C and H environments only.

The ERC biplot, which accounted for 39% of the variation in the GxE table showed that the majority of the C and H environments fell within a mega-environment characterized by the vertex genotype N29 (Figure 1c). This genotype was released to the industry for its characteristically high ERC%. The M environments were scattered throughout the biplot, forming indistinct clusters. However, the vertex genotype N37 formed a mega-environment composed primarily of M environments in the top left hand corner of the biplot. The formation of a dominant mega-environment for ERC may be linked to the larger ratio of G to GxE

observed from variance components analyses (Table 2), which indicates that ERC is influenced to a larger extent by genetic effects rather than the GxE interaction effects.

Although environments tended to cluster according to the ratoons harvested within a trial, frequently, a given trial showed diverse responses between ratoons. These patterns are illustrated in Figure 2, which shows the biplot for TERC (same as Figure 1a, but with full names of the environments indicated). Environments CEM30, CEM32, CEM34 and CEM36 were located within the N31 mega-environment while the other ratoons from this site (CEM31, CEM33, CEM35) were located in different mega-environments. Genetic correlations between these environments confirmed that CEM30, CEM32, CEM34 and CEM36 were positively correlated while CEM31, CEM33 and CEM35 were negatively correlated to the former ratoons (Table 3). Similarly, environments HES10, HES11, and HES12 formed part of the N29 mega-environment, while the next harvest from this site (HES13) was located in the N36 mega-environment. Table 3 confirms the negative correlations between HES13 and the other ratoons from that particular trial. Other examples may be identified through further interrogation of Figure 2. Similar observations were also made in the TCANE and ERC biplots (not shown). This differential grouping of ratoons from a trial site demonstrates the trial x ratoon interaction, which was also observed to be highly influential from the variance components analysis (Table 2). In addition to differential responses of ratoons within trials, it was also frequently observed that trials from a particular location grouped into different mega-environments (Figure 2). For example, ratoons from CAK1, CAK2 and CAK3 were located in different mega-environment sectors. Table 3 shows positive correlations between ratoons of CAK1 and CAK2, while correlations between those trials and ratoons of CAK3 were negative.



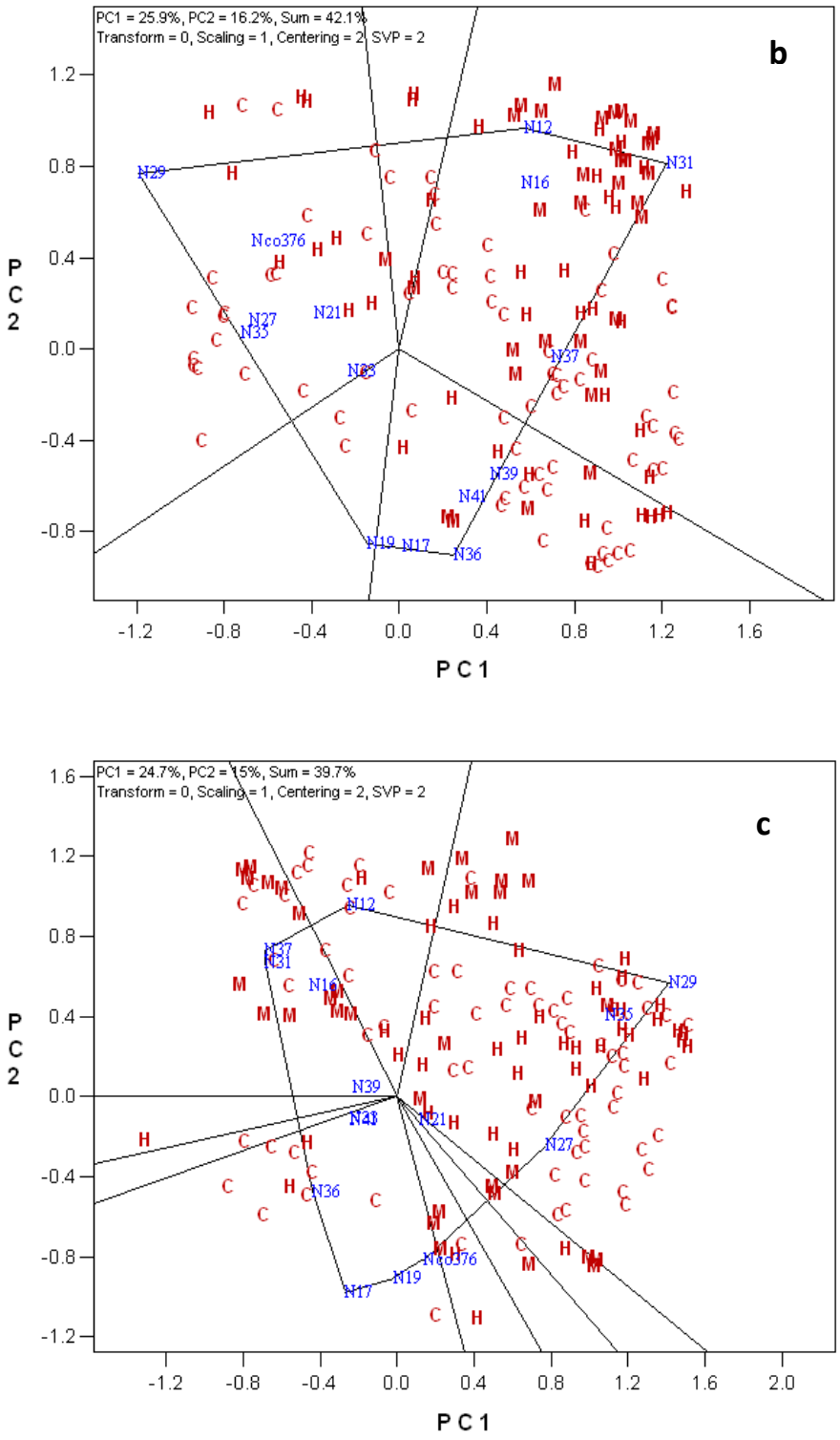


Figure 1. Polygon views of GGE biplots for TERC (a), TCANE (b) and ERC (c) based on data of 15 genotypes tested over 153 environments. For better visualization of regional overlapping and separation, the environments are abbreviated by their first letters only (in red), while the genotypes are indicated in blue.

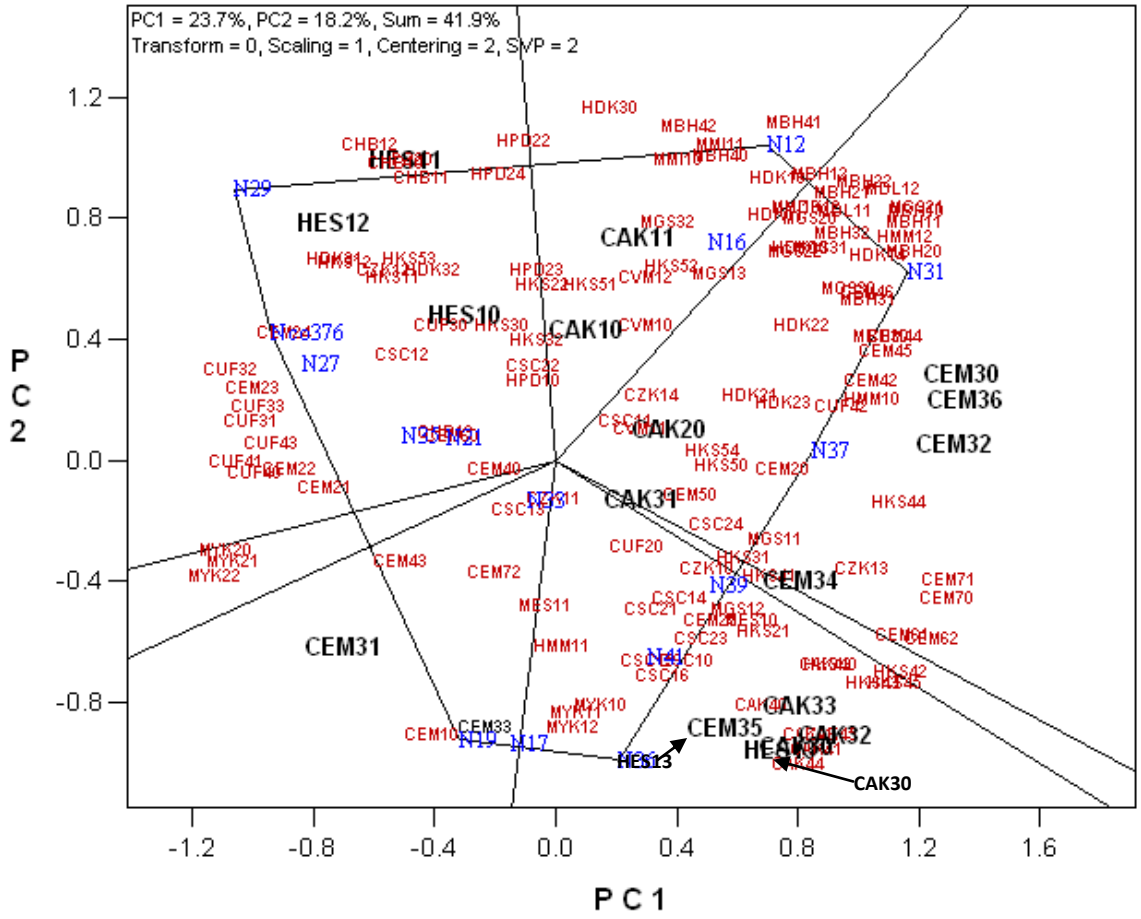


Figure 2. Polygon view of a GGE biplot for TERC based on data of 15 genotypes tested over 153 environments. Environments are shown in red and genotypes are shown in blue. Environments cited in text are shown in bold, black font.

Table 3. Genetic correlations between selected environments showing the variable relationships between ratoons of the same trials and between trials within the same location.

	CEM30	CEM31	CEM32	CEM33	CEM34	CEM35	CEM36
CEM30	1						
CEM31	**-0.75****	1					
CEM32	**0.91****	**-0.46**	1				
CEM33	**-0.56***	**0.79****	**-0.22**	1			
CEM34	**0.37**	**0.11**	**0.61***	**0.23**	1		
CEM35	**-0.04**	**0.44**	**0.24**	**0.61***	**0.68****	1	
CEM36	**0.98****	**-0.65****	**0.96****	**-0.44**	**0.47**	**0.11**	1
	HES10	HES11	HES12	HES13			
HES10	1						
HES11	**0.61***	1					
HES12	**0.32**	**0.73****	1				

HES13	**-0.39**	**-0.75****	**-0.52***	1			
	CAK10	CAK11	CAK20	CAK30	CAK31	CAK32	
CAK10	1						
CAK11	**0.54***	1					
CAK20	**0.79****	**0.55***	1				
CAK30	**-0.27**	**-0.55***	**-0.01**	1			
CAK31	**-0.76****	**-0.41**	**-0.72****	**0.22**	1		
CAK32	**-0.51***	**-0.57***	**-0.25**	**0.91****	**0.51***	1	
CAK33	**-0.59***	**-0.51***	**-0.31**	**0.84****	**0.49***	**0.89****	
**p<0.01							
* p<0.05							

*p<0.01 **p<0.05

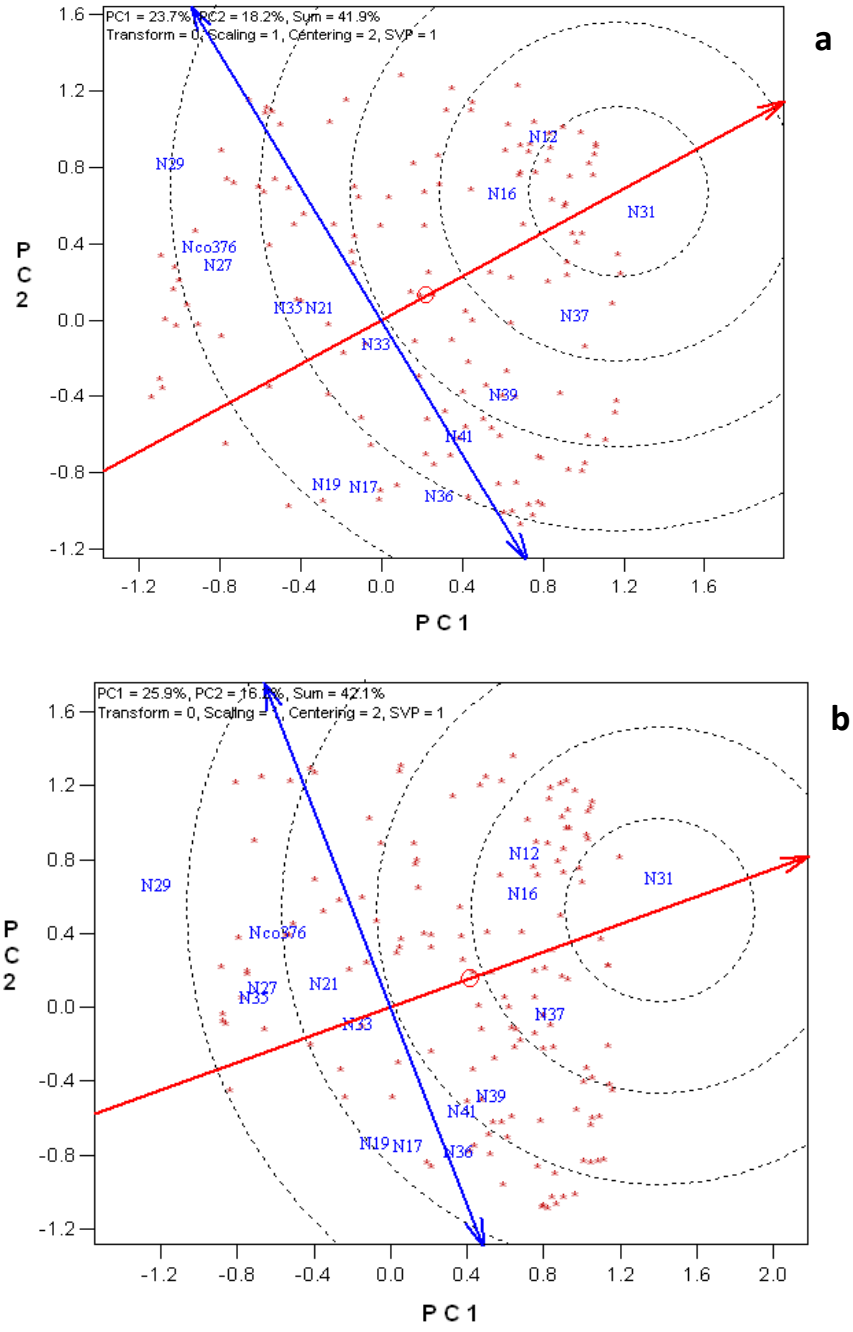
Genotype performance and stability

Figure 3a is a different view of the TERC biplot highlighting the average performance and stability of the different genotypes. In this view, the average TERC of the genotypes are indicated by the projections of their markers onto a derived axis (red line with single arrowhead). When perpendicular lines are drawn from the variety markers to this axis, the TERC rankings of varieties can be seen. When comparing the biplot results with the mean TERC of the varieties, the biplot approximated the ranking of the genotypes very well (correlation of 0.78). Thus, genotype N31 had the highest mean TERC, followed by N12, and then N16, while N19 had the lowest average TERC. The stability of the genotypes is measured by the projections of their markers on the other derived axis (blue line with double arrowhead). Perpendicular lines are drawn from the variety markers to this axis and the points at which they intersect this axis then define their stability. The greater the absolute distance of the intersections from the origin, the less stable a variety is. Thus, genotypes like N33, N31 and N21 were the most stable, while genotypes such as N29, N36 and N17 were the most responsive. The center of the concentric circles in the biplot represents the location of a hypothetical 'ideal' genotype (i.e. high yield and stability). Genotype N31 was located the closest to this point, suggesting that it may be the ideal genotype to utilise as a control when selecting for TERC in selection programmes.

Performance and stability of genotypes based on TCANE can be seen in Figure 3b. The biplot and genotype rankings are very similar to the TERC biplot, with genotype N31 showing the highest TCANE and exhibiting stability, i.e closest to the ideal genotype. The correlation between biplot ranks and genotype means was 0.88, suggesting that the projections of the genotype markers were a good approximation of the average genotype performance. Once again, N31 was located closest to the ideal genotype, suggesting that it would be a good candidate as a control during selection for TCANE.

In terms of ERC (Figure 3c), genotype N29 was ranked as the top performer, followed by N35 and N27, while the lowest ranked genotype was N17 (correlation between biplot ranks and means was 0.88). Genotypes N41, N36, N39 and N21 were the most stable for ERC. Genotype N31, which was stable in terms of TCANE (Figure 3a), was responsive in terms of ERC. In contrast, genotype N36 was responsive for TCANE (Figure 3a) and stable for ERC (Figure 3b). Genotypes N33 and N21 were stable for both TCANE and ERC; however, their mean

performance was not impressive for both traits. Genotype N29 was the closest to the ideal genotype for ERC, suggesting that it would be a good candidate for a control genotype when selecting specifically for ERC.



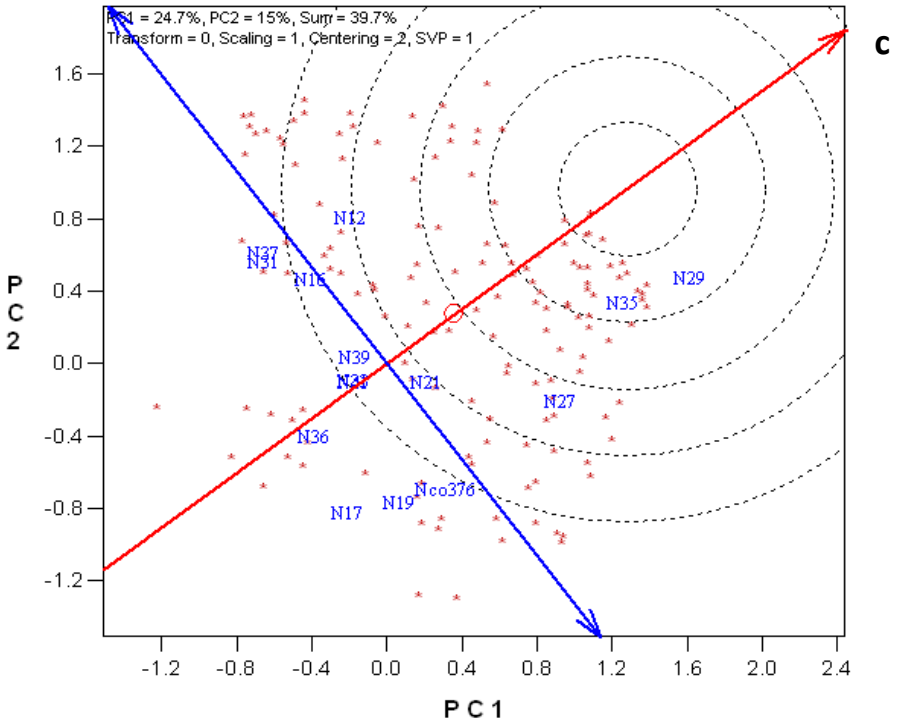


Figure 4. The 'means vs. stability' view of GGE biplots for TERC (a), TCANE (b), and ERC (c) based on data from 15 genotypes tested over 153 environments. Environments are indicated by an asterisk, while genotypes are indicated by their names.

Discussion

The results of this study have provided valuable insights into the nature of the GxE interactions characterising the rainfed part of the industry. The large component of variation accounted for by the genotype x trial interaction relative to the genotype x ratoon interaction highlights that there may not be much value in extending trials to longer ratoons. It follows that, to increase responses to selection, greater emphasis should be placed on sampling more trial sites than on testing ratoons within a trial site. This confirms results obtained in other studies on sugarcane in Australia (Mirzawan *et al.* 1994, Jackson *et al.* 1991). This trend was also evident in the biplots where, in general, environments that clustered close together were ratoons of the same trial. However, frequently, ratoons of the same trial also showed diverse responses (Figure 2), showing that seasonal/ratoon differences could have an effect on the categorization of a particular environment into a mega-environment. The grouping of ratoons from a trial into different mega-environment sectors suggests the existence of influential climatic (soil characteristics remain constant), biotic, or management factors that can produce genotypic rank changes between ratoons. Further analyses of the factors characterizing such environments may help explain the reasons for these differential ratoon responses within trials. The existence of 'subsets of environments' or 'environment types' that are characterized by similar biotic and abiotic stresses may be a possibility, and these should be investigated for the improvement of recommendation strategies.

The occasional differential grouping of trials from the same location into different mega-environments suggests rapid changes in the values of relevant environmental factors for environments that were geographically close. This may be due to the use of different experimental fields within a location, where factors such as cropping history, management practices (the majority of trials were conducted on commercial farms), and variations in soil physical and chemical properties brought about by topographic position have large confounding effects. These are further complicated by the effects of different times of harvest and harvest ages, thereby giving rise to large and complex GxE interactions that can only be interpreted through proper characterization of trial sites. Other studies have also addressed the issue of such genotype x management interactions for grain yield of wheat (Cooper *et al*., 2001).

Biplot analysis has shown that selection strategies for TERC in the rainfed parts of the sugar industry should be targeted at four mega-environments, which do not necessarily conform to the traditional regional sub-divisions currently in use. Mega-environments consisted of a mix of C, H, and M environments which discriminated genotypes similarly, implying that certain environments, although geographically distinct, had a similar effect on genotypic responses. This was particularly true of environments belonging to the C and H regions, as the M environments tended to cluster fairly tightly, with only few outliers. The six M environments that were outliers (Figure 2) were actually trials conducted in frost pockets in the midlands that were harvested on a 12-month cutting cycle as opposed to the conventional 18-24 month cycle for that region. This suggests a possible dominant role of harvest age in determining environmental groupings. Further analysis of the biplot also showed that environments representative of the five selection programmes (coastal high potential, coastal low potential, hinterland, midlands humic, midlands sandy) did not cluster together to form five separate mega-environments. This may imply the need for possible adaptation of the current selection strategy to target the mega-environments identified in this study.

In addition to the implications on selection, the results of this study have provided insight on the redundancy of trial sites for the purposes of post-release evaluation. Environments tightly clustered within a mega-environment should be similar with regard to genotypic performance (Yan and Tinker, 2005), suggesting that one or more environments could be removed without much loss of information. In this study, a range of M environments clustered tightly within the N31 sector in Figure 1a. These environments included those representative of the ‘sandy’ and ‘humic’ conditions which are currently evaluated in the midlands and for which two separate selection programmes currently exist. The appropriateness of the sandy vs humic subdivision in the Midlands has never been tested, and the results of this study suggest the need for such an evaluation to optimise resources and improve the efficiencies of both selection and evaluation. Another example is the tight clustering of ratoons from the trials CUF3, CUF4 and CEM3 on the left hand side of Figure 2, which suggests that at least one of those trial sites could be removed for post-release evaluation purposes.

The analysis of genotype performance and stability identified N31 as the ideal genotype for TERC and TCANE, while N29 was identified as the ideal genotype for ERC. These results suggest that these genotypes may be ideal candidates for controls in selection trials to improve genetic gains. Currently, however, genotype NCo376 is used as a control in all plant breeding selection trials. The poor relative performance of NCo376 in terms of performance (TERC,

TCANE and ERC) and stability are an indication of successful selection gains achieved through the release of newer genotypes. In order to sustain future gains, the use of alternate, higher yielding and stable controls such as N31 must be considered. The variance components ratio of genotype to GxE was higher for ERC than for TERC and TCANE, suggesting that gains from selection are likely to be more rapid for ERC relative to the other two traits. This has indeed been the case in the South African sugar industry, where the average ERC of new commercial releases have increased substantially relative to TCANE. Additionally, adoption of new varieties in the South African sugar industry has not occurred at the same rate as variety release, leading to large areas dominated by single (older) varieties such as NCo376. The results of this study demonstrate the potential benefits of adopting newer varieties.

Conclusions

This study was the first to evaluate the variance components and mega-environments associated with METs conducted in the rainfed regions of the South African sugar industry. The results have demonstrated the existence of large and complex GxE interactions which may be exploited for improved selection and evaluation. Important findings of the study were:

- The identification of larger trial x ratoon interactions relative to genotype x ratoon interactions, which have direct impact on the design of the selection and evaluation network.
- The existence of four mega-environments that do not correspond to the current regional sub-divisions used in the rainfed parts of the industry.
- A discrepancy between the current five selection programmes and the target environments comprising the rainfed regions.
- The existence of differential mega-environment groupings of ratoons within trials and trials within locations, highlighting the need to characterise trials according to environmental factors.
- The evaluation of relative performance and stability of genotypes for use as controls when selecting for different yield traits.

Future studies will involve a complete characterization of environments used in this study in terms of climatic variables, soil characteristics, and management/crop factors (age and time of harvest). Such studies will also allow for the analytical interpretation of GxE interactions which has implications in terms of developing more targeted breeding programmes, better choice of environments for selection and evaluation, and will provide insight into the genetic and physiological makeup of genotypes for further physiological studies.

REFERENCES

Anon (2009). GenStat® Executable Release 12.1. Lawes Agricultural Trust, Rothamstead Experimetnal Station, Harpenden, UK. Clarendon Press, London, UK.

Ceretta S and van Eeuwijk FA (2008). Grain yield variation in malting barley cultivars in Uruguay and its consequences for the design of a trials network. *Crop Science* 48: 167-180.

Cooper M, Woodruff DR, Phillips IG, Basford KE and Gilmour AR (2001). Genotype-by-mangement interactions for grain yield and grain protein concentration of wheat. *Field Crops Res* 69: 47-67.

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

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

de la Vega AJ.and Chapman SC (2010). Mega-environment differences affecting genetic progress for yield and relative value of component traits. *Crop Sci* 50: 574-583.

Finlay KW and Wilkinson GN (1963). The analysis of adaptation in a plant-breeding programme. *Aust J Agric Res* 14: 742-754.

Gauch HG (1992). Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier, Amsterdam.

Gauch HG and Zobel RW (1997). Identifying mega-environments and targeting genotypes. *Crop Science* 37: 311-326.

Gilmour AR, Thompson R and Cullis BR (1995). Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 51: 1440-1450.

Glaz B and Kang MS (2008). Location contributions determined via GGE biplot analysis of multienvironment sugarcane genotype-performance trials. *Crop Science* 48: 941-950.

Jackson PA, Horsley D, Foreman J, Hogarth DM and Wood AW (1991). Genotype x environment (GE) interactions in sugarcane variety trials in the Herbert. *Proc Aust Soc Sug Cane Technol* 13: 103-109.

Mirzawan PDN, Cooper M, DeLacy IH and Hogarth DM (1994). Retrospective analysis of the relationship among the test environments of the southern Queensland sugarcane breeding programme *Theoretical and Applied Genetics* 88: 707-716.

Nuss KJ (1998). Aspects considered in the search for new farms for the experiment station. *Proc S Afr Sug Technol Ass* 72: 42-45.

Patterson HD and Thompson R (1971). Recovery of interblock information when block sizes are unequal. *Biometrika* 58: 545-554.

Queme JL, Orozco H and Melgar M (2010). GGE biplot analysis used to evaluate cane yield of sugarcane (*Saccharum* spp.) cultivars across sites and crop cycles. *Proc Int Soc Sug Cane Technol* 27: 1-7.

Sharma RC, Morgounov AI, Braun HJ, Akin B, Keser M, Bedoshvili D, Bagei A, Martius C and van Ginkel M (2010). Identifying high yielding stable winter wheat genotypes for irrigated environments in central and west Asia. *Euphytica* 171: 53-64.

Shukla GK (1972). Some statiscal aspects of partitioning genotype-environment components of variability. *Heredity* 29: 237-245.

Smith AB, Cullis BR and Thompson R (2005). The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. *Journal Agric Sci* 143: 449-462.

Yan W (2001). GGEbiplot – A Windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agronomy Journal* 93: 1111-1118.

Yan W and Hunt LA (2001). Interpretation of genotype x environment interaction for winter wheat in Ontario. *Crop Science* 41: 19-25.

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

Yan W and Tinker NA (2005). Biplot analysis of multi-environment trial data: Principles and applications. *Can J Plant Sci* 86: 623-645.

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