Field Crops Research 124 (2011) 392–399

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Interpretation of genotype × environment interactions of sugarcane: Identifying significant environmental factors

S. Ramburan[a,*], M. Zhou[a], M. Labuschagne[b]

[a] South African Sugarcane Research Institute, Private Bag X02, Mount Edgecombe, Durban, South Africa
[b] Division of Plant Breeding, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein, South Africa

ARTICLE INFO

Article history:
Received 25 May 2011
Received in revised form 19 July 2011
Accepted 21 July 2011

Keywords:
AMMI
Biplot
Genotype × environment
Mega-environment
Sugarcane

ABSTRACT

An understanding of the causes of genotype × environment (G × E) interactions is essential for the implementation of efficient selection and evaluation networks. Currently, studies involving the interpretation of sugarcane (*Saccharum* spp.) G × E interactions are limited. The objective of this study was to investigate the relative influence of environmental factors on the G × E interactions of sugarcane under rainfed conditions in South Africa through a comprehensive analysis of a multi-environment trial (MET) dataset. Fifteen commercial cultivars were evaluated over 147 environments (trial × ratoon combinations) across the coastal (C), hinterland (H) and midlands (M) regions of the sugar industry. Environments were characterized according to five site covariates (soil depth, clay percentage, organic matter percentage, nitrogen mineralization category, and total available moisture) and nine seasonal covariates (time of harvest, age at harvest, average daily heat units, solar radiation, rainfall, evaporation, and three derived water stress indices).

Additive main effects and multiplicative interaction (AMMI) biplots for cane yield (TCANE), estimated recoverable crystal percent (ERC) and tons ERC (TERC) revealed overlapping of C and H environments, while M environments formed unique clusters characterized by specific cultivar adaptabilities. Principal components analysis (PCA) allowed visualization of the covariates determining the regional separation patterns. AMMI interaction principal components axes (IPCA) 1 and 2 scores were correlated to the covariates and showed that harvest age, temperature, and water stress were mainly responsible for separation of M environments from C and H environments on the TCANE and TERC biplots. Time of harvest was identified as an important covariate influencing ERC G × E patterns in the C and H regions. The third water stress index (based on a ratio of observed yields to simulated irrigated yields) was a dominant factor influencing G × E patterns within the C and H regions and was identified as a superior indicator of water deficient environments for future studies. The M trials were characterized by shallower soils with lower total available moisture and greater variability in this regard compared with the C and H trials. Nitrogen mineralization category, organic matter percent, and clay percent were not significantly correlated to IPCA scores, while soil depth was identified as a major site selection criterion in the M region. The M region should be treated as a single mega-environment, while the C and H regions could be combined for future interpretive studies, where covariates should be summarized within growth phases. The results of this study will assist in restructuring the MET network through exploitation and targeting of the relevant environmental factors within the different regions.

1. Introduction

Sugarcane production in South Africa extends along the north eastern regions of the country, across a wide range of agro-climatic conditions and management regimes. The crop is harvested and processed over a nine-month period from April to December, and depending on production conditions, reaches maturity between 12 and 24 months after planting/ratooning. The rainfed parts of the industry are partitioned into three broad regions characterized by different production conditions. The coastal (C) region is characterized by areas with low altitude (<200 m above sea level), higher temperatures, and high topographic, soil, and rainfall variability. The midlands (M) region is characterized by higher altitudes (>600 m asl), lower temperatures, flatter topography, and lower soil variability, while the hinterland (H) region is a transitional region between the midlands and the coast. Commercial genotypes (cultivars) are developed for these regions from five separate selection programmes representative of different production

* Corresponding author. Tel.: +27 31 5087444; fax: +27 31 5087597.
E-mail address: Sanesh.Ramburan@sugar.org.za (S. Ramburan).

0378-4290/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2011.07.008

conditions. This is further supplemented by trial-based post-release evaluation under commercial conditions. Despite the above efforts to minimize genotype × environment (G × E) interactions and improve recommendations, frequent deviations have been observed between pre- and post-release cultivar performance. This may be due to the lack of representativeness of selection sites and the limited understanding of the factors responsible for G × E interactions. In a recent analysis of a multi environment trial (MET) dataset spanning the rainfed regions, Ramburan and Zhou (in press) demonstrated large and complex G × E interactions. The study, which employed biplot analysis (Yan et al., 2000), showed differential mega-environment groupings of ratoons from the same trials and trials from the same locations, suggesting large spatial and temporal fluctuations in major yield determining environmental factors. This highlighted the need to investigate the nature of such G × E interactions for exploitation within the breeding and evaluation networks.

G × E interaction has long been a barrier to efficient plant improvement through its impairment of efforts to select superior genotypes and its negative impact on heritability. For sugarcane, the majority of G × E studies have been empirical in nature, focusing mainly on quantifying G × E interactions and genotype stability (Tai et al., 1982; Queme et al., 2005; Jackson et al., 2007), identifying mega-environments (Jackson et al., 1991; Mirzawan et al., 1994; Queme et al., 2007), and quantifying sources of variation for resource allocation (Bissessur et al., 2010, Rattey and Kimbeng, 2001). In comparison, fewer sugarcane studies have focused on the interpretation of G × E interactions to understand the nature and causes thereof. Such an understanding may benefit sugarcane improvement initiatives and allow for the exploitation of G × E interactions rather than its avoidance or acceptance. An understanding of the factors responsible for G × E interactions can be used to establish breeding objectives, formulate recommendation domains, contribute to ideotype design, and identify ideal test conditions (Yan and Hunt, 2001).

Approaches currently used to gain an understanding of the causes of G × E interactions can be categorized into two strategies. The first involves the use of factorial regression models based on two-way G × E tables with concomitant variables which could either be environmental factors, genotypic traits, or combinations thereof (Baril et al., 1995). The second strategy involves the correlation of genotypic or environmental scores derived from additive main effects and multiplicative interaction (AMMI) analysis to genotypic or environmental covariates (Van Eeuwijk et al., 1995). Both strategies, although different in approaches, have been shown to produce similar results (Vargas et al., 1999). Many studies have utilized the second strategy due to the production of principal components analysis (PCA)-based biplots, which allow for easy visualization of responses and relationships. In addition, biplots produced from AMMI analysis can be enriched with covariates to allow for easier interpretation, provided there are strong correlations between AMMI IPCA scores and the covariates. Voltas et al. (1999) used this method to identify genotypic and environmental covariates influencing G × E interactions for grain filling in barley. The three covariates identified were subsequently included into a factorial regression model to interpret G × E interactions. Van Oosterom et al. (1996) related AMMI IPCA scores to environmental covariates and found that factors such as the mean maximum temperature 10 days after flowering and the changes in a water satisfaction index during grain filling were some of the causes of G × E interactions of pearl millet. Van Eeuwijk and Elgersma (1993) found that four environmental covariates showed strong correlations to AMMI environment IPCA scores and these corresponded to the covariates identified by factorial regression, in an analysis of G × E interactions of ryegrass. De Vita et al. (2010) used this approach to investigate environmental and genotypic causes of yield variation of wheat in Italy, and identified water availability as a key factor. Vargas et al. (1999) showed that this approach produced results that were comparable with those of factorial regression and partial least squares regression in their analysis of two wheat datasets. The methodology can also be adapted to understand genetic causes of G × E interactions. Annicchiarico et al. (2010) investigated the adaptation of lupin landraces by correlating AMMI IPCA scores of different germplasm pools to mean values of different morphophysiological traits.

Interpretive studies of sugarcane G × E interactions are currently limited to a single analytical study that investigated causes of G × E interactions in Australian selection trials (Jackson et al., 1995). Subsequent interpretive studies of sugarcane G × E interactions involved the straightforward characterization of METs (in a single-factor approach) according to soil type (del Blanco et al., 2010; Glaz and Kang, 2008), regional characteristics (Bissessur et al., 2010), and time of harvest (Gilbert et al., 2006). However, since Jackson et al. (1995), no further attempts have been made to investigate the nature of sugarcane G × E interactions within a MET framework through combined analysis of the major environmental drivers of crop growth. This approach should allow for a better understanding of crop performance in different environments and can only be achieved through a detailed characterization of the soil and climatic factors associated with METs. The complexity of sugarcane G × E interactions in South Africa necessitates an investigation into the nature thereof, leading to more informed breeding objectives, better test site selection, development of appropriate recommendation domains, and sugarcane ideotype design. The objective of this study was to investigate the relative influence of different environmental factors on the G × E interactions of sugarcane under rainfed conditions through a comprehensive analysis of a MET dataset.

2. Materials and methods

2.1. MET dataset

The trial dataset used in this study comprised 43 trials (32 post-release cultivar evaluation trials and 11 advanced plant breeding selection trials) grown in 18 different locations, harvested during the period 1999–2009. The number of ratoons harvested from each trial varied from one to six. An environment was defined as a trial × ratoon combination, resulting in a total of 147 environments spanning the C, H, and M regions. Post-release trials were conducted on commercial fields while selection trials were established on selection stations. 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 a location were numbered in the order of establishment. Environments were coded according to the region, location, site, and ratoon number, respectively. For example, "CEM21" referred to a first ratoon crop, of the second trial conducted at Empangeni, in the coastal region. Further examples are discussed in the analysis by Ramburan and Zhou (in press) and no further elaboration is required here. Genotypic data for cane yield (TCANE), estimated recoverable crystal percent (ERC) and tons ERC/ha (TERC) of the 15 cultivars most common to all trials were analyzed.

All trials were conducted as RCBD's with four–six replicates. Trial plots consisted of five or six rows that were between 8 and 10 m long, spaced 1–1.2 m apart. Weed, fertilizer and pest and disease management proceeded as per commercial farm practice. 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 TCANE. A 12-stalk sample was taken from each plot

Table 1
Descriptions of nine seasonal and five site covariates measured over 43 trials and 147 environments.

Covariate	Description	Mean	Range
Seasonal covariates			
TT	Thermal time in heat units using a base temperature of 10 °C (Heat units/day)	10.78	6.74–13.97
RAD	Average daily solar radiation ($MJ/m^2/s/day$)	15.77	12.99–17.89
RAIN	Average daily rainfall (mm/day)	2.56	0.98–4.57
EVP	Average daily A-pan evaporation (mm/day)	4.37	2.83–8.33
WSI1	Water stress index 1. Simulated rainfed evapotranspiration vs. simulated irrigated evapotranspiration	0.48	0.15–0.98
WSI2	Water stress index 2. Simulated rainfed cane yield vs. simulated irrigated cane yield	0.43	0.08–0.99
WSI3	Water stress index 3. Observed trial mean yield vs. simulated irrigated cane yield	0.61	0.18–1.32
AGE	Age at harvest (months)	15.69	10.72–25.97
TOH	Time of harvest (1–9). Coded as month 1 (April)–month 9 (December)	N/A	N/A
Site covariates			
Depth	Effective rooting depth (mm)	667.44	300–1200
Clay	Clay percentage (%)	30.02	5–65
OM	Organic matter percentage (%)	2.16	0.2–4.54
Ncat	Nitrogen mineralization category (1–4)	N/A	N/A
TAM	Total available moisture (mm)	74.35	30–192

to determine a range of milling quality characteristics, including ERC, which is an estimate of the recoverable value of consignments delivered to sugarcane mills. The ERC was calculated as:

$$ERC = aS - bN - cF$$

where S = sucrose % in cane; N = non-sucrose % in cane; F = fibre % in cane and a, b, c are industry determined factors representing the loss of sucrose during processing (0.978, 0.535, and 0.018 respectively), and are derived from previous milling seasons. The TERC was calculated as a product of TCANE and ERC.

2.2. Environment characterization

For each trial in the dataset, stable site characteristics were determined through soil profile analyses and conventional soil sampling procedures to develop five *site covariates*. These included the effective rooting depth (Depth), clay percent (Clay), organic matter percent (OM), and N mineralization category (Ncat) (1 = low, 2 = medium, 3 = high, 4 = very high) (Anon., 2000). The total available moisture (TAM) for each site was subsequently estimated from Clay and Depth, from a formula developed by Van Antwerpen et al. (1994).

Apart from the stable site covariates, nine other covariates that varied seasonally (hence termed *seasonal covariates*) were also developed. For each environment (trial × ratoon combination) weather data were summarized from weather stations that were closest to the trial sites, and used to calculate the following covariates: average daily thermal time (base 10 °C) in heat units (TT); average daily solar radiation (RAD) in $MJ\,m^{-2}\,s^{-1}$, average daily rainfall (RAIN) in mm, average daily evaporation (EVP) in mm, and three water stress indices (WSI1, WSI2, and WSI3) that ranged from 0 to 1. The WSI1 was calculated by using the TAM, crop start/ratoon dates, and relevant weather station data as inputs to run the Canesim crop growth model (Singels and Donaldson, 2000), which estimated the actual crop evapotranspiration (EVT_{act}) for the respective crop cycle. The same model scenarios were also run for the same cycle as if fully irrigated to simulate unstressed crop growth and estimate the potential evapotranspiration (EVT_{pot}) had there been irrigation. The WSI1 was subsequently calculated as the ratio of EVT_{act}:EVT_{pot}, where a high WSI1 represented no moisture stress, while values closer to zero represented higher levels of stress. The second WSI (WSI2), was based on cane yields, and was calculated as a ratio of simulated rainfed cane yield:simulated irrigated cane yield. The third WSI (WSI3) was calculated as a ratio of observed trial mean yield:simulated irrigated yield. For each environment, the age at which the crop was harvested (AGE) and the time of harvest (TOH) was also recorded as additional covariates. Details of the five site and nine seasonal covariates, together with their means and ranges are indicated in Table 1.

2.3. Statistical analyses

2.3.1. G × E interactions

The two-way table of G × E means for TERC, TCANE and ERC were analyzed using the AMMI model:

$$Y_{ij} = \mu + g_i + e_j + \Sigma\lambda_k\,\alpha_{ik}\,\delta_{jk} + R_{ij} + \varepsilon \quad (1)$$

where Y_{ij} is the value of the ith genotype in the jth environment; μ is the grand mean; g_i is the mean of the ith genotype; e_j is the mean of the jth environment; λ_k is the singular value for principal component (PC) axis k; α_{ik} and δ_{jk} are the PC scores for axis k of the ith genotype and jth environment, respectively; R_{ij} is the residual and ε is the error term (Gauch, 1992). AMMI2 biplots were produced to visualize the G × E interactions for TERC, TCANE and ERC. AMMI2 biplots are characterized by the projection of the genotype and environment IPCA1 and IPCA2 scores onto a two-dimensional biplot. The interpretations of G × E interactions from an AMMI2 biplot are similar to other forms of biplot analysis. Genotypes and environments are represented by points. The distance from the origin of the biplot to these points represents the amount of interaction that is exhibited by the respective genotype or environment. The angles between two genotype vectors correspond to their correlation. In general, acute angles represent positive correlations, right angles represent no correlation, and obtuse angles represent negative correlations (Voltas et al., 1999). Similar interpretations hold true for environments. The magnitude of interactions between genotypes and environments can be interpreted by their respective vector direction, where similar directions indicate positive interactions and vice versa (Gauch, 1992).

In order to study the correlation structures of the environments in relation to the nine seasonal covariates, a principal components analysis (PCA) was conducted on the environment × covariate two-way table. The biplot of the first two PC axes was used to assess the main differences between the environments, identify the covariates driving such differences, and to reveal relationships and redundancies between the covariates. The PCA provided an overview of the environmental factors characterizing the different regions. However, in order to investigate the effects of the environmental factors on the G × E interactions, the covariates were further correlated to the AMMI IPCA1 and IPCA2 environment scores. Significant correlations suggested that any separation of the environments on the AMMI2 biplots was attributed to the relevant covariate, thereby highlighting the importance of that covariate to the G × E interactions. Hence, the results of the correlation analy-

sis, together with the differential grouping of environments in the biplots allowed for biological interpretations of the factors driving G × E interactions.

2.3.2. G × trial interactions

In the above analysis, an environment was considered as a trial × ratoon combination characterized by varying seasonal covariates. Each trial, however, was also characterized by stable/fixed/immovable site covariates usually related to soil factors. Therefore, the effects of the five site covariates could not be evaluated using the above procedures, as ratoons of the same trial would be characterized by the same site covariates, resulting in overlapping of these environments on the PCA biplot. Consequently, a second approach to the analysis was adopted, where the phenotypic performance of a genotype was averaged across ratoons of a trial, and the AMMI analysis was conducted on the G × trial two-way table of means. In order to study the correlation structures of the trials in relation to environmental factors, the nine seasonal covariates (averaged across ratoons, within a trial), together with the five site covariates were analyzed using PCA, as above. This PCA allowed for the comparison of the relative effects of site vs. seasonal covariates on the separation of the trials used in the study. The effects of the covariates on the G × trial interactions were subsequently evaluated through correlation analysis, as above. This approach represented an evaluation of the repeatable component of G × E, while the former approach represented an analysis of a mixture of repeatable and non-repeatable G × E. All statistical analyses were conducted using Genstat® Version 12.1 statistical software (Anon., 2009) and biplots were produced using Canoco® for Windows Version 4.51 (Ter Braak and Smilauer, 2003).

3. Results

3.1. G × E interactions

3.1.1. AMMI analysis

The AMMI analysis of variance results are presented in Table 2. The analysis was conducted using genotypic means for each environment averaged over replicates and the pooled error was determined from the individual trial errors. The results indicate that the genotype and environment main effects were significant for all three variables. As a percentage of total sums of squares, the environment accounted for 39.9, 40.4 and 42.8% of variation, while genotype accounted for 3.7, 6.1, and 3.8% of variation in TERC, TCANE, and ERC, respectively. The G × E interaction accounted for 56, 53.3, and 53.1% of total variation for TERC, TCANE and ERC, respectively.

Fig. 1a shows the AMMI2 biplot for TERC, which accounted for 52.6% of the G × E interaction and showed considerable separation of the M environments (i.e. environments in the M region) from the C and H environments. There was a substantial overlap between the C and H environments suggesting that trials within these two regions may discriminate cultivars similarly. Isolated groups of M environments, which were located in the upper and lower left quadrants of the biplot represented trials conducted in frost pockets and harvested at 12-months of age. The C and H environments exhibited substantial deviation along the IPCA2 axis, while the M environments were spread more explicitly along the first axis. Three groups of cultivars were apparent: N12, N16, N31 and N37 grouped together and showed better adaptability to the M environments in general; N19, N17, N41, N36, N39 and N33 grouped together in the upper left quadrant and did not appear to have any specific regional affinity; N21, N35, N27, NCo376 and N29 grouped together in the bottom left quadrant and were correlated to some of the C and H environments. The four cultivars that

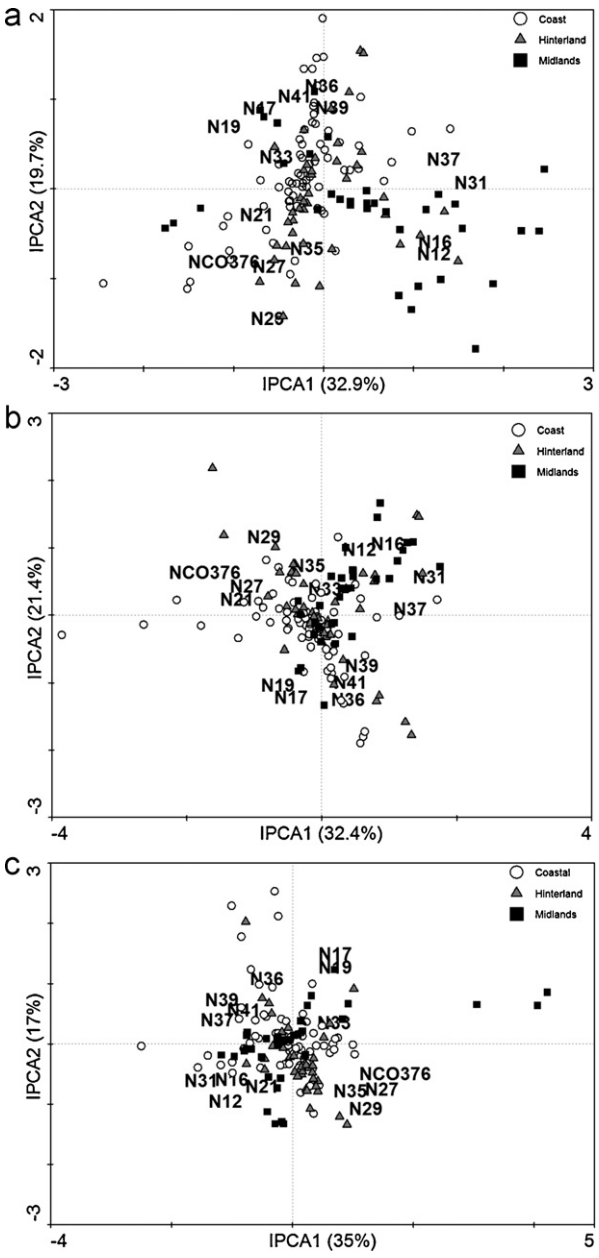

Fig. 1. AMMI2 biplots for TERC (a), TCANE (b) and ERC (c) of 15 genotypes evaluated over 147 environments.

grouped together with the M environments are typically harvested older than 15 months of age. This suggests that the regional grouping may be due to the harvest age effect, rather than a true positive interaction.

The biplot for TCANE depicted in Fig. 1b, which accounted for 53.8% of the G × E interaction, revealed substantial clustering of the environments closer to the biplot origin, suggesting that environments were less interactive for TCANE than for TERC (Fig. 1a). However, the regional groupings were similar to that of the TERC biplot, with the C and H environments showing considerable overlap, separate from most of the M environments. The H and M environments were spread more explicitly along the IPCA2 axis, while the C environments showed greater separation along the first axis. Cultivar groupings were also similar to the TERC biplot.

The ERC biplot (52% of G × E interaction) in Fig. 1c displayed considerable overlapping of the C, H, and M environments, which tended to cluster close to the biplot origin. Three ratoons from

Table 2
Analysis of variance for yield (TCANE), estimated recoverable crystal (ERC), and tons ERC/ha (TERC) of 15 sugarcane genotypes evaluated in 147 environments.

SOV	df	TERC		TCANE		ERC	
		SS	MS	SS	MS	SS	MS
Genotype (G)	14	2647	189**	231,862	79,161**	920	65.7**
Environment (E)	146	28,766	197**	1,521,952	49,826**	10,129	69.4**
G × E	2044	40,518	16.9**	2,007,361	812**	12,545	5.3**
Error (pooled)		91.6	3.2	4684	172	29.7	0.9

df, degrees of freedom; SS, sums of squares; MS, mean squares.
** $P < 0.001$.

the trial in a frost pocket in the midlands formed a cluster on the right side of the biplot. No other environments showed significant deviations along the IPCA1 axis. Four C environments, and one H environment exhibited high IPCA2 scores, however, in general there was tight clustering of environments along the IPCA2 axis as well. The cultivar groupings differed slightly from those observed in the TCANE and TERC biplots: N37, N39, N41, and N36 grouped together in the top left quadrant; N17 and N19 formed another group in the top right quadrant; NCo376, N35, N29 and N27 grouped in the bottom right quadrant; while the remaining cultivars formed a group in the bottom left quadrant. There was no clear affinity of cultivar groups to specific regions, as observed with TERC (Fig. 1a) and TCANE (Fig. 1b).

3.1.2. Relationship between environments based on PCA

Fig. 2 shows the PCA based on the environment × covariate two-way table, which accounted for 51.2% of the total variation. Separation of the C, H, and M regions were observed along the second PC axis, with most of the M environments forming a cluster at the bottom, while most of the C and H environments overlapped towards the middle and upper sector of the biplot. Some H environments formed part of the cluster of M environments in the bottom of the biplot, and were positively correlated to AGE. The covariates that had the largest effects on environmental separation (as indicated by the lengths of their vectors) included AGE, TT, WSI1, and WSI2. This was followed by RAIN, which was negatively correlated to EVP and RAD, while TOH and WSI3 showed minor effects on environment separation. The TT, RAD and EVP covariates were positively correlated and were associated to a larger extent with the C and H environments in general. The WSI1 (based on simulated rainfed vs. simulated irrigated evapotranspiration) and WSI2 (based on simulated rainfed vs. simulated irrigated cane yields) were positively correlated, as the Canesim model calculates cane yield as a function of evapotranspiration. Consequently, either one of these covariates could be used in future studies. Although correlated to the other two water stress indices, the WSI3 (based on actual trial yield vs. simulated irrigated yield) showed a slightly different response, and was closely associated with RAIN. The M environments were characterized by higher AGE as well as lower TT, RAD, EVP and RAIN. Additionally, the negative correlation of the M environments to TOH suggests greater bias towards harvesting such trials early in the season.

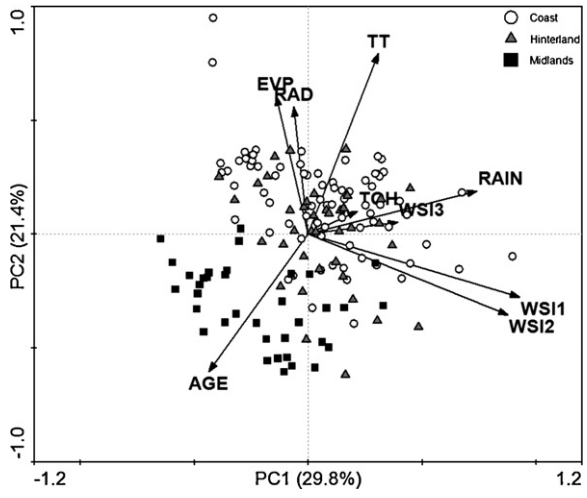

Fig. 2. PCA biplot summarizing the relationships among 147 environments and nine environmental covariates.

3.1.3. Interpretation of G × E interactions

The correlations between the AMMI IPCA scores and each of the nine seasonal covariates are indicated in Table 3. These correlations must be interpreted jointly with the environmental separation patterns observed in the AMMI biplots in Fig. 1. The effects of the covariates on the G × E interactions are analyzed by interpreting the positions of environments (IPCA1 and IPCA2 scores) on the AMMI biplots and relating those positions/patterns to any observed correlations to the covariates. Most of the significant correlations were observed between the covariates and the IPCA1 scores. The AGE, TT, WSI1 and WSI3 covariates showed significant correlations with IPCA1 scores for all three variables. The TERC and TCANE AMMI biplots in Figs. 1a and b showed that most of the M environments were characterized by higher IPCA1 scores compared with the C and H environments. The positive correlations of AGE to the TERC (0.58) and TCANE (0.37) IPCA1 scores suggest that the M environments were characterized by greater AGE. This is confirmed by the direction of the AGE vector in the PCA biplot in Fig. 2. Similarly, the significant negative correlations of RAIN, TT, WSI1, WSI2, and WSI3 to the TERC and TCANE IPCA1 scores suggest that the M environments were characterized by lower daily rainfall and temperature, and higher water stress (low WSI values) compared with the C and H environments. The three M environments with low IPCA1 scores in Fig. 1a represent ratoons of a trial in a frost pocket (valley bottom), usually characterized by deeper soils with better water holding capacity and less water stress (high WSI values). Fig. 1a showed a greater spread of the C and H environments compared to the M environments along the IPCA2 axis. Consequently, the strong positive correlation (0.55) between these IPCA2 scores and WSI3 shown in Table 3 suggest large variability in water stress within these regions, highlighting its importance for selection and evaluation. The significant correlations of AGE, RAD, TOH, TT and WSI3 to the TCANE IPCA2 scores, and the substantial deviation of the H environments along this axis in Fig. 1b, highlights the importance of these covariates as discriminating factors within this region.

The ERC AMMI biplot did not reveal clear separation of the C, H, and M regions, thereby preventing clear associations between the environment groups and the covariates. There was greater deviation of environments along the IPCA2 axis compared with the IPCA1 axis (ignoring the three frost pocket environments). The WSI3 showed a weak (0.28) but significant ($P < 0.001$) correlation to the ERC IPCA1 scores, while TOH was the only covariate that showed a significant correlation to the ERC IPCA2 scores. Therefore, the deviation of the C and H environments along the IPCA2

Table 3
Correlation coefficients between nine seasonal and five site covariates and interaction principal component 1 (IPCA1) and 2 (IPCA2) environment scores derived from AMMI analysis of cane yield (TCANE), estimated recoverable crystal (ERC) and tons ERC/ha (TERC).

Covariate	G × E analysis						G × trial analysis					
	IPCA1			IPCA2			IPCA1			IPCA2		
	TERC	TCANE	ERC	TERC	TCANE	ERC	TERC	TCANE	ERC	TERC	TCANE	ERC
AGE	0.58***	0.37***	−0.20**	0.16ns	0.33***	0.02ns	0.65***	−0.51***	−0.29ns	−0.06ns	0.08ns	−0.34*
EVP	−0.12ns	−0.03ns	0.07ns	−0.08ns	−0.09ns	−0.00ns	−0.21ns	0.17ns	0.04ns	−0.10ns	−0.18ns	0.09ns
RAD	0.09ns	−0.01ns	−0.14ns	0.11ns	0.18*	−0.06ns	0.12ns	0.02ns	−0.19ns	0.09ns	0.09ns	−0.09ns
RAIN	−0.29***	−0.25**	−0.08ns	0.02ns	−0.02ns	−0.10ns	−0.44**	0.42**	0.18ns	0.23ns	0.13ns	0.19ns
TOH	−0.14ns	0.01ns	0.19*	0.00ns	−0.22**	0.36***	−0.15ns	−0.01ns	0.19ns	0.09ns	−0.00ns	−0.01ns
TT	−0.24**	−0.25**	−0.23**	−0.19*	−0.19*	−0.07ns	−0.23ns	0.35*	−0.23ns	−0.08ns	−0.09ns	0.12ns
WSI1	−0.25**	−0.25**	0.17*	0.16ns	0.09ns	−0.07ns	−0.34*	0.29ns	0.31*	0.32*	0.26ns	0.14ns
WSI2	−0.18*	−0.22**	0.09ns	0.15ns	0.12ns	−0.11ns	−0.24ns	0.22ns	0.22ns	0.28ns	0.26ns	0.14ns
WSI3	−0.31***	−0.44***	0.28***	0.55***	0.40***	−0.04ns	−0.27ns	0.28ns	0.45**	0.58***	0.42**	−0.06ns
Clay	–	–	–	–	–	–	−0.16ns	0.08ns	0.23ns	0.1ns	0.14ns	0.28ns
Depth	–	–	–	–	–	–	−0.33*	0.30*	0.32*	0.04ns	−0.17ns	−0.06ns
Ncat	–	–	–	–	–	–	0.12ns	−0.12ns	0.05ns	0.26ns	0.34*	0.04ns
OM	–	–	–	–	–	–	0.24ns	−0.28ns	0.04ns	0.25ns	0.37*	−0.02ns
TAM	–	–	–	–	–	–	−0.32*	0.34*	0.24ns	0.19ns	0.07ns	0.07ns

ns = not significant.
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

axis in Fig. 1c is an indication of the importance of TOH as an ERC determinant in these two regions. Although EVP and RAD showed moderate effects on the separation of the M environments from the C and H environments in the PCA (Fig. 2), these covariates showed no significant correlations to the AMMI IPCA1 scores, and only a single significant ($P < 0.05$) correlation to IPCA2 scores for TCANE (Table 3).

3.2. G × trial interactions

In this approach, the phenotypic data of the genotypes were averaged over ratoons within a trial and the two-way tables of G × trials were analyzed using AMMI. The AMMI biplots for the G × trial interactions (not shown) revealed similar patterns of responses observed in the G × E biplots. Genotypic groupings mimicked the G × E biplots, while trials within regions showed similar deviations as the environments (within regions) in the previous analysis. The second PCA analysis i.e. with site covariates included and seasonal covariates averaged over ratoons, accounted for 44.4% of the variation and revealed slightly different patterns to the first PCA analysis (Fig. 3). Once again, most of the M trials formed a group and were positively correlated to AGE. All site covariates were positively correlated to each other, with Clay and TAM accounting for most of the variability (longest vectors). Either one of these site covariates could therefore be used as a group representative for further evaluations. Additionally, WSI3 (actual trial mean yield vs. simulated irrigated yield) was more strongly associated with the site covariates than the other two water stress indices. This suggests its possible superiority as a representative index to characterize water availability, which is essentially defined by the site covariates. The site covariates appeared to be negatively associated with a large proportion of the H trials, suggesting that these trial sites are dominated by poorer soils. Also, RAIN, TT, and TOH were once again correlated and negatively associated with the M environments, while EVP and RAD accounted for very little variation.

The correlations of the G × trial AMMI IPCA scores to the site and season covariates are also shown in Table 3. The AGE covariate showed strong significant correlations to the IPCA1 scores for TERC (0.65) and TCANE (−0.51), once again highlighting its important role in discriminating regions and trials. The RAIN covariate showed stronger correlations with the G × T IPCA1 scores for TERC and TCANE compared with the G × E IPCA1 scores for the same variables, while TT only showed a significant correlation with the TCANE IPCA1 score. The WSI3 showed stronger correlations than WSI1 in terms of ERC IPCA1 scores, TERC IPCA2 scores, and TCANE IPCA2 scores. These stronger correlations again highlight the superiority of WSI3 for use as an indicator of water deficient environments in this study. From all site covariates, DEPTH showed significant correlations with IPCA1 scores for the three variates and interpretation of this correlation with the AMMI biplots suggested that most of the M trials were characterized by shallower soils. The biplots also showed that most of the deviation along the IPCA1 axis was observed by the M trials, suggesting that sites in this region were the most variable in terms of soil depth. The Ncat and OM showed significant ($P < 0.05$) correlations with the TCANE IPCA2 scores, suggesting that deviations of trials along this axis, which were evident within all three regions, were due to differences in Ncat and OM. The TAM caused separation of trials along the TCANE and TERC IPCA1 axis, and interpretations from the biplots suggested that M trials were characterized by lower TAM.

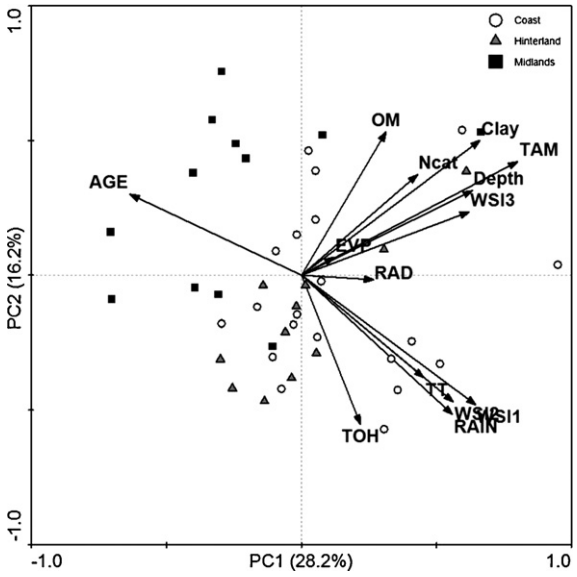

Fig. 3. PCA biplot summarizing the relationships among 43 trials, nine seasonal covariates, and five site covariates.

4. Discussion and conclusions

The results of this study showed that the M environments/trials were unique in their ability to discriminate cultivar performance in terms of TERC and TCANE. The M environments regularly formed distinct clusters, separate from the C and H environments, on both the AMMI and PCA biplots. The overlap of the C and H environments on the AMMI biplots suggests that these regions discriminated cultivars similarly, while the PCA biplots showed that the environments in these regions were also characterized by similar climatic and soil features. These regions should be considered as a single mega-environment within the rainfed selection program and for any future studies involving the interpretation of G × E interactions. The similarities between environments in the C and H regions were also highlighted by Ramburan and Zhou (in press), where genotype + genotype × environment (GGE biplot) analysis was performed on the same dataset. The biplots produced from the GGE biplot analyses differ slightly from that of AMMI analyses. The discrepancy lies in the treatment of the G × E matrix prior to performing singular value decomposition. The GGE biplot technique utilizes environment centered data (matrix minus environment means), while AMMI uses the matrix of residuals (matrix minus genotype and environment means). Despite this difference, the techniques allowed for similar conclusions to be drawn. AMMI was chosen as the analysis of choice in this study due to the demonstrated successes of interpreting environmental covariates through its use (De Vita et al., 2010; Van Eeuwijk et al., 1995; Van Oosterom et al., 1996; Voltas et al., 1999).

The approach used in this study identified factors that separated rainfed environments. However, the correlations between AMMI IPCA scores and covariates were generally weak (strongest correlation of 0.65). This limited the further use of the technique to enrich the AMMI biplots with the covariates (Van Eeuwijk and Elgersma, 1993), which could have allowed for easier interpretations. Nevertheless, the useful inferences made about the nature of the different environments and regions will inform future site selection for breeding and evaluation. The variability and significance ($P < 0.01$) of water stress within the C and H regions suggest that improved genetic tolerance to limited water availability may be achieved through selection in contrasting environments within these regions. The results also suggested that the M region was characterized by greater variability in soil depth. This is in contrast to local knowledge of regional soil characteristics, as greater topographic variability characteristic of the C and H regions were expected to produce greater variability in soil depth. The discrepancy may be linked to biased trial site selection in the C and H regions, where the most homogeneous, accessible, high potential fields were conventionally chosen. Consequently, more emphasis should be placed on selecting trial sites with contrasting soil characteristics within the C and H regions.

The PCA of the environment × covariate matrix identified AGE as a dominant factor influencing the separation of the M region from other regions in the rainfed parts of the industry. The M region is characterized by higher altitude, which is normally associated with lower temperatures and slower growth rates, thereby requiring longer crop cycles. The negative correlation between AGE and TT is depicted in Fig. 2, which also shows that the M environments are characterized by lower daily RAD, RAIN, and EVP. These are typical indications of an overall lower climatic potential of cropping environments. However, this is not the case with the M environments, where lower climatic potential is compensated for by the longer cutting cycle (AGE). Ramburan and Zhou (in press) and other studies have demonstrated higher yields in M environments. The present study has shown that these higher yields are related to increased harvest age rather than superior climatic potential of the region.

In this study, the influence of climatic factors and water stress was captured using the average daily values for the entire crop cycles (plant/ratoon to harvest). These averages may not be appropriate (and may be misleading), especially when considering the within-season weather variations and their influence on particular growth phases. For example, high water stress during the early growth phase of tillering may negatively affect stalk population and final cane yield. Favorable moisture conditions in later growth phases will improve the seasonal average; however, the effects of early water stress may go undetected. Radiation, evaporation and rainfall did not show clear, consistent associations with IPCA scores. This lack of correlation may have also been due to the use of seasonal averages, as high levels during one growth phase may have been offset by lower levels during another phase. A more accurate approach would be to characterize climatic and water stress covariates within individual growth phases to determine their specific effects on G × E interactions. This approach has been used in other crops such as barley (Voltas et al., 1999), ryegrass (Van Eeuwijk and Elgersma, 1993) and wheat (Voltas et al., 2005). No attempts have been made to interpret sugarcane G × E interactions using this approach, and future studies in this regard are recommended.

G × E interactions may be categorized as being repeatable or non-repeatable, depending on the likelihood of particular G × E interaction events being observed again (DeLacy et al., 2010). Our initial analysis (G × E) represented an investigation into the combined influence of repeatable and non-repeatable G × E interactions, while the second approach to the analysis (G × Trial) represented an evaluation of the repeatable component of G × E interactions. The second approach was not successful at identifying site factors that correlated with IPCA scores, as most covariates were not significant. This was unexpected as Clay, in particular, is generally considered as an important site factor causing variation in cultivar responses in the sugar industry. The lack of frequent correlations between IPCA scores and covariates such as Clay, Ncat, and OM suggest that these may not be as influential on G × E interactions as previously thought. Stronger correlations were observed between IPCA scores and the covariates relating to soil physical characteristics i.e. Depth and TAM. Based on their influence on the repeatable G × Trial interaction, these covariates should be given priority when selecting contrasting sites for selection and evaluation. Furthermore, their mutual correlation shown in Fig. 3, suggests that either one may be used as a representative covariate for future studies (TAM was derived from Depth measurements).

With respect to water stress indices, WSI3 was identified as the most promising at detecting genotypic responses to environmental differences in water stress, due to its correlations with IPCA scores, as well as its correlations with the site covariates in Fig. 3. In this respect (correlation with site covariates), WSI3 best captured the combined influence of the site covariates, suggesting its potential use in future studies as an integrated covariate that represents the ability of soil to provide moisture. The WSI1 and WSI2, which were based on simulated evapotranspiration and cane yields, respectively, did not correlate as strongly with IPCA scores as WSI3. This is likely due to WSI3 being derived from observed trial mean yields, while the other two indices were based completely on simulated responses and the inaccuracies usually associated with such simulations.

In this study, the categorization of environments into the C, H, and M regions were pre-determined and based on industry knowledge. The joint use of AMMI and PCA successfully separated the regions (the M region specifically) according to their environmental characteristics and the genotypic responses, thereby confirming local knowledge. The patterns of genotypic responses to environments in the AMMI analyses were explained by the environmental characterization based on seasonal and site covariates. This suggests that the methodology could have potential benefits to

sugarcane G × E studies, particularly in areas where little is known about the mega-environment constitution of a crop area. This is in contrast to Ceretta and Van Eeuwijk (2008), who found that differential barley cultivar responses between two environments (using GGE Biplot) were not fully explained by an environmental characterization based on PCA of meteorological variables. Additionally, this study highlighted the value of environmental characterization of sugarcane field trials and demonstrated the potential for integrated use of crop models to interpret G × E interactions. Future initiatives will focus on optimizing the use of crop models in G × E interpretive studies through evaluations of effectiveness of different weather data sources, comparisons of simulated vs. observed yields, and using simulated yields as hypothetical probe genotypes (commonly studied genotypes for which detailed physiological and yield parameters are known).

In addition to the effects of some covariates on regional separation, substantial variation was also observed within regions i.e. environments within regions showed deviation along IPCA axes. Attempts will be made, within regions, to group environments based on either seasonal covariates or G × E patterns. These analyses will also involve the use of genotypic covariates (traits) to help investigate the responses of sugarcane phenotypic groups to environmental factors. Such an approach will complement this study and provide a clearer understanding of variation within regions to inform trial site selection and regional specific breeding objectives.

References

Anon., 2009. GenStat® Executable Release 12.1. Lawes Agricultural Trust, Rothamstead Experimental Station/Clarendon Press, Harpenden, U.K./London, U.K.

Anon., 2000. Information Sheet 7.1: Recommendations for Nitrogen. The South African Sugar Association Experiment Station, Mount Edgecombe, Durban, South Africa, pp. 2–3.

Annicchiarico, P., Harzic, N., Carroni, A.M., 2010. Adaptation, diversity, and exploitation of global white lupin (*Lupinus albus* L.) landrace genetic resources. Field Crops Res. 119, 114–124.

Baril, C.P., Denis, J.B., Wustman, R., Van Eeuwijk, F.A., 1995. Analysing genotype by environment interaction in Dutch potato variety trials using factorial regression. Euphytica 82, 149–155.

Bissessur, D., Ramnawaz, C., Ramdoyal, K., 2010. Use of historical multi-location sugar cane variety trials data to identify relationships among environments in terms of genotype responses. Proc. Int. Soc. Sugar Cane Technol. 27, 1–8.

Ceretta, S., Van Eeuwijk, F.A., 2008. Grain yield variation in malting barley cultivars in Uruguay and its consequences for the design of a trials network. Crop Sci. 48, 167–180.

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

del Blanco, I.A., Glaz, B., Edme, S.J., 2010. Improving efficiency of sugarcane genotype selection in Florida. Crop Sci. 50, 1744–1753.

De Vita, P., Mastrangelo, A.M., Matteu, L., Mazzucotelli, E., Virzi, N., Palumbo, M., Lo Storto, M., Rizza, F., Cattivelli, L., 2010. Genetic improvement effects on yield stability in durum wheat genotypes grown in Italy. Field Crops Res. 119, 68–77.

Gauch, H.G., 1992. Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs. Elsevier, Amsterdam.

Gilbert, R.A., Shine, J.M., Miller, J.D., Rice, R.W., Rainbolt, C.R., 2006. The effect of genotype, environment and time of harvest on sugarcane yields in Florida, USA. Field Crops Res. 95, 156–170.

Glaz, B., Kang, M.S., 2008. Location contributions determined via GGE biplot analysis of multienvironment sugarcane genotype-performance trials. Crop Sci. 48, 941–950.

Jackson, P., Horsley, D., Foreman, J., Hogarth, D.M., Wood, A.W., 1991. Genotype × environment (GE) interactions in sugarcane variety trials in the Herbert. Proc. Aust. Sugar Technol. Assoc. 13, 103–109.

Jackson, P., McRae, T., Hogarth, M., 1995. Selection of sugarcane families across variable environments. II. Patterns of response and association with environmental factors. Field Crops Res. 43, 119–130.

Jackson, P., Chapman, S., Rattey, A., Wei, X., Cox, M., 2007. Genotype × region interactions, and implications for sugarcane breeding programs. Sugar Cane Int. 25, 3–7.

Mirzawan, P.D.N., Cooper, M., DeLacy, I.H., Hogarth, D.M., 1994. Retrospective analysis of the relationships among test environments of the southern Queensland sugarcane breeding programme. Theor. Appl. Genet. 88, 707–716.

Queme, J.L., Orozco, H., Ovalle, W., Melgar, M., 2005. Analysis of genotype-by-environment interaction for sugarcane based on the AMMI model. Sugar Cane Int. 23, 21–24.

Queme, J.L., Crossa, J., Orozco, H., Melgar, M., 2007. Analysis of genotype-by-environment interaction for sugarcane using the sites regression model (SREG). Proc. Int. Soc. Sugar Cane Technol. 26, 764–769.

Ramburan, S., Zhou, M. Investigating sugarcane genotype × environment interactions under rainfed conditions in South Africa using variance components and biplot analysis. Proc. S. Afr. Sugar Technol. Assoc., in press.

Rattey, A.R., Kimbeng, C.A., 2001. Genotype by environment interactions and resource allocation in final stage selection trials in the Burdekin district. Proc. Aust. Soc. Sugar Cane Technol. 23, 136–141.

Singels, A., Donaldson, R.A., 2000. A simple model of unstressed sugarcane canopy development. Proc. S. Afr. Sugar Technol. Assoc. 74, 151–154.

Tai, P.Y.P., Rice, E.R., Chew, V., Miller, J.D., 1982. Phenotypic stability analyses of sugarcane cultivar performance tests. Crop Sci. 22, 1179–1184.

Ter Braak, C.J.F., Smilauer, P., 2003. Canoco® for Windows Version 4.51. Biomteris Plant Research International, Wageningen, The Netherlands.

Van Antwerpen, R., Meyer, J.H., Johnston, M.A., 1994. Estimating water retention of some Natal sugar belt soils in relation to clay content. Proc. S. Afr. Sugar Technol. Assoc. 68, 75–79.

Van Eeuwijk, F.A., Elgersma, A., 1993. Incorporating environmental information in an analysis of genotype by environment interaction for seed yield in perennial ryegrass. Heredity 70, 447–457.

Van Eeuwijk, F.A., Keizer, L.C.P., Bakker, J.J., 1995. Linear and bilinear models for the analysis of multi-environment trials. II. An application to data from the Dutch maize variety trials. Euphytica 84, 9–22.

Van Oosterom, E.J., Mahalakshmi, V., Bidinger, F.R., Rao, K.P., 1996. Effect of water availability and temperature on the genotype-by-environment interaction of pearl millet in semi-arid tropical environments. Euphytica 89, 175–183.

Vargas, M., Crossa, J., Van Eeuwijk, F.A., Ramirez, E., Sayre, K., 1999. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype × environment interaction. Crop Sci. 39, 955–967.

Voltas, J., Van Eeuwijk, F.A., Sombrero, A., Lafarga, A., Igartua, E., Romagosa, I., 1999. Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley. I. Individual grain weight. Field Crops Res. 62, 63–74.

Voltas, J., Lopez-Corcoles, H., Borras, G., 2005. Use of biplot analysis and factorial regression for the investigation of superior genotypes in multi-environment trials. Eur. J. Agron. 22, 309–324.

Yan, W., Hunt, L.A., Sheng, Q., Szlavnics, Z., 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Sci. 40, 597–605.

Yan, W., Hunt, L.A., 2001. Interpretation of genotype × environment interaction for winter wheat in Ontario. Crop Sci. 41, 19–25.