Abstract

This investigation was done to study GE interaction over twelve environments for seed yield in 18 genetically diverse genotypes. Grain yield performances were evaluated for three years at four locations in Iran using a randomized complete block design. The first two principal components (IPC1 and IPC2) were used to create a two-dimensional GGE biplot that accounted percentages of 49% and 20% respectively of sums of squares of the GE interaction. The combined analysis of variance indicated that year and location were the most important sources affecting yield variation and these factors accounted for percentages of 50.0% and 33.3% respectively of total G+E+GE variation. The GGE biplot suggested the existence of three lentil mega-environments with winning genotypes G1, G11 and G14. According to the ideal-genotype biplot, genotype G1 was the better genotype demonstrating high mean yield and high stability of performance across test locations. The average tester coordinate view indicated that genotype G1 had the highest average yield, and genotypes G1 and G12 recorded the best stability. The study revealed that a GGE biplot graphically displays interrelationships between test locations as well as genotypes and facilitates visual comparisons.

Keywords: GGE biplot, genotype × environment interaction, multi-environment trials, principal component

Introduction

Plant breeders perform multi-environment trials (MET) to evaluate new improved genotypes across test environments (several locations and over years), before a specific genotype is released for production to supply growers. In such experiments, genotype × environment (GE) interaction is a commonly evaluated (Annicchiarico 2002; Kang, 1998; Karimizadeh et al., 2012a; Yan et al., 2007). A GE interaction refers to differential ranking of genotypes across environments and may complicate the selection process and recommendation of a genotype for a target environment (Ebdon and Gauch, 2002; Gauch, 2006). It may also reduce the selection efficiency in different breeding programs because in a GE interaction, measured traits are less predictable and cannot be interpreted using main effects (genotype or environment) and need more analysis (Gauch et al., 2008). GE interaction is one of the most important reasons for the failure or decreased efficiency of breeding efforts to serve small resource-poor farmers in arid and semi-arid areas (Ceccarelli et al., 1996). Plant breeders perform MET to select favorable genotypes based on both mean yield and performance stability; and to determine whether a test environment is homogeneous or should be divided into various mega-environments (Gauch, 2006; Yan and Kang, 2002).

Different statistical models were used to describe GE interaction and facilitate genotype recommendations in MET. These models have been classified as univariate versus multivariate approaches or parametric versus nonparametric methods (Flores et al., 1998; Karimizadeh et al., 2012b; Lin et al., 1986). Parametric approaches (univariate and multivariate) are based on statistical assumptions considering the distribution of a dataset. Parametric univariate statistics of stability measurements are mostly related to linear regression analysis, variance components or related methods. Alternative strategies are nonparametric stability statistics or analytical clustering procedures that make no specific modeling assumptions about datasets (Lin and Butler, 1990; Huehn, 1996). Several multivariate procedures have been proposed to explore GE interaction including principal component analysis (PCA), additive main effects and multiplicative interactions (AMMI) and genotype plus GE interaction biplot (GGE) analysis (Yan, 2002).
The multiplicative GE interaction is far too complex to be summarized by one or two stability parameters using univariate measures of stability. Multivariate statistical approaches explore multidirectional aspects of GE interaction and attempt to extract more information from GE interaction components (Gauch and Zobel, 1996; Gauch et al., 2008; Sabaghnia et al., 2008).

In total phenotypic yield variation, an environment main effect describes more observed variation; while a genotype main effect and GE interaction is usually smaller (Yan, 2002). According to Cooper and Hammer (1996) the relevant variance components in genotype evaluation at MET are genotype main effect (G) and GE interaction, which are regarded as repeatable sources. Therefore Yan et al. (2000) proposed using both G and GE effects instead of only GE interaction for yield stability analysis. To facilitate use of this model, the biplot approach (Gabriel, 1971) was used to display the GGE of a MET dataset (Yan et al., 2000). Detailed accounts of different models have been presented by Yan and Kang (2002) and Yan et al. (2007). A GGE biplot as a data visualization tool is able to graphically demonstrate a GE interaction pattern. It is an effective tool to identify a mega-environment, genotype evaluation based on the both yield and stability; and evaluation of test environments from a discrimination aspect. A GGE biplot indicating both genotype and environment based on a Site Regression (SREG) model have been used to demonstrate a GE interaction pattern as well as possible (Yan and Tinker, 2005; Yan et al., 2007).

A GGE biplot has been used to study a GE interaction and yield stability analysis of different crops in semiarid and semiarid areas of Iran Dehghani et al. (2006) for barley, Sabaghnia et al. (2008) for lentil, Dehghani et al. (2009) for maize. The objectives of this investigation were: (1) to apply a GGE biplot model to evaluate the magnitude of effect of GE interaction on grain yield of 18 lentil genotypes tested across five locations, and (2) to evaluate relationships between test environments for identification of favorable genotypes within the identified mega-environment for lentil production areas in semiarid conditions.

**Materials and methods**

Seventeen improved lentil genotypes with one cultivar check (‘Gachsaran’) were used as experimental materials (Tab. 1). The research was performed in a completely randomized block design with four replications across four test location containing Kermanshah, Gorgan, Gachsaran and Shirvan; and during three years from 2003 to 2005 growing period. Sowing was done manually in rows 25 cm apart. Each plot was 4 m² (1 × 4 m plots consisting of 4 rows) in size and the harvested plot size was 1.75 m² (two 3.5 m rows at the center of each plot). The test locations were selected as samples of lentil growing areas of Iran having variation in latitude, rainfall, soil type, temperature and other agro-climatic factors (Tab. 2).

**Results and discussion**

Based on the combined analysis of variance, a significant effect of year (Y) and location (L) was not observed on the grain yield of lentil genotypes in the tests but Y × L interaction was significant. Yield data were subjected to combined analysis of variance and GGE biplots were constructed using the GGEbiplot software (Yan, 2001) and vector view biplots to visualize correlations between locations, a vector view biplot was made. These graphic analyses were done using the GGEbiplot software (Yan, 2001) and a statistical analysis of variance for the SREG model was done using the SAS codes program as cited by Burgueno et al. (2001).

**Results and discussion**

Based on the combined analysis of variance, a significant effect of year (Y) and location (L) was not observed on the grain yield of lentil genotypes in the tests but Y × L interaction was significant. Yield data were subjected to combined analysis of variance and GGE biplots were constructed using the GGEbiplot software (Yan, 2001) and vector view biplots to visualize correlations between locations, a vector view biplot was made. These graphic analyses were done using the GGEbiplot software (Yan, 2001) and a statistical analysis of variance for the SREG model was done using the SAS codes program as cited by Burgueno et al. (2001).

**Results and discussion**

Based on the combined analysis of variance, a significant effect of year (Y) and location (L) was not observed on the grain yield of lentil genotypes in the tests but Y × L interaction was significant. Yield data were subjected to combined analysis of variance and GGE biplots were constructed using the GGEbiplot software (Yan, 2001) and vector view biplots to visualize correlations between locations, a vector view biplot was made. These graphic analyses were done using the GGEbiplot software (Yan, 2001) and a statistical analysis of variance for the SREG model was done using the SAS codes program as cited by Burgueno et al. (2001).
**L interaction effect was highly significant, possibly due to changes in environmental conditions, which vary from one environment (year × location combination) to another (Tab. 3).** Genotype demonstrated significant effect, which could be due to changes in genotype characteristics, varying from one genotype to another. The interaction effect of genotype with location (G × L) and the three way interaction of factors (G × L × Y) were significant, but G × Y interaction was not significant (Tab. 3). Environment significantly explained about 70% (21.06, 14.90 and 34.45% for year, location and their interactions, respectively) of the total sum of squares due to G + E + GE interaction (Tab. 3). However, the partitioning of variance components for environment revealed that both predictable (locations) and unpredictable (year) components were important sources of variation. When GE interaction is due to variation in predictable factors, a plant breeder has the choice of either developing specific genotypes for selected environments or broadly adapted genotypes that can perform well under variable conditions (Delghan I et al., 2006). Anyhow, when GE interaction results from unpredictable sources, a plant breeder needs to develop stable genotypes that can perform reasonably well under a range of environmental conditions.

A remarkable grain yield variation explained by environments (60%) indicated that environments tested in the study were diverse, with large differences among environmental effects causing the most variation in grain yields of lentil genotypes. The G × L × Y or GE interaction significantly explained 23% of the G + E + GE variation in grain yield. The relatively large magnitude of the GE interaction sum of squares was about two times larger than that for genotype, indicating that there were sizeable differences in responses according to genotype across environments. Kang and Pham (1991) indicated that GE interactions minimize the usefulness of genotype by confounding yield performance. To better investigate GE interaction, Becker and Leon (1988) showed that assessment of yield stability across many locations and years could increase both repeatability and heritability of studied characters such as grain yield. The significance of GE interaction indicated that differential genotype expression across environments depends on the reaction of genotype on changing environmental conditions across locations and years. Therefore, these results from semiarid conditions of Iran have confirmed that although grain yield has depended on genetic potential, this is modified under environmental factors and their conditions like cold regions of Canada (Yan and Rajcan, 2003).

Results of combined ANOVA for the yearly datasets are shown in Tab. 4, which provides a general picture of the relative magnitudes of effects of genotype, location and the interaction (G × L) due to G + L + GL variations. Location was the most important source of yield variation (70.85 and 63.79% of the G + L + GL variation for the first and second year (2003 and 2004, respectively), except in the last year when it accounted for 28%. The relatively large yield variation due to location, which is disjointed to genotype evaluation and mega-environment identification (Gauch and Zobel, 1996), justifies selection of site regression statistical model as the suitable tool for investigation of the multi-environment trials dataset. Regarding the magnitudes of G × L interaction in comparison to genotype main effect, suggests the possible existence of different mega-environments. PC1 and PC2 derived by subjecting the location-centered yield to singular value decomposition (SVD), which make up a genotype plus GL interaction (GGL) biplot, explained from 49% and 20% of the total G + GL (Fig. 1).

According to the biplot shown in Fig. 1, the corner genotypes that are the most responsive ones, can be visually determined. These corner genotypes were G1, G6, G9, G10, G11 and G14. In this figure, locations are divided
Based on Fig. 2, it is possible to assess both mean yield and stability performance through a biplot. An average tester coordinate (ATC) horizontal axis passes through the biplot origin and the average location and the oval show the positive end of the ATC horizontal axis. The average yields of genotypes are estimated by projections of their markers onto the ATC horizontal axis. Thus, genotype G1 had the highest average yield, and G6 had the lowest (Fig. 2). Stability of each genotype is explored by its projection onto the ATC vertical axis. The smaller the absolute length of projection of a genotype, the more stable it is. Thus, genotypes G9 and G10 were the least stable and genotypes G1, G2, G5, G7, G8, G12, G15 and G18 were the most stable. However, considering both mean yield and stability performance, genotypes G1 following to G12, G2, G15, G5, G18 and G7 could be regarded as the most favorable.

Tab. 4. Genotype (G), location (L), and genotype × location (GL) variance in multi-environment trials

<table>
<thead>
<tr>
<th>SOV</th>
<th>df</th>
<th>2003 Mean Squares</th>
<th>% of L+G+GL</th>
<th>2004 Mean Squares</th>
<th>% of L+G+GL</th>
<th>2005 Mean Squares</th>
<th>% of L+G+GL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location (L)</td>
<td>3</td>
<td>8694670.96**</td>
<td>70.85</td>
<td>3520783.12**</td>
<td>63.79</td>
<td>905614.83**</td>
<td>28.34</td>
</tr>
<tr>
<td>Replication / L</td>
<td>12</td>
<td>44458.28</td>
<td></td>
<td>53893.05</td>
<td></td>
<td>16103.70</td>
<td></td>
</tr>
<tr>
<td>Genotype (G)</td>
<td>17</td>
<td>174604.69*</td>
<td>8.06</td>
<td>177570.51*</td>
<td>18.23</td>
<td>129364.71**</td>
<td>22.94</td>
</tr>
<tr>
<td>G × L</td>
<td>51</td>
<td>512205.39*</td>
<td>21.09</td>
<td>58377.74*</td>
<td>17.98</td>
<td>91595.27*</td>
<td>48.72</td>
</tr>
<tr>
<td>Error</td>
<td>204</td>
<td>42910.01</td>
<td></td>
<td>25026.00</td>
<td></td>
<td>27202.73</td>
<td></td>
</tr>
</tbody>
</table>

* and **, respectively significant at the 0.01 and 0.5 probability level and non-significant.

**Fig. 1. GGE biplot identification of winning genotypes and their related mega-environments**

**Fig. 2. GGE biplot of mean and stability of 18 lentil genotypes for yield and specific genotype × environment interactions**
Fig. 3. Vector view of GGE biplot for relationships among locations

Fig. 3 provides a summary of interrelationships between test locations. The location vectors are lines that connect the biplot origin and the markers of test locations and the angle between them is related to the correlation coefficient. In other words, the cosine of the angle approximates the correlation coefficient between related locations (Kroonenberg, 1995). According to the angles of test location vectors, the four locations are grouped into two major groups. One group includes Shirvan, Gachsaran and Kermanshah while the other group involves Gorgan. This result is relatively coincident with the geographic pattern belonging to different location types, (warm or cold versus moderate climatic conditions respectively). Similarly this tool was used for evaluation of interrelationships between 18 lentil genotypes (Fig. 4). The overall picture of interrelationships between genotypes indicated that there were different genotype groups. In other words, these studied genotypes had diverse characteristics in terms of performance for grain yield and stability. This could be related to different geographical sources of the parent plants (Tab. 2). It is interesting that none of the genotypes had strong positive associations with the most favorable genotype (G1). Only genotype G14 had a relatively significant moderate correlation with G1. However, most of the above results on location or genotype can be verified from original data but some of them are not consistent with original data. These discrepancies are predicted because the GGL biplot explained only 69% rather than 100% of the total variation due to G+L+GL sources. Therefore all data contain some error, but according to Yan and Hunt (2002) and Yan and Tinker (2005) as the biplot shows and makes conclusions on the overall pattern of the whole dataset, these predictions are probably more reliable than individual observations.

In Fig. 5, the center of the concentric circles is where an ideal genotype (high mean yield and the most stable one) should be located. In other words, projection of the ideal genotype on the ATC horizontal axis is equal to the longest vector of all genotypes and its projection on the ATC vertical axis is obviously zero (it is absolutely stable). Therefore, a smaller the distance from genotype to the virtual ideal genotype represents an ideal genotype. Therefore, genotype G1 following to genotypes G11, G12 and G14 were closest to the concentric center, but genotype G6 was the closest to this position (Fig. 5). Also, genotypes G2, G5, G15 and G18 do not seem to be meaningfully different from other genotypes such as G8, G10, G16 and G17 that were apparently inferior. Similar to the ideal genotype, it is possible to define ideal location or environment for ranking of test locations according to their discriminating ability and suitability of representation. Centers of the concentric circles in Fig. 6 are where an ideal location should be. The projection of ideal location on the ATC horizontal axis is equal to the longest vector of all locations (the most discriminating location) and its projection on the ATC vertical axis was zero, meaning that it

Fig. 4. Vector view of GGE biplot for relationships among genotypes

Fig. 5. GGE biplot of ideal genotype and comparison of the genotypes with the ideal genotype
are excellent tools for visual MET data analysis in plant breeding programs.

Conclusions

Three distant mega-environments were identified for lentil producing areas of Iran including Shrvan with G11 as the most favorable, Gorgan with G14 as the most favorable and Gachsaran and Kermanshah with G1 as the most favorable. Genotype G1 demonstrated a high mean yield as well as stability presenting a suitably good plant material for future breeding programs. Furthermore the GGE biplot procedure was an effective tool for visual interpretation of the complex GE interaction and yield stability for studies applied to plant breeding programs.

Acknowledgments

We wish to thank Dr. Wei-Kai Yan (Eastern Cereal Oilseed Research Center of Agriculture and Agri-Food Canada) for making available a time-limited version of GGEbiplot as “testBiplotxlsx”. Contributions of the cooperators of the Iran’s Dryland Agricultural Research Institute research stations are also gratefully acknowledged.

References


Ebdon JS, Gauch HG (2002). Additive main effect and mul-


Kang MS (2002). Quantitative genetics, genomics, and plant breeding. Wallingford, UK, CABI.


