Analysis of Multi-environment Grain Yield Trials in Mung Bean *Vigna radiate* (L.) Wilczek Based on GGE Biplot in Southern Ethiopia

A. Asfaw1*, F. Gurum1, F. Alemayehu1, and Y. Rezene2

ABSTRACT

The objective of this research was to evaluate and to quantify the magnitude of the genotype×environment interaction effects on mung bean grain yield and to determine the winning genotype for the test environments. Seven mung bean genotypes were tested at three locations for over two years. The grain yield data for each environment (location year combination) was first subjected to analysis of variance using generalized linear model. Mean grain yields of genotypes for the environments were computed to generate a genotype and environment two-way table data for the GGE biplot analysis. The analysis revealed the presence of significant genotype x environment interactions for grain yield. Location effect explained more than 60% of the total grain yield variation. GGE biplot analysis depicted the adaptation pattern of genotypes at different environments and discrimination ability of testing environments. MH-96-4, shown to have the potential of combining high yield with stable performance, can be recommended for production in mung bean growing ecologies in southern Ethiopia.

Keywords: GGE-biplot, Mung bean, *Vigna radiate*, Yield stability.

INTRODUCTION

Mung bean (*Vigna radiate*) is a warm season annual grain legume. The optimum temperature range for good production is 27-30°C (Imrie, 1998). Mung bean is a quick crop, requiring 75–90 days to mature. It is a useful crop in drier areas and has a good potential for crop rotation and relay cropping with cereals using residual moisture. Smallholder farmers in drier marginal environments in Ethiopia grow mung bean. As compared to other pulse crops, its production in Ethiopia is very negligible. However, for resource poor farmers in drier marginal environments it has been an important grain legume. These farmers need a variety that is of maximum production and stable yield in their environments.

 Breeders usually undertake a series of genotypes evaluation across locations and over years before a new genotype is released to be produced by growers. In such genotype/variety evaluation trials, Genotype×Environment Interaction (GEI) is a common phenomenon (Kang and Gauch, 1996; Yan and Kang, 2003; Ceccarelli et al., 2006). GEI refers to the differential ranking of genotype among locations or years (Fernandez, 1991). It may complicate the process of selection and recommendation of superior genotypes to target environments (Magari and Kang, 1993; Ebdon and Gauch, 2002). It may also reduce the efficiency of breeding programs (Comstock and Moll, 1963). This is because in the presence of GEI, yield is less predictable and cannot be

1 Ethiopia Awassa Agricultural Research Center, P. O. Box: 6, Awassa, Ethiopia.
2 Corresponding author, e-mail: asrat.asfaw@gmail.com
3 Ethiopia Areka Agricultural Research Center, P. O. Box: 79, Soddo, Ethiopia.
interpreted as based on genotype and environmental means alone (Ebdon and Gauch, 2002). It is also one of the main reasons for the failure of formal breeding to serve small resource-poor farmers in the marginal fragile environments (Ceccarelli et al., 2006).

However, scientist around the globe have been trying to exploit GEI rather than ignoring it. The use of statistical models to explain GEI and facilitate cultivar recommendations is among the strategies adopted by scientific communities. The various statistical methodologies have been extensively reviewed and published (Lin et al., 1986; Becker and Leon, 1988; Crossa, 1990; Flores et al., 1998; Hussein et al., 2000; Ferreira et al., 2006). The different methodologies have been broadly classified as univariate parametric/non-parametric and multivariate parametrics. Parametric analyses are based on statistical assumptions regarding the distribution of genotypic, environmental and GEI effects. Parametric measures of phenotypic stability are mostly related to variance components or related statistics. These stability estimates are of suitable properties under certain statistical assumptions as based on the normal distribution of errors and interaction effects, but may not perform well if these assumptions are violated by such factors as the presence of outliers. The alternative, nonparametric or analytical clustering makes no specific modeling assumptions when relating environments and phenotypes. Several procedures have been proposed based on comparing the ranks of genotypes in each environment, with genotypes of similar ranking, across environments, being considered as stable (Ferreira et al., 2006). The multivariate methods include Principal Component (PC), Additive Main effects and Multiplicative Interactions (AMMI) as well as Genotype plus Genotype×Environment interaction (GGE) analysis (Ferreira et al., 2006; Zoble et al., 1998; Gauch, 2006). Detailed accounts on different models have been given by different authors (Becker and Leon, 1988; Zobel et al., 1998; Gauch, 2006; Yan et al., 2007). More recently, GGE biplots which show both genotypes and environments as based on Site Regression (SREG) model have been advocated to describe GEI pattern (Yan and Tinker, 2006; Yan et al., 2007). GGE biplot captures both genotype main effects and genotype x environment interaction effects, which are two important sources of variations relevant to genotype evaluation (Yan et al., 2001). The objectives of this study were to evaluate and to quantify the magnitude of the GEI and describe the which-won-where pattern using GGE biplot.

**MATERIALS AND METHODS**

Data analyzed in this study were obtained from mung bean regional variety trials, conducted for two years (2004 and 2005) at three locations in South Nations, Nationalities and Peoples (SNNP) regional state of Ethiopia. The locations were Awassa, Gofa and Inseno represented mung bean production regions in southern Ethiopia (Table 1). Awassa is at an elevation of 1,694 m above sea level in the southern rift valley of Ethiopia with clay loam fluvisol whereas Inseno is located at an elevation of 1,829 m above sea level in the central rift valley, furnished with a clay eutric fluvisol. Gofa is situated in an elevation of 1,400 m above sea level with pellic vertisol near to the Omo river valley. Awassa and Inseno represented tepid to cool sub-humid agro-ecology whereas Gofa represented hot to warm humid agro-ecology (MOA, 1998). The experimental genotypes were MH-98-3, MH-97-7, MH-97-6, MH-85-11, Asha, and MH-96-4, all breeding lines, obtained from Haryana Agricultural University of Hisar, India. The landrace genotype, Gofa-local was used as a check in all trials. At each environment (year by location combination) a randomized complete block design replicated three times was used. The plots were comprised of six rows of 4 m long with between and within row spacing of 30 and 10 cm, respectively. The central four rows
were harvested for grain yield assessment. Grain yield was adjusted for 10% seed moisture before conversion to kg ha\(^{-1}\) for statistical analysis.

### Statistical Analysis

The analysis of variance was performed using a SAS PROC GLM procedure in SAS version 9.1 (SAS Institute, 2003). Grain yield (kg ha\(^{-1}\)) per location per year was analyzed separately with the homogeneity of error variance tested according to Bartlett (1937) prior to performing the combined analysis. The fixed effect three-way Analysis of Variance (ANOVA) model that includes additive terms for the main effects of replications, blocks, genotypes, locations, and years together with an extra additive term that accounts for all the possible interaction effects of genotype, location and year was employed. The ANOVA model employed for an analysis of the data is:

\[
X_{ijk} = \mu + G_i + L_j + Y_k + (GL)_{ij} + (GY)_{ik} + (LY)_{jk} + (GLY)_{ijk}
\]

where \(X_{ijk}\) is the mean yield over \(r\) replications of the \(i\)th genotype in the \(j\)th location in year \(k\), with the right hand side of the equation giving grand mean yield \(\mu\) and respective main and interaction effects of the genotypes, locations and years. The magnitude of variance components was computed as the percentage of total variation to find out how the main and interaction effects explain the variations in yield. Mean grain yields of genotypes for the combinations of the 3 locations and 2 years, treated as six environments, were computed to generate a genotype and environment two-way table data for the biplot analysis. The GGE biplot software (Yan, 2001) was employed to generate graphs showing (i) “which-won-where” pattern, (ii) ranking of genotypes on the basis of mean yield and stability, and (iii) an evaluation of test environments (Yan et al., 2007).
RESULTS AND DISCUSSION

Analysis of Variance

The results from analysis of variance (Table 2) revealed significant (P< 0.05) genotype×environment interaction. The environment effect was highly significant (P< 0.01). The genotypes, locations and years’ main effect as well as Genotype×Year (GY) interactions were highly significant (P< 0.01). The interaction effects for Genotype×Location (GL) and Genotype×Location×Year (GLY) were significant (P< 0.05), whereas Location×Year (LY) effect was not significant. The location and year main effects explained most (up to 77 %) of the total variation. The contribution of genotypes’ main effect as well as GL, GY, LY and GLY interaction effects appeared to be negligible. Locations’ main effect by itself explained more than 61.52 % of the total grain yield variations. Moreover, the significant (P< 0.05) GL effects demonstrated that genotypes responded differently to different locations, confirming the importance of testing mung bean genotypes at multi-locations in southern Ethiopia.

When genotypes are tested in multi-location yield trials, a cross over genotype by environment interaction most often occurs (Ceccarelli et al., 2006). Such an interaction results from changes in relative ranking of genotypes from one environment to another. This complicates cultivar recommendation in breeding programs. The significance of main and interaction effects for majority of the sources of yield variations in mung bean revealed the importance of further analysis for adaptation pattern, genotypes’ response and their stability for better exploitation of the genotype by environment interaction.

Mean yield levels of different genotypes at three locations for 2004 and 2005 are depicted in Table 3. Relatively better yield was obtained in 2005 as compared with 2004. The most favourable location among the tested sites and in 2005 was Gofa. Inseno on the other hand was the least conducive location as in 2004, with a yield of 132 kg ha⁻¹. The highest yield across environments obtained from by MH-96-4 (908.1 kg ha⁻¹).

GGE Biplot Analysis

The first two principal components explained 90.4% of the total Genotype plus Genotype by Environment (G+GE) variation (Figures 1-3). This demonstrated a biplot constructed by plotting the first Principal Component (PC1) scores of genotypes and

Table 2. Analysis of variance of grain yield of seven mung bean genotypes grown at three locations in 2004 and 2005.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>Explained(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>125</td>
<td>34267680.26</td>
<td>2705638.45</td>
<td>78.31</td>
</tr>
<tr>
<td>Environment (E)</td>
<td>5</td>
<td>26837291</td>
<td>5367458.00**</td>
<td>61.52</td>
</tr>
<tr>
<td>Location (L)</td>
<td>2</td>
<td>21082332.63</td>
<td>10541166.31**</td>
<td>16.27</td>
</tr>
<tr>
<td>Year(Y)</td>
<td>1</td>
<td>5575077.84</td>
<td>5575077.84**</td>
<td>0.52</td>
</tr>
<tr>
<td>Location×Years (LY)</td>
<td>2</td>
<td>179878.37</td>
<td>89939.19ns</td>
<td></td>
</tr>
<tr>
<td>Replication (Environment)</td>
<td>12</td>
<td>1867206.5</td>
<td>155600.55**</td>
<td></td>
</tr>
<tr>
<td>Genotype (G)</td>
<td>6</td>
<td>727617.78</td>
<td>121269.63**</td>
<td></td>
</tr>
<tr>
<td>Genotype×Environment (GE)</td>
<td>30</td>
<td>2510990.162</td>
<td>83699.67*</td>
<td></td>
</tr>
<tr>
<td>Genotype×Location (GL)</td>
<td>12</td>
<td>893108.36</td>
<td>74425.70*</td>
<td></td>
</tr>
<tr>
<td>Genotype×Year (GY)</td>
<td>6</td>
<td>813364.71</td>
<td>135560.79**</td>
<td></td>
</tr>
<tr>
<td>Genotype ×Location×Year (GLY)</td>
<td>12</td>
<td>804518.42</td>
<td>67043.20*</td>
<td></td>
</tr>
<tr>
<td>Pooled Error</td>
<td>72</td>
<td>2324575.60</td>
<td>32285.77</td>
<td></td>
</tr>
</tbody>
</table>

*, ** Significance at respectively 5% and 1% level probability.
Table 3. Mean yield (kg ha\(^{-1}\)) of mung bean genotypes tested at three locations in 2004 and 2005.

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>2004</th>
<th>2005</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Awassa</td>
<td>Gofa</td>
<td>Inseno</td>
</tr>
<tr>
<td>Gofa-local</td>
<td>431.6ac</td>
<td>704.1a</td>
<td>123.6a</td>
</tr>
<tr>
<td>MH-98-3</td>
<td>492.1ab</td>
<td>1062.5ac</td>
<td>150.6b</td>
</tr>
<tr>
<td>MH-97-7</td>
<td>508.3ab</td>
<td>1504.1b</td>
<td>154.1b</td>
</tr>
<tr>
<td>MH-97-6</td>
<td>507.0ab</td>
<td>1662.5b</td>
<td>172.9b</td>
</tr>
<tr>
<td>MH-85-11</td>
<td>516.4ab</td>
<td>1079.1c</td>
<td>113.1a</td>
</tr>
<tr>
<td>Asha</td>
<td>374.9c</td>
<td>745.8a</td>
<td>89.5c</td>
</tr>
<tr>
<td>MH-96-4</td>
<td>551.3b</td>
<td>1350.0bc</td>
<td>118.0a</td>
</tr>
<tr>
<td>Mean</td>
<td>483.1</td>
<td>1158.3</td>
<td>131.7</td>
</tr>
<tr>
<td>CV</td>
<td>10.8</td>
<td>17.7</td>
<td>10.5</td>
</tr>
</tbody>
</table>

Means followed by similar letters are not significantly different at the 0.05 probability level based on LSD values; Underlined values are highest yields at each test environments (Location×Year).

Figure 1. GGE biplot based on environment-focused Singular Value Partitioning (SVP= 2) showing "which-won-where". The environments are indicated as AW-04 for Awassa 2004, AW-05 for Awassa 2005, GF-04 for Gofa 2004, GF-05 for Gofa 2005, IN-04 for Inseno 2004 and IN-05 for Inseno 2005. Genotypes are denoted by G1 to G7 where G1= Gofa-local; G2= MH-98-3; G3= MH-97-7; G4= MH-97-6; G5= MH-85-11; G6= Asha, G7= MH-96-4.
Asfaw et al.

Figure 2. The “discriminating power vs. representativeness” view of the GGE biplot. The data were not transformed (Transform= 0), not scaled (Scaling= 0), and were environment-centered (Centering= 2). The biplot was based on genotype-focused Singular Value Partitioning (SVP= 2). Environment and genotype names as of Figure 1.

Figure 3. The “mean vs. stability” view of the GGE biplot as based on genotype focused Singular Value Partitioning (SVP= 1). The data were not transformed (Transform= 0), not scaled (Scaling =0), but environment-centered (Centering= 2). An ideal cultivar is at the center of the innermost circle. Environment and genotype names as of Figure 1.

developments against their respective scores for second Principal Component (PC2) scores adequately capturing the environment-centered data. Moreover, the large yield variation due to location (Table 2) justified the selection of Site Regression (SREG) analysis model for Multi-Environment (MET) data (Yan et al., 2000). GGE biplot analysis was hence used for which-won-where analysis, test environment and genotype evaluation for mung bean MET data.
Which-Won-Where

The “which-won-where” pattern of the GGE biplot (Yan et al., 2000) is the most suitable tool for mega-environments analysis in variety trials (Yan et al., 2007). The “which-won-where” pattern of MET data is represented by a polygon formed by connecting the markers of genotypes that are further from a biplot origin, and a set of lines drawn from the biplot origin perpendicular to each side of the polygon. The perpendicular lines to the polygon sides divide the polygon sectors, each having its own winning cultivar which is the vertex genotype for that sector (Yan et al., 2000). Six out of the seven genotypes located in the vertex formed a six-sided polygon having six possible sectors (Figure 1). The vertex genotype for each sector is the one that yielded the highest for the environments filling within that sector. Four of the sectors had no environments. The six environments fell into two sectors delineated by years 2004 and 2005 with different winning genotypes. Awassa 2004 stood as intermediate between the two sectors indicating the existence of one mega-location classification repeatable over years for mung bean. Genotype G7 (MH-96-4) was the winning genotype at all locations in the favorable year 2005 and at Awassa in 2004. Genotype G4 (MH-97-6) was the winner genotype at Gofa and Inseno in cropping year 2004 which was relatively not conducive for mung bean genotypes to express their potentials. The vertex genotypes G1 (Gofa-local) and G6 (Asha) had no environment in their sector. The two genotypes were not the highest yielding ones at any of the test environments. G5 (MH-85-11) is located near to the plot origin and hence was less responsive than the vertex genotypes. The genotypes within the polygon and located nearer to plot origin are less responsive than vertex genotypes (Yan et al., 2001). The mung bean MET data did not strongly indicate the presence of different mega-environments, which is defined as the group of locations that consistently share the most suitable set of genotypes across years (Yan and Rajcan, 2002). Yan et al. (2007) stressed the need for data from multiple years to decide whether the target region can be divided into different mega-environments. It would therefore be impossible for our data set to confirm the existence of mega-environments.

Test Environment Evaluation

An ideal environment should be both discriminating of the genotypes and representative of the mega-environment (Yan et al., 2007). Figure 2 is a GGE biplot which is based on environment-focused scaling (Yan, 2002), with the singular values entirely partitioned into the environment scores (SVP= 2) making it appropriate for studying the relationships among test environments. In the biplot, the line that connects the environment marker to the biplot origin is proportional to the standard deviation of the genotype mean in the environment when the data is not standardized (Scaling= 0). Environments with longer vectors are more discriminating of the genotypes whereas environments with very short vectors are little or not informative on the genotype difference (Yan, 2002; Yan et al., 2007). Accordingly, Gofa in 2004, and Awassa in 2005 provided more information regarding the genotype differences whereas Inseno in 2004 provided little information concerning the genotype differences. Representativeness of the test environment is visualized by the angle formed between the environment vector and abscissa of average environment axis (the line passing through the biplot origin and the average environmental coordinate). The smaller the angle, the more representative the environment is (Yan and Tinker, 2006; Yan et al., 2007). The most representative location for mung bean grain yield was Awassa 2004. The ideal test environment (characterized by the combined ability of a
location to discriminate among genotypes in the study and to represent other locations in the overall environment of interest) was not very much clear with the present MET data indicating the need for more multi-location sites and annual data.

**Genotype Evaluation**

An ideal genotype should possess both high mean performance and high stability within a mega-environment (Yan et al., 2007). In Figure 1, the grouping of the test locations for mung bean yield delineate the locations by years. Awassa in year 2004 was the average environment of the trial (Figure 2) suggesting the three locations could be considered as one mega-environment. The mean performance and stability across environments were presented in Figure 3. For such an evaluation Yan et al., 2000 and 2001 indicated the high correlation ($r > 0.95$) between genotypes’ yield (averaged over environments) and their PC1 scores as a requirement. Such near-perfect correlations between genotypes mean yield and PC1 scores may not be always met. For instance, the correlation for this data was 0.83. In such instances Yan and Rajcan (2002) proposed an alternative, the use of a mean-environment coordinate system created by drawing a mean environments’ axis line that passes through the biplot origin and the mean environment marker for evaluation of both genotypes and environments. The projections of the genotype markers on the average environment axis are proportional to the rank-two approximation of the genotype means representing the main effects of the genotypes. The correlation between the measured and predicted mean performance (based on AEC) was 0.989, warranting for mean and stability analysis with this data set. The arrow shown on the axis of the Average Environment Coordinate (AEC) abscissa points in the direction of higher mean performance of the genotypes (Figure 3) and, consequently ranks the genotypes with respect to mean performance (Yan et al., 2007). Yield ranking of the genotypes as based on position relative to the end of the mean-environmental axis was recorded as: G7 (MH-96-4), G4 (MH-97-6), G2 (MH-98-3), G3 (MH-97-7), G5 (MH-85-11), G1 (Gofa local) and G6 (Asha). The projection of genotype marker onto the AEC approximates the genotype stability. The stability ranking of the genotypes based on the increasing absolute difference between genotype marker and AEC axis was G7 (MH-96-4), G5 (MH-85-11), G2 (MH-98-3), G6 (Asha), G1 (Gofa local), G3 (MH-97-7) and G4 (MH-97-6). MH-96-4 combining a high mean yield with stable performance was qualified as the most suitable genotype among the others, evaluated for production in mung bean growing ecologies in southern regions of Ethiopia.

**CONCLUSIONS**

Mung bean is important grain legume in semi-arid ecologies. It is mostly grown by vulnerable and risk-averse smallholder farmers in southern Ethiopia. Smallholder growing conditions in southern Ethiopia are myriad. Breeding programs should deliver these farmers robust germplasms that fit their wide range of environmental conditions. This requires testing of genotypes for quantification of the importance of $G \times E$ interaction, mainly genotype $x$ location, genotype $x$ year and genotype$\times$location$\times$year as well as a determination of the winning genotypes for the test environments. Several statistical methods have been proposed for increasing the chance of exploiting positive $G \times E$ interaction and supporting the breeding program decision in variety selection and recommendation. GGE biplot is among the many that effectively quantifies $G \times E$ interaction and provides meaningful interpretation of multi-environmental trial data. The application of GGE biplot to mung bean multi-environmental grain yield trial facilitated the visual comparison and
identification of the winning genotype in relation to the test environment.

ACKNOWLEDGEMENTS

Thanks go to Southern Agricultural Research Institute (SARI) for support of the field research. The support of technical assistants of pulse improvement teams of Awassa and Areka Research Centers is duly acknowledged.

REFERENCES

تجزیه عملکرد دانه ماس (Vigna radiate (L.) Wilczek) در شرایط محیطی مختلف با استفاده از روش پای پلات GGE در جنوب اتیویی

چکیده

هدف این تحقیق آرزیابی اثر مقایل محیط × زنونیب در عملکرد دانه ماس و همجنین تعیین بهترین زنونیب ماس در محیط های مورد بررسی در جنوب اتیویی می باشد. برای این منظور داده های عملکرد دانه ماس در محیط های مختلف شامل مکان، سال و رشدگر آنها مورد تجزیه و ارائه نشان داده شد. اثرات ترکیب یک عامل به طرف محسوس در نمود. سپس از میان این عملکرد زنونیب ها در محیط های مختلف جهت ایجاد جدول دو طرفه محیط و زنونیب و همجنین تعیین بهترین پای پلات GGE استفاده شد. اثر مقایل معنی داری بین زنونیب و محیط برای عملکرد دانه مشاهده شد و اثر محیط به شیب از 46 درصد از کل تغییرات عملکرد دانه را توجه نمود. همجنین پای پلات GGE به طرف هر عامل محیطی و همجنین مشابه تر تا نمایانه می‌باشد. همچنین به طرف زنونیب به طرف محدود 96-4 به عملکرد و پایداری بالامی تواده به عنوان یک زنونیب مناسب جهت کاشت در مناطق جغرافیایی قابل کشت ماس در جنوب اتیویی توصیه گردید.