Application of Linear Mixed Model to Multi-location Variety Trial in the Presence of Missing Plots: A Case of Barley Crop in Ethiopia

Diriba Tadese [1] and Girma Taye[2]

[1]Ethiopian Institute of Agricultural Research, Addis Ababa, Ethiopia

[2]School of Public Health, College of Health Sciences, Addis Ababa University, Ethiopia

Abstract

Background: Despite the presence of both fixed and random effects in most agricultural experiments, many researchers often use the conventional analysis of variance (ANOVA) or general linear model (GLM) that provide reliable outputs if and only if all the effects are fixed. Disregarding the presence of random effects will inevitably jeopardize precision of the analysis and thus will end up with wrong conclusions. The objective of this study was to lay a ground for a thorough understanding and proper application of the mixed effects model for a robust analysis of data obtained from multi-location field trials and conduct stability analysis under conditions of different patterns of missing observations. More often, stability analysis is among the basic objectives of multi-environment trial in barley breeding program.

Method: For this study, data were obtained from the Barley Research Program of the Ethiopian Institutes of Agricultural research (EIAR) whereby 15 barley genotypes were tested across 15 environments (five locations and three years) with four replications in each environment. Separate analysis of data was done for each environment using ANOVA, while combined analysis of data from all environments was done using GLM. Classical GLM and linear mixed-effects model (LMM) were employed in order to estimate missing observations of different patterns; which then were used to perform stability analysis using additive main effects and multiplicative interaction (AMMI) method.

Results and conclusion: Combined ANOVA results revealed that there were significant differences between genotypes, environments and their interactions. In the presence of missing data, LMM was found to be more robust than GLM to estimate the effect of genotypes. The significant interactions indicate that the genotypes responded differently across different environments. The mean grain yield of genotypes averaged over environments indicated that the yield estimate for genotype EH1847 was the highest (3575.13 kg/ha). AMMI 2 biplot indicates that genotypes EH1877, EH1893, EH1900, HB120 and F2SXS133 were adapted to a wider range of environments than the remaining genotypes. The study also showed that stability changes slightly in the presence of missing observations. Based

on the fitted model, environment-13 (Env13) was predicted as the winner environment among all other environments.

***Key Words**: Analysis of variance, fixed and random effects, general linear model, mixed models, AMMI stability*

1. Introduction

Crop improvement is a lengthy process and involves several stages in breeding programs to develop genotypes with superior grain yields of high quality that are adaptable over a wide range of environments. Genotype by environment interaction (GxE) is one of the main challenges encountered in selection of desirable genotypes since phenotypes are determined by the combined-effects of the environment and genetic makeup which interact with one another.

Different statistical approaches have been developed to analyze genotype by environment interactions especially for the purpose of determining yield stability over environments. Among these approaches, AMMI model is the popular one (Eberhart and Russell, 1966).

Missing values for some genotype-environment combinations are commonly encountered in multi-environment trials. Although quite a lot of researches have been done to analyze the stability of genotypes of different crops using data exhibiting different patterns of missing observations, almost all employed fixed effects models (Demeke and Girma, 2010; Dibaba and Girma, 2010).

The objective of this study was to compare the performance of different statistical methods using data obtained from barley multi-environment experiment (i.e., 15 barley genotypes tested across five locations and three years) involving missing observations based on estimation of the effect of missing observations on identification of most stable genotypes using mixed effects model.

2. Materials and Methods

The data for this study were obtained from Ethiopian Institutes of Agricultural Research (EIAR). Table 1 presents the names of the genotypes and the environments under which they were tested at the National Variety Trials (NVT). The environments considered in this study are combinations of five locations (Hollota, Adet, Sgonder, Asasa and Bekoji) and three years (2005, 2006 and 2007). The experimental design used to test the 15 genotypes in each environment was a complete randomized block design (CRBD). The genotypes were selected from a preliminary variety trial conducted in 2004.

In barley breeding programs, especially in national variety trials, suitable candidates of environments are selected via evaluating GxE interactions as well as stability analysis. However, this investigation is aimed at exploring a suitable model that best explains the genotype by environment interactions.

Table 1: List of barley genotypes used for investigation, their codes and the environments under which they were tested

Genotype name	Code	Environments
EH1847	t1	Env1 (Holeta-2005)
EH1900	t2	Env2 (Adet-2005)
EH1505	t3	Env3 (Sgonder-2005)
EH1893	t4	Env4 (Asasa-2005)
IBON4499	t5	Env5 (Bekoji-2005)
EH1877	t6	Env6 (Holeta-2006)
EH1869	t7	Env7 (Adet-2006)
EH1551	t8	Env8 (Sgonder-2006)
EH18472	t9	Env9 (Asasa-2006)
EH1864	t10	Env10 (Bekoji-2006)
IBON2796	t11	Env11 (Holeta-2007)
F2SXS133	t12	Env12 (Adet-2007)
BNEth019	t13	Env13 (Sgonder-2007)
HB120	t14	Env14 (Asasa-2007)
Beka	t15	Env15 (Bekoji-2007)

2.1 Fixed effects model

Analysis of variance in a fixed effects model setting is best suited to situations where the researcher applies one or more treatments to the subjects under the experiment to investigate the presence of treatment effects. This allows the experimenter to estimate the range of values of the response variable that the treatment would generate in a population as a whole. The basic assumption behind fixed effects model is that the outputs of the analyses are relevant only to the entire population under study and no generalizations are made about other populations that are not included in the study. Hence, the entire variability discerned between effect sizes is due to sampling error alone, and. the effect sizes are weighted only by the within-study variance which is often called the error variance.

The linear form of the fixed effects factor for a single experiment can be written as:

$$Y_{ij} = \mu + \alpha_i + \gamma_j + \varepsilon_{ij} \qquad i = 1, 2, \ldots, a;\ j = 1, 2, \ldots, b \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where Y_{ij} is the yield of i^{th} variety in the j^{th} block, μ is an overall mean, α_i is the effect of the i^{th} variety, γ_j is the effect of j^{th} block, a is the total number of varieties, and b is the total number of blocks. We assume that $\sum_{i=1}^{a}\alpha_i = 0$ and $\varepsilon_{ij} \sim \text{NID}(0, \sigma^2)$ so that $Y_{ij} \sim N(\mu + \alpha_i + \gamma_j, \sigma^2)$. Both treatment and block effects are assumed to be fixed as a classical combined analysis of variance.

In matrix notation, Equation (1) can be written as:

$$Y = X\beta + \varepsilon \quad \text{..... (2)}$$

where X is the design matrix,

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{a1} \\ \vdots & \ddots & \vdots \\ y_{1b} & \cdots & y_{ab} \end{bmatrix}, \beta = \begin{bmatrix} \mu \\ \alpha \\ \gamma \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_a \end{bmatrix}, \gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_b \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_{11} & \cdots & \varepsilon_{a1} \\ \vdots & \ddots & \vdots \\ \varepsilon_{1b} & \cdots & \varepsilon_{ab} \end{bmatrix}$$

Under the assumption that $\varepsilon \sim \text{MVN}(0, V)$, where $V = \sigma^2 I_n$, the ordinary least-squares (OLS) estimator of β is BLUE.

2.2 Linear mixed-effects model

Mixed-effects model contains experimental factors of both fixed and random-effect types, with appropriately different interpretations and analysis for the two types. In this study, G (genotype) is a fixed effect, while E (environment, which is a combination of locations and years), genotype by environment interaction and block/replication were considered as random factors.

The model can be written as:

$$Y_{ijk} = \mu + \alpha_i + \gamma_{j(k)} + E_k + \alpha E_{ik} + \varepsilon_{ijk} \quad \text{..... (3)}$$

where Y_{ijk} is the yield of i^{th} treatment (variety) from j^{th} block in k^{th} environment, α_i is the effect of the i^{th} treatment, $\gamma_{j(k)}$ is the effect of j^{th} random block (the replication-within-environment effect), E_k is the random effect of k^{th} environment, αE_{ik} is the interaction effect of i^{th} treatment with k^{th} random environment, and ε_{ijk} is random error.

The mixed effects model can be written in matrix notation as:

$$Y = X\beta + Zu + \varepsilon \quad \text{..... (4)}$$

where Y is a matrix of observations of the response variable, X is the design matrix of fixed variables, β is a vector of fixed effects parameters, Z is the design matrix of random variables, u is a vector of random effects parameters. Here we assume that $E(u)=0$, $cov(u)=G$, $E(\varepsilon)=0$, $cov(\varepsilon)=R$, $cov(u,\varepsilon)=0$, and both u and ε are normally distributed. Thus, the variance-covariance matrix of Y is:

$$V = ZGZ' + R \quad \text{.....} \quad (5)$$

and can be estimated by setting up the random-effects design matrix Z and by specifying the covariance structures for $G \sim qxq$ and $R \sim nxn$ (Little et al. 1996).

2.3 Missing observations

The original intent of almost all designed experiments is to obtain a balanced and orthogonal data set. However, non-balanced and non-orthogonal data may arise from unforeseen circumstances. Data imbalance refers to having different numbers of observations in individual levels of factors or combinations of levels of different factors in experiments. Due to missing values, blocks and genotypes can also be non-orthogonal, that is, the proportion of occurrence of treatments in each block may not be the same. Statistical analysis and hypothesis testing for mixed effect models remain valid if the missing-completely-at-random (MCAR) assumption is true (Verbeke and Molenberghs, 2000). In other words, if the MCAR assumption is satisfied, then missing data are independent of the design.

Hence, linear mixed model (LMM) is very flexible and capable of fitting a large variety of data sets including agricultural research data. The true and full use of LMM analysis accommodates many complex situations in real experiments (e.g., unbalanced data and long-term experiments) (Yang, 2009). The strategy of using an average information matrix is shown to be computationally convenient and efficient for estimating variance components by restricted maximum likelihood (REML) in the mixed linear model.

2.4 The AMMI method

One of the main deficiencies of the combined analysis of variance of multi-location yield trials is that it does not explore any underlying structure within the observed non-additives (genotype by environment interaction). The additive main effects and multiplicative interaction (AMMI) method integrates analysis of variance and principal component analysis into a unified approach (Bradu and Gabriel, 1978; Gauch, 1988). It can be used to analyze multi-location trials (Gauch and Zobel, 1988; Crossa et al., 1990). AMMI analysis first fits the additive main effects of genotypes and environments

by the usual analysis of variance and then describes the non-additive part, genotype-environment interaction, by principal component analysis. Hence, the AMMI method can be used for:

- Model diagnosis: AMMI is more appropriate at the initial stages of statistical analysis of yield trials since it provides an analytical tool for diagnosing other models as sub-cases when these are better for a particular data set (Bradu and Gabriel, 1978; Gauch, 1988).
- Clarify genotype-by-environment interactions: AMMI summarizes patterns and relationships of genotype and environments (Kempton, 1984; Gauch and Zobel, 1988; Crossa et al., 1990).
- Improve the accuracy of yield estimates: AMMI increases the number of replicates by a factor of two to five (Gauch and Zobel, 1988; Crossa et al., 1990). Such gains may be used to minimize costs by reducing the number of replications, to include more treatments in the experiment, or to improve efficiency in selecting the best genotypes.

The AMMI model is a hybrid model involving both additive and multiplicative components of two-way data structure that enables the breeder to get precise prediction on genotypic potentiality and environmental influences on it (Bradu and Gabriel, 1978). AMMI uses ordinary ANOVA to analyze the main effects (the additive part) and principal component analysis (PCA) to analyze the non-additive residual left out by the ANOVA.

The AMMI model is given by:

$$Y_{ij} = \mu + G_i + E_j + \sum_{k=1}^{n} \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \qquad i = 1, 2, \ldots, g;\ j = 1, 2, \ldots, h \quad \ldots\ldots\ldots (6)$$

where Y_{ij} is the mean yield of i^{th} genotype in the j^{th} environment, μ is the grand mean, G_i and E_j are the genotype and environment deviations from the grand mean, respectively, α_{ik} and γ_{jk} are the genotype and environment principal components scores for axis k, n is the maximum number of multiplicative terms (or $n = \text{rank}(X)$), λ_k is the k^{th} singular value of X (square root of the k^{th} eigenvalue of $X'X$) with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$. We impose the restrictions:

$$\sum_{i=1}^{g} \alpha_{ik} = \sum_{j=1}^{h} \gamma_{jk} = 0,\ \sum_{i=1}^{g} \alpha_{ik}^2 = \sum_{j=1}^{h} \gamma_{jk}^2 = 1 \text{ and } \sum_{i=1}^{g} \alpha_{ik} \alpha_{ik'} = \sum_{j=1}^{h} \gamma_{jk} \gamma_{jk'} = 0 \ ,\ k \neq k'.$$

In a more general setting, $s\ (< n)$ non-null eigenvalues are kept producing a reduced AMMI model, denoted by $AMMI - s$, given by:

$$Y_{ij} = \mu + G_i + E_j + \sum_{k=1}^{s} \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij} + \varepsilon_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

where s indicates the number of multiplicative terms necessary for an adequate description of genotype by environment interaction, and $\rho_{ij} = \sum_{k=(s+1)}^{n} \lambda_k \alpha_{ik} \gamma_{jk}$ is the residual not accounted for by the retained multiplicative terms for the interaction.

3. Results and Discussion

Descriptive statistics revealed that, on average, genotype EH1847 had a higher yield as compared to the remaining genotypes over the study years and performed higher in 2006. In fact, most of genotypes had a higher yield in 2006 when compared with the average performance in 2005 and 2007. The reason behind might be the fairly better distribution of rainfall in the 2006 production seasons.

The existence of genotype by environment interaction (GEI) was examined by plotting the mean of each genotype in all environments as depicted in Figure 1. One can clearly note from the graph that the yields of the genotypes exhibit interactions across the test environments with the presence of possible crossover interaction.

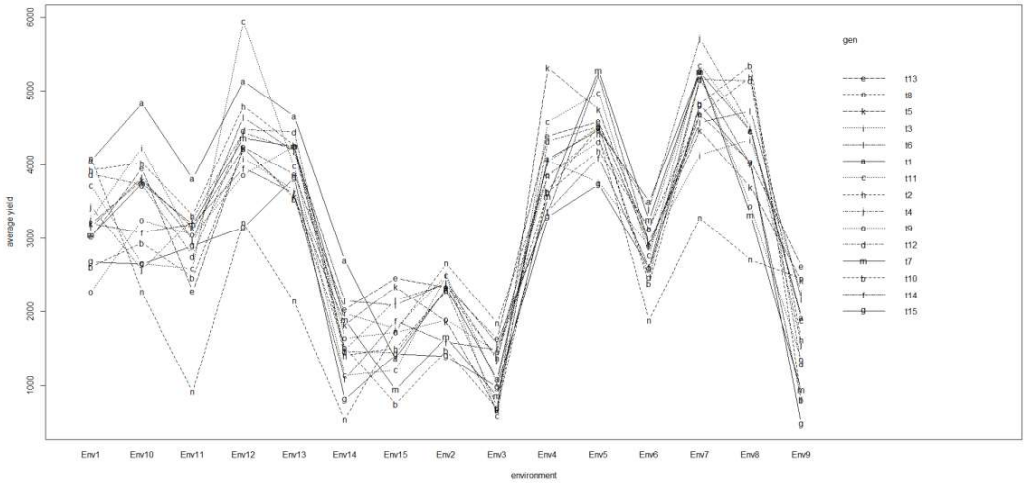

Figure 1: LS-mean graphical display of genotype by environment interaction

The combined analysis of variance (Table 2) revealed that there were significant differences among environments ($p < 0.0001$) and genotypes ($p < 0.0001$). This is an indication of the presence of variability in genotypes as well as combinations of test-locations and years defining the different environments under which the genotypes were tested. The GxE interaction effect was also significant ($p < 0.0001$) reflecting differential responses of the genotypes to the various environments.

Table 2: Combined ANOVA of barley yield data by GLM procedures

Source	DF	SS	% SS	MS	F Value	Pr > F
Replication within	45	52648157	3.051985	1169959	4.49	<.0001
Genotypes	14	69489927	4.028293	4963566	19.06	<.0001
Environments	14	1.23E+09	71.39871	87975785	337.81	<.0001
Env * Gen	196	2.09E+08	12.12852	1067462	4.1	<.0001
Error	629	1.64E+08	9.495954	260429		
Total	898	1.73E+09				

The variance components computed by restricted maximum likelihood (REML) were all significant at the 1% level (Table 3). During computation, environment and interaction effects were considered as random factors including replications. The estimated variance component of the random environment (1436507) was relatively large compared to other random factors of the linear mixed effect model. This indicates significant variability between environments of barley breeding ecology.

Table 3: Covariance parameters of barley yield data by REML of LMM

Covariance Parameters	Estimate	Standard Error	Z Value	Pr > Z
Environment	1436507	555385	2.59	0.0048
Replication within Environment	60734	16494	3.68	0.0001
Env * Gen	201963	27229	7.42	<.0001
Residual	260422	14684	17.74	<.0001

3.1. Mixed models and general linear model in the presence of missing observations

One challenge of the generalized linear model with missing observations is that it is unable to estimate least-square means when the maximum number of missing observations reaches four for a specific genotype at a specific environment. In fact, the effect can be estimated using the classical linear regression model for fewer than four missing observations for one genotype in one environment, albeit with small number of samples of a given genotype than the other genotypes. In this study, GLM was unable to compute an estimate for the first genotype due to the presence of missing observations (that is, EH1847 was totally damaged in environment-13 in all replications). In contrast, the mixed model does compute estimates since this method tries to find an estimable linear combination over all fixed and random effects, as if they were all fixed.

When genotype EH1847 was not missing at 13th environment, its average yield from combined analysis of variance was 3575.133, and thus, was considered as the best genotype among the candidate varieties. When mixed model is applied with the same number of missing observations in

Env13, still the same genotype would have the highest mean yield of 3553.88 kg/ha. In contrast, GLM was unable to estimate the effect.

Table 4 presents the Gollob F-test (Gollob, 1968) showing the significance of each AMMI terms. The first column shows the sum of squares corrected by the number of replicates (SSAMMI); the second column shows the percent of the genotype by environment interaction sum of squares explained by each AMMI term (PORCENT); and the third column shows the cumulative percent of the genotype by environment interaction sum of squares explained until the i^{th} AMMI term (CUM). The other columns show the degrees of freedom of each AMMI term (DF), their mean squares (MS), their F-values (F_AMMI) and the probability level of the F-test for each AMMI term. We can see that the first six interaction principal components (IPCA) are significant at the 1% level. The first interaction principal component (IPCA1) explains 37.7% of the variation in the total multiplicative interaction effect (GxE). This figure is 16.5% for the second principal component, and goes down for the rest of the AMMI model. Computer simulations have shown that the Gollob F-test is very liberal and can result in many multiplicative terms judged significant (Cornelius et al., 1996).

Table 4: Gollob F-test for the AMMI terms

IPCA	SSAMMI	PORCENT	CUM	DF	MS	F_AMMI	PROBF
IPCA1	78990304	37.7022	37.702	27	2925567	9.13681	0.00000
IPCA2	34562780	16.4969	54.199	25	1382511	4.31771	0.00000
IPCA3	29691692	14.1719	68.371	23	1290943	4.03173	0.00000
IPCA4	22777698	10.8718	79.243	21	1084652	3.38747	0.00000
IPCA5	14790199	7.0594	86.302	19	778431.5	2.43111	0.00064
IPCA6	11885960	5.6732	91.975	17	699174.1	2.18359	0.00389
IPCA7	6432594	3.0703	95.046	15	428839.6	1.33931	0.17249
IPCA8	3884993	1.8543	96.900	13	298845.7	0.93332	0.51758
IPCA9	2487129	1.1871	98.087	11	226102.7	0.70614	0.73334
IPCA10	1937990	0.9250	99.012	9	215332.2	0.6725	0.73423
IPCA11	1273364	0.6078	99.620	7	181909.1	0.56812	0.78210
IPCA12	562617.2	0.2685	99.888	5	112523.5	0.35142	0.88142
IPCA13	233191.3	0.1113	100	3	77730.43	0.24276	0.86650
IPCA14	734.21	0.0004	100	1	734.21	0.00229	0.96182
IPCA15	29.92	0.0000	100	-1	-29.92	-0.00009	.

Figure 2 is AMMI 2 biplot of the first interaction principal component (IPCA1) against IPCA2 of 15 barley genotypes in 15 environments with no missing observations. This plot is often used in stability analysis, that is, genotypes that are closer to the center have wide stability while genotypes far from the center have high interaction effect with the environment. From the figure one can observe that genotypes t6, t4, t2, and t14 are close to the center, and thus, seem more stable compared to the other genotypes. In contrast, genotypes t8, t11, t13 and t10 are more or less far from the center, and hence, have high contribution for the interaction effects.

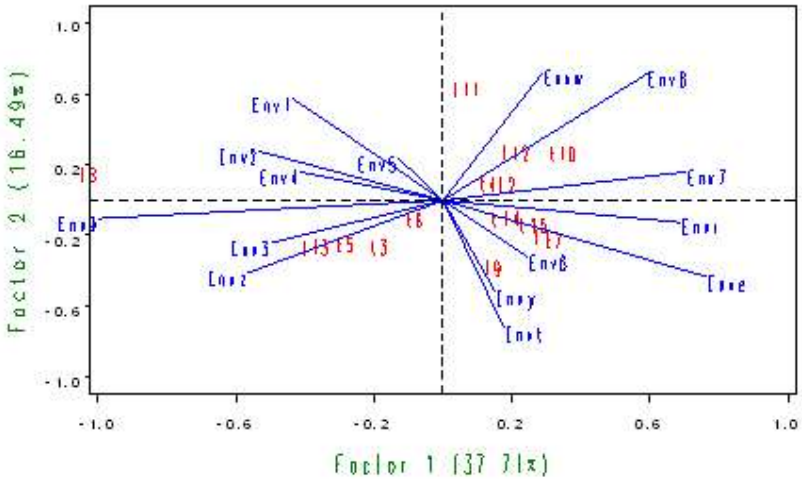

Figure 2: AMMI 2 biplot for grain yield (kgha^{-1}) of 15 barley genotypes in 15 environments

3.2. Discussion

In agricultural experimentation, a large number of genotypes are normally tested over a wide range of environments (locations, years, growing seasons, etc.) and the underlying statistical and genetic theories used to model this system may be rather complicated. The occurrence of genotype-by-environment interaction (GEI) effects further complicates the selection of superior genotypes for a target population of environments (Yang, 2009). In the absence of genotype-by-environment interaction, a superior genotype in one environment may be regarded as superior at all environments, whereas the presence of genotype by environment interaction only points to a particular genotype being superior in particular environments.

The genotype-by-environment interaction is an important aspect of both plant breeding and the introduction of new crop cultivars. It is used to explore environmental and genetic variations as well as the variation due to interaction effects (Annicchiarico, 2002). In this paper, analysis of multi-environment trials of barley was carried out with data from fifteen genotypes evaluated across fifteen environments. With the classical analysis of GLM procedure, environmental variation explains about 71.39% of total variation, followed by variance due to interactions (12.13%) and genotypes (4.03%).

This calls for the selection of genotypes for specific environments due to the interaction effects. Hence, crop yield is a complex trait that is influenced by a number of environmental factors exerting direct or indirect influence.

Stability analysis of genotype by environment interaction was done by AMMI model which is a powerful method of assessing the interaction and stability of genotypes from multi-environment. AMMI is essentially effective when the assumption of linearity of responses of genotypes to environmental changes is not fulfilled. The interaction principal component analysis indicated that the first six AMMI interaction principal components are significant and explain 91.73% of the total variation due to interaction. In particular, the first and second principal components together explain about 54.2% of the total interaction effects. The IPCA1 versus IPCA2 biplot shows that some genotypes are more stable than others. The values explained by each principal component changed slightly due to missing values for some genotypes in some environments.

Comparison of mixed model and classical GLM under different numbers of missing observations showed that LMM is preferable to the GLM. When one or two missing observations for different genotypes at different environments were considered, the classical GLM estimates the effect of the genotype which contains missing values with less degree of freedom while LMM estimates the same with equal degrees of freedom as with original data by randomly considering the remaining observations. When observations on genotype EH1847 were missing from env-13 in the whole block/replications, GLM was unable to estimate its effect while there was no estimability problem with LMM. Yang (2009) asserts that LMM is more flexible than the classical analysis of variance in handling missing observations and unbalanced data set.

4. Conclusion and Recommendations

The objective of this paper was to compare the average yield potential of different barley genotypes under different methods of estimation and with different patterns of missing observations. To meet this objective, fifteen barley genotypes were tested across fifteen environments with randomized complete block design. GLM combined analysis of variance indicated that the effect of environment accounted for more than 70% of the total sum of squares. This is probably an indication that environment is the major factor that influences yield performance of barley in Ethiopia. Moreover, the presence of significant genotype-by-environment interaction is an indication that the relative rankings of genotypes are likely to change from one location to another and/or from one year to another.

Comparison of mixed effects model and GLM under different patterns of missing observations showed that the former is better than the latter in estimating average yield of genotypes. The result affirmed that when there was a total damage of a genotype in an environment, it was not possible to

estimate the effect of that particular genotype using GLM. In contrast, no such problems were faced with LMM. Therefore, for making estimations of effects in the presence of missing observations in a data set, it is better to use the LMM approach.

Based on the findings of this research the following recommendations were forwarded:

- Researchers should consider the pattern of missing values in data sets since it is an important issue in stability analysis. In particular, to estimate the effects of genotypes in the presence of missing observations, the use of linear mixed model is recommended (than the classical ANOVA) as it faces no estimability problem.
- The selection process of good performing and stable genotypes is complicated by the presence of genotype-by-environment interaction. Hence, it is imperative to have a proper understanding of the effects of genotypes on various environments on variety evaluation.

References

1. Annicchiarico, P. (2002). Genotype x Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations. FAO, Rome.
2. Bradu, D. and Gabriel, K.R. (1978). The Biplot as a Diagnostic Tool for Models in Two Way Table. *Technometrics*, 20: 47-68.
3. Cornelius, P.L., Crossa, J. and Seyedsadr, M.S. (1996). Statistical Tests and Estimators of Multiplicative Models for Genotype-By-Environment Interaction. pp. 199-234. In M.S. Kang and H.G. Gauch (ed.) Genotype-by-environment interaction. CRC Press, Boca Raton, FL.
4. Crossa, J, 1990. Statistical Analysis of Multiplication Trials. *Advances in Agronomy*, 44:55-85.
5. Demekew Lake and Girma Taye. (2010). Applications of LMM to Incomplete Block Design. *Journal of the Ethiopian Statistical Association*, Vol. 19.
6. Dibaba Bayisa and Girma Taye. (2010). Application of Spatial Mixed Model in Agricultural Field Experiment. *Journal of the Ethiopian Statistical Association*, Vol.19.
7. Eberhart, S.A. and Russell, W.A. (1966). Stability Parameters for Comparing Varieties. *Crop science*, 6:36-40.
8. Gauch, H.G. and Zobel, R.W. (1988). Predictive and Postdictive Success of Statistical Analyses of Yield Trials. *Theoretical and Applied Genetics*, 76:1-10.
9. Gauch, H.G. (1988). Model Selection and Validation for Yield Trials with Interaction. *Biometrics*, 44:705-715.

10. Gollob, H.F. (1968). A Statistical Model which Combines Features of Factor Analytic and Analysis of Variance Techniques. *Psychometrika*, 33(1):73-115.

11. Kempton R. A. (1984). The Use of Biplots in Interpreting Variety by Environment Interactions. *Journal of Agricultural Science*, 103:123-135.

12. Little, R.C., Milliken, G.A., Stroup, W.W. and Wolfinger, R.D. (1996). SAS for Mixed Models. SAS Institute, Inc., Cary, NC.

13. Yang, R. (2009). Biplot Analysis of GXE Interactions. *Crop Science*, 49:1564-1576.

14. Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. Springer, Berlin.