Application of Mixed Model and Spatial Analysis Methods in Multi-Environmental and Agricultural Field Trials

By

Asnake Worku Negash

Submitted in fulfillment of the academic

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Applied Statistics

in the

School of Mathematics, Statistics & Computer Science

University of KwaZulu-Natal

Pietermaritzburg

2015

Dedication

This thesis is dedicated to

My wife Messeret Sahilu

and

My son Edomiyas Asnake

Declaration

The research work described in this thesis was carried out in the School of Mathematics, Statistics and Computer Sciences, University of KwaZulu-Natal, Pietermaritzburg, under the supervision of Professor Henry Mwambi and Professor Temesgen Zewotir.

I, Asnake Worku Negash, declare that this thesis is my own, unaided work. It has not been submitted in any form for any degree or diploma to any other University. Where use has been made of the work of others, it is duly acknowledged.

April, 2015

_____ _____

Mr Asnake Worku Negash Date

_____ _____

Professor Henry Mwambi Date

_____ _____

Professor Temesgen Zewotir Date

Acknowledgment

I would like to thank my supervisors Prof. Henry Mwambi and Prof. Temesgen Zewotir for their tireless guidance, encouragement and supervision in the preparation of this thesis. In addition, I extend special thanks to Dr. Girma Taye for all support during my study. I would like to thank the staff members and postgraduate students in the School Mathematics, Statistics and Computer Science for the hospitable environment they provided during my study.

I would like to thank Ethiopian Institute of Agricultural Research (EIAR) and UKZN College of Agriculture, Engineering and Science for providing the data and financial support for my study respectively. I would especially like to express my gratitude to my wife Meserte Sahilu, my son Edomias and my family for all the social, moral and logistical support throughout my life. I hope that I will be able to make up for my absence.There are too many people close to me who were supporting me during this study period, specially looking after my family to mention some of them, Mr Tesfaye Gezahege, Mrs Mekdes Atela , Mr Samson Mogese, Mrs Kongit Girma , Mr Hashim Dari, Mrs Sabat Kore and Mr Fedhesa Tadese .

I would like to pass my gratefulness to my colleguse Dr Dawit Getent, Kassahun Mengist, Dr Amelework Beyene, Dr Selessh Fanta and his family, Dr Surafel L. Tilahun, Dr Netsante Bacha, Dr Asnakech Tekaleg, Dr Demeissew Abakemal, Mr Ermias Abate, Mrs Hirut Getinet, Mr Fikadu Gurumu, Mrs Tigest Shiferaw, Mis Mizan Tesfay, I will not mention you all by names, except to say, thank you and I really love you.

Note

The following papers have been published from this thesis.

1 Mixed model with spatial variance-covariance structure for accommodating Of local stationary trend and its influence on multi-environmental crop variety trial assessment *(Spanish Journal of Agricultural Research 2014 12(1): 195-205eISSN: 2171-9292, Available online at www.inia.es/sjar ISSN: 1695-971-X http://dx.doi.org/10.5424/sjar/2014121-4926)*

2 Additive main effects and multiplicative interactions model (AMMI) and genotype main effect and genotype by environment interaction (GGE) biplot analysis of multi-environmental wheat variety trials. *(African Journal of Agricultural Research vol. 8(12), pp. 1033-1040, 4 April, 2013 DOI: 10.5897/AJAR2012.6648 ISSN 1991-637X ©2013 Academic Journals http://www.academicjournals.org/AJAR)*

Abstract

Agricultural experimentation involves selection of experimental materials, selection of experimental units, planning of experiments, and collection of relevant information, analysis and interpretation of the results. An overall work of this thesis is on the importance, improvement and efficiency of variety contrast by using linear mixed mode with spatial-variance covariance compare to the usual ANOVA methods of analysis. A need of some considerations on the recently widely usage of a bi-plot analysis of genotype plus genotype by environment interaction (GEE) on the analysis of multi-environmental crop trials. An application of some parametric bootstrap method for testing and selecting multiplicative terms in GGE and AMMI models and to show some statistical methods for handling missing data using multiple imputations principal component and other deterministic approaches.

Multi-environment agricultural experiments are unbalanced because several genotypes are not tested in some environments or missing of a measurement from some plot during the experimental stage. A need for imputation of the missing values sometimes is necessary. Multiple imputation of missing data using the cross-validation by eigenvector method and PCA methods are applied. We can see the advantage of these methods having easy computational implementation, no need of any distributional or structural assumptions and do not have any restrictions regarding the pattern or mechanism of missing data in experiments.

Genotype by environment (G×E) interaction is associated with the differential performance of genotypes tested at different locations and in different years, and influences selection and recommendation of cultivars. Wheat genotypes were evaluated in six environments to determine the G×E interactions and stability of the genotypes. Additive main effects and multiplicative interactions (AMMI) was conducted for grain yield of both year and it showed that grain

yield variation due to environments, genotypes and (G×E) were highly significant. Stability for grain yield was determined using genotype plus genotype by environment interaction (GGE) biplot analysis. The first two principal components (PC1 and PC2) were used to create a 2-dimensional GGE biplot. Which-won where pattern was based on six locations in the first and five locations in the second year for all the twenty genotypes? The resulting pattern is one realization among many possible outcomes, and its repeatability in the second was different and a future year is quite unknown. A repeatability of which won-where pattern over years is the necessary and sufficient condition for mega-environment delineations and genotype recommendation.

The advantages of mixed models with spatial variance-covariance structures, and direct implications of model choice on the inference of varietal performance, ranking and testing based on two multi-environmental data sets from realistic national trials. A model comparison with a $\chi 2$-test for the trials in the two data sets (wheat and barley data) suggested that selected spatial variance-covariance structures fitted the data significantly better than the ANOVA model. The forms of optimally-fitted spatial variance-covariance, ranking and consistency ratio test were not the same from one trial (location) to the other. Linear mixed models with single stage analysis including spatial variance-covariance structure with a group factor of location on the random model also improved the real genotype effect estimation and their ranking. The model also improved varietal performance estimation because of its capacity to handle additional sources of variation, location and genotype by location (environment) interaction variation and accommodating of local stationary trend. The knowledge and understanding of statistical methods for analysis of multi-environmental data analysis is particularly important for plant breeders and those who are working on the improvement of plant variety for proper selection and decision making of the next level of improvement for country agricultural development.

Table of Content

List of Tables

List of figures

Abbreviations

AEC	Average Environment Coordination
AEZs	Agro-ecological zones (AEZs)
AMMI	Additive Main effects and Multiplicative Interaction
EARO	Ethiopian Agricultural Research Organization
CSA	Central Statistical Agency
ECEA	Economic Commission for Africa
EIAR	Ethiopian Institute of Agricultural Research
GEE	Genotype by environment interaction
IB	Incomplete block
IPCA	Interaction Principal Component Axes
MAR	Missing at random
MCAR	Missing completely at random
MET	multi-environmental trials
MNAR	Missing not at random
MEYT	Multi-Environment Yield Trials
RCB	Randomized block design
SVD	Singular Value Decomposition

Chapter 1
Introduction

Agriculture is a major industry in Ethiopia and there is a continual need to increase productivity. State-owned plant breeding programs play an important role aimed at developing varieties which are higher yielding, resistant to disease and have quality characteristics. Agricultural field trials involve selection of experimental materials, selection of experimental units, planning of experiments, and collection of relevant information, analysis and interpretation of the result. There is no perfect experimental design nor are there any perfect analysis procedures known to account for all variations encountered in practice. Hence, the quality of the results depends on how well the selected experimental design or the method of analysis helps in estimation of spatial variations. Design based approaches fail to account for such variation, especially, when blocks cannot be laid out successfully. As a result, the experimental mean square error of classical analysis may be severely inflated (Warren and Mendez, 1982). High mean square error affects sensitivity of the test thus leading to inappropriate conclusions. An alternative would be to consider a modeling approach to account for spatial variations, in addition to the experimental design employed.

In view of decreasing arable land and differing agro-ecology within countries, and also increase in population and the subsequent rise in demand for agricultural produce are expected to be greater in regions where production is already insufficient, in particular in Sub-Saharan Africa. The necessity and demand to increase agricultural production represents a huge challenge to local farming systems given it must come mainly from increased yield per unit area in addition to the limited extension of cultivated land in the country. Failure to account for spatial variation during experimentation has an impact on development of improved technology. There will not be an increase in crops' area in the future due to the limited area of arable land. Therefore, the best

strategy of increasing production of crop is by increasing productivity per unit area using improved production technology. On the other hand, reliable improved technologies can be evaluated and adopted if proper designing and modeling is done by accounting for spatial variation.

The recommendation of new varieties for farmers or commercial use is complicated by the fact that not all varieties respond in the same way to change in the environment, a phenomena known as variety by environment (V×E) interaction. Selection of high yielding and disease resistance varieties is a typical activity in variety trial programs. Crop breeding programs run at different stages, often from nursery to national variety trial with the following major (Aweke, 2005) objectives:

- To generate new cultivars which are superior to the standard ones
- To explore and adopt new technologies to increase food production
- To study the effect of different factors of production on yield and quality and identify the best combination of these factors which can optimize the yielding ability of crop
- To generate information about basic understanding of factors of production to be used by farmers and researchers
- To develop recommendations regarding adaptability of selected varieties in different agro-ecologies to assist users.

To full fill these objectives research needs to cover several crops grown in a range of agro-ecological zones under different conductions. Agro-ecological zones (AEZs) are geographical areas exhibiting similar climatic conditions that determine their ability to support rained agriculture. At a regional scale, AEZs are influenced by latitude, elevation, and temperature, as well as seasonality, and rainfall amounts and distribution during the growing season. In the light of scarce resources and high field variability, therefore, effort should be exerted to achieve the above goals through improved design and modeling approaches that account for spatial variation and variety by environment (V×E) interaction.

In breeding programs development and selection of new varieties requires a number of stages. At the beginning, namely the nursery stage, several breeding lines are included which require a large size experiment. At later stages where few and promising varieties are selected; only a small experimental set-up is required. This, therefore, calls for an optimum design and effective modeling approach. At each stage in the breeding program it is vital that selection decisions are made with minimal error. It is also important, particularly at the final stages, to accurately predict the future yield gains of new varieties relative to existing commercial varieties. Selections and predictions are based on the analysis of yield data from the appropriate series of trials (commonly referred to as to as multi-environmental trials (MET)). The aim of plant breeding is to develop varieties for wide-spread small holders and commercial use, so that in the MET, attention is focused on varieties which have both high overall mean yield across trials and consistent (stable) performance see (Piepho, 1996) for example.) Varieties which excel in some trials but have ordinary performance otherwise are also of interest since there may be potential for making environment specific recommendations.

Irrespective of whether broad or specific recommendations of prime interest, it is important that an analysis of MET data provides both a measure of overall yield performance for each variety and information about the magnitude and nature of V×E interactions. In terms of the latter it is useful to distinguish between the two sources of V×E interaction, namely heterogeneity of genetic variances between environments (that is, change in scale) and lack of genetic correlation between environments (change in variety rankings) (Cockerham, 1963, Cooper et al., 1996, Robertson, 1959).An understanding of the latter, also known as cross-over interactions, is particularly important since it is this type of V×E interaction which can complicate selection and recommendation.

To full fill the above five objectives research needs to cover several crops grown in a range of agro ecologies each under different conditions. In the light of

scarce resources and high field variability, therefore, effort should be exerted to achieve the above goals through improving design and modeling approaches that account for spatial variation and V×E interactions. In crop improvement programs the major interest is in detecting differences among varieties or treatments and on the method of modeling employed. To optimize the use of resources, the best method of describing and accounting for field variability should be used so that variety effects and differences between them are quantified as precisely as possible. This requires improvement in both design and modeling.

A majority of research trials in sub-Saharan Africa are planned as randomized complete block design (RCBD) regardless of the size of experiments (Aweke, 2005). Randomized complete block designs have been widely used in agricultural field experiments compared to other known designs such as Latin square and families of incomplete block designs. The major reasons for the popularity of this design are its computational convenience; flexibility and efficiency given the required assumptions are fulfilled. A review done to assess the type of designs used in early days shows the importance of randomized complete block design (RCBD) in the history of field experimentation. Most of these studies show that the layout of these field trials were RCBDs with or without split-plot. These studies include (Atkinson et al., 2007, Cochran and Cox, 1957, and Fisher, 1992). Based on a survey conducted on the type of design used in field trials in Ethiopia (EARO, 1996), it was found that about 78% of the designs used in breeding and agronomy research programs are RCBD with and without split-plot. More recently it was found that large proportions of experiments (90%) were conducted as RCBD. The obvious reason for using RCBD so frequently are its simplicity and practical convenience for layout in a field. Such designs are often applied at all levels of research, including nursery, regional variety trial, and national variety trial. This resulted in squeezing a large number of treatment combination into a

complete block, which led to loss of homogeneity within the block. This invalidated assumptions for tests of hypothesis affecting the analysis results.

The major problem of designing large field trials in sub-Saharan African is the presence of considerable field variability. Often this is compounded by the presence of erratic rainfall conditions and a moisture stressed environment. In a study of the stability of major crops in Ethiopia (Taye et al., 2000) it was shown that variability in soil and environment is very high even within a very small area. Variability increase with the size of experiments and consequently the size of block. With the increases in the size of blocks, the possibility for controlling variability within a block diminishes leading to loss of precision. The advances in recent modeling approaches have even questioned the homogeneity of small blocks. However small a block may be there could be considerable spatial correlation between neighboring plots, regardless of direction.

1.1 Some statistical approaches to the analyses of multi-environmental trial (MET) data

Multi-environmental trial (MET) data are often analyzed using a two-stage approach in which variety means are first estimated separately for an overall analysis. Individual trials have traditionally been analyzed using complete or incomplete block analysis, but recently the more efficient spatial techniques have been adopted. With the rapid development in computing technology and sophistication in the areas of computer science and mathematical statistics, the computational aspect is no longer a problem and it is only necessary to look for improved methods of designing and modeling large field trials to enhance precision of estimation and treatment composition.

The classical method for the combined analysis of a series of agricultural experiments is an Analysis of Variance (ANOVA) which has a long history of application (Caliński et al., 2009, Cockerham, 1963, Dias and Krzanowski,

2003, So and Edwards, 2009, Smith et al., 2005, Yates and Cochran, 1938). Such analysis account variation for varieties (V), environments (E), V×E interaction and within trial error (the pooled error mean square from analysis of individual trials). Apart from the within trial errors, all other effects in the ANOVA are classically regarded as fixed so are estimated using least squares. Effects may be assumed random (Cockerham, 1963, Comstock and Moll, 1963, Kelly et al., 2007a, Masjkur, 2012) in which case the associated variance components are estimated by equating ANOVA mean squares with their expectation.

The ANOVA approach can only be used when the data are complete, that is, when the same set of varieties have been tested in every trial. Often, this is not the case, particularly if the data set spans several years since varieties are continually moving through the breeding program. When the data are incomplete, variance components from random effects can be estimated using residual maximum likelihood (REML) (Patterson and Thompson, 1971). This procedure provides the same variance estimates as ANOVA method when the data are complete. The model underlying the analysis is usually a mixed model, that is, it comprises a mixture of fixed and random effects. There are many examples of fixed model analysis of V×E means in the literature, including Cullis et al. (1996a), Cullis et al. (1996b), Frensham et al. (1997), Frensham et al. (1998), Patterson and Silvey (1980), Patterson et al. (1977) and Talbot (1984).

In the early mixed model approaches to the analysis of MET data, variety effects were regarded as fixed. Cullis et al. (1996a), Cullis et al. (1996b) Frensham et al. (1997) and Frensham et al. (1998) treat variety effect as random whilst Patterson and Abugoomu (1992) consider both alternatives. The V×E interactions are always regarded as random.

The two-stage analysis is an approximation to the analysis of individual plot data from all trials. When individual trials are analyzed using orthogonal

analysis such as randomized complete block there is no loss of information in the two-stage approach provided appropriate measures of within trial error variation are carried through into the second stage mixed model (through the use of weights). However, when more efficient, but non-orthogonal analyses such as spatial or incomplete block analysis can be applied in two-stage analysis may provide a poor approximation to the analysis of individual plot data. Frensham et al. (1997) present a weighted mixed model for the second stage analysis which aims at providing a reasonable approximation to the fully efficient spatial mixed model of Cullis et al. (1998a). In Cullis et al. (1998a) individual plot data are analyzed and allowance is made for a separate spatial covariance structure and error variance for each trial. Two-stage approaches such as that by Frensham et al. (1997) arose from necessity, since historically individual plot data were often unavailable and mixed model software was unable to handle the size of some MET data-set. With the advancement of electronic storage of individual plot data and sophisticated mixed model software such as SAS, GenStat and ASReml, the efficient analysis of individual plot data analysis is usually feasible and is the preferred approach.

The standard mixed model provides an estimate of the magnitude of V×E interaction (as reflected in the variance component) but does not provide any insight into its structure. Some researchers have addressed this by extending the model to include interactions with external covariate information, such as rain fall and soil acidity for the environments, and disease resistance and maturity for the varieties. Freeman and Perkins (1971) and Knight (1970) used this approach within the framework of a mixed effects model. An alternative method for examining the V×E interaction involves the use of multiplicative models. In these models V×E interaction is explained in terms of environmental and varietal parameters which have been estimated from the data. The earliest model of this kind is the regression on the mean model, in which yields for individual varieties are regressed on the mean yield of all varieties in an environment.

Many researchers used the general class of multiplicative model originally proposed by Mandel (1971) and popularized under the banner of AMMI (Additive Main effects and Multiplicative Interaction) (Gauch, 1992b). These models involve singular value decomposition (SVD) of the matrix of residuals from a two-way ANOVA with variety and environment main effects. V×E interaction is there by decomposed into a number of multiplicative terms from the SVD. The conjecture is that most of the V×E interaction can be explained by the first few terms from the SVD and that these have some meaningful interpretation.

The aims of AMMI analysis clearly have merit but the approaches have many disadvantages compared with the mixed model approach. The AMMI models are fixed-effects models. But there are strong arguments for the use of random variety and V×E interaction effects. There are deficiencies in terms of handling unbalanced data, the assumption of a common error variance for all trials and inability to accommodate spatial variation within trials.

1.2 Objectives

The overall objectives of this thesis are:

- To study the data and thereby investigate the impact and remedies of missingness in such multi-environmental crop variety trial data.
- To assess the efficiency of variety contrast by using spatial-variance covariance.
- To examine the usage of a bi-plot analysis of genotype plus genotype by environment interaction (GEE) on the analysis of multi-environmental crop trials.
- To explore some parametric bootstrap method for testing and selecting multiplicative terms in GGE and AMMI models.

The knowledge and understanding of statistical methods for analysis of multi-environmental data analysis is particularly important for plant breeders and those who are working on the improvement of plant variety for proper selection and decision making of the next level of improvement for country agricultural development.

The thesis is organized into seven chapters. The study background and the objectives of the study is presented in Chapter 1. Chapter 2 deals with the description of the data as well as the theoretical and practical method of missing data handling mechanisms for multi-environmental crop variety trial data. Chapter 3 presents the additive main effects and multiplicative interactions model (AMMI) and genotype main effect and genotype by environment interaction (GGE) biplot analysis of multi-environmental trials. In Chapter 4 parametric bootstrap methods for testing and selecting multiplicative terms in GGE and AMMI Models are presented. Chapter 5 deals with spatial analysis of the field experiment. Chapter 6 focuses on the mixed model approach with spatial variance–covariance structure for accommodating local stationary trend and its influence on multi-environmental crop variety trial assessment. In Chapter 7, the discussions and conclusions as well as implications and avenues for future research are presented.

Chapter 2
The Data and Missingness Assessment

2.1 The Data

Ethiopia is the second largest wheat and barley producer country in Sub-Saharan Africa after South Africa (GAIN, 2012). Wheat and barley are cultivated on 1.605 and 1.02 million hectares and account for 12.94% and 8.22% of the grain crop area, with an annual production of 3.4 and 1.91million metric tons respectively. Wheat contributes about 12.94% and barley 8.22% of cereal production in the country (CSA, 2013/14). Interims of area, wheat and barley ranks fourth and fifth after teff, maize and sorghum. These crops are widely grown by subsistent farmers and one-third of cereal farm households are dependent on wheat and barley farming (Shiferaw et al., 2014).

According to ECEA (2008) the major wheat and barley producing regions in Ethiopia includes Oromia, Amhara, Southern Nations and Nationalities Peoples' Region (SNNPR) and Tigray. The data used in this thesis are from a study carried out between 2004 and 2008 in six different research stations in Ethiopian Agricultural Research Institute National Variety Trials for Bread Wheat and Barley Trial of 2004-2008. The locations are labelled as loc1 (Kulumsa), loc2 (Adet), loc3 (Bekoji), loc4 (Sinana), loc5 (Holeta) and loc6 (DeberZeit) that is, five research station from Oromiya region, which include the wheat belts in East Africa and one station from Amhara region. All the trials in each location were laid out as a randomized complete block (RCB) design with four replicates. Most data sets were obtained from the Ethiopian Agricultural Research Institute for further analysis consists from 5-40% missing values in different locations with different proportion.

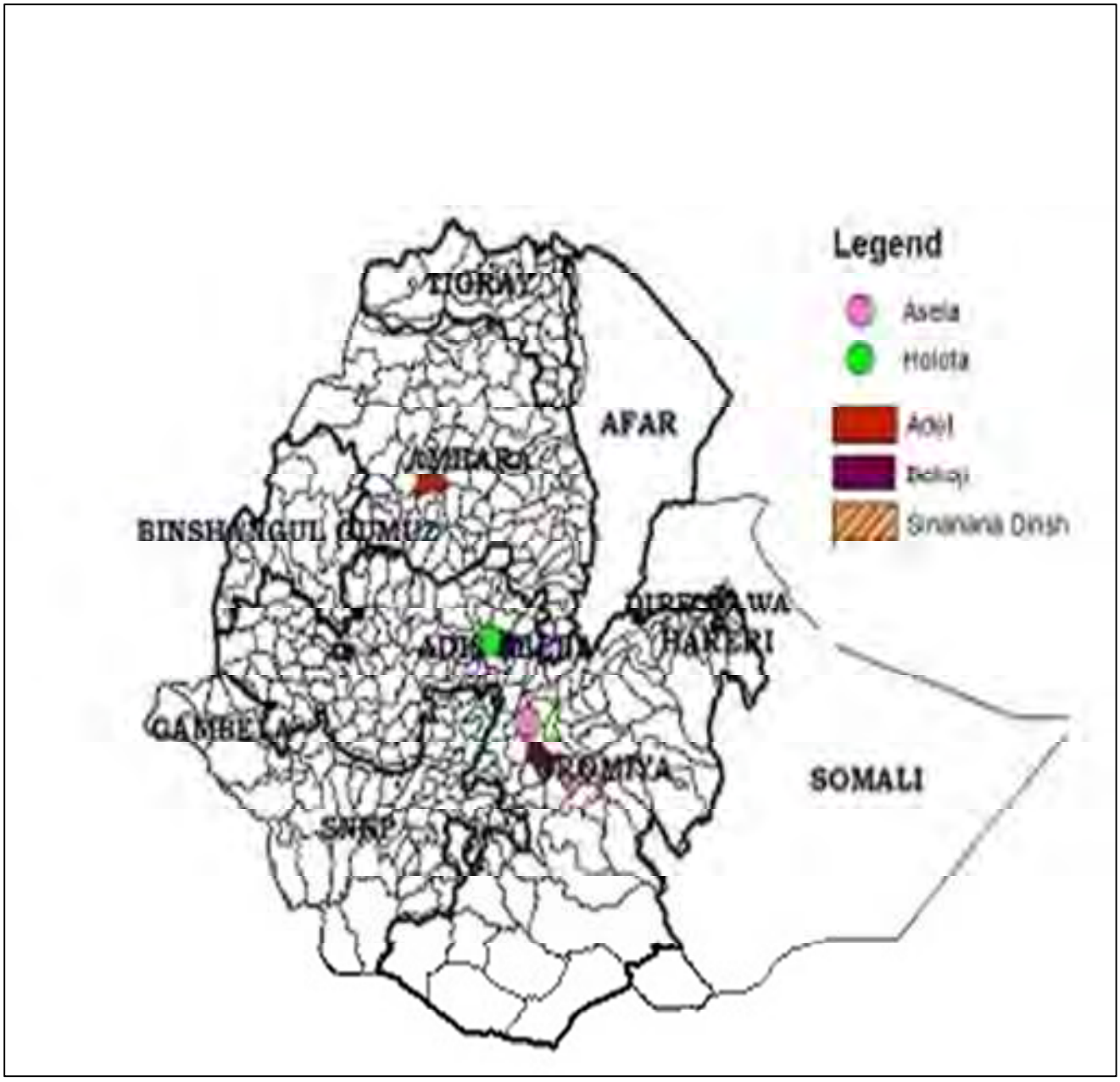


Figure 2. 1 Location of the six research stations in the Oromya and Amhara regions

Missing data observations due to natural and man-made reason create difficulties for widely applicable agricultural experiment data analysis methods such as AMMI, GGE, regular grid REML, row-column design analysis which are only applicable for a balanced case. Twenty bread wheat genotypes were evaluated in each of the above locations (environments) in a randomized complete block design with four replications. These Twenty genotypes are coded from G1-G20. Similarly 25 barley varieties were tested in five locations (environments).

2.1.1 The impact of missing data

In plant breeding, multi-environment trials are important for testing the general and specific adaptations of cultivars. Even in well-controlled trials or studies, missing data always occur in multi-environmental crop data analysis. Missing data can be due to weather issues, dead or damaged plants, incorrect data measurement or transcription, and many other situations that arise when working with real data. Missing data may bias statistical analysis results, such as in the estimation of confidence intervals, reduce statistical power and bias parameter estimation.

In the case of missing data, the loss of information produces unbalanced designs that lose their symmetry (i.e. the balance) of the design. With this loss of symmetry goes the simplicity of the analysis as well. And as more and more values are missing, the analysis becomes more and more complex. Therefore, hypothesis tests of interest such as those for the difference between the treatments may need special theoretical development. Sometimes, if the proportions of missing data are large, some parametric functions are not estimable and wrong calculation of the degrees of freedom for the sums of squares may cause inappropriate inferences and poor conclusions about the experiment. A possible solution could be to repeat the experiment under similar conditions and in this way to obtain new values for the missing

observations. However, this solution, although ideal, might not be viable in terms of available time and money. Dodge (1985) and Little and Rubin (2014) present two of the most common approaches used to solve this problem. Dodge (1985) presents theoretical considerations for an analysis based only on the observed data, while Little and Rubin (2014) describe a large number of imputation methods in order to fill the empty cells.

2.1.2 Missing Data Patterns

Missing data patterns provide important information about the amount and structure of missing data. Through examination of the missing data pattern, the missingness can be characterized as arbitrary or a more specialized pattern of missing data such as monotone missing data. Nonmonotone (arbitrary) missing data are used to describe a missing data pattern that have missingness interspersed among full data values while monotone missing data are a pattern in which the missing data exists at the end (reading from left to right) of the data record with no gaps between full and missing data. In other words, once a variable has missing data, all variables to the right of the missing data variable in a rectangular data array are also missing. This is an important distinction due to the manner in which missing data are imputed, moving from left to right across the rectangular data array of columns and rows. The implication for the imputation step and selection of imputation method is that a monotone missing data pattern allows the analyst more flexibility in selecting an appropriate imputation technique.

Analysis of existing missing data patterns is a critical first step in planning the overall imputation. Another important consideration in planning an imputation is the type of variables (numeric or character) that either require imputation or will contribute to the imputation process. Careful attention to the variable type will help ensure that the imputation is done correctly. Knowledge of the variables with missing data as well as variables used during the imputation will

allow the analyst to make correct decisions about how to set up the imputation.

	Monotone			Nonmonotone		
Obs	Z_1	Z_2	Z_3	Z_1	Z_2	Z_3
1	O	O	O	O	M	O
2	O	M	M	O	M	M
3	O	M	M	O	M	M
4	O	M	M	O	M	M
5	O	O	M	O	O	M
6	O	O	M	O	O	M
7	O	O	M	O	O	M
8	O	O	O	O	O	O

Fig 2.1 Monotone and Nonmonotone patterns of missingness (O=observed, M=missing)

It may be necessary to reorder variables and/or individuals to change from non-monotone pattern to monotone missing data pattern. Assumptions and patterns of missingness are used to determine which methods can be used to deal with missing data (Yuan, 2005). When working with PROC MI of SAS/STAT, one has more flexibility in selecting the imputation method. Therefore, starting with a dataset with a non-monotone missing pattern, it is desirable to attempt to convert it to have a monotone pattern. This may be possible by simply reordering the variables, but this is not always guaranteed to succeed. In this case, we can impute just enough values to make n monotone and then apply one of the methods of monotone pattern.

2.2 Missingness data mechanisms

Missing data can introduce bias into studies and obscure implication might be imbedded in the missingness. Therefore, it is important that appropriate and effective methods available to resolve the problems of missing data be used. The impact of missing data and the ways to handle incomplete data depend much upon the patterns of incompleteness. A set of definitions for missing data

mechanisms has been provided by Little and Rubin (2014), including missing at random (MAR), missing completely at random (MCAR), and non-ignorable missing data (MNAR). Data are said to be missing at random (MAR) when the probability that responses are missing depends on the set of observed responses, but is unrelated to the specific missing values that, in principle, should have been obtained.

Let Y= complete data matrix, Y^O= observed part of Y, Y^M = missing part of Y and R is missing data indicator matrix where R_{ij}=1 for missing, 0 for observed, i and j are indicating the position of Y in the row and column of a matrix. Then $P(R|Y, \varphi) = P(R|Y^O, \varphi)$ for all Y^M , φ. Where φ denotes unknown parameter. MCAR occurs when the missing values on variable Y are independent of all other observed variables and the values of Y itself. When the probability that response are missing is unrelated to either the specific values that in principle, should have been obtained or the set of observed responses. MCAR is a special case of MAR, and occurs when the distribution doesn't depend on observed data, either.

In notation, $P(R|Y, \varphi) = P(R| \varphi)$ for all Y, φ. The distinction between MCAR and MAR is that missingness cannot depend on observed values of the dependent variable Y^O in MCAR, but can in MAR. Therefore, the test of MCAR is based on analysis involving Y^O (Muthén and Khoo, 1998). The second pattern "missing at random" (MAR) provides a more realistic condition. MAR and MCAR are both ignorable when the parameters governing the missing data process are not related to the parameters of interest, and therefore it is not required to model the missingness as part of the estimation process.

Data are said to be Not Missing At Random (NMAR) when the probability that responses are missing depends on both the set of observed responses and the specific missing values that, in principle, should have been obtained. Sometimes it is referred as Missing At Not Random (MANR) or Missing Not At Random (MNAR).Since the probability of missing data are related to at least

some elements of Y^M, NMAR is often referred as non-ignorable missingness. The term non-ignorable refers to the fact that missing data mechanism cannot be ignored. When missingness is non-ignorable, it means that we cannot predict future unobserved responses, conditional on past observed responses; instead, we need to incorporate a model for the missingness mechanism. It is common for bi-factorial experiments to have only one observation per cell and additionally to have missing data. An example of this situation is in multi – environment experiments, where the cultivars are studied in different locations or environments, and each cell presents the mean of each factor level combination.

Most statistical packages, like SAS exclude observations with any missing variable values from the analysis. In effect this assumes that the missing data mechanism is MCAR which us too restrictive as earlier stated. Although using only complete cases is simple, information that is in the incomplete cases is lost. Excluding observations with missing values also ignores the possible systematic difference between the complete cases and incomplete cases, and the resulting inference might not be applicable to the population of all cases, especially with a smaller number of complete cases.

2.3 Methods of handling missing data.

In this section, we describe some of the most commonly used methods and discuss the characteristics of the method to yield valid analysis. Some important references in the field can be found in Allison (2001), Demirtas and Schafer (2003), Fitzmaurice et al. (2012), Hedeker and Gibbons (1997), Hedeker et al. (2007) and Little and Rubin, (2014) .

2.3.1 Complete-Case Analysis (CCA)

One approach to handling missing values is to simply omit all cases with missing values at any measurement occasion. This is called a Complete-Case Analysis (CCA). The advantage of this method is that it can be used for any kind of statistical analysis and no additional special computational methods are required. However, it will yield unbiased estimate of mean response trends only when the missingness is MCAR. When the missing data are not MCAR, the results from CCA may be biased because the complete case can be unrepresentative of the full population. Also, it can result in a very substantial loss of information by deleting all case with missing value, and this may lead to reduced statistical precision and power. If the missing-data problem can be resolved by discarding only a small part of the sample, then the method can be quite effective. But, CCA is very problematic and is rarely an acceptable approach to the analysis. This method can be done using PROC REG or PROC FACTOR in SAS.

2.3.2 Available Case Analysis

Another approach to handling missing values is Available Case Analysis. This is a general term for a variety of different methods that use the available information to estimate means and covariance. It can readily incorporate vectors of repeated measures of unequal length in the analysis. The popular method in available case analysis is pair-wise deletion method (Peugh and Enders, 2004). In this method, a covariance (or correlation) matrix is computed where each element is based on the full number of cases with complete data for each pair of variables. The attempt is to maximize sample size by not requiring complete data on all variables in the model. In general, Available Case Analysis is more efficient than CCA because it incorporate the partial information obtained from those who are missing. The disadvantage of this method is that the sample base changes from variable to variable according to the pattern of

missing data and the resulting correlation matrix might not be positive definite. This variability in the sample base creates practical problems such as the determination of sample size and degree of freedom. Also, it yields biased estimates of treatment comparisons unless missing data are MCAR. This method can be done using PROC CORR in SAS.

2.3.3 Single Imputation

Third approach to handling missing values is Single Imputation. This is a method that involves replacing an incomplete observation with complete information based on an estimate of the true value of the unobserved variable. It is widely used in practice because the analysis is straight forward once imputation is done. The obvious disadvantage of single imputation is that imputing a single value treats that value as known, and thus without special adjustments, single imputation cannot reflect sampling variability. Single imputation does not reflect the uncertainty about the predictions of the unknown missing values, and the resulting estimated variances of the parameter estimates are biased toward zero. This clearly leads to inflated type I error in subsequent analyses.

One of the most widely used single imputation method to handling missing value is mean imputation (Little and Rubin, 2014). This method is to fill in any missing values with mean of the non-missing values. It therefore assumes that the mean of the variable is the best estimate for any observation that has missing value on the variable. Even though it is simple to impute, this strategy can severely distort the distribution for the variable, leading to complication with summary measures including underestimates of the standard deviation. Also, the missing values require being MCAR as an assumption. Therefore, mean imputation is unacceptable in most applications. This method can be done using PROC STANDARD in SAS.

A more appealing method to carry out single imputation to handle missing values is Expectation Maximization (EM) algorithm. EM algorithm is an iterative algorithm that finds the parameters which maximize the log likelihood when there are missing values. It relies on the relationship between missing data and the unknown parameters of a data model (Dempster et al., 1977). A disadvantage of EM algorithm is that its rate of convergence can be painfully slow when there is a large fraction of missing values. Each iteration of EM consists of an E step (expectation step) and M step (maximization step). Given a set of parameter estimates, E-step calculates the conditional expectation of the complete data log likelihood given the observed data and the parameter estimates. Suppose θ^t is the current estimate of the parameter θ.

Then, $Q(\theta|\,\theta^t) = \int g(\theta|\,Y) f\,(Y^M/Y^O, \theta = \theta^t) dY^M$.

where $g(\theta|\,Y)$ is the complete data log likelihood. Given complete data log likelihood, the M step finds the parameter estimates to maximize the complete data log likelihood from E step.

$$Q\left(\theta^{(t+1)}/\theta^t\right) \geq Q(\theta/\theta^t)\ for\ all\ \theta.$$

And, these two steps are iterated until the iteration converges.

2.3.4 Multiple Imputation

The general statistical theory and framework for managing missing information has been well developed since Rubin (1987) published his pioneering treatment of multiple imputation methods for nonresponse in surveys. The most popular imputation and more robust method to handling missing value is multiple imputation (MI) (Little and Rubin, 2014). The method is valid under the ignorable assumption. MI provides a useful strategy for dealing with data sets that have missing values. Instead of filling in a single value for each missing value, a multiple imputation procedure replaces each missing value with a set

of plausible values that represent the uncertainty about the right value to impute. MI replaces each missing item with two or more acceptable values, representing a distribution of possibilities. The advantage of the method is that once the imputed data set have been generated, the analysis can be carried out using procedures in virtually any statistical package, which makes the analysis simple. Also, the inferences such as standard error, p-value, valid confidence intervals for parameters, etc. obtained from MI are generally valid because they incorporate uncertainty due to missing values. The MI can be highly efficient even if the number of imputation is relatively small, especially when between-imputation variance is not too large. However, there are some disadvantages in MI. First, since we impute some values into missing values, missing value individuals are allowed to have varying probability. Thus, individual variation is being ignored. Secondly, the uncertainty inherent in missing values is ignored because the analysis doesn't distinguish between the observed and imputed values. At last, the only disadvantage of MI over single imputation is that it takes more work to create the imputations and analyse the results. However, from SAS® version 8.2, the procedures PROC MI and PROC MIANALYZE, have been developed which improve the computing environment and save time to analyse and space to store data.

The multiple imputation inference involves three distinct phases: (a) The missing data are filled in m times to generate m complete data sets (b) The m complete data sets are analysed by using standard procedures (c) The results from the m complete data sets are combined for the inference. PROC MI creates imputed data sets for incomplete multivariate data. It uses methods that incorporate appropriate variability across the m imputations. SAS multiple imputation procedures assume that the missing data are ignorable. Once the m complete data sets are analysed by using standard procedures such as PROC REG, PROC GLM or PROC MIXED, then PROC MIANALYZE can be used to generate valid statistical inference about these parameters by combining results from m complete data sets. There are three imputation mechanisms in

PROC MI. The method of choice depends on the type of missing data pattern. For monotone missing data patterns, either a regression method or propensity score method can be used. For an arbitrary missing data pattern, a Markov chain Monte Carlo (MCMC) method can be used. Without the detail of theoretical methods, Regression method is fitted for each variable with missing values with previous variables as covariates. Propensity Score method is that observations are grouped based on propensity scores, and an approximate Bayesian bootstrap imputation is applied to each group. At last, MCMC constructs a Markov chain long enough for distribution of the elements to stabilize to a common distribution (Yuan, 2000).

A statistical package that has several features that allow the user to get inside the imputation process and evaluate the reasonableness of the resulting model imputation is mi package in R (R Development Core Team 2010). These features include: choice of predictors, models, and transformations for chained imputation models; standard and binned residual plots for checking the fit of the conditional distributions used for imputation; and plots for comparing the distributions of observed and imputed data. mi uses an algorithm known as a chained equation approach (Raghunathan, et al., 2001, van Buuren and Oudshoorn, 2000). The user can specifies the conditional distribution of each variable with missing values conditioned on other variables in the data, and the imputation algorithm sequentially iterates through the variables to impute the missing values using the specified models. The procedure to obtain sensible multiply imputed datasets approach requires four steps: setup, imputation, analysis, and validation. Each step is divided into sub steps as follows:

1. Setup
    - Display of missing data patterns.
    - Identifying structural problems in the data and preprocessing.
    - Specifying the conditional models.
2. Imputation

- ˆIterative imputation based on the conditional model.
- Checking the tit of conditional models and checking to see if the imputed values are reasonable.
- Checking the convergence of the procedure.

3. Analysis
    - Obtaining completed data.
    - Pooling the complete case analysis on multiply imputed datasets.
4. Validation
    - Sensitivity analysis.
    - Cross validation.
    - Compatibility check.

At first glance, it may seem more complicated to conduct multiple imputations using mi compared to other available imputation software. However this is because of outline four steps that other packages have traditionally ignored. mi is designed for both novice and experienced users. For the novice users, mi has a step-by-step interactive interface where users choose options from the given multiple choices and a graphical user interface (GUI) where users click buttons. For more experienced users, mi has simple commands that users can use to conduct a multiple imputation.

Another multiple missing data imputation methods, which are not so much affected by missing data patterns and assumptions is principal component analysis (PCA) based imputation method. Principal component analysis (PCA) methods basically allow performing PCA on incomplete data and thus may also be used for missing value estimation. When doing PCA one assumes that the data are restricted to a subspace of lower dimensionality. PCA aims to extract these structures there by filtering noise out. If only the most significant loadings (eigenvectors also referred to as principal components) are used for projection this can be written as:

$$X = 1 \times \bar{x}^T + TP^T + V \ . \qquad (2.1),$$

where the term $1 \times \bar{x}^T$ represents the original variable averages, X denotes the observations, $T = t_1, t_2, \cdots, t_k$ k the latent variables or scores, $P = p_1, p_2, \cdots, p_k$ the transformation matrix (consisting of the most significant eigenvectors of the covariance matrix) and V are the residuals.

Missing values may be estimated by projecting the scores back into the original space using $\hat{X} = 1 \times \bar{x}^T + TP^T$. Optimally, this produces an estimate of the missing data based on the underlying correlation structure, there by ignoring noise. This will only produce reasonable results if the residuals V are sufficiently small, implying that most of the important information is captured by the first k components.

In order to calculate the transformation matrix P one needs to determine the covariance matrix between variables or alternatively calculate P directly via singular value decomposition (SVD). In both cases, this can only be done on complete matrices. However, an approximation may be obtained by use of different regression methods.

SVD imputation, implements the SVDimpute algorithm as proposed by Troyanskaya et al. (2012). The idea behind the algorithm is to estimate the missing values as a linear combination of the k most significant loadings when PCA is applied considering variables as observations (eigengenes). The algorithm works iteratively until the change in the estimated solution falls below a certain threshold. Each step the eigengenes of the current estimate are calculated and used to determine a new estimate.

An optimal linear combination is found by regressing an incomplete variable against the k most significant eigengenes. If the value at position j is missing, the jth value of the eigengenes is not used when determining the regression coefficients. SVDimpute seems to be tolerant to relatively high amount of missing data (>10%).

Probabilistic PCA (ppca), combines an EM approach for PCA with a probabilistic model (Stacklies et al., 2007). The EM approach is based on the assumption that the latent variables as well as the noise are normal distributed. PPCA defines a likelihood function such that the likelihood for data far from the training set is much lower, even if they are close to the principal subspace. This allows improving the estimation accuracy. PPCA is tolerant to amounts of missing values between 10% to 15%. If more data are missing the algorithm is likely not to converge to a reasonable solution.

Bayesian PCA (bpca), Similar to probabilistic PCA, Bayesian PCA uses an EM approach together with a Bayesian model to calculate the likelihood for a reconstructed value. This approach seems to be tolerant to relatively high amounts of missing data (>10%) (Stacklies et al., 2007). Scores and loadings obtained with Bayesian PCA slightly differ from those obtained with conventional PCA. This is because BPCA was developed especially for missing value estimation and is based on a variational Bayesian framework (VBF), with automatic relevance determination (ARD). In BPCA, ARD leads to a different scaling of the scores, loadings and eigenvalues when compared to standard PCA or PPCA.

The algorithm does not force orthogonality between loadings. However, the (Stacklies et al., 2007), found that including an orthogonality criterion made the predictions worse. They also state that the difference between “real” and predicted Eigenvalues becomes larger when the numbers of observations are smaller, because it reflects the lack of information to accurately determine true loadings from the limited and noisy data. As a result, weights of factors to predict missing values are not the same as with conventional PCA, but the missing value estimation is improved. BPCA was proposed by Oba et al. (2003).

Inverse non-linear PCA (NLPCA), is especially suitable for data from experiments where the studied response is non-linear. NLPCA is based on training an auto-associative neural network composed of a component layer which serves as the \bottle-neck", a hidden non-linear layer and an output layer corresponding to the reconstructed data. The loadings can be seen as hidden in the network. Missing values in the training data are simply ignored when calculating the error during back-propagation. Thus NLPCA can be used to impute missing values in the same way as for conventional PCA. The only difference is that the loadings P are now represented by a neural network. A shortcoming of the current implementation is that there is no reasonable stop criterion. The quality of the estimated solution depends on the number of iterations. This should in most cases be somewhat between 500 and 1500.

Nipals PCA (Nonlinear Estimation by Iterative Partial Least Squares) is an algorithm at the root of PLS regression which can execute PCA with missing values by simply leaving those out from the appropriate inner products. It is tolerant to small amounts (generally not more than 5%) of missing data.

2.4 Data Imputation in multi-environment trials

Genotypes by environment (G×E) experiments are unbalanced because several genotypes are not tested in some environments or missing of a measurement from some plot during the experimental stage. A common way of analyzing this type of study is by imputing the missing values and then applying established procedures on the complete dataset and fitting matrix (observed + imputed), for example, the additive main effects and multiplicative interaction model (AMMI) or factorial regression (Arciniegas-Alarcón and Dias, 2009, Gauch, 2006, Romagosa et al., 2008, van Eeuwijk et al., 2005, Van Eeuwijk et al., 2007). An alternative approximation is to

work with the incomplete data using a mixed model with estimates based on maximum likelihood (M. S. Kang et al., 2004).

Several imputation methods have been suggested in the literature to solve the problem of missing values in multi-environmental trials. One of the first was made by Frensham et al. (1998), Gauch Jr and Zobel (1990), who suggested imputing the missing values iteratively by minimizing the residual sum of squares and doing the G×E analysis on the completed table, reducing the degrees of freedom by the number of missing values. This work was developed by Gauch Jr and Zobel (1990), who made the imputations using the EM algorithm and the AMMI model or EM-AMMI. Some variants of this procedure using multivariate statistics (cluster analysis) were described in Godfrey et al. (2002), and Godfrey (2004). Raju (2002) propose the EM-AMMI algorithm by treating the environments as random and suggested applying a robust statistic to the missing values in the stability analysis. Wasito (2003) proposed the imputation to be made in incomplete two-way tables using linear functions of the rows (or columns).

Other studies recommended by van Eeuwijk and Kroonenberg (1998) as having good results in the case of missing values for G×E experiments have been developed by several authors (Calinski et al., 1992, Denis, 1991, Denis and Baril, 1992, Wold, 1978). These authors found that using imputations through alternating least squares with bilinear interaction models or AMMI estimates based on robust sub models could give results as good as those found with the EM algorithm. Additionally, Caliński et al. (1999) introduced an algorithm that combines the singular value decomposition (SVD) of a matrix with the EM algorithm, obtaining results very useful for experiments in which the alternating least squares have some problems, for instance, convergence failures (Piepho, 1995). Recently, Bergamo et al. (2008), proposed a distribution-free multiple imputation method that was assessed by Arciniegas-Alarc´on (2008) and compared by Arciniegas-Alarcón et al. (2010) with

algorithms that use fixed effects models in a simulation study with real data. Meanwhile, a deterministic imputation method without structural or distributional assumptions for multi environment experiments was proposed by Arciniegas-Alarcón et al. (2010). The method uses a mixture of regression and lower-rank approximation.

Some studies for analysis of multi-environment experiments with missing values can be found in the literature. For example, methodologies for stability analysis have been studied by Raju and Bhatia (2003), Raju et al. (2009) and Raju et al. (2006). Recently, Pereira et al. (2007), Rodrigues et al. (2011), Rodrigues (2012) assessed the robustness of joint regression analysis and AMMI models without the use of data imputation.

Given the historical information about data imputation in experiments, and specifically in two-factor G×E experiments, the objective of the next section is focus an application of a deterministic imputation algorithm without distributional or structural assumptions, using an extension of the cross-validation by eigenvector method presented by Bro et al. (2008) and principal component analysis (PCA) based imputation method.

2.4.1 Data Imputation Using the Cross-Validation by Eigenvector Method.

The cross-validation method was presented by Bro et al. (2008) to find the optimum number of principal components in any data set that can be arranged in a matrix form. In this approximation, principal component analysis (PCA) models are calculated with one or several samples left out and the model is used to predict these samples. The method used cross-validation "leave-one-out" and the same study showed it to be more efficient than other well-known methodologies used in multivariate statistics, such as those presented by Eastment and Krzanowski (1982a), Wold (1978). Because of this finding,

Arciniegas-Alarcón et al. (2011) used the method to determine the best AMMI models in (G×E) experiments.

This methodology works through different stages. The first step is, consider the $n \times p$ matrix $\mathbf{X}$ with elements x_{ij} ,$(i = 1,\ldots,n;\ j = 1,\ldots,p)$. The matrix is divided into disjoint groups, each group is deleted in turn (leave-one-out), and a PCA model ($\mathbf{Z}$,$\mathbf{P}$) is obtained from the remainder by solving

$$min\left\|\mathbf{X}^{(-i)} - \mathbf{Z}\mathbf{P}^{T}\right\|_{m}^{2} \tag{2.2}$$

with $m \leq (n-1\,, p-1)$. Here $\mathbf{X}^{(-i)}$ represents the matrix after deleting the i^{th} group (leave-one-out), $\|.\|^2$ defines the squared Frobenius norm, $\mathbf{P}^{\mathrm{T}}\mathbf{P}=\mathbf{I}$, and $\mathbf{Z}$, $\mathbf{P}$ are scores and loadings matrices with dimensions $(n-1) \times m$ and $p \times m$ respectively, where p is the number of columns and m is the number of components. Note that, in this method the deleted group corresponds to the i^{th} row of $\mathbf{X}$ and according to Smilde A. et al. (2004) the model (1) can be rewritten in terms of the singular value decomposition (SVD)

$$\mathrm{X}^{(\mathrm{i})} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathrm{T}} = \sum_{\mathrm{k}=1}^{\mathrm{m}} \mathbf{U}_{\mathrm{k}}\mathrm{d}_{\mathrm{k}}\mathbf{V}_{\mathrm{k}}{}^{\mathrm{T}} \tag{2.3},$$

where $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_{\mathrm{m}}]$, $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}, \cdots, \mathbf{V}_{\mathrm{m}}]$, $\mathbf{D} = \mathrm{diag}[\mathrm{d}_1, \mathrm{d}_2, \cdots, \mathrm{d}_{\mathrm{m}}]$, $\mathbf{Z} = \mathbf{U}\mathbf{D}$ and $\mathbf{P} = \mathbf{V}$

A second step is a procedure to estimate the score

$$\mathbf{t}^{(-\mathrm{j})\mathrm{T}} = \mathbf{x}_{\mathrm{i}}{}^{(-\mathrm{j})\mathrm{T}}\mathbf{P}^{(-\mathrm{j})\mathrm{T}}\left(\mathbf{P}^{(-\mathrm{j})\mathrm{T}}\mathbf{P}^{(-\mathrm{j})}\right)^{-1} \tag{2.4},$$

where $\mathbf{P}^{(-\mathrm{j})\mathrm{T}}$is the loading matrix found in step 1 with the j^{th} row excluded. $\mathbf{x}_{\mathrm{i}}{}^{(-\mathrm{j})\mathrm{T}}$ is a row vector containing the i^{th} row of $\mathbf{X}$ except the Jth element.

A third step, estimate the element x_{ij} by $\hat{x}_{ij}{}^{(m)} = \mathbf{t}_{\mathrm{i}}{}^{(-\mathrm{j})\mathrm{T}}\mathbf{P}_{\boldsymbol{j}}{}^{\mathrm{T}}$. $\mathbf{P}_{\mathrm{j}}$ is the j^{th} row of $\mathbf{P}$ the last Step which includes finding the prediction error of the $(\mathrm{ij})^{\mathrm{th}}$ element, $\mathrm{e}_{\mathrm{ij}}{}^{(\mathrm{m})} = \mathrm{x}_{\mathrm{ij}} - \hat{\mathrm{x}}_{\mathrm{ij}}{}^{(\mathrm{m})}$ and obtaining the criterion value

$$\mathrm{PRESS(m)} = \sum_{i=1}^{n}\sum_{j=1}^{p}\left(e_{ij}^{(m)}\right)^2 \qquad (2.5)$$

In order to make the imputation of missing values in the matrix from (G × E) experiments, a change in the imputation method is proposed following the work of Arciniegas-Alarcón et al. (2010), Bergamo et al. (2008), Krzanowski, (1988) and Smilde A. et al. (2004) and using the singular value decomposition of a matrix (Good, 1969).

Initially, suppose that $n \geq p$ and the matrix **X** has several missing values; in the case $n < p$, the matrix should first be transposed. The missing values are replaced by their respective column means $\bar{x}_j$, and after this has been done the matrix is standardized by columns, subtracting $\bar{x}_j$ and dividing by $\bar{s}_j$ (where $\bar{x}_j$ and $\bar{s}_j$ represent, respectively , the mean and the standard deviation of the j^{th} column). The eigenvector procedure using the SVD in expressions (2.2)–(2.4) is applied to the standardized matrix to find the imputation of the (i,j) element, denoted by $s_j\hat{x}_{ij}^{(m)}$. After the imputation, the matrix must be returned to its original scale, $x_{ij} = \bar{x}_{ij} + s_j\hat{x}_{ij}^{(m)}$.

At this point the matrix does not have any missing values, but the imputations are rather basic and need to be refined. In the initial stage of imputation an iterative scheme is advocated, iterations continuing until the imputations achieve convergence (i.e., there is stability in successive imputed values), but Caliński et al. (1999) showed that this convergence is not always necessary when using a method that combines the EM algorithm with SVD. Therefore, taking this into account, It is possible to fix in advance the number of iterations between 0 and 3, as well as permitting the process to run until convergence has been achieved. As regards to the computing effort, convergence can depend strongly on the size of matrix analyzed and also on the data structure (size of correlations, proportion of missing values, etc.). But for instance, the SVD method of Hastie et al. (1999), convergence is achieved

usually between 5 and 6 iterations, and in the Bergamo et al. (2008) method it is achieved in between 20 and 50 iterations maximum.

On the other hand, the data imputation depends directly on (2.2) and (2.3). Equation (2.2) needs prior choice of the number of components (m) to extract from the SVD. Bergamo et al. (2008) and Krzanowski (1988) took $m = min\{n-1, p-1\}$ with the objective of using the maximum amount of available information, but Hedderley (1995) asserted that if the estimation is based on the choice of a unique fixed number of dimensions, some of the lower dimensions may be essentially random. This can influence the imputation within an iterative scheme and can lead to the estimates becoming trapped in a cycle, hence preventing convergence. To solve this problem, Josse et al. (2011) suggested including a test to check on the convergence rate, and in case a specific criterion is not being attained the number of dimensions should be reduced. Another option that has satisfactory results, suggested by Josse et al. (2011) to choose an optimum m, is through cross-validation based uniquely on the observed data. However, the computational cost of this option is likely to be high.

Taking into account all the above mentioned in this sub-section, for imputation of each missing value of the matrix **X** the value of m in (2.2) is allowed to be different in each SVD calculated and is chosen according to the criterion used by Arciniegas-Alarcon et al. (2010). Thus, m is chosen such that $((\frac{\sum_{k=1}^{m} {d_k}^2}{\sum_{k=1}^{\min(n-1,p-1)} {d_k}^2})) \approx 0.75$. Moreover, in (3.3), the Moore-Penrose generalized inverse can be used instead of the classic inverse matrix as was studied in cross-validation by Dias and Krzanowski (2003). Five imputation methods using the cross-validation by eigenvector method have been assessed. These five imputations are denoted Eigenvector0, Eigenvector1, Eigenvector2, Eigenvector3, and Eigenvector where the number indicates the number of iterations used while in the case of Eigenvector the process is iterated until convergence is achieved in the imputations.

These imputation methods are all deterministic imputations, and they have the advantage over other stochastic imputation methods (parametric multiple imputations) that the imputed values are uniquely determined and will always yield the same results when applied to a given data set. This is not necessarily true for the stochastic imputation methods (Bello, 1994).

2.5 Comparison Criteria.

In general, the objective after imputation is to estimate model parameters from the complete table of information. One of the models frequently used in genotype-by-environmental trials is the AMMI (Gauch, 1992b, Gauch, 2006), and for this reason the methods of missing data imputations mentioned in the previous sections can be compared through the genotypic and environmental parameters of the fitted AMMI models using the root mean squared predictive difference (RMSPD) (Dias and Krzanowski, 2003).

The AMMI model is first briefly presented here and also in detail in the next chapters. The usual two-way ANOVA model to analyze data from genotype-by-environment trials is defined by

$$y_{ij} = \mu + a_j + b_j + ab_{(ij)} + e_{ij} \qquad (2.6),$$

$(i = 1,\ldots,n; j = 1,\ldots,p)$ where, $\mu, a_j, b_j, ab_{(ij)}$ and e_{ij} are respectively, the overall mean, the genotypic and environmental main effects, the genotype-by-environment interaction, and an error term associated with the i^{th} genotype and j^{th} location. It is assumed that all effects except the error are fixed effects. The following re-parameterization constraints are imposed:$\sum_i (ab)_i = \sum_j (ab)_j = \sum_i a_i = \sum_j b_j = 0$.

The AMMI model implies that interactions can be expressed by the sum of multiplicative terms. The model is given by

$$y_{ij} = \mu + a_j + b_j + \theta_1 \alpha_{i1} \beta_{j1} + \theta_2 \alpha_{i2} \beta_{j2} + \cdots + e_{ij} \qquad (2.7),$$

where, θ_i, α_{ij} and $\beta_{ij} (l = 1,2,\ldots, min(n-1, p-1)$ are estimated by the SVD of the matrix of residuals after fitting the additive part. θ_i is estimated by the i^{th} singular value of the SVD, α_{ij} and β_{ij} are estimated by the genotypic and environmental eigenvector values corresponding to θ_i.

Alternating regressions can be used in place of the SVD García-Peña and Dias, (2009); depending on the number of multiplicative terms, AMMI family model that was used in this chapter for example, if no component is considered significant by the procedure of, we have the AMMI0 model that contains only the additive effects of genotypes and environments, without GEI. If a component is considered significant, we have the AMMI1 model, which contains a component that explains GEI, beyond the additive genotypes × environments effects. For two significant components, we have the AMMI2model that contains two components, which explain GEI, beyond the additive genotypes × environments effects, and so forth. These models may be called AMMI0, AMMI1, and so forth.

AMMI0 $$y_{ij} = \mu + a_j + b_j + e_{ij}$$

AMMI1 $$y_{ij} = \mu + a_j + b_j + \theta_1 \alpha_{i1} \beta_{j1} + e_{ij}$$

AMMI2 $$y_{ij} = \mu + a_j + b_j + \theta_1 \alpha_{i1} \beta_{j1} + \theta_2 \alpha_{i2} \beta_{j2} + e_{ij}$$

An inherent requirement of the AMMI model is prior specification of the number of multiplicative components (Dias and Krzanowski, 2006, García-Peña and Dias, 2009, Hedderley, 1995). Rodrigues (2012) made an exhaustive analysis of the related literature and concluded that usually two or three components can be used because, in general, one component is not enough to capture the entire pattern of response (main effect and interaction effect between genotypes and environment) in the data, but with more than three components there are obvious visualization problems, and a huge quantity of noise is liable.

A comparison criteria used in this section the root mean squared predictive difference (RMSPD) are defined as follows

$RMSPD(gen) = \sqrt{\frac{\sum_{i=1}^{NG}(a_i - \hat{a}_i)^2}{NG}}$	$RMSPD(env) = \sqrt{\frac{\sum_{j=1}^{NG}(b_i - \hat{b}_i)^2}{NG}}$
$RMSPD_l(envmul) = \sqrt{\frac{\sum_{h=1}^{i}\sum_{i=1}^{NE}(\beta_{jh} - \hat{\beta}_{jh})^2}{(NG)i}}$	$RMSPD_l(genmul) = \sqrt{\frac{\sum_{h=1}^{i}\sum_{i=1}^{NG}(\alpha_{ih} - \hat{\alpha}_{ih})^2}{(NG)i}}$

Here $RMSPD(gen)$ represents the RMSPD among the estimated parameters for genotype main effects from the original data a_j and the corresponding parameters obtained from the completed data sets by imputation $\hat{a}_j$. $RMSPD\,(env)$ represents the RMSPD among the estimated parameters for environments main effects from the original data b_j and the corresponding parameters obtained from the completed data sets by imputation $\hat{b}_j$. $RMSPD_l(genmul)$ represents the equivalent RMSPD for the pairs of estimated parameters of genotype multiplicative components $\alpha_{ih}, \hat{\alpha}_{ih}$. $RMSPD_l(envmul)$ represents the equivalent RMSPD for the pairs of estimated parameters of environments multiplicative components $\beta_{jh}, \hat{\beta}_{jh}$. In the statistics, NG represents the number of genotypes, NE the number of environments, and $l = 2\ or\ 3$ depending on the considered model AMMI2 or AMMI3.

The best imputation method is the one with the lowest values of RMSPD in each case. Summarizing, in each imputed data set with missing values, an application of the methods Eigenvector, Eigenvector0, Eigenvector1, Eigenvector2, and Eigenvector3 and, then, in the completed data (observed + imputed) a fitted AMMI2, AMMI3 models for the calculation of the respectively RMSPD statistics. In order to visualize any differences more readily, the

RMSPD values were standardized and the comparison was made directly. Note that because of the standardized scale, the values of the statistics can be either positive or negative. Figure 2.2 shows the $RMSPD(gen)$ distribution on the standardized scale for the BRVII data set, showing each imputation method and each percentage. It can be seen that the Eigenvector distribution is left asymmetric i.e most standardized RMSPD values are less than 0, and this asymmetry increases as the missing values percentage increases.

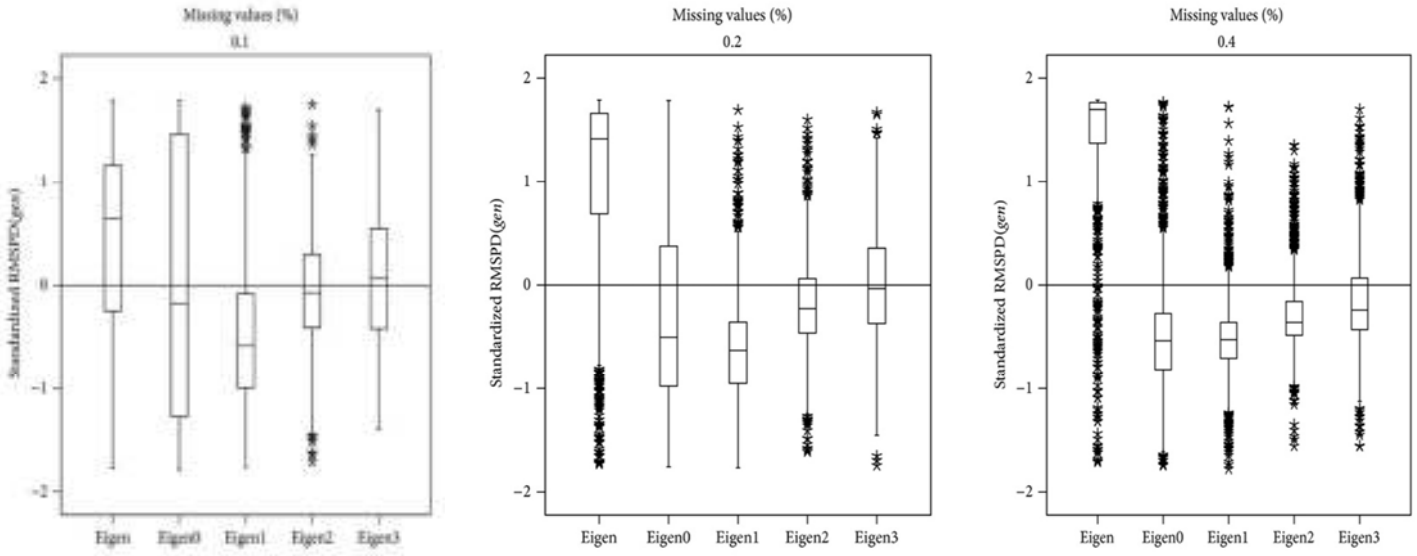


Figure 2. 2 Box plot of the RMSPD(gen) distribution in BRVII data set.

In general, the Eigenvector distribution has values above zero and when the number of missing values increases, it is concentrated above one. This means that this method had the biggest differences among the additive genotypic parameters of the real and completed (by imputation) data.

The best method according to $RMSPD(gen)$ is Eigenvector1, the method with just one iteration. This method has the smallest median for the 0.1% and 0.2% percentages. In the 0.4% percentage the medians of Eigenvector0 and Eigenvector1 are practically the same in the figure, but Eigenvector1 continues be preferable because it has the smallest dispersion. So, Eigenvector1 gave the

smallest differences between the additive genotypic parameters of the real and completed data.

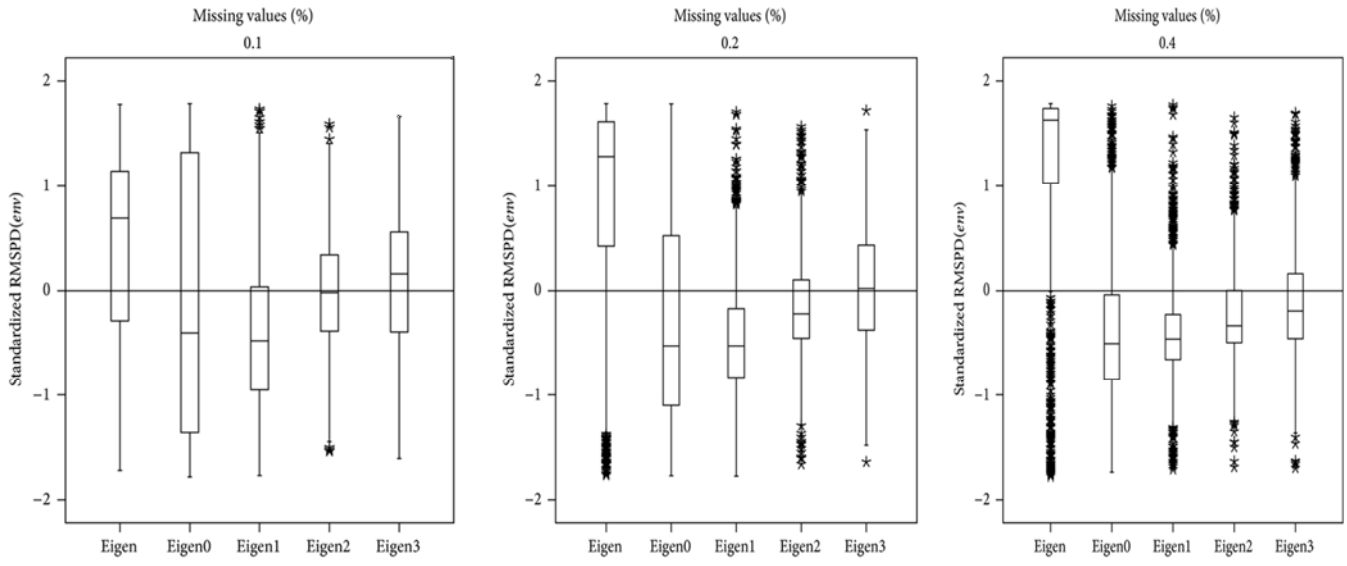


Figure 2. 3 Box plot of the RMSPD(env) distribution in BRVII data set.

Figure 2.3 shows the $RMSPD\ (env)$ on the standardized scale for the BRVII data set. It shows very similar behaviour to that of $RMSPD\ (gen)$.Again the Eigenvector method presents the biggest differences among the additive environment parameters of the real and completed data because of the algorithm that maximizes the $RMSPD\ (env)$. In this case, the $RMSPD\ (env)$ is minimized with Eigenvector0 and Eigenvector1, and in all the percentages of missing values the two have nearly equal medians. However, Eigenvector1 has the smallest dispersion and that makes this again the method of choice.

The box plot figures from RMSPD was useful in determining the best imputation method for $RMSPD\ (gen)$, and $RMSPD\ (env)$the and distributions, but in the case of $RMSPD\ (genmul)$, and $RMSPD\ (envmul)$, a more formal analysis can be used to compare the distributions; for instance the Friedman

nonparametric test and, if this is significant, then the Wilcoxon test (Sprent and Smeeton; 2001). Table 2.1 shows the Friedman test statistics. It can be seen that a significant difference exists among the imputation methods for the 0.1% and 0.2% percentage of missing values, but with 0.4% the five methods have equivalent results. After the general test, it is necessary to make multiple pairwise comparisons for the two lower percentages.

Table 2. 1 Test statistic for the standardized BRVII data set.

	RMSPD2(genmult)		RMSPD3(genmult)		RMSPD2(envmult)		RMSPD3(envmult)	
Perc	Friedman	P-value	Friedman	P-value	Friedman	P-value	Friedman	P-value
0.1%	15.62	0.0036	34.48	0.0000	34.93	0.0000	30.49	0.0000
0.2%	10.78	0.0291	11.36	0.0227	16.71	0.0022	11.11	0.0254
0.4%	2.84	0.5847	2.55	0.6345	4.94	0.2925	5.94	0.2033

Application of PCA method of multiple imputations for a missing data on mult-environmental crop trial data are also showing the advantage of these methods. The independence on missing data pattern or some specific assumptions are main advantage of PCA imputation method. In most case of small plot size mult-environmental crop variety trial data it is a common phenomenon missing a data at plot level without showing any specific type of pattern. Multiple imputation of such a missing data using these PCA methods is necessary, for a reason of most statistical applications on multi-environmental data analysis such as AMMI and GGE are applicable only for a balanced data case.

A random deletion of 5% data for a data set BW01RVII was performed and a PCA methods are applied for multiple imputation. The random deletion process form a data matrix $\boldsymbol{X}\,(n \times p)$ was conducted as follows. Random numbers between 0 and 1 were generated in R with the runif function (R Development Core Team 2010). For a fixed r value $(0 < r\,1)$, if the $(pi + j)$th random number was lower than r , then the element in the $(i + 1,\ j)$ position of the

matrix was deleted $(i = 0,1,2,\cdots,n; j = 1,\cdots,p)$. The expected proportion of missing values in the matrix will be r. This technique was used with $r = 0.05$ (i.e, 5%). The comparison and the performance of these PCA imputation (PCA, PPCA, BPCA, SVDimput, Nipals PCA and Nlpca) method is tested by a graphical eigenvalue structure as obtained with different methods and a simple linear regression analysis fit of original data set and the new imputed data (original + imputed).

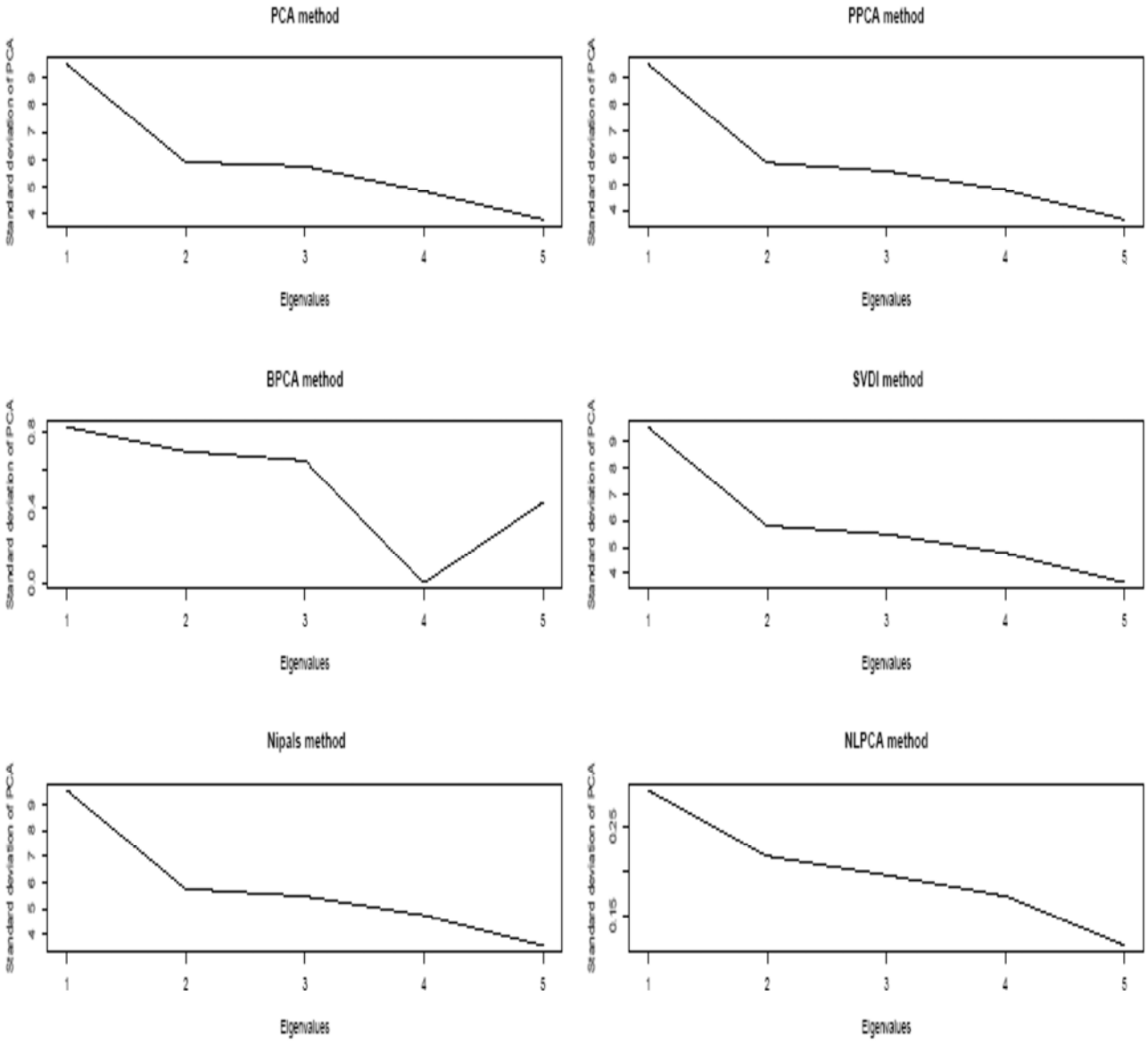
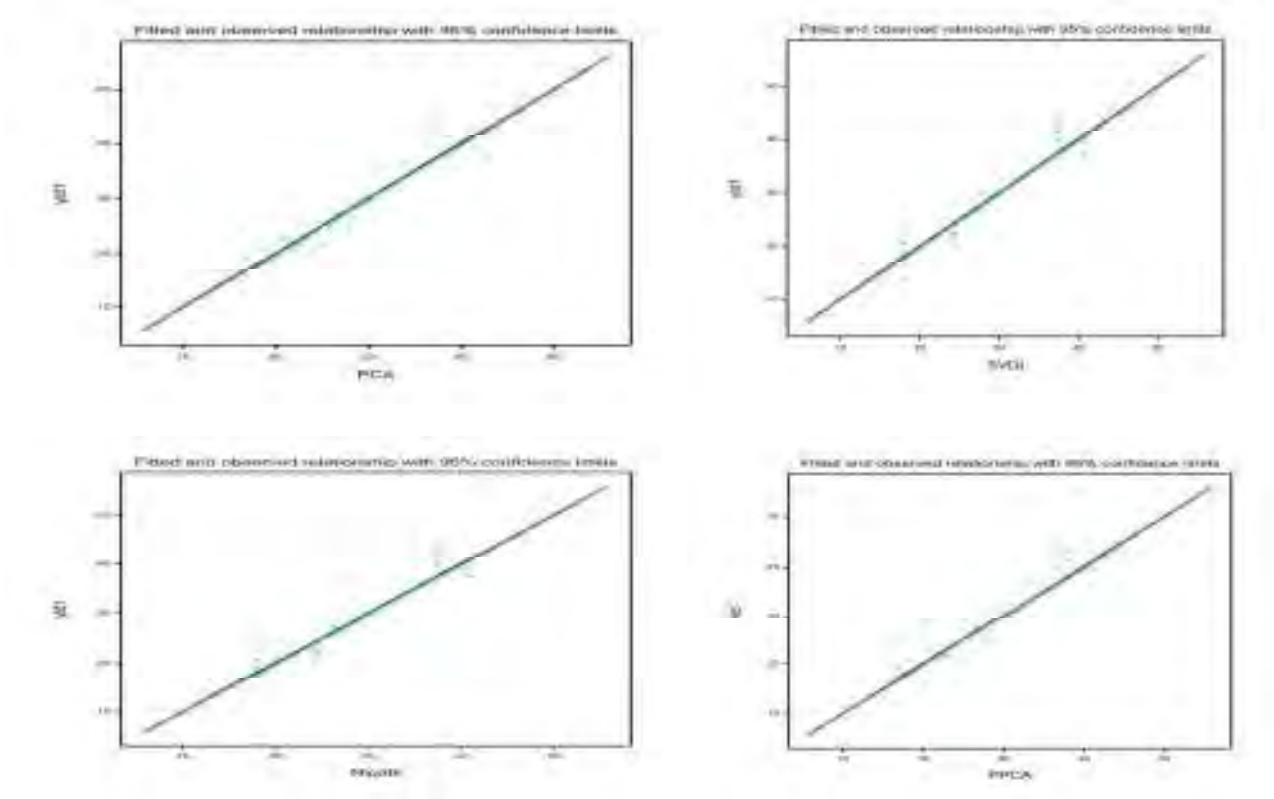


Figure 2.4 Eigenvalue structure as obtained with PCA, PPCA, BPCA, SVD, Npals and NLPCA methods

Figure 2.4 shows a plot of the eigenvalue structure. If most of the variance is captured with few loadings PCA is likely to produce good missing value estimation results. For BW01RVII data set all methods show similar eigenvalues. One can also see that most of the variance is already captured by the first loading, thus missing data estimation is likely to work fine on this data. For BPCA the eigenvalue are scaled differently for reason of it is based on variational bayesian framework (VBF), with automatic relevance determination (ARD) . To get an impression of the correctness of estimation it is a good idea to plot the scores/loadings obtained with classical PCA and one of the probabilistic methods against each other. A simple regression fit Figure 2.5 for original data set (BW01RVII) and imputed values of PCA methods show the similarity on most of the multiple imputation of (PCA, SVDimput, PPCA,Nipals, BPCA and Nipca) with a Percentage variance accounted for 98-99.1



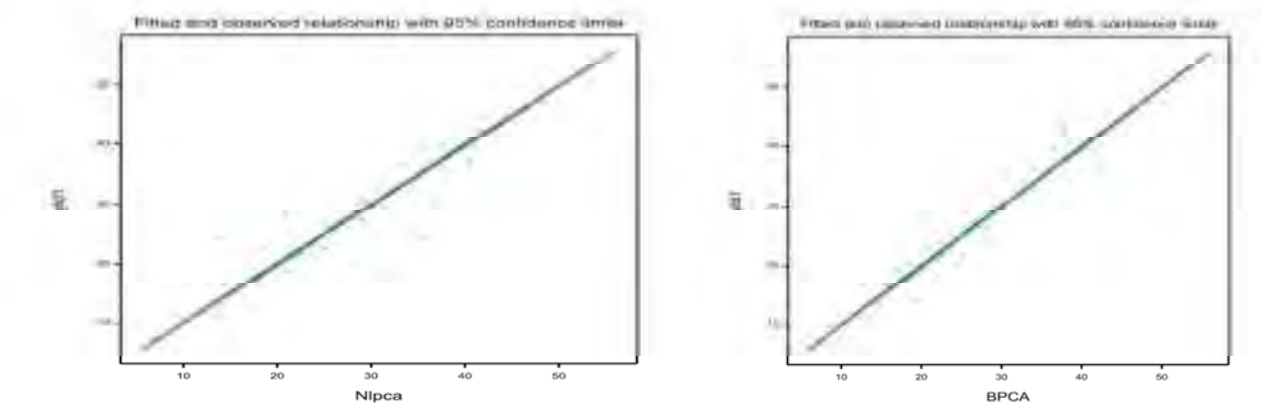


Figure 2. 5 Plot for fitted and observed values.

For a summary of this chapter, a multi-environmental small size plot crop variety trial which is known usual for missing data at a plot level. The modern treatment of missing values suggests multiple imputations as an alternative to find the standard error of the statistics of interest (van Eeuwijk and Kroonenberg, 1998), but in the case of deterministic imputation a solution well known and tested with success can be applied. This is the proportional bootstrap method proposed by Bello (1993), in which the proportion of present and missing values that appear in each bootstrap sample is exactly equal to the proportion that appear in the original incomplete data.

Another aspect that can be of interest is the mechanism producing the missing data. Generally, in situations that involve the assessment of several genotypes in different environments, missing observations follow one of the definitions proposed by Little and Rubin (2014), namely, missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). Values missing completely at random can occur, for example, when plants are damaged due to uncontrollable factors in the experiments, or by incorrect data measurement or transcription. In this case the cause of the missing value is not correlated with the variable that has it. However, in the genotypes test program in which the cultivars are chosen during each year, using only the

observed data without considering the missing values, the missing mechanism is clearly random MAR, (Piepho and Möhring, 2006). The last type of missing, MNAR, can be seen usually when the same subset of genotypes can be missing in some environments of the same sub region, because the plant breeder in the location does not like these genotypes. So, a genotype missing in one environment possibly will be missing too in other environments. In these cases, the mechanism that produces missing values is naturally not at random. The present chapter has focused exclusively on the MCAR mechanism, and further research is needed to study the remaining mechanisms.

Finally, the proposed methods in this chapter have easy computational implementation, but one of the main advantages is that a cross validation or PCA methods of imputation do not make any distributional or structural assumptions and do not have any restrictions regarding the pattern or mechanism of missing data in experiments. Once we see the options and methods of handling missing data of multi-environmental trial, the next two chapters focus on the commonly applicable methods of analysis AMMI and biplot techniques which only applicable for a balanced (data without missing).

Chapter 3

Biplot Analysis of Multi-Environmental Trials

3.1 Introduction

A demand for agricultural product increase from time to time. To meet this demand various crop improvement programmes have been initiated by the Ethiopian Institute of Agricultural Research (EIAR). Under any crop improvement programme a sample of promising genotypes are performance-tested asserted each year at a number of sites, representing major crop growing areas with the a view to identify genotypes which possess the dual qualities of high yield capacity and low sensitivity to adverse change in environmental condition. One of the important focuses in the current chapter is to assess the performance of improved genotypes in multi environment (multi-location, multi-year or both) trials.

Multi-Environment Yield Trials (MET) are conducted for different crops throughout the world (Dehghani et al., 2006, Yan and Kang, 2002, Yan and Rajcan, 2002) not only to identify high yielding cultivars but also to identify sites that best represent the target environment (Yan, 1999, Yan et al., 2001, Yan et al., 2000). As usual in MET a number of genotypes are tested over a number of sites and years to see adaptation of the crop. But, it is often difficult to determine the pattern of genotypic responses across environments without the use of appropriate analytical and statistical tools such as additive main effects and multiplicative interactions (AMMI) and genotype main effect and genotype × environment interaction (GGE) biplot (Gauch, 1992b, Gauch and Zobel, 1996, Yan et al., 2000, Yan and Tinker, 2006a) for graphical display of data.

The measured yield of each cultivar in each test environment is a result of genotype main effect (G), and environment main effect (E) and genotype by environment (G×E) interaction (Yan and Kang, 2003). Though, E mostly accounts for about 80% of the total yield variation; it is only G and G×E interaction that are relevant to cultivar evaluation and mega environment classification (Kaya et al., 2006, Rao and Polignano Prabhakaran, 2005, Yan et al., 2000, Yan and Kang, 2002, Yan, 2002). AMMI and GGE models are analyses of variance and singular value decomposition (SVD) based statistical analyses often applied to yield trial studies for visualizing the data. The methods helps in understanding complex genotype by environment (G×E) interactions, determining which genotype has been best in which environments, and also helping in grouping environments with the same winner (or similar winners) into mega-environments.

Understanding genotype by environment interaction (GEI) helps plant breeders to design better breeding strategies. Therefore, the objectives of this chapter are to evaluate the yield performance and stability of genotypes in relation to environment (location) on year to year basis. Secondly the study will examine the possible existence of different mega environments and the wining genotype for each mega environments and consistency of genotype performance on a year to year basis.

3.2 The Model

Plant breeding programs commonly analyse the existence of a genotype by environment interaction in a two-way table (genotypes by environments).This type of table features multi-environment trials (MET), where it is important to test general and specific adaptation of genotypes. The genotypes are influenced by different environmental conditions and may show significant variation in the yield performance in relation to other genotypes. This type of behaviour is known as GEI.

The AMMI and GGE method, recommended by Crossa (1990) and Ferraudo and Perecin (2014), is nothing more than a combination between the usual univariate analysis of variance (ANOVA) and principal components analysis (PCA), which can be treated directly through the mathematical technique called singular value decomposition (SVD). AMMI and GGE are relies, initially, on the estimation of additive effects of genotypes and environments by the method of conventional variance analysis.

The residuals obtained from this matrix constitute the interactions matrix where the GEI effects are estimated, considered multiplicative, using PCA. Interims of effects the basic model for a multi-environment trial can be written as

$$Y_{ijl} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijl}. \quad (3.1),$$

where Y_{ijl} is the measured yield value of the i^{th} genotype in the jth environment and lth replicate, μ is the grand mean, α_i is the main effect of the ith genotype, β_j is the main effect of jth environment, γ_{ij} is interaction between i^{th} genotype and jth environment and ϵ_{ijl} is random error. Were we assume that $\epsilon_{ijl} \sim$ indep $N(0, \delta_j{}^2)$. The ranges of indices are $i = 1, 2,..., 20$ $j = 1, 2,..., 6$ $l = 1, 2, 3, 4$

Thus the cell mean for the model is

$$\mu_{ij} = E(Y_{ijl}) = \alpha_i + \beta_j + \gamma_{ij} \quad (3.2)$$

In GGE biplots genotype plus genotype × environment (*G + GE*) interaction are studied together and to achieve this G+GE effect is separated out from the observed mean and eventually the model becomes (omitting the random error)

$$\mu_{ij} - (\mu + \beta_j) = \alpha_i + \gamma_{ij}. \quad (3.3)$$

However in the case of the AMMI model, the effect of genotypes is also separated out only genotype × environment (GE) interaction is studied for biplot, and eventually the model becomes

$$\mu_{ij} - (\mu + \beta_j + \alpha_i) = \gamma_{ij}. \quad (3.4)$$

The mathematical expressions for partitioning of G+GE for GGE biplots and GE for AMMI models are similar except a difference in model formulation. The G+GE for GGE and GE for AMMI effects are partitioned into multiplicative terms by using the singular value decomposition (SVD) as

$$\mu_{ij} - \mu - \beta_j = \lambda_1 \xi'_{i1} \eta_{j1} + \lambda_2 \xi'_{i2} \eta_{j2} + \gamma_{ij} \text{ and}$$

$$\mu_{ij} - \mu - \beta_j - \alpha_i = \lambda^*_1 \xi'^*_{i1} \eta^*_{j1} + \lambda^*_2 \xi'^*_{i2} \eta^*_{j2} + \gamma_{ij}^*. \quad (3.5)$$

respectively, where λ_1 (λ^*_1) and λ_2 (λ^*_2) are the singular values (SV) for the first and second principal component (PC1 and PC2), ξ'_{i1} (ξ'^*_{i1}) and ξ'_{i2} (ξ'^*_{i2}) are eigenvectors of genotype i for PC1 and PC2, η_{j1} (η^*_{j1}) and η_{j2} (η^*_{j2}) are eigenvectors of environment j for PC1 and PC2 and γ_{ij} (γ_{ij}^*) is the residual not explained by PC1 and PC2 for genotype i in environment j. The PC1 and PC2 eigenvectors cannot be plotted directly to construct a meaningful biplot before the singular values are partitioned into the genotype and environment eigenvectors. To generate a biplot that can be used in visual analysis of MEYT data, the SVs have to be partitioned into the genotype and environment eigenvectors so that Eq. (3.5) can be written in the form of

$$\mu_{ij} - \mu - \beta_j = \sum_{l=1}^{k} g'_{il} e'_{lj} + \gamma_{ij} \text{ and } \mu_{ij} - \mu - \beta_j - \alpha_i = \sum_{l=1}^{k} g'_{il} e'_{lj} + \gamma^*_{ij}. \quad (3.6),$$

where g'_{il} and e'_{lj} are called PCl scores for genotype i and environment j, respectively. In a biplot, genotype i is displayed as a point defined by all g'_{il} values, and environment j is displayed as a point defined by all e'_{lj} values (l = 1

and 2 for a two-dimensional biplot). Singular-value partitioning is implemented by

$$g'_{il} = \lambda_l^{f_i} \xi_{il} \; and \; e'_{lj} = \lambda_l^{1-f_i} \eta_{lj} . \qquad (3.7),$$

where f_i is the partition factor for PCl. Theoretically, f_i can be anything between 0 and 1 although 0.5 is so far the most commonly used partition factor (Yan, 2002). In this chapter we have use a value of 0.5 to give equal importance to both genotype and environment.

3.3 Result and Discussion of Graphical Statistical Methods Based on GGE Biplot Analysis

The data used in the current chapter are from a study carried out between 2004 and 2005 in six different research stations in Ethiopia. The locations consist of loc1 (Kulumsa), loc2(Adet), loc3 (Bekoji), loc4 (Sinana), loc5 (Holeta) and loc6 (DeberZeit). Twenty bread wheat genotypes were evaluated in each of the above locations (environments) in a randomized complete block design with four replications. These Twenty genotypes are coded from G1- G20.

The AMMI analysis of variance of grain yield (Table 3.1) showed significant effects of genotype, environment (location) and Genotype by Environment interaction. Location explained 84.65% of the total (G + E + GE) variation of year 2004 and 70.63% for year 2005, whereas the genotype by environment interaction and genotype captured 12.5% and 0.0029% of year 2004 and 15.34% and 14.03% for year 2005, respectively. The magnitude of genotype by environment interaction as compared to genotype suggested a possible existence of different mega environments in year 2004. The partitioning of GGE sum of squares through the GGE biplot analysis showed that PC1 and PC2 accounted 43.21% and 26.43% of GGE sum of squares of year 2004 and 58.01% and 22.14% for year 2005, respectively. The two principal components

explained a total of 69.6% and 80.16 % variation in the two years respectively. Nonetheless agricultural biplot literature provides no guidance concerning how much of the total variability accounted for by the first two principal components are considered adequate (Sabaghnia et al., 2012b, Yang et al., 2009b). This result revealed that there was a differential yield performance among wheat genotypes across testing environment (location) due to the presence of genotype by environment interaction.

Table 3.1 ANOVA table for AMMI model.

	Year 2004					Year 2005				
Source	df	SS	MS	F	F-prob	df	SS	MS	F	F-prob
Total	479	54590	114			399	27188	68.1		
Treatments	119	41599	349.6	10.2	<0.001	99	19806	200.1	9.93	<0.001
Genotypes	19	1187	62.5	1.82	0.01944	19	2779	146.3	7.26	<0.001
Environments	5	35212	7042.4	99.8	<0.001	4	13988	3497.1	31.97	<0.001
Block	18	1270	70.6	2.06	0.00706	15	1641	109.4	5.43	<0.001
Interactions	95	5200	54.7	1.6	0.00134	76	3038	40	1.98	0.00003
IPCA	23	2035	88.5	2.58	0.00012	22	1459	66.3	3.29	<0.001
IPCA	21	1588	75.6	2.21	0.00193	20	897	44.9	2.23	0.00227
Residuals	51	1577	30.9	0.9	0.66493	34	682	20.1	1	0.47979
Error	342	11721	34.3			285	5742	20.1		

Note. The block source of variation refers to blocks within environments. (IPCA) Interaction Principal Component Axes

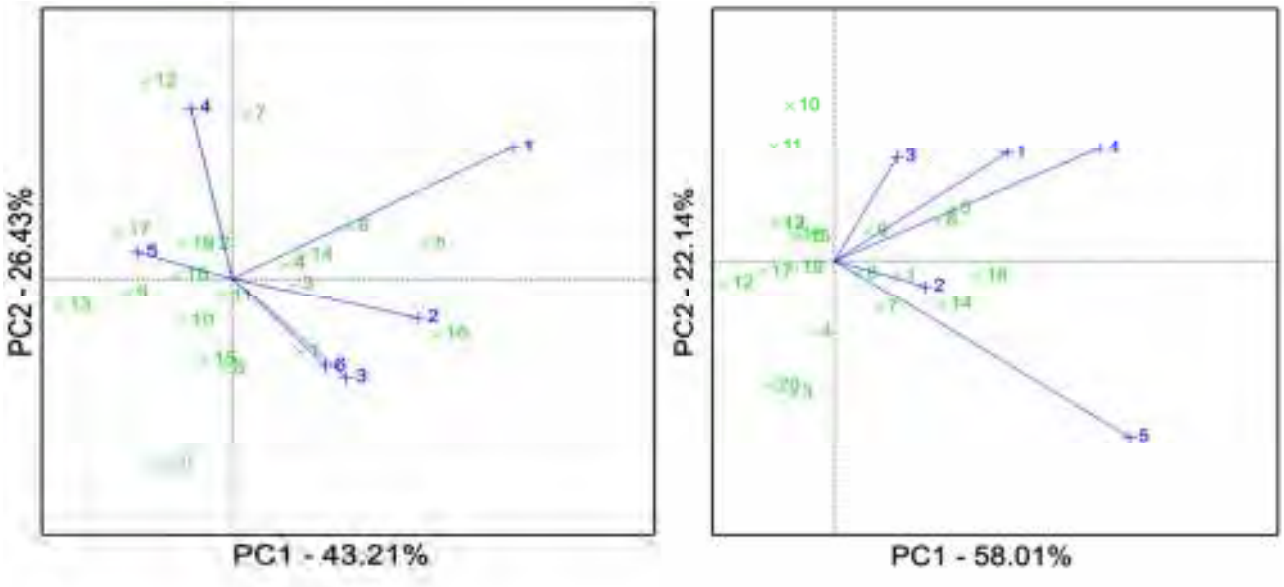
3.3.1 Relationships Among Test Environments

GGE biplot, which was based on environment focussed scaling, was used to estimate the pattern of environments (locations) as shown in Fig 3.1. Environment PC1 score had both negative and positive scores indicating that there was a difference in rankings of yield performance among genotypes across environments leading to cross-over G × E interactions.

Like PC1, the environment PC2 scores had both positive and negative values. This gave rise to crossover, leading to inconsistent genotype yield performance

across environment (locations). To visualize the relationship between environments, lines are drawn to connect the test environments to the biplot origin known as environment vectors. The cosine of the angle between two environments is used to approximate the correlation between them as described and used in Dehghani et al. (2009), Dehghani et al. (2010), Kaya et al. (2006), Yan and Tinker (2006a). For example locations 2, 3 and 6 were positively correlated (an acute angle), location 1 and 5 were negatively correlated (an obtuse angle), and location 1 and 4 were not correlated (a right angle) in year 2004.The presence of wide obtuse angle (i.e., strong negative correlations) among test environments is an indication of high cross over GEI (Yan and Tinker, 2006a).

The distance between two environments measures their dissimilarity in discriminating the genotype, thus the six locations in (Figure 3.1(a)) fell into 4 apparent groups where locations 2,3 and 6 form the first group while lactations 1,4 and 5 each of them separately form their own group. The presence of close associations among some test locations in year 2004, suggest that the same information about genotypes could be obtained from fewer test locations, and hence the potential to reduce test cost (Choukan, 2010, Tukamuhabwa et al., 2012). If two test locations are closely correlated consistently across years, one of them can be drooped without loss of much information about the genotypes.



Figure(a) Figure(b)

Figure 3. 1 Scatter plot of environments (a) year 1 (b) year 2.

However, in reality the correlation consistency for formation of a group between the locations vary from year to year as it shown in Figure 3.1. This inconsistency is much reflected on location 2 and 5, which form their own group on the first year but it forms same group in the second year. Clearly Figure 3.1 (a) and Figure (b) show differing genotype and environment structure. However it should be noted that data in 2005 had only five of the location in 2004.

3.3.2 Discriminating Ability and Representativeness of the Test Environment

GGE biplot discriminating ability and representativeness is an important measure of the testing environments. The concentric circles on the biplot as shown in Figure 3.2 help to visualize the length of the environment vectors, which is proportional to the standard deviation within the respective environments and is a measure of the discriminatory ability of the environments. Therefore, among the six environments, E1 and E4 were most

discriminating (informative) and E5 least discriminating in year 1; whereas in year 2 (Figure 3.2) E5 and E4 are most discriminating and E2 was least-discriminating. Test environments that are consistently non-discriminating (non-informative) provide little information on the genotypes and, therefore, should not be used as test environments.

The average environment (represented by the small circle at the end of the arrow) has the average coordinates of all test environments, and AEA or ATA (Average-Environment Axis or Average-Tester-Axis) is the line that passes through the average environment and the biplot origin (Yan, 2002). A test environment that has a smaller angle with the AEA is more representative of other test environments. Thus, E1 and E4 are most representative whereas E5 and E3 least representative in their respective year. Test environments (locations) that are both discriminating and representative (e.g., E1) are good test environments for selecting generally adaptable genotypes. Discriminating but non-representative test environments like E3 are useful for selecting specifically adapt-able genotypes if the target environments can be divided into mega-environments or they are useful for culling unstable genotypes if the target environment is a single mega-environment. A mega-environment is defined as a group of locations that consistently share the same best cultivar(s) (Yan and Rajcan, 2002).

This definition involves several essential elements: 1) mega-environments are defined by different winning cultivars, noting that different genotypes can be equally adapted to the same mega-environment and that a mega-environment may need different types of genotypes to stabilize the overall production; 2) mega-environment is a concept of geographical locations; and 3) the cultivar-location interaction pattern should be repeatable across years.

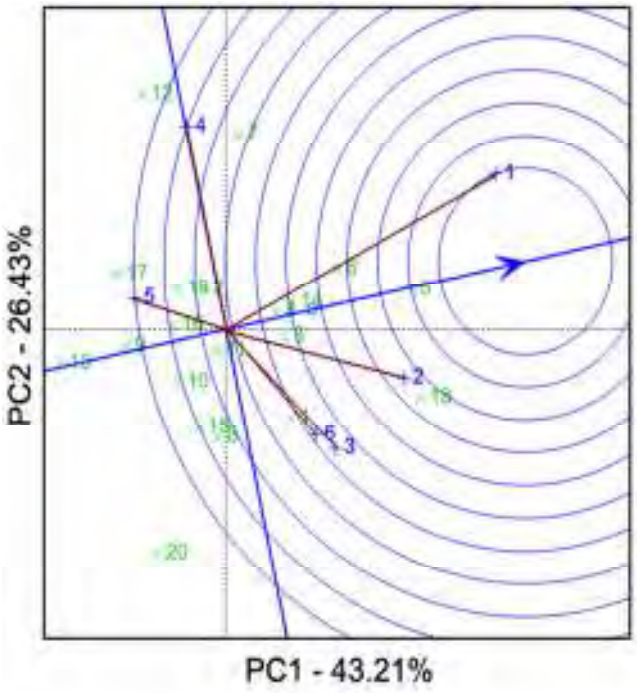


Figure (a)

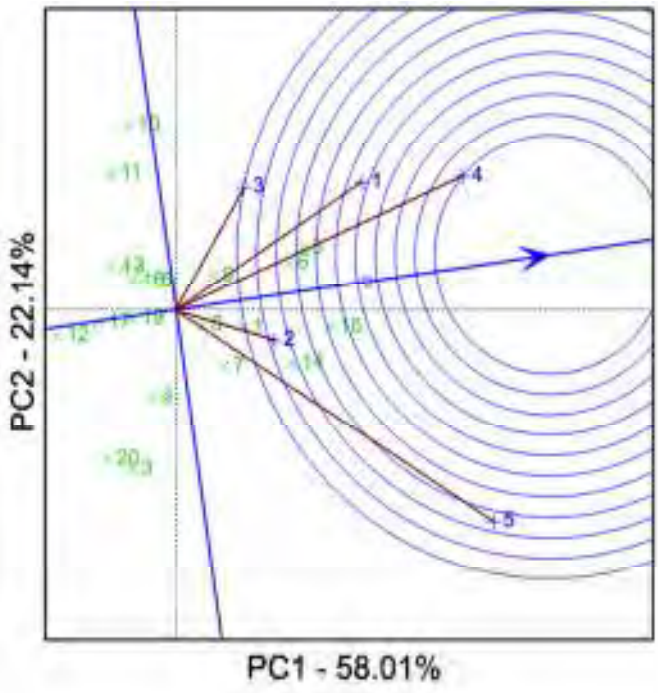


Figure (b)

Figure 3. 2 GGE biplot based on environment-focused scaling for comparison of the environment with ideal environment (a) year 1 (b) year 2.

3.3.3 Ranking Genotypes Relative to the Ideal Genotype

An ideal genotype should have the highest mean performance and be absolutely stable (i.e performs the best in all environments). Such an ideal genotype is defined by having the greatest vector length of the high yielding genotypes and with zero GEI, as represented by an arrow pointing to it (Figure 3.3). Although such an ideal genotype may not exist in reality, it can be used as a reference for genotype evaluation (Yan and Tinker, 2006a). A genotype is more desirable if it is located closer to the ideal genotype. Thus, using the ideal genotype as the centre, concentric circles were drawn to help visualize the distance between each genotype and the ideal genotype. Because the units of both PC1 and PC2 for the genotypes are the original unit of yield in the genotype-focused scaling (Figure 3.3), the units of the AEC abscissa (mean yield) and ordinate (stability) should also be in the original unit of yield. The

unit of the distance between genotypes and the ideal genotype, in turn, will be in the original unit of yield as well. Therefore, the ranking based on the genotype-focused scaling assumes that stability and mean yield are equally important (Farshadfar E. et al., 2012, Yan, 2002). Figure 3.3 revealed that G5, which fell into the centre of concentric circles, was the ideal genotype in terms of higher yielding ability and stability, compared with the rest of the genotypes. In addition, G6 and G14, located on the next consecutive concentric circle, may be regarded as desirable genotypes.

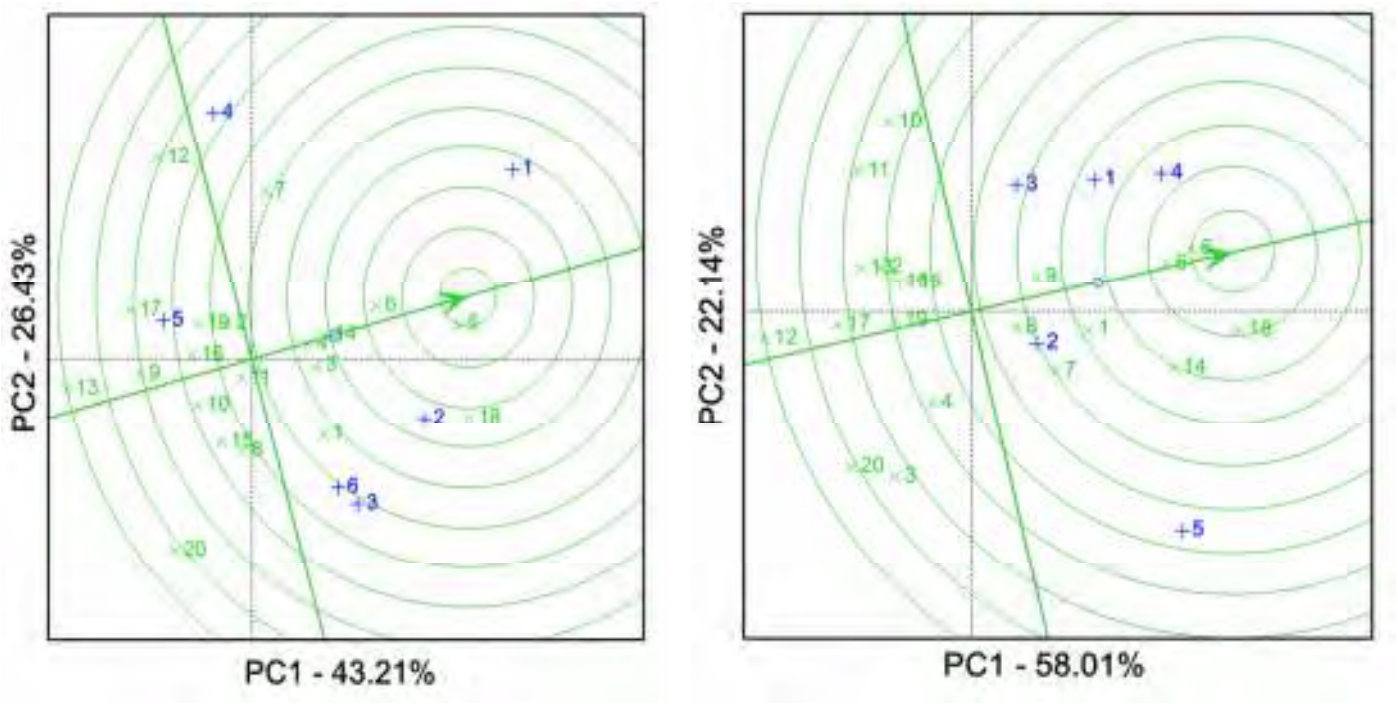


Figure (a) Figure (b)

Figure 3. 3 GGE biplot based on genotype-focused scaling for comparison of the genotype with ideal genotype (a) year 1 (b) year 2.

3.3.4 Mean Performance and Stability of the Genotypes

Yield performance and stability of genotypes were evaluated by an average environment coordination (AEC) method (Farshadfar et al., 2011). Within a

single mega-environment, genotypes should be evaluated on both mean performance and stability across environments. Figure 3.4(a) gives the average environment coordination (AEC) view of the GGE biplot. The single-arrowed line is the AEC abscissa, it points to higher mean yield across environments. Thus, G5, G18, G6 and G14 had the highest mean yield. The non-arrowed line is the AEC ordinate; it points to greater variability (poorer stability) in either direction. Thus, G12 and G20 ware highly unstable and below average yield, whereas G4 and G14 highly stable, were followed by G5, G6, and G3 with had above average yield in the first year.

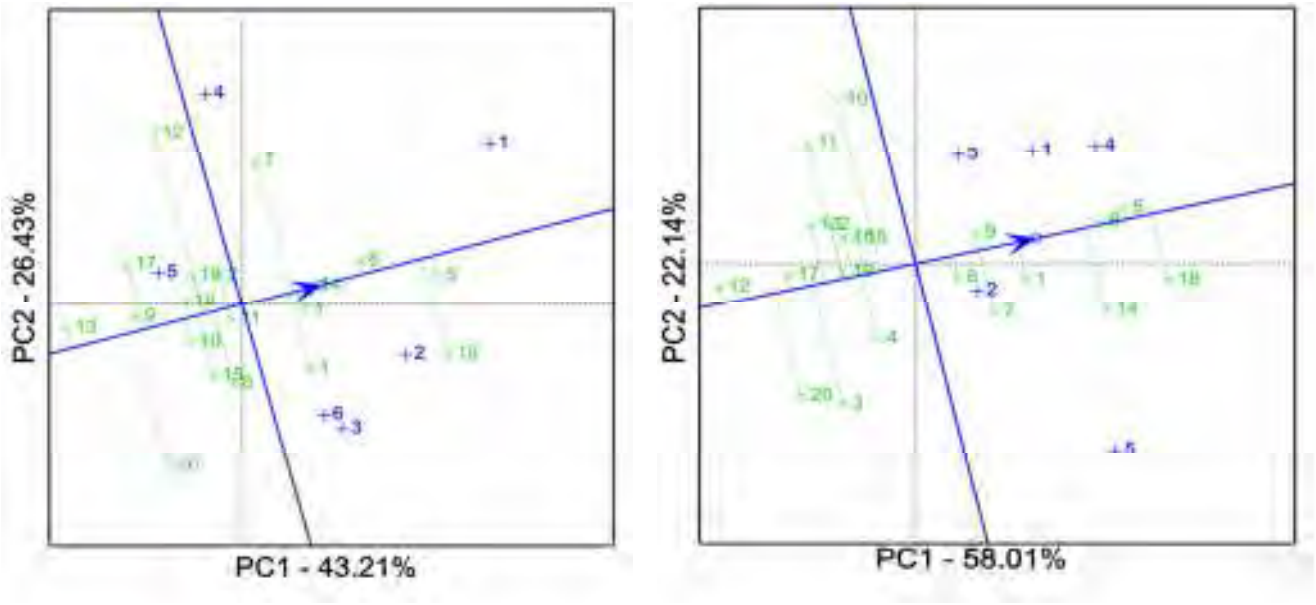


Figure (a) Figure (b)

Figure 3. 4 GGE biplot based on environment-focused scaling for mean performance and stability of the genotypes (a) year 1 (b) year 2.

The mean performance and stability of these 20 genotypes in five locations (environment) in the second year of the trial shows some variation from the

first year as it shown in Figure 3.4(b). However G6, G5 G4 and G18 were relatively high yielding and stabile genotypes in both trial years.

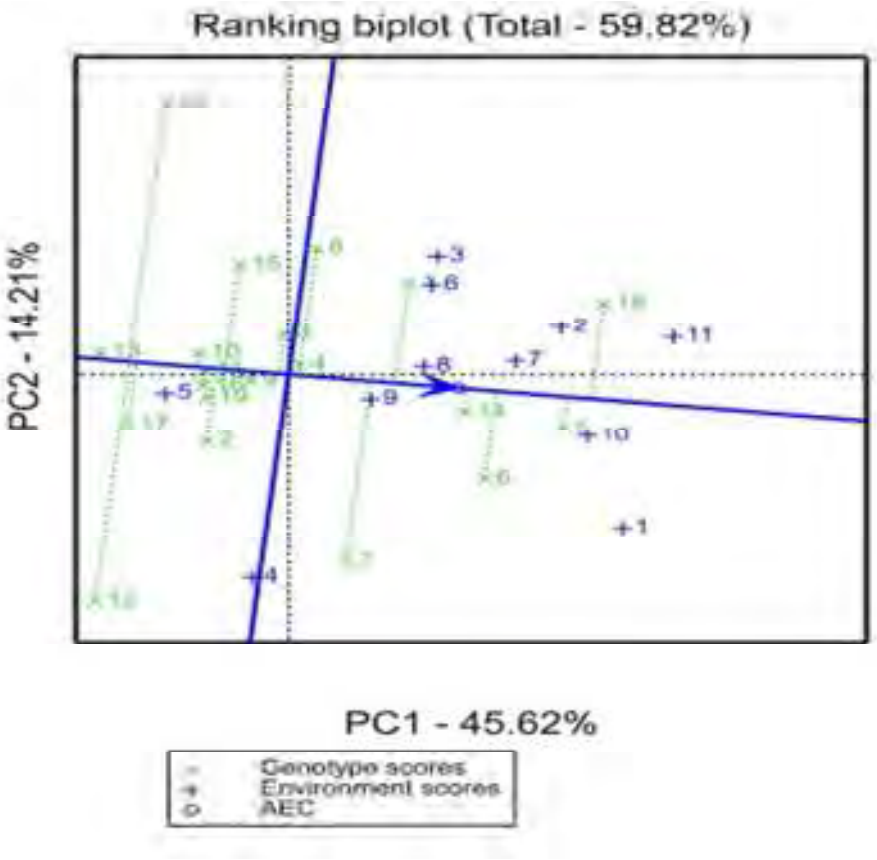


Figure 3.5 GGE biplot based on environment-focused scaling for mean performance and stability of the genotypes for the combined data.

A comparison of the year by year data analysis (Fig 3.4 a and b) and a combined data analysis (Fig 3.5) on the mean performance and stability of genotypes for a combined data analysis shows the presence of some variation. A low mean performing and low stable genotypes like G20 and G12 are the most affected on a combined data analysis. However G6, G5 G4 and G18 were relatively high yielding and stabile genotypes in both year by year and combined data analysis of GGE.

3.3.5 Which Genotype Won Where and Mega Environments With GGE biplot

One of the most attractive features of a GGE biplot is its ability to show the which-won-where pattern which-won-where pattern (specific genotypes recommended to specific environments) of a genotype by environment data set (Yan, 2002). Many researchers find this use of a biplot intriguing, as it graphically addresses important concepts such as crossover genotype-environment interaction, mega environment differentiation, specific adaptation, etc as discussed in Yan and Tinker (2006a) and Gower et al. (2011). Crossover GEI is the differential response of genotype to divers environments, when genotype ranks change from one environment to another. The polygon is formed by connecting the markers of the genotypes that are further away from the biplot origin such that all other genotypes are contained in the polygon.

Genotypes located on the vertices of the polygon performed either the best or the poorest in one or more locations since they had the longest distance from the origin of biplot. The perpendicular lines are equality lines between adjacent genotypes on the polygon, which facilitate visual comparison of them. For example in Figure 3.6, the equality line between G5 and G18 in 2004 indicates that G5 was better in E1, whereas G18 was better in E2, E3 and E6. An interesting feature of this view of a GGE biplot is that the vertex genotype(s) for each sector has higher (sometimes the highest) yield than the others in all environments that fall in the sector (Gauch et al., 2008b, Yan, 2002).These six equality lines divide the biplot into six sectors, and the environments fall into four of them (Figure 3.6). This pattern suggests that the target environment may consist of four different mega-environments and that different cultivars should be selected and deployed for each.

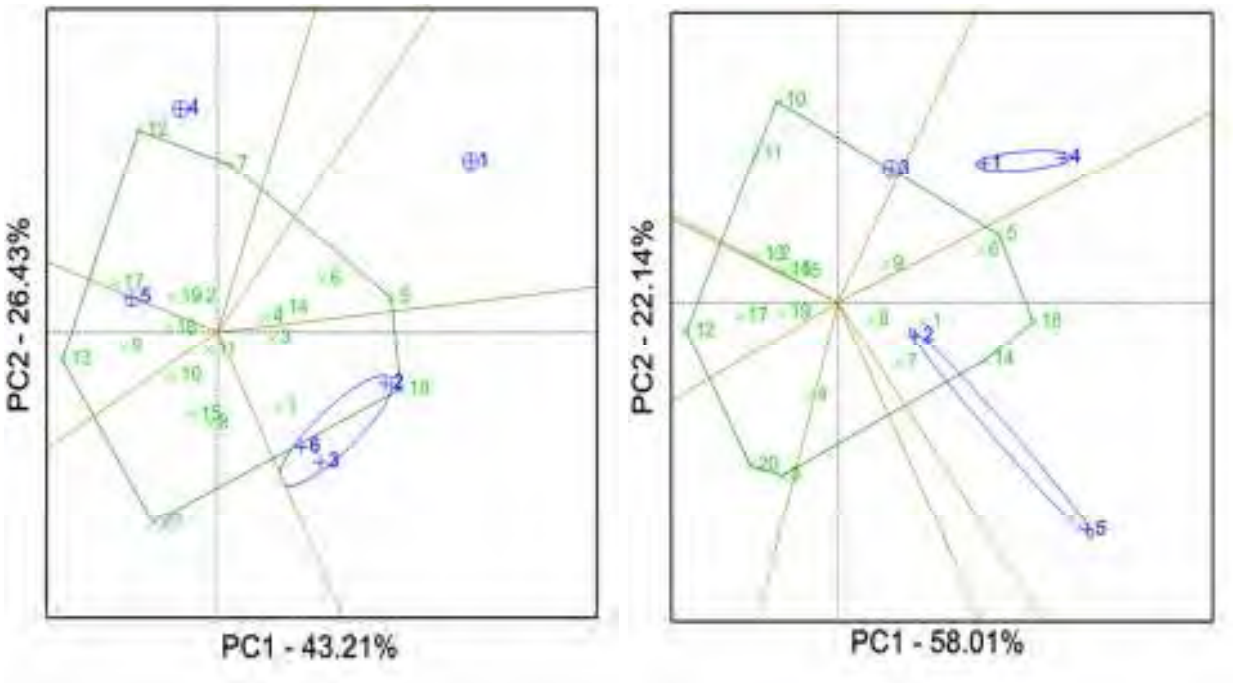


Figure (a) Figure (b)

Figure 3.6 The which-won-where view of the GGE biplot to show which genotypes performed best in which environment (a) year 1 (b) year 2.

In which-win-where GGE biplot for the second year Figure 3.6(b), eight equality lines divide the biplot into eight sectors and the five locations fell into three of them. The mega-environment classification of these five trial location is different from the first year. This difference leads to a different wining genotype in different locations (environment) across a year.

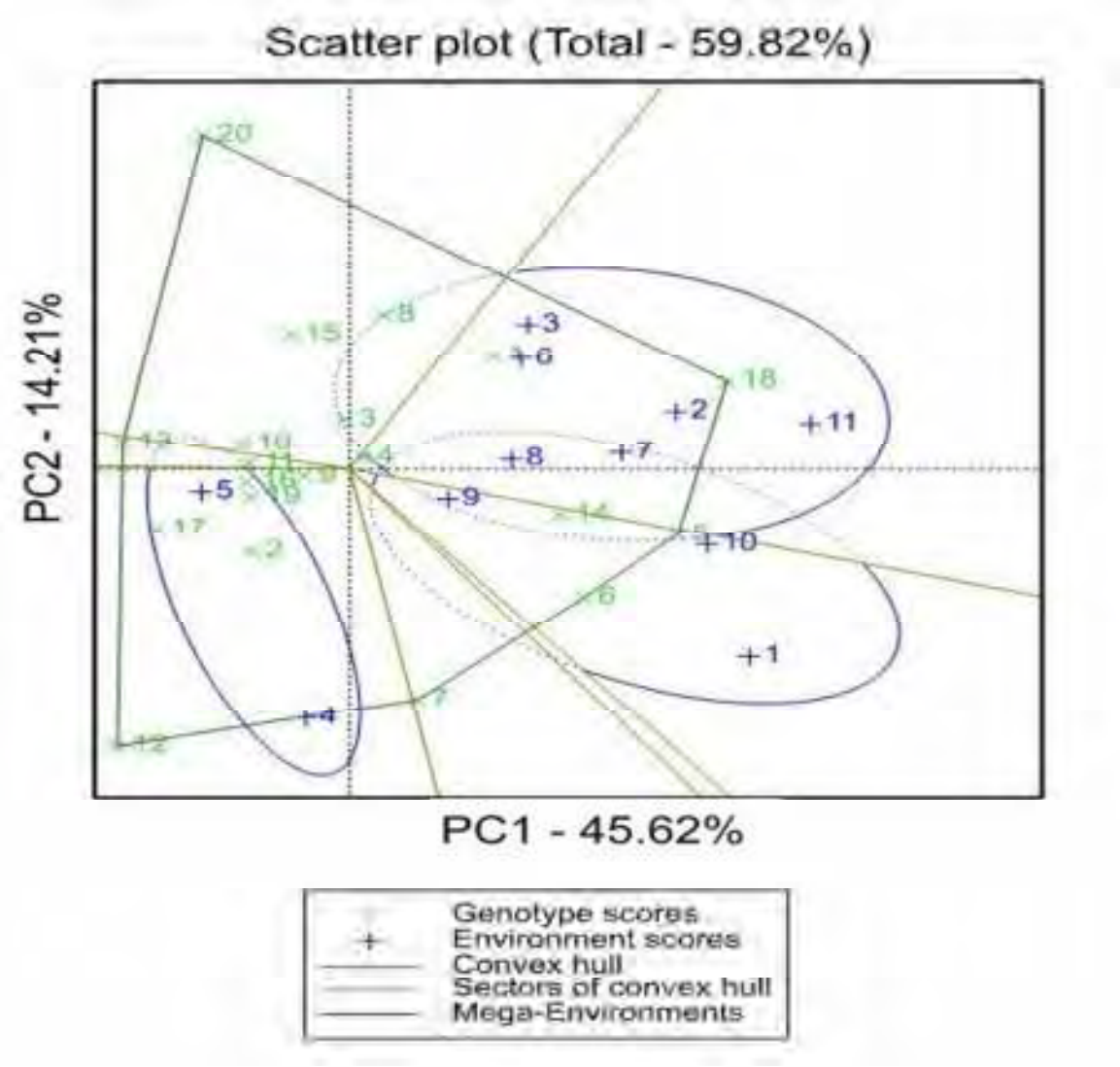


Figure 3.7 The which-won-where view of the GGE biplot to show which genotypes performed best in which environment of a combined data of year1 and year 2.

Fig 3.7 the which-win-where GGE biplot for the two year combined data clearly indicate the presences of year to year variation of on the classification of mega-environments. For example E1 and E7, E2 and E8, E3 and E9, E4 and E10, E5 and E11 all these pair are the same geographical locations, but most of them are classified into different mega-environments. This difference leads to a different wining genotype in different locations (environment) across a year.

These differences on the classification of geographical location and which genotype win where of a separate and combined data analysis of GGE indicate a need of the decision of analysis procedure (separate or combined) and careful recommendations of crop varieties to small holder farmers or large scale farms should be based on the performance of genotypes in different locations.

4 Conclusions

The GGE biplots of MEYT data allow visualizing the inter relationship among genotypes including the ranking of genotypes based on both mean performance and stability, inter-relationship among environments, and interaction between genotypes and environments including the which-won-where pattern. The result of this chapter indicated that wheat yield performance was highly influenced by the environment effect followed by the magnitude of GEI and genotype. The AMMI result shows that, total yield variation accounted by the genotype increased from 0.0029% in first year to 14.03% in the second year which had almost equal effect with the G × E interactions. These two years repeated over location data analysis result; which-win-where pattern, yield performance and stability of genotype indicate that repeatability pattern over years is the necessary and sufficient condition for mega-environment delineation and genotype recommendation for small holder farmers and commercial farm crop producers in terms of adaptability of different environment and yield performance. Decision making based on one year data should be done with caution. One lop of AMMI and GGE in agricultural biplot literature provides no guidance concerning how much of the total variability accounted for by the first two principal components are considered adequate or how many principal components (PC) are important for testing multiplicative terms, this will be addressed in the next chapter.

Chapter 4
Testing and Selecting Multiplicative Terms in GGE and AMMI Models

4.1 Introduction

When differences between genotypes depend on environment, then there exists a genotype-by-environment interaction is present. Genotype-by-environment interaction is studied in many branches of biology, not least in agriculture. In plant breeding and crop variety experimentation, cultivars or potential cultivars are commonly investigated at several environmentally different locations. For analysis of such data, two bi-additive models are commonly used (Denis and Gower,1994). These include genotype main effects and genotype-by-environment interaction effects (GGE) model (Yan and Kang, 2002, Yan et al., 2000) and the additive main effects and multiplicative interaction (AMMI) model (Gauch, 1988, 1992b). Both GGE and AMMI explore a matrix of genotype-by-environment means, using a combination of analysis of variance (ANOVA) and singular value decomposition. With GGE, singular value decomposition is performed on the matrix of residuals from a one-way ANOVA with fixed effects for environments. With AMMI, singular value decomposition is performed on residuals from a two-way ANOVA with fixed effects of genotypes and environments. Cornelius et al. (1996) called the GGE model a sites regression model (SREG). For a discussion about the relative merits of GGE and AMMI, see Gauch (2006), Gauch et al. (2008b), Yan and Tinker (2006a) and Yan et al. (2007).

The present chapter considers GGE and AMMI models with fixed main effects and normally distributed errors, although extensions have been made to mixed models (Piepho, 1998a, Piepho, 1999, Smith et al., 2001b, Yan and Kang, 2002), other distributions and three-way interactions (Eeuwijk and

Kroonenberg, 1998). The data used in the current chapter are data set we used in chapter three (BW00RVTI) and additional one data set (BW01RVII) of 25 barley varieties were tested in five locations (environments) in 2007/8. All the trials in each location were laid out as a randomized complete block (RCB) design with four replicates. The locations consist of loc1 (Kulumsa), loc2 (Adet), loc3 (Bekoji), loc4 (Sinana), and loc5 (Holeta).

The result of the singular value decomposition is often presented in a biplot illustrating the first two multiplicative terms of the singular value decomposition as explained in chapter three. With GGE, such a biplot presents a rank-two approximation of the sum of genotype effects and genotype-by-environment interaction effects, which is a useful and popular tool for breeders (Yan and Tinker, 2006a). With AMMI, genotype-by-environment interaction is studied separately from main effects of genotypes. Figure 4.1 is a biplot for an AMMI analysis of a dataset with twenty genotypes (G1–G20) investigated in 11 environments (E1–E11). Points near the origin have small interaction effects, and points near each other have similar interaction effects (Gauch, 1992b). Yang et al. (2009b) discussed the validity of the biplot as a statistical method for analysis of genotype-by-environment interaction. One of their main concerns was the frequent lack of statistical hypothesis testing for determining the number of multiplicative terms. Figure 5.1 illustrates the first two multiplicative terms, but it is possible that more or fewer terms give a better description of the interaction. In practice, researchers would like to separate fixed genotype-by-environment interaction from random noise. This main objective of this chapter is : (1) to show a method for significance testing of multiplicative terms (2) to show statistical hypothesis testing for determining the number of multiplicative terms in GGE and AMMI models.

Testing for interaction in non-replicated two-way layouts goes back to Tukey (1949), who introduced the one-degree-of-freedom test for additivity, and Mandel (1961), who proposed row-specific regression on additive column

effects. Finlay and Wilkinson (1963) and Yates and Cochran (1938) proposed genotype-specific regression of yield on site means. Such models can be fitted using nonlinear regression (Ng and Grunwald, 1997, Piepho, 1999a) and Mandel, (1971) proposed the AMMI model. Johnson and Graybill (1972) derived a likelihood ratio test for the first multiplicative term of the AMMI model.

Based on their work, Marasinghe (1985) and Schott (1986) proposed a sequential testing procedure for all terms. This procedure tests the $(K+1)^{th}$ multiplicative term as if it were the first term in a problem with the numbers of rows and columns reduced by Cornelius et al. (1996) presented the approximately F-distributed JG/SM test statistic, which is built on the contributions by Johnson and Graybill (1972), Marasinghe (1985) and Schott (1986). Cross-validation is another option for selecting the number of multiplicative terms (Dias and Krzanowski, 2006 and Dias and Krzanowski, 2003).

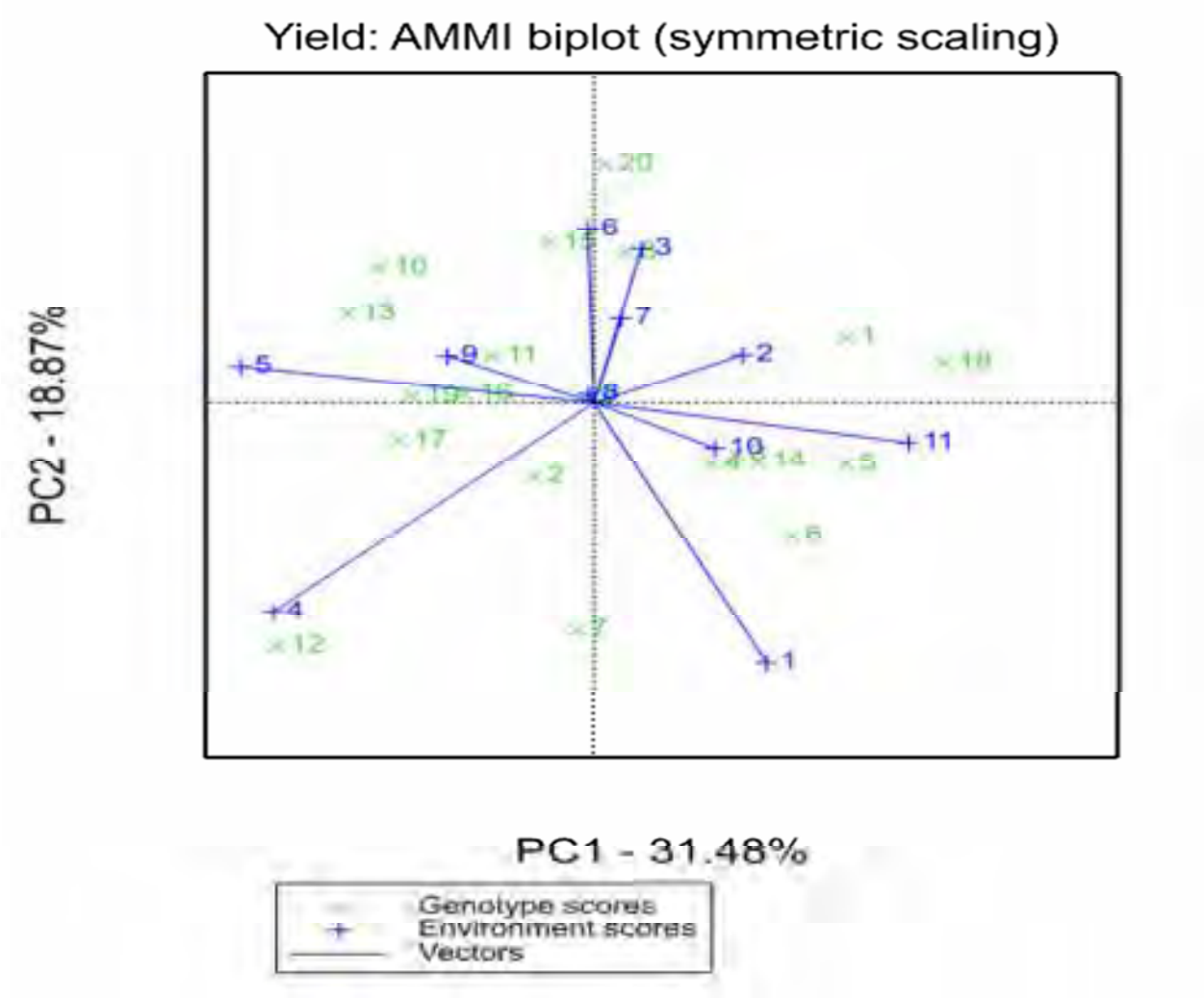


Figure 4. 1 AMMI biplot for the BW00RVTI ALL data

For replicated data, Gollob (1968) proposed an F-test for selecting the number of multiplicative terms. According to this method, the mean square of the multiplicative term is divided by the error mean square and compared with an F-distribution, similarly as with ANOVA. The Gollob (1968) approximation F-test assume that $n\hat{\lambda}^2{}_k \Big/ \sigma^2$ is distributed as chi-square variable, and he suggests using the statistics

$$F = \hat{\lambda}^2{}_m \Big/ (f_1 MS(Error\ mean))$$

against an F-distribution with $f_1 = J + I - 1 - (2m)$ and $JI(n-1)$ degrees of freedom, or $J(I-1)(n-1)$ degrees of freedom if blocks are present, to test the m^{th} multiplicative term of the model for significance.

It is well-known that this method is too liberal (Cornelius et al., 1996). Cornelius (1993), Cornelius et al (1996), and Piepho (1995), proposed and investigated various other F-tests for replicated data. In these tests, an error mean square is calculated from replicates within genotype-by-environment combinations. When there is only a single observation for each cell of the genotype-by-environment table, these methods do not apply.

The present chapter considers using resampling for the problem of separating fixed genotype-by-environment interaction from random noise. Following Mandel (1971), it is suggested that the fixed interaction be modeled by the first K terms of the singular value decomposition. Random noise is estimated as the remainder when K is smaller than the maximum possible number of terms. The proposed methods for selecting K use the parametric bootstrap technique (Efron and Tibshirani, 1993). With this approach, the distributions of the observed data are assumed to belong to a parametric family, and the expected value follows a statistical model, specifically the GGE or AMMI models. The model parameters are estimated from the observed data, and random samples are generated from the estimated model.

Based on these so-called bootstrap samples, distributions of test statistics or other statistics can be approximated. In the present application, the distribution of a test statistic, T, for the significance of the $(K+1)^{th}$ multiplicative term is approximated and used as a reference distribution for calculation of the p-value. When K=0, this test statistic is the same as the likelihood ratio statistic used by Johnson and Graybill (1972), and when K>0, the test statistic is the same as the likelihood ratio statistic that was derived by Yochmowitz and

Cornell (1978) and advocated by Schott (1986). When K >1, an exact test based on T is not possible, since the distribution of T depends upon unknown singular values (Schott, 1986). In the context of hypothesis testing, these singular values are nuisance parameters that are not specified by the null hypothesis.

A parametric bootstrap method presented in the present chapter resolves this problem by substituting estimates for unknown parameters. By this method, simulation of the null distribution is made possible. Utilizing an approximate result for distributions of eigenvalues (Muirhead, 1978), a simplified version of parametric bootstrap method is derived. This parametric bootstrap method is particularly easy to apply, because it uses standard normally distributed values as bootstrap samples. Thus, with this method it is not necessary to estimate the parameters; it suffices to assume that errors are normally distributed. The proposed bootstrap methods for GGE and AMMI can with small adjustments also be used for the completely multiplicative model (COMM), and the genotypes regression model (GREG) (Cornelius et al., 1996). Furthermore, the methods may be used for testing components in principal component analysis (PCA).

4.2. Methods

4.2.1 Statistical Model

Assume that J genotypes have been investigated in I environments. Let y_{ij} denote the observed mean yield of the j th genotype in the ith environment ($i = 1,2,\cdots,I$; $j = 1,2,\cdots,J$). Let Y denote the I × J matrix of observations y_{ij} i.e. $\boldsymbol{Y} = \{y_{ij}\}$. GGE and AMMI models can be written in the form $\boldsymbol{Y} = \boldsymbol{A} + \boldsymbol{E}$, where **A** denotes an additive part and **E** denotes a matrix of interactions and residual errors. Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots \alpha_I)^T$ be a vector of environment main effects, and $\beta = (\beta_1, \beta_2, \cdots, \beta_J)^T$ a vector of genotype main effects. Let $\mathbf{1}_I$ denote an I-vector of

ones,$\mathbf{1}_J$ a J-vector of ones, and $\boldsymbol{J}_{I\times J} = \mathbf{1}_I \mathbf{1}_J{}^T$ and I × J matrix of ones. Including an overall mean, μ, the additive part of the GGE model is

$$\boldsymbol{A} = \mu \boldsymbol{J}_{I\times J} + \boldsymbol{\alpha}\mathbf{1}_J{}^T \tag{4.1}$$

and the additive part of the AMMI model is

$$\boldsymbol{A} = \mu \boldsymbol{J}_{I\times J} + \boldsymbol{\alpha}\mathbf{1}_J{}^T + \mathbf{1}_I \beta^T \tag{4.2}$$

In this chapter, it will be assumed that $\mathbf{E}$ can be written as

$$\boldsymbol{E} = \boldsymbol{\theta}_k + \boldsymbol{R} = \boldsymbol{U}_k \Lambda_k \boldsymbol{V}^T{}_k + \boldsymbol{R} \tag{4.3},$$

where $\boldsymbol{\theta}_k$ models interaction and $\boldsymbol{R} = \{r_{ij}\}$ is a matrix of independent $N(0\,;\,\sigma^2)$ distributed errors r_{ij}. In (3), $\boldsymbol{U}_k \Lambda_k \boldsymbol{V}^T{}_k$, is the singular value decomposition of $\boldsymbol{\theta}_k$.

The subscript (k) indicates that the rank of $\boldsymbol{\theta}_k$ is k. Thus, the singular value decomposition of $\boldsymbol{\theta}_k$ comprises multiplicative terms that are not 0. It will be assumed that the diagonal elements of Λ_k are sorted in decreasing order. The singular values will be denoted by $\lambda_1, \lambda_2, \cdots, \lambda_k$, i.e. $\Lambda_k = diag(\lambda_1, \lambda_2, \cdots, \lambda_k)$. The I×k matrix of left-singular vectors is $\boldsymbol{U}_k = (\gamma_1, \gamma_2, \cdots, \gamma_k)$, and the J×k matrix of right singular vectors is as $\boldsymbol{V}_k = (\delta_1, \delta_2, \cdots, \delta_k)$. In scalar form, GGE and AMMI models can be written $y_{ij} = \mu + \alpha_i + \gamma_{i1}\lambda_1\delta_{j1} + \gamma_{i2}\lambda_2\delta_{j2} + \cdots + \gamma_{ik}\lambda_{1k}\delta_{jk} + r_{ij}$ and $y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{i1}\lambda_1\delta_{j1} + \gamma_{i2}\lambda_2\delta_{j2} + \cdots + \gamma_{ik}\lambda_{1k}\delta_{jk} + r_{ij}$ respectively.

The general mean, μ, can be estimated as the average

$$\hat{\mu} = \sum_{i=1}^{I}\sum_{j=1}^{J} {}^{y_{ij}}/_{(IJ)},$$

and the row and column effects as $\hat{\alpha}_i = \sum_{j=1}^{J} {}^{y_{ij}}/_{(J - \hat{\mu})}$ and $\hat{\beta}_j = \sum_{i=1}^{I} {}^{y_{ij}}/_{(I - \hat{\mu})}$, respectively. Define $\hat{A}$ as the least squares estimator of A in (4.1) or (4.2) when $\hat{\mu}$, $\hat{\alpha}_i$ and $\hat{\beta}_j$ respectively. Let $\hat{\mathbf{E}} = \hat{\mathbf{Y}} + \hat{\mathbf{A}}$. Let M denote the rank of $\hat{\boldsymbol{E}}$. Generally,

$M = min\,(I\,,\ J-1)$ in GGE analysis, whereas $M = min\,(I-1\,,\ J-1)$ in AMMI analysis. Through singular value decomposition, $\hat{\mathbf{E}} = \hat{\mathbf{U}}\,\hat{\Lambda}\,\hat{\mathbf{V}}^{\boldsymbol{T}}$, where $\hat{U} = (\hat{\gamma}_1, \hat{\gamma}_2, \cdots \hat{\gamma}_m)$ is an I×M matrix of estimated left singular vectors, $\hat{\Lambda} = \text{diag}\,(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_M)$ is an M×M diagonal matrix of estimated singular values sorted from largest to smallest, $\hat{\mathbf{V}} = \left(\hat{\delta}_1, \hat{\delta}_2, \cdots, \hat{\delta}_k\right)$ is a J×M matrix of estimated right- singular vectors.

Consider testing the null hypothesis $H_o: k = K$ against the alternative hypothesis $H_1: k > K$ note that k is the actual number of multiplicative terms, where K is the assumed number of multiplicative terms under the current null hypothesis. The ratio

$$\mathrm{T} = \frac{\hat{\lambda}_{\mathrm{K+1}}{}^{2}}{\sum_{\mathrm{k=K+1}}^{\mathrm{M}} \hat{\lambda}_{\mathrm{k}}{}^{2}} \tag{4.4}$$

can be used as a test statistics for H_o. Large values of T present evidence against H_o. For $K = 0$, Johnson and Graybill (1972) provided simulation-derived critical values for selected rows and columns.

4.2.2 JG/SM –statistics

Yochmowitz and Cornell (1978) and Schott (1986) proposed using the test statistic (4) for the problem of selecting the number of multiplicative terms. In (4), the squared singular values are distributed as eigenvalues of a Wishart matrix (Johnson and Graybill, 1972). Cornelius et al. (1996) showed, based on Johnson and Graybill (1972), Marasinghe (1985) and Schott (1986), that (4.4) can be transformed into an approximately *F*-distributed statistic. To test the K^{th} term, the approximate F-test is computed as follows:

$$Q\ =\ ((\,M-K)T-1)/(M-K-1),$$

$$c_1 = \frac{u_1 - L + K}{(L-K)(M-K+1)},$$

$$c_2 = \frac{(M-K)(L-K){u_2}^2 - 2{u_1}^2}{(L-K)^2(M-K)(L-K)+2)(M-K-1)^2},$$

$$a = (c_1(1-c_1) - c_2)c_1/c_2,$$

$$b = (c_1(1-c_1) - c_2)/c_2,$$

$$\mathrm{F}_{JG/SM} = \frac{(b\mathrm{Q})}{\big(a(1-Q)\big)}.$$

For $M = min\,(I, J-1)$ and $L = max(I, J-1)$. For AMMI, $M = min\,(I-1, J-1)$ and $L = max(I-1, J-1)$. The distribution of $\mathrm{F}_{JG/SM}$ is approximated by F with df= (2a, 2b) (Cornelius et al. (1996). u_1 and u_2 are calculated according to the approximating functions given by Liu and Cornelius (2001) for mean and standard deviation, respectively, of the first squared singular value divided by the error variance. The values a and b are functions of expressions that approximate the first two moments, estimated through Monte Carlo simulation, of eigenvalues of Wishart matrices (Liu and Cornelius, 2001).

4.2.3 Parametric bootstrap method

Forkman and Piepho (2014) propose a parametric bootstrap for computing the p-value associated with the JG/SM statistic. According to this approach (Efron and Tibshirani, 1993), the null distribution is simulated using parameter estimates under the null hypothesis. First, the model parameters, including the variance, are estimated, and then a large number, B, of samples are drawn from the fitted model using the estimates. For each sample, the test statistic, T, is calculated. The obtained distribution of simulated test statistics approximates the true sampling distribution of T under H_0 and can be used for estimation of the p-value.

In this application, the first K terms from the singular value decomposition are taken to represent the true fixed interaction ($\boldsymbol{\theta}_k$) in all B bootstrap simulations. Thus, ($\boldsymbol{\theta}_k$) in (4.3) is estimated as

$$\boldsymbol{\theta}_k = \begin{cases} 0 & if\ k = 0 \\ \sum_{K=1}^{K} \hat{\gamma}_k \hat{\lambda}_k \hat{\delta}_k{}^T & if\ K > 0 \end{cases} \qquad 4.5$$

To simulate residual error, random noise is added. Specifically, a random matrix ${R_b}^B$ of independent $N(0\,, \hat{\sigma}_{(k)}{}^2)$ distributed errors is added to $\widehat{\boldsymbol{\theta}}_k$. The superscript, B, indicates that ${R_b}^B$ is a bootstrap version of R, whereas the subscript, b, indicates that ${R_b}^B$ is the bth bootstrap sample. The variance, $\hat{\sigma}_{(k)}{}^2$, which is an estimate of σ^2, can be derived as follows.

Consider an I-by-J two-way layout with additive main effects of rows and columns, additive fixed effects of interactions, and independent additive errors with variance σ^2. When there is no replication within cells, a two-way ANOVA without interaction is commonly fitted. In this case, the expected residual sum of squares equals the unknown interaction sum of squares plus $(I-1)(J-1)\,\sigma^2$ (Searle et al., 1992, 2008). Similarly, as relevant for GGE analysis, when a one-way ANOVA with row effects is performed, the expected residual sum of squares equals the unknown sum of squares due to column and interaction effects plus $I(J-1)\,\sigma^2$. Under H_o, the sum of squares of true interaction effects is 0 when K= 0 and $\sum_{K=1}^{k} \sum {\lambda_K}^2$ when $1 \le K \le M-2$. It is proposed that the observed error sum of squares from the additive model, i.e. the model without multiplicative terms, be equated to the expected error sum of squares under H_o , that is

$$\sum_{K=1}^{k} \sum {\lambda_K}^2 + v\sigma^2, \qquad 1 \le K \le M-2 \qquad (4.6),$$

where $v = I(J-1)$ for GGE, and $v = (I-1)(J-1)$ for AMMI. When fitting the additive model, the error sum of squares always equals $\sum_{K=1}^{M} \hat{\lambda}_K{}^2$. Hence, by substituting $\hat{\lambda}_K{}^2$ for $\lambda_K{}^2$ in (4.6), the estimating equation $\sum_{K=1}^{M} \hat{\lambda}_K{}^2 = \sum_{K=1}^{k} \hat{\lambda}_K{}^2 + v\hat{\sigma}_{(k)}{}^2$, $1 \le K \le M-2$, is obtained. Solving for $\hat{\sigma}_{(k)}{}^2$, the error variance σ^2 in (4.3) can be estimated as

$$\hat{\sigma}_{(k)}{}^2 = \frac{1}{v} \sum_{k=K+1}^{M} \hat{\lambda}_K{}^2, \quad 1 \le K \le M-2$$

where v is defined as in (4.6).

An approximate version of the above parametric bootstrap method can be derived as follows. Let ($\boldsymbol{L}_J$) be a $J \times (J-1)$ matrix such that $\boldsymbol{L_J L^T}_J = \boldsymbol{I}_J - \left(1/_J\right) J_{J\times J}$. Then $\widehat{\boldsymbol{E}}$ equals $\boldsymbol{Y L_J L^T}_J$ and $\boldsymbol{L_J L^T}_J \boldsymbol{Y L_J L^T}_J$ in GGE and AMMI analysis, respectively. For the AMMI model with $J < I$ Johnson and Graybill (1972) showed that $\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}^T/\sigma^2$ is non-centrally Wishart distributed with $I-1$ degrees of freedom, scale matrix $\boldsymbol{I}$, and noncentrality matrix $\boldsymbol{L^T}_J \boldsymbol{\theta^T}_{(k)} \boldsymbol{\theta_k L_J}/\sigma^2$, that is $\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}^T/\sigma^2$ is $W_{J-1}\left(I-1, \frac{\boldsymbol{L^T}_J \boldsymbol{\theta^T}_{(k)} \boldsymbol{\theta_k L_J}}{\sigma^2}\right)$. When $J > 1$ $\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}^T/\sigma^2$ is $W_{I-1}\left(I-1, \frac{\boldsymbol{L^T}_I \boldsymbol{\theta^T}_{(k)} \boldsymbol{\theta_k L_I}}{\sigma^2}\right)$. The positive eigenvalues of $\widehat{\boldsymbol{E}}^T\widehat{\boldsymbol{E}}/\sigma^2$ are the same as those of $\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}^T/\sigma^2$ and equal $\hat{\lambda}_1{}^2/\sigma^2$, $\hat{\lambda}_2{}^2/\sigma^2$, $\cdots, \hat{\lambda}_M{}^2/\sigma^2$ provided that H_o is correct and $\hat{\lambda}_1{}^2/\sigma^2$, $\hat{\lambda}_2{}^2/\sigma^2$, $\cdots, \hat{\lambda}_M{}^2/\sigma^2$ are large, the asymptotic joint distribution of these eigenvalues can be approximated by the joint distribution of the eigenvalues of a central Wishart matrix distributed as $W_{J-1-K}(I-1-K, \boldsymbol{I})$ (Marasinghe, 1985, Muirhead, 1978). As a result, the distribution of T in (4.4) may be approximated by the distribution of the ratio of the first eigenvalue to the sum of all eigenvalues of a central Wishart matrix. Since the joint distribution of the eigenvalues of a $W_{J-1-K}(I-1-K, \boldsymbol{I})$ distributed matrix is equal to the joint distribution of the squared singular values of a random $(I-1-K) \times (J-1-K)$ matrix of independent standard normal values is large, the parametric bootstrap procedure proceeds as follows

1. For $b = 1, 2, \cdots, B$, where B is large, the following can be done.
   a. Sample a $(D - K) \times (J - 1 - K)$ matrix $\hat{E}_B^{\,B}$ of independent $N(0;1)$ distributed errors, where $D = I - 1$ in AMMI analysis and $D = I$ in GGE analysis.
   b. Subject $\hat{E}_B^{\,B}$ to singular value decomposition and use the obtained singular values for calculation of bootstrap samples T_b according to the right hand side of (4), here using K= 0 (i.e. compute T_b as the ratio of the first squared singular value to the sum of all squared singular values).
2. Estimate the p-value as the observed frequency of T_b larger than T computed from the data.

4.2.4 Sequential F-test and Cross-validation method

A study by Gollob (1968) and Wold (1978) defined degrees of freedom associated with parameters in multiplicative models as the number of parameters minus the number of constraints. With this definition, the number of degrees of freedom, DF_K, needed to estimate the K^{th} interaction are $I + J - 2K$ and $I + J - 2K - 1$ term is in GGE and AMMI analysis, respectively. Consequently, the residual error degrees of freedom, DF_{Residual}, is $I(J-1) - \sum_{K=1}^{K}(I + J - 2k) = (I - K)(J - 1 - K)$ and $(I-1)(J-1) - \sum_{K=1}^{K}(I + J - 2k - 1) = (I - 1 - K)(J - 1 - K)$ in GGE and AMMI analysis, respectively. (Gollob, 1968) proposed an F-test with denominator mean squared error calculated from replicates within genotype-by-environment combinations. This test was not applicable to the examples of the following section, since these examples used means or non-replicated data. For this reason, sequential F-test statistics were instead calculated as ratios between the mean square due to fixed interaction, i.e. the multiplicative terms, and residual mean square.

Dias and Krzanowski (2003, 2006) also proposed cross-validation for choosing the number of multiplicative interaction terms. Based on their work, the following is presented. Denote by $\hat{\mathbf{E}}^{(-\mathbf{j})}$ the result of deleting the j^{th} column of $\hat{\mathbf{E}}$

and after this subtracting row means. Denote by $\hat{\mathbf{E}}^{(-\mathbf{i})}$ the result of deleting the ith row of $\hat{\mathbf{E}}$ and after this, in case of AMMI-analysis, subtracting column means. Let $\hat{\mathbf{U}}^{(-\mathbf{i})}$ and $\hat{\boldsymbol{U}}^{(-\mathbf{j})}$ denote the matrix of left-singular vectors, $\hat{\Lambda}^{(-\mathbf{i})}$ and $\hat{\Lambda}^{(-\mathbf{j})}$ the matrix of singular values, $\hat{\mathbf{V}}^{(-\mathbf{i})}$ and $\hat{\mathbf{V}}^{(-\mathbf{j})}$ the matrix of right-singular vectors, of $\hat{\mathbf{E}}^{(-\mathbf{i})}$ and $\hat{\mathbf{E}}^{(-\mathbf{j})}$, respectively. Write $\hat{\mathbf{U}}^{(-\mathbf{i})} = \{\hat{u}_{st}{}^{(-\boldsymbol{i})}\}$, $\hat{\mathbf{U}}^{(-\mathbf{j})} = \{\hat{u}_{st}{}^{(-\boldsymbol{i})}\}$, $\hat{\mathbf{V}}^{(-\mathbf{i})} = \{\hat{v}_{st}{}^{(-\boldsymbol{i})}\}$ and $\hat{\mathbf{V}}^{(-\mathbf{j})} = \{\hat{v}_{st}{}^{(-\boldsymbol{j})}\}$. Denote by $\hat{\lambda}_k{}^{(-\mathrm{i})}$ and $\hat{\lambda}_k{}^{(-\mathrm{j})}$ the kth diagonal element of the diagonal matrix $\hat{\Lambda}^{(-\mathbf{i})}$ and $\hat{\Lambda}^{(-\mathbf{j})}$ respectively. Define $\hat{e}_{ij}$ through $\hat{\mathbf{E}} = \{\hat{e}_{ij}\}$ and let $\hat{e}^{c}{}_{ij(k)} = \sum_{k=1}^{k} \hat{u}_{ik}{}^{(-\boldsymbol{j})} \sqrt{\hat{\lambda}_k{}^{(-\mathrm{i})} \hat{\lambda}_k{}^{(-\mathrm{j})}}\, \hat{v}_{jk}{}^{(-\mathrm{j})}$ where c indicates that $\hat{e}^{c}{}_{ij(k)}$ is a cross-validation prediction of $\hat{e}_{ij}$. Following Eastment and Krzanowski (1982a), when calculating $\hat{e}^{c}{}_{ij(k)}$, the sign of $\hat{u}_{ik}{}^{(-\boldsymbol{j})} \sqrt{\hat{\lambda}_k{}^{(-\mathrm{i})} \hat{\lambda}_k{}^{(-\mathrm{j})}}\, \hat{v}_{jk}{}^{(-\mathrm{j})}$ was set equal to the sign of $\hat{\gamma}_k \hat{\lambda}_k \hat{\delta}_k{}^{T}$. This is the cross-validation method, because due to deletion of rows and columns before singular value decomposition, $\hat{e}_{ij}$ is not used in $\hat{e}^{c}{}_{ij(k)}$. With $\hat{e}^{c}{}_{ij(k)}$ as prediction of $\hat{e}_{ij}$, the prediction residual sum of squares is

$$PRESS_k = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} (\hat{e}^{c}{}_{ij(k)} - \hat{e}_{ij})^2$$

where $\hat{e}^{c}{}_{ij(0)} = 0$. Eastment and Krzanowski (1982), proposed using

$$W = \left(\frac{PRESS_{k-1} - PRESS_k}{DF_k}\right) / \left(\frac{PRESS_k}{DF_{Residuals}}\right)$$

as a decision rule for model section. According to their suggestion, the optimum value for K is the largest value of K at which W is greater than 1.

4.3 Application of parametric bootstrap methods

In this section, a parametric bootstrap method is compared with the cross-validation method, the sequential F-test and the JG/SM-test, using three datasets from the GGE and AMMI BW00RVTI and BW01RVII datasets of mean yields from two wheat and one barely trial. The two first multiplicative terms of

the AMMI analysis of the BW01RVII ALL datasets were shown in Figure 4.1. Principal component axis 1 (PC1) displays the values of $\hat{\gamma}_1\sqrt{\hat{\lambda}_1}$ and $\hat{\delta}_1\sqrt{\hat{\lambda}_1}$, whereas axis 2 (PC2) displays the values of $\hat{\gamma}_2\sqrt{\hat{\lambda}_2}$ and$\hat{\delta}_2\sqrt{\hat{\lambda}_2}$. GGE and AMMI analysis can be used to determine which genotypes are performing similarly (e.g., G4 and G14) or dissimilarly (e.g., G12 and G18) in varying environments, and to classify environments into groups of environments that are similar with regard to performance of genotypes. Table 4.1 and Table 4.2 present results of AMMI and GGE analyses, respectively.

Table 4.1 AMMI analysis of BW00RVTI and the BW01RVII datasets (sum of squares, test statistics for the cross-validation method(W), the sequential F-test (F),the JG/SM-test and the parametric bootstrap test (T), probability value for the sequential F-test, the JG/SM-test, and the simple parametric tests. Bootstrap test p-values where computed using B=100000 bootstraps samples. Under H_0, the model has K terms)

	AMMI								
Term	Data set	Sum of squares	Test statistics				Probability value		
K+1	BW00RVTI Year1	$\hat{\lambda}_{K+1}^{\ 2}$	W	F	G	T	F-test	JG/SM	Simple Bootstrap
1		508.8582	-0.2181	2.0135	1.20	0.3914	0.0131	0.0273	0.0269
2		397.0113	0.4086	2.4460	1.69	0.5018	0.0048	0.0436	0.0486
3		203.2656	0.2400	1.7931	1.06	0.5157	0.0705	0.4193	0.4161
4		98.8242	0.0203	0.9468	0.08	0.5176	0.5469	0.9960	0.9910
1	BW00RVTI Year2	364.7664	0.1272	2.2680	1.53	0.4803	0.0075	0.1144	0.1170
2		224.2592	0.2653	2.2359	1.60	0.5681	0.0188	0.1439	0.1473
3		140.8219	0.1814	4.2159	4.31	0.8259	0.0029	0.0145	0.0112
1	BW00RVTI (year 1 & 2)	693.6966	0.3676	2.6576	1.56	0.3148	0.0001	0.0148	0.0199
2		415.9191	-0.0794	1.9881	1.04	0.2754	0.0062	0.4073	0.3900
3		345.0615	0.0412	2.1493	1.18	0.3153	0.0040	0.2421	0.2329
4		232.7798	0.1980	1.8439	0.94	0.3107	0.0236	0.5688	0.5542
5		182.9161	0.1001	1.9195	1.01	0.3542	0.0241	0.4697	0.4545
1	BW01RVII	454.4874	0.0622	3.0491	2.06	0.4813	0.0001	0.0092	<.0001
2		224.4913	0.0701	2.1482	1.28	0.4584	0.0066	0.2372	0.0009
3		167.5780	0.1889	3.0019	2.29	0.6317	0.0009	0.0333	0.0960
4		97.6916	0.1535	6.48E+28	0	1	0.0001	<.0001	<.0001

Table 4.2 GGE analysis of BW00RVTI and the BW01RVII datasets (sum of squares, test statistics for the cross-validation method(W), the sequential F-test (F),the JG/SM-test and the parametric bootstrap test (T), probability value for the sequential F-test, the JG/SM-test, and the simple parametric tests. Bootstrap test p-values where computed using B=100000 bootstraps samples. Under H_0, the model has K terms)

	GGE								
Term	Data set	Sum of squares	Test statistics				Probability value		
K+1	BW00RVTI Year1	$\hat{\lambda}_{K+1}^{\ 2}$	W	F	G	T	F-test	JG/SM	Simple Bootstrap
1		9207.5563	0.7943	32.5601	31.6	0.9114	0.0001	<.0001	<.0001
2		494.4637	0.1312	3.0264	2.34	0.5522	0.0005	0.0085	0.0145
3		203.3296	0.0437	1.7484	1.04	0.5070	0.0737	0.4397	0.4675
4		98.8785	-0.0043	0.8894	0	0.5002	0.5977	0.9998	1
1	BW00RVTI Year2	3671.7824	0.2871	15.5584	15.99	0.8626	0.00010	<.0001	<.0001
2		346.2378	0.0419	2.4873	1.93	0.5920	0.0079	0.0672	0.0834
3		207.2377	0.0532	5.9063	6.70	0.8684	0.0003	0.0021	0.0020
1	BW00RVTI (year 1 & 2)	13843.8972	1.1645	43.2276	10.28	0.8800	0.0001	<.0001	<.0001
2		671.2481	-0.0794	2.9413	6.40	0.3555	1.86E-05	<.0001	0.0085
3		350.4476	0.0436	1.9247	8.47	0.2879	0.0103	<.0001	0.5128
4		255.3868	0.0581	1.7437	5	0.2947	0.0325	<.0001	0.7205
5		187.49399	-0.0135	1.5799	15.44	0.3067	0.0777	<.0001	0.8582
1	BW01RVII	3671.7824	0.2871	15.5584	15.99	0.8626	0.0001	<.0001	<.0001
2		346.2378	0.0419	2.4873	1.93	0.5920	0.0079	0.0672	0.0841
3		207.2377	0.0532	5.9063	6.70	0.8684	0.0003	0.0021	0.0017

The tables show tests of the first three to five multiplicative terms were tested. The p-value of the parametric bootstrap test was derived using B=100,000 bootstrap samples.

The parametric bootstrap tests and the JG/SM-tests gave some similar, but not identical, p-values. The sequential F-test was generally much more liberal. In most cases, the cross-validation W-statistic was smaller than 1 although the parametric bootstrap and JG/SM-tests indicated significant effects at level 0.05. Following Eastment and Krzanowski (1982a) decision rule, the cross-validation method agreed with the parametric bootstrap and JG/SM-tests.

Using the bootstrap tests, an AMMI model with two terms was appropriate for the BW00RVTI (year1) and the BW01RVII datasets. For the BW00RVTI (year 2) dataset, there were no significant patterns in the interaction. For the BW00RVTI All dataset, AMMI PC1 and PC 2 captures 31.5% and 19 % of the genotype-by-environment interaction, respectively (Table 4.1). The parametric bootstrap methods indicate that the first term (PC1) is significant, but the second term (PC2) is not. Since $H_o: k = 1$ cannot be rejected, the interaction pattern displayed by the vertical axis in Figure 4.1 is not larger than one would expect by chance. For example, the difference between genotypes G12 and G18 should not be overemphasized. In this case, a biplot that illustrates genotype and environment means on the horizontal axis and PC1 on the vertical axis (Gauch, 1992b) is appropriate. Forkman and Piepho (2014) propose that terms are tested sequentially until a non-significant result is obtained. With this decision rule, the final model for the BW00RVTI All dataset contains a single multiplicative term.

The problem of testing multiplicative terms in GGE and AMMI analysis is complicated, because estimated squared singular values are not chi-squared

distributed (Schott, 1986). The F-distribution has been proposed as reference distribution for various approximate test statistics based on ratios of mean squares (Cornelius et al., 1996), and the question has been how to calculate numbers of degrees of freedom (Gauch, 1992b, Mandel, 1971). Using resampling methods, the degrees-of-freedom problem of identifying the actual reference distribution can be circumvented. The reference distribution is simulated, which enables approximate inference. Forkman and Piepho (2014) proposed parametric bootstrap methods for testing multiplicative terms in GGE and AMMI models. The results of the simulation study indicated that these methods can be used to select the number of multiplicative terms to be retained in the model. According to Bradley (1978), the empirical level should not deviate from the nominal level by more than 10 %. Under this rule, the Type I error rate should, at nominal level 0.05, not exceed 0.055. The parametric bootstrap methods fulfill this requirement, but the JG/SM-test does in general not.

The JG/SM-test was proposed together with a sequential testing procedure (Schott, 1986), under which terms are tested sequentially, beginning with K=0 and continuing with $K = 1,2,\dots,M-2$ as long as H_0 is rejected. The $(K+1)th$ term should not be tested unless the Kth term is significant. Through this procedure, the requirement that the first K singular values be large is expected to be fulfilled. Since a simple parametric bootstrap method makes use of the same approximation, a sequential testing procedure can be recommended for this method as well. Equivalently, if all multiplicative terms are tested and the results compiled in a table similar to Table 4.1, a forward selection procedure may be applied when deciding on which terms to retain in the final model. If models are tested sequentially until a non-significant result is obtained, then computed p-values are not strictly correct provided they have not taken into account the probabilities that the tests are performed at all.

The parametric bootstrap methods of the present section can be used for the problem of selecting principal components in PCA. Due to the demonstrated fine performance with regard to power and probability of Type I errors by Forkman and Piepho (2014), a parametric bootstrap methods for testing multiplicative terms in GGE and AMMI analyses is recommended. A parametric bootstrap methods which is easier to program and computationally more effective, it is advisable to determine the number of multiplicative terms to retain in the model through a forward-selection procedure.

The previous chapters demonstrated performance with regard to methods of missing data imputation and AMMI and GGE statistical data analysis methods which are dependent on a balanced data case and also do not consider spatial variability in the experiment. The next two chapters cover a case of unbalanced data analysis and spatial variations with experimental plot through a mixed model approach.

Chapter 5
Spatial analysis of a field experiment

5.1 Introduction

Crop breading data from field experiment exhibit spatial variation, so-called because it is a function of the location of the plot in the field (Gilmour et al., 1997a) present a method of analysis in which spatial variation is modelled, resulting in estimates of treatment effects can be attributed to small scale which have greater accuracy and precision than more traditional methods such as RCB and IB see Gleeson and Cullis (1987), for example. Gilmour et al. (1997a) partition spatial variation into two type of smooth spatial trend (local and global) and extraneous variation.

Local trend effects can be attributed to small scale soil depth and fertility fluctuations. Global trend effects could be non-stationary trend across the field. Extraneous variation is often linked to the management of the trial. An example is the effect of harvesting in a serpentine manner up and down the rows in the field, with plots harvested in 'up' direction being consistently lower or higher yielding than plots harvested in 'down' direction. Global trend and extraneous variation are accommodated in the model by including appropriate terms such as design factors and polynomial functions of the spatial co-ordinates of the field plots. Local stationary trend is accommodated using a covariance structure.

The decomposition of error variation provides a more plausible approach than the original spatial methodology of Gleeson and Cullis (1987) and Cullis and Gleeson (1991a) in which error variation as a whole was modelled using a covariance structure. It is assumed that an individual experiment consists of n plots which are laid out in the field as rectangular array of r rows and c

columns (n=rc). The data $\mathbf{y}_{(n\times 1)}$ are ordered correspondingly (as rows within columns). The model for $\mathbf{y}$ is given by

$$\boldsymbol{y} = \boldsymbol{X\tau} + \boldsymbol{Zu} + \boldsymbol{\varepsilon} \qquad 5.1,$$

where $\boldsymbol{\tau}^{(t\times 1)}$ and $\boldsymbol{u}^{(b\times 1)}$ are the vectors of fixed and random effects respectively.

$\boldsymbol{X}^{(n\times t)}$ and $\boldsymbol{Z}^{(n\times b)}$ are associated design matrices for these two types of effects respectively, the former assumed to be of full column rank. The vector of residual is given by $\boldsymbol{\varepsilon}^{(n\times 1)}$. It is assumed that the joint distribution of $(\boldsymbol{U}, \boldsymbol{\varepsilon})$ is Gaussian with zero mean and variance matrix

$$\begin{bmatrix} \boldsymbol{G}(\boldsymbol{\gamma}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}(\emptyset) \end{bmatrix}$$

Where γ and $\emptyset$ are vectors of variance parameters. The marginal distribution of the data are thus Gaussian with mean $\boldsymbol{X\tau}$ and $\boldsymbol{H} = \boldsymbol{ZGZ'} + \boldsymbol{R}$ variance matrix. The vector of errors $\boldsymbol{\varepsilon}$ is decomposed into a vector $\boldsymbol{\xi}^{(n\times 1)}$ of spatial trend effects and a vector of independent errors denoted by . The spatial process is assumed to be second order stationary with var[$\boldsymbol{\xi}$]=$\sigma^2\boldsymbol{\Sigma}(\boldsymbol{\alpha})$ where $\boldsymbol{\Sigma}$ denotes the spatial correlation matrix which is a function of parameters α and σ^2 the associated variance . The independent measurement error has a variance $\boldsymbol{\sigma_\eta}^2$. Thus the errors in 5.1 are given by

$$\boldsymbol{\varepsilon} = \boldsymbol{\xi} + \boldsymbol{\eta} \qquad 5.2$$

With associated variance matrix

$$\boldsymbol{R} = \boldsymbol{\sigma^2\Sigma}(\boldsymbol{\alpha}) + \boldsymbol{\sigma_\eta}^2\boldsymbol{I_n}$$

Traditional methods of analysis such as RCB and IB are a special case of 5.1. For an RCB analysis $\boldsymbol{U}$ contains replicate effects, $\boldsymbol{\eta}$ is omitted and $\boldsymbol{\Sigma} = \boldsymbol{\sigma^2 I_n}$ is simply the trial error variance, the REML estimate of which is identical to the error mean square from an ordinary ANOVA. For an IB analyses with recovery

of inter block information, $\boldsymbol{U}$ contains effects for replication and within replicates and blocks within replicates, $\boldsymbol{\eta}$ is omitted and $\boldsymbol{\Sigma = I_n \sigma^2}$ is within block error variance. Treatment effects will be included in either $\boldsymbol{\tau}$ or $\mathbf{u}$ depending on the aims of the experiment.

The key in the Gilmour et al. (1997a) approach to spatial analysis is the identification of an appropriate variance structure for plot errors. There is no longer a dichotomy between spatial analysis and traditional methods such as RCB and IB. The latter provides a legitimate error variance model which would be adopted in the spatial approach if found to be consistent with the data. This is rarely the case, however.

5.2 Local spatial trend

Local trend reflects the fact that in the absence of design effects, data from plots which are close together are more similar than those which are further apart. Thus the element of $\boldsymbol{\xi}$ are correlated, the correlation being a function of spatial distance between plots. Let $\boldsymbol{\Sigma} = \{\boldsymbol{\rho_{ij}}\}$ where $\{\boldsymbol{\rho_{ij}}\} = [\boldsymbol{\xi_i}, \boldsymbol{\xi_j}]$ is the spatial correlation between plot i and j. Since field experiments are arranged as rectangular arrays a two-dimensional co-ordinate system is required to define the location of each plot. Let $\boldsymbol{s_i} = (\boldsymbol{s_{ir}}, \boldsymbol{s_{ic}})$ denote the spatial location of the ith plot in the field, where s_{ir} and s_{ic} are the row and column co-ordinates respectively. The spatial correlation between $\boldsymbol{\xi_i} = \boldsymbol{\xi_i}(\boldsymbol{s_i})$ and $\boldsymbol{\xi_j} = \boldsymbol{\xi_j}(\boldsymbol{s_j})$ can then be written as

$$\boldsymbol{\rho_{ij}} = \boldsymbol{V}(\boldsymbol{s_i}, \boldsymbol{s_j}\,;\, \boldsymbol{\alpha})$$

Where the correlation function $\mathbf{V}$ depends on the vector of unknown parameters α. Since the process for $\boldsymbol{\xi}$ is second order stationary the correlation between two plots depends on only on the distance between them. Thus

$$\boldsymbol{\rho_{ij}} = \boldsymbol{V}\big(\boldsymbol{s_i}, \boldsymbol{s_j}\,;\, \boldsymbol{\alpha}\big) = \boldsymbol{V}(\boldsymbol{l_{il}};\, \boldsymbol{\alpha})$$

Where $\boldsymbol{l_{ij}} = (\boldsymbol{l_{ilr}}, \boldsymbol{l_{ilc}}) = \boldsymbol{s_i} - \boldsymbol{s_j}$. It is further assumed that the two–dimensional process is separable so that the correlation function is given by the product of the correlation function for each dimension. The separability assumption is computationally convenient and with field trials see Cullis and Gleeson (1991b), and Martin (1990) for example.

Thus

$$\boldsymbol{V}(\boldsymbol{l_{ij}}; \boldsymbol{\alpha}) = \boldsymbol{V_r}(\boldsymbol{l_{ijr}}; \boldsymbol{\alpha_r})\boldsymbol{V_r}(\boldsymbol{l_{ijc}}; \boldsymbol{\alpha_{rc}})$$

Where $\boldsymbol{V_r}$ and $\boldsymbol{V_c}$ are the correlation functions from the rows and columns respectively. Correspondingly, the variance matrix for $\boldsymbol{\xi}$ can be written as

$$var(\boldsymbol{\xi}) = \boldsymbol{\sigma^2}\Sigma(\alpha) = \boldsymbol{\sigma^2\Sigma_c}(\boldsymbol{\alpha_c}) \otimes \boldsymbol{\sigma^2\Sigma_r}(\boldsymbol{\alpha_r})$$

Where $\boldsymbol{\Sigma_r}$ and $\boldsymbol{\Sigma_c}$ are r×r and c×c correlation matrices for row and columns respectively.

Many forms for **V** are possible. Zimmerman and Harville (1991a) give an example used in geostatistical applications including the exponential model, which, for single dimensions is given by

$$V(l_{ij}\,;\alpha) = \exp(-\alpha_r \left|l_{ijr}\right|^p - \alpha_c \left|l_{ijc}\right|^p \ . \qquad 5.3$$

The model with p=1(referred to by Cullis et al. (1998a), as the directional exponential covariance (DEC) model) is particularly important for field experiments.

In a field experiments plots are often of equal size and laid out in a contiguous array so that the distance between plots can be measured simply in terms of row and column numbers. Let l^*_{ijr} be the difference in row number between plots i and j so that l^*_{ijr} has possible values 0,1,…, (r-1). Define l^*_{ijc} similarly. If d_r and d_c are the actual directions respectively then $l_{ijr} = d_r l^*_{ijr}$ so that

$exp\left(-\alpha_r \left|l_{ijr}\right|\right) = \rho r^{|l^*_{ijr}|}$ where $\rho r = \exp(-\alpha_r d_r)$. The function in 3.4 with p=1 is then given by

$$V(l_{ij}\,;\alpha) = \rho r^{|l^*_{ijr}|}\rho c^{|l^*_{ijr}|} \qquad 5.4,$$

where ρr and ρc are by definition positive. If this restriction is lifted 5.4 is the correlation function for a separable autoregressive process of order 1 (AR1 ×AR1). Cullis and Gleeson (1991b) proposed this as a plausible correlation structure for the spatial trend. The parameters $\boldsymbol{\alpha} = (\rho r, \rho c)$ are known as autoregressive correlation coefficients. Many other forms of spatial correlation matrix (**Σ**) are possible, Gleeson and Cullis (1987) for example gives the subclass of $ARIMA\,(p_1, d_1, q_1) \times ARIMA(p_2, d_2, q_2)$ models with $P_1 = P_2 = 0$ and $q_1 = q_2 = 0,\ 1$ and $d_1 = d_2 = 0,\ 1$ fitted most uniformity data sets. Experience by Cullis et al. (1998a) has shown, however, that the AR1×AR1model (or variant with an identity matrix for one of the dimension) usually provides an adequate variance structure for local spatial trend.

5.3 Global trend and extraneous variation

The determination of an appropriate variance structure for separable spatial trend process and the detection of extraneous effect are made possible through the use of graphical diagnostics. Two key graphs are of estimated residuals against row or column number and the three-dimensional graph of the sample variogram.

Global trend, non-stationary global trend in the row direction, say, can be displayed in the residual plot as a smooth trend (linear or non-linear) over row number for each column. A sample variogram which fails to reach a plateau in the row/ column direction is the evidence of global trend. To see this consider, without loss of generality, an experiment in which all plots are laid out in a single column. Assume that the errors are functionally related to the row co-

ordinates of the plots. If the relationship is linear, for example, then in the absence of measurement error

$$\varepsilon_i = \xi_i + \beta s_{ir}$$

where β is a regression coefficient. Assume that the second order stationary trend process ξ has an AR1 structure. The variogram for ε non-zero displacement l^*_{ijr} is given by

$$\omega(l_{ijr}) = \tfrac{1}{2} E\left[(\varepsilon_i - \varepsilon_j)^2\right] = \tfrac{1}{2} E\left[\left((\xi_i - \xi_j)^2 + \beta(s_{ir} - s_{jr})^2\right)\right]$$

$$= \tfrac{1}{2} E\left[(\xi_i - \xi_j)^2\right] + \tfrac{1}{2} \beta^2 (s_{ir} - s_{jr})^2$$

The first term is the variogram for AR1 process ξ so that it can be written as

$$\omega(l_{ijr}) = \sigma^2\left\{1 - \rho r^{|l^*_{ijr}|}\right\} + \frac{1}{2}\left\{\beta d_r l^*_{ijr}\right\}^2$$

The first term in this equation tends to σ^2 as the displacement l^*_{ijr} increases but the second term keeps increasing. Hence the variogram for the non – stationary process ε never reaches a plateau.

Historically non-stationarity of this type was corrected by differencing the data see Gleeson and Cullis (1987) for example but this complicates the analysis. Gilmour et al. (1997a) recommend an alternative approach which involves the fitting of polynomial functions or cubic smoothing splines to row and/or column coordinates of the plots. Thus the non-stationary global trend is explicitly modelled.

Extraneous variation that comes as a result of management of field trials involves procedures which are aligned with the rows and columns. Examples are the sowing and harvesting of plots. Certain procedures may result in row and column effects (systematic and or random) in the data. Gilmour et al. (1997a) identifies this extraneous variation to distinguish it from smooth trend.

Extraneous variation may be evident from an examination of residual plots. It may be more clearly seen, however, using the sample variogram.

The inclusion of a measurement error term can be justified on both biological and statistical grounds as it constitutes lack of fit about the smooth spatial model. The philosophy of inclusion of a measurement error has been adopted by several authors including Wilkinson et al. (1983a) in their smooth trend plus independent error model for field experiments. The need for measurement error may be revealed in the sample variogram.

5.4 Correlation modelling

For a mixed model of equation 5.1, each element of the residual vector ε follows a normal distribution with mean zero and variance σ^2. Equivalently, we can say that the vector ε follows a multivariate normal distribution with a mean vector of zeros, and variance-covariance matrix $\sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Likewise, each element of the vector of random effects u_i follows a normal distribution with mean zero and variance $\gamma_i\sigma^2$. Again, equivalently, we can say that the vector u_i follows a multivariate normal distribution with mean vector of zeros, and variance-covariance matrix $\gamma_i\sigma^2\boldsymbol{I}$.

In correlation modelling, the equation of the mixed model remains the same, but the vectors of random effects $\boldsymbol{U_i}$ now follow multivariate normal distributions with a variance-covariance matrix $\gamma_i\sigma^2,\boldsymbol{G}$, where γ_i is a variance ratio and the matrix $\boldsymbol{G}$ is defined using a correlation model (or it may remain the identity matrix $\boldsymbol{I}$ if the effects are independent, i.e. uncorrelated, as in traditional model). Likewise, the residual vector $\boldsymbol{\varepsilon}$ now follows a multivariate normal distribution with variance $\sigma^2\boldsymbol{R}$, where the matrix $\boldsymbol{R}$ may be defined using a correlation model. (Again $\boldsymbol{R}$ remains the identity matrix $\boldsymbol{I}$ if the effects are independent.) If we write the value in the correlation matrix $\mathbf{C}$ (either $\boldsymbol{G}$ or

R) C_{ij} in row i and column j as C_{ij}, the most useful models for spatial modelling can be defined as follows in table 5.1.

In a two-dimensional spatial model, a correlation model is fitted to a random term field row and field column, where field row and field column are factors representing row and column positions up-and-down or from side-to-side of the whole field (rather than within replicates). Usually a separable correlation model is fitted, in which the correlation between the plots at coordinates (i, j) and (k, l) is the product of a correlation from a model defined on the rows of the experiment, and a correlation from a model defined on the columns of the experiment: i.e. correlation cr_{ik} between rows $(i - k)$ apart × correlation cr_{jl} between columns $(j - l)$ apart with the correlations cr_{ik} and cr_{jl} being defined by one of the models above.

Table 5.1 Four spatial models and their parameter structure.

Identity	$C_{i,i} = 1$ $C_{i,j} = 0$, for $i \neq j$
Auto-regressive order 1 (AR1)	$C_{i,i} = 1$ $C_{i+k,i} = \varphi^k$
Auto-regressive order 2 (AR2)	$C_{i,i} = 1$ $C_{i+1,i} = \varphi_1/1 - \varphi_2$ $C_{i,j} = \varphi_1 C_{i-1,j} + \varphi_2 C_{i-2,j}$ $i > J + 1, -1 < \varphi_1, \varphi_2 < 1$ $\lvert\varphi_1 + \varphi_2\rvert < 1, \varphi_2 + \varphi_1 < 1, \varphi_2 > -1$
Power-distance	$C_{i,i} = 1$ $C_{i+k,i} = \varphi^d$ $d = \lvert i - j\rvert$

Separable correlation models are often represented using the direct product symbol ⊗: so a model constructed from two AR1 models is written as AR1⊗AR1.

5.4.1 Correlation model application to BW01RVII data set

A spatial correlated model was fitted for data set BW01RVII (location1-5) with fixed term variety, and then considered the error model by specifying the row and column factors and selecting one of the four correlation model (identity, power, AR1 or AR2). In addition we include a random row or column effect or/and a linear trend across arrow or/and a column. The functional form of three models with one of the four correlation structures (identity, power, AR1 or AR2) applied in this chapter are as follows:

$$\boldsymbol{Y} = \boldsymbol{X}\beta + r + c + r \times c + \varepsilon \quad 5.5$$

$$\boldsymbol{Y} = \boldsymbol{X}\beta + b_r T + b_c T + r \times c + \varepsilon \quad 5.6$$

$$\boldsymbol{Y} = \boldsymbol{X}\beta + b_r T + b_c T + r + c + r \times c + \varepsilon \quad 5.7$$

where β is either a vector of variety effect or variety and block effect r and c are the random row and column effects, $b_r T$ and $b_c T$ represents a fixed linear trend.

The results for each location are given in Table 5.2 and Appendix A. These tables show the fact that the F-statistics in the spatial analysis is larger than the F-statistics from the analysis of variance of randomized complete block design suggests that the spatial analysis capacity to model the fertility of the field more effectively for all five different locations. The assessment of the effect of extending or simplifying the random model by a more complicated correlation structure is performed through Akaike information criteria (AIC) values between the models.

The best model from is the one with the smallest AIC value and it is the fixed model of Constant + entry + lin_row + lin_col and random effect of row × column with an auto-regressive order 2 (AR2) correlation for location 1 and location 5 and auto-regressive order 1 (AR1) for location 4 . An identity and auto-regressive order 1 (AR1) correlation structures with a fixed effect model

constant +entry and random model row + column + row × column are the best models for location 2 and location 3 respectively. An additional fit of a linear trend across rows (lin_row) and columns (lin_col) on the fixed effect model for location 1, location4 and location 5 improve the fit and smoothens of the variaogram. The addition of lin_row or lin_column and correlated structures improve the analysis which can lead to increase in efficiency of genotypes comparisons and decision making for variety recommendations. The presence of best model difference between the locations is also an indication of the pre-question of doing one type of analysis for the different data sets emanating from different locations that is a usual practice of most plant breeders. The next chapter addresses the spatial variation of experimental plots and variety ranking in a better way using more complicated spatial isotropic and anisotropic covariance structures.

Table 5.2 Results from the application of the linear mixed correlation models with four correlation structures to the BW01RVII data set in Location 1.

Model for location1	Correlation structures	AIC	Fixed term	df	F-value	p-value
Response yield Fixed model: Constant +entry Random model: row + column +row.column (5.5)	identity	358.81	Entry(genotypes)	24	2.04	0.013
	Power	354.96	Entry(genotypes)	24	2.70	0.002
	auto-regressive order 1 (AR1)	354.96	Entry(genotypes)	24	2.70	0.002
	auto-regressive order 2 (AR2)	345.74	Entry(genotypes)	24	3.73	<0.001
Response yield Fixed model: Constant + entry + lin_row + lin_col Random model: row.column (5.6)	identity	353.42	Entry(genotypes)	24	2.05	0.010
			lin_row	1	11.63	0.001
			lin_col	1	8.01	0.006
	Power	349.66	Entry(genotypes)	24	2.61	0.002
			lin_row	1	4.60	0.050
			lin_col	1	4.72	0.048
	auto-regressive order 1 (AR1)	349.66	Entry(genotypes)	24	2.61	0.002
			lin_row	1	4.60	0.050
			lin_col	1	4.72	0.048
	auto-regressive order 2 (AR2)	341.39	Entry(genotypes)	24	3.59	<.001
			lin_row	1	1.07	0.379
			lin_col	1	2.53	0.158
Response :yield Fixed model: Constant + entry + lin_row + lin_col Random model: row + column + row.column (5.7)	identity	353.68	Entry(genotypes)	24	2.24	0.005
			lin_row	1	2.80	0.236
			lin_col	1	8.79	0.004
	Power	352.92	Entry(genotypes)	24	2.60	0.002
			lin_row	1	2.49	0.245
			lin_col	1	5.20	0.039
	auto-regressive order 1 (AR1)	352.92	Entry(genotypes)	24	2.60	0.002
			lin_row	1	2.49	0.245
			lin_col	1	5.20	0.039
	auto-regressive order 2 (AR2)	344.95	Entry(genotypes)	24	3.74	<.001
			lin_row	1	0.90	0.428
			lin_col	1	2.79	0.153

*(ANOVA model Entry(genotypes) F-value 1.96, P-value 0.015)

Chapter 6
Local stationary trend and its influence on multi-environmental crop variety trial assessment

6.1 Introduction

National multi-environmental yield trials (MET), allow assessment of the potential yield performance of different varieties across a range of environments (locations and possibly over years, as well as combination of the two). These trials play an important role in crop variety evaluation in breeding programs and varietal recommendations for plant production. It is therefore vital that the statistical methods used to design the studies and analyse data from national yield trial evaluation programs are as accurate, efficient and informative as possible. Although the development of statistical methods for analysing variety trial data has a long history, due to the complexity of varietal and environmental interactions there is no specific model that is generally suitable for analysing combined data sets from national trials. Spatial variability often exists in field experiments due to factors such as moisture, fertility, pH and structure of the soil, as well as the pressure of diseases and pests (Davidoff and Selim, 1988, Scharf and Alley, 1993, Stroup, 2002, Wu and Dutilleul, 1999). Multi-environment crop variety trials and field evaluations are a particularly well-known example of this. Failure to effectively control for spatial variability greatly increases the risk of misleading interpretations or erroneous inferences (Mo and Si, 1986, Stroup, 2002, Yang et al., 2004).

Historically, the analysis of variance (ANOVA), along with randomised block designs (including complete, incomplete blocks), has been used to deal with the spatial variability of these trials. Numerous studies have shown that such design-based control of the spatial variation of field trials are often not optimal and results in poor analysis efficiency (Yang et al., 2004). Statistical procedures that account for spatial variation between plots within trials have been

proposed to address the topic of modelling spatial variation in crop evaluation trials using polynomial trend analysis, nearest neighbour analysis and a model with correlated errors.

The problem with the ANOVA method as a means to analyse multi-environmental crop variety trials is that it requires the assumption of homogenous variance–covariance structures across locations or environments. This homogeneity of variance and covariance may be unrealistic in many circumstances (Kempton, 1984, Piepho, 1999a). As a result, a range of more complex and informative models that can account for variance or/and covariance heterogeneity have been proposed for analysing MET data Stefanova and Buirchell, (2010). While other models are available, the problem of how the models should be assessed and which model is more suitable for a given trial's data has not been solved. This restricts the applicability of the models and model choice. Therefore, a linear mixed model approach with flexible spatial variance–covariance structures is proposed. Correspondingly, model-based approaches for analysing field trials that focus on the need to control spatial variation have been put forward. These approaches include nearest neighbour adjustment (NNA) analysis and its modifications (Bartlett, 1978, Clarke and Baker, 1996, Cullis and Gleeson, 1991b, Yang et al., 2004). Other options include linear mixed models with spatial covariance structures such as those used in geostatistics (Gilmour et al., 1997a, Stroup, 2002, Zimmerman and Harville, 1991a). The efficiency of spatial approaches has been compared with the no spatial analyses found in the literature (Brownie and Gumpertz, 1997, Hong et al., 2005, Smith et al., 2001a, Wu and Dutilleul, 1999, Yang et al., 2004).

However, most comparisons of efficiency in the literature appear to focus on the nearest neighbour adjustment (including its modification or extensions) and/or the linear mixed model with one special covariance structure (usually the first order autoregressive model, AR(1)) against the analysis of variance of

block designs (Aweke, 2005). There have been few comparisons of mixed models with different spatial covariance structures. Now a migration seems to be taking place from the NNA to a fully-fledged mixed model analysis with different spatial components for spatial variability because of the flexibility, simplicity of use and other advantages of mixed model analysis (Piepho et al., 2008). Recently, linear mixed models have become well developed, and range from simple variance component models that provide information similar to ANOVA, to models with complex variance–covariance structures that aim to explore complex sources of variability and better accommodate interactions. Specifically, different analytical models can be cast in a unified mixed modelling framework (Denis et al., 1997, Piepho, 1998a, Piepho, 1999).Within such a framework, different models can be handled as mixed models with different variance–covariance structures. Thus candidate models can be assessed and selected for MET data analyses, which can result in high accuracy when estimating variety effects and testing for differences between them.

Within advanced experimental designs, many spatial methods were proposed for adjusting the spatial trend (Bartlett, 1978, Gilmour et al., 1997a, Gleeson, 1997, Piepho, 1999, Schwarzbach, 1984, Wilkinson et al., 1983a, Williams, 1986). A common feature of these methods is that plots that are closer together are assumed to have a higher correlation than plots farther apart. Via such models the precision of genotypic value estimates can be improved through both blocking and the adjustment of spatial trend in one or two dimensions.

With regard to the practical application of the linear mixed model with a spatial component, various unsolved problems must be dealt with. Among other issues, these are concerned with the selection of a suitable covariance model, *i.e.*, a model with criteria that form the basis for a user's choice of whether or not to use a spatial model at all. Another point in this regard is the fact that the covariance parameters are unknown in practice and the estimated values

based on observed data have to be used. In this case the statistical tests about the fixed effects of linear mixed models are generally not exact and their degrees of freedom are obtained by an estimation process Kenward and Roger (1997). For some types of mixed models, the available methods for approximating degrees of freedom have been well examined (Schaalje et al., 2002, Spilke et al., 2004, Spilke et al., 2005). For mixed models with spatial covariance structures, however, the use of the approximation methods has to be undertaken with care. In addition of the approximation, further consideration has to be given to the question of what influence the various spatial models have on the statistical tests used for, ranking and selection of lines in cultivar trial evaluations, apart from on efficiency vis-a-vis standard errors for line effects estimates. In MET, the local spatial tendency within trials and the residual heterogeneity between trials can be jointly modelled in the context of linear mixed models. By using a two-dimensional coordinate system at each trial, it is possible to define the plot location in a field, for example by specifying the latitude and longitude of plot centres (Casanoves et al., 2005, Casanoves et al., 2013).

The main objective and present contribution of this paper were (1) to highlight the advantages of mixed effect models in the data analysis of a national MET; (2) to show the importance of several main spatial variance–covariance structures, and direct implications of model choice for the inference of varietal performance, ranking and testing based on two data sets from real national trials by comparing blocking without spatial effect (ANOVA) model and a model with a block and spatial effect; the mixed models with spatial variance-covariance structure models were fitted using restricted maximum likelihood (REML) approach; and finally (3) we were able to compare parameter estimates, ranking of varieties and comparing estimates, ranking order and tests of varietal effects between the ANOVA model with only block effects and the mixed effects model with a block effect with selected spatial variance–covariance structure.

6.2 Material and methods

Linear mixed models have become well developed, and range from simple variance component models that provide information similar to ANOVA, to models with complex variance–covariance structures that aim to explore or better accommodate interactions. Specifically, different analytical models can be cast in a unified mixed modelling framework (Denis et al., 1997, Piepho, 1998,1999). Within such a framework, different models with specific variance–covariance structures can be formulated. Thus candidate models can be assessed and selected for MET data analyses, which result in high accuracy when estimating and testing varietal effects. Although there are already some general reviews of crop breeding analysis and variety evaluation trials (Davidoff and Selim, 1988, Smith et al., 2001a, 2005), as well as studies on the analysis of MET data using the mixed models (Bartlett, 1978, Kelly et al., 2007a, Piepho, 1997, Piepho and Möhring, 2010, Stefanova and Buirchell, 2010), most references just contain some examples for demonstration, or contain just one specific type of mixed model in data analysis.

Both traditional block design ANOVA models and spatial effect models can take the general form of the linear mixed model:

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{Z}u + e \tag{6.1}$$

where y stands for the vector of observations, X is a matrix of constants associated with the fixed effects contained in the vector ß, ß is a vector of unknown fixed effects, Z is a matrix of constants associated with the random effects, u is a vector of random effects, and e is a vector of random residual errors. The random effects are assumed to be distributed as multivariate normal (MVN) or more precisely $\mathbf{u}$ ~MVN (0, $\mathbf{G}$) and the residual errors (e) distributed as MVN (0, $\mathbf{R}$). It follows that the vector of observations is distributed as $\mathbf{y}$~MVN($\mathbf{X}\beta, \mathbf{V}$) where $\mathbf{V} = \mathbf{ZGZ}'+\mathbf{R}$, The matrix $\mathbf{G}$ is the covariance matrix among random effects, $\mathbf{R}$ is the covariance matrix among the random

residual errors, and **V** is the covariance matrix of **y**. For block designs, block effects may be regarded as fixed or random effects. A random block analysis makes additional use of the so-called inter-block information and is generally the preferred approach (Littell et al., 2006). In this article, block effects will be considered random in a combined analysis of data from different location. In this situation, **u** is the vector of block effects, and **Z** corresponds to the block effect design.

For analysis of variance models for block designs, block effects are assumed to be iid ~ N $(0,\sigma^2{}_b)$, and residual errors are assumed to be iid N~ $(0,\sigma^2)$, where iid denotes independent and identically distribution, and $\sigma^2{}_b$and $\sigma^2$are variance components of blocks and residual errors, respectively. Hence, $\mathbf{G} = \mathbf{I}_b\sigma^2{}_b$and $\mathbf{R} = \mathbf{I}_n\sigma^2$, where I_b is an identity matrix whose dimension equals the number of blocks, I_n is an identity matrix whose rank equals the number of observations. The main feature of analysis of variance models for block designs is that random variables located in the same block have the same covariance regardless of the extent of spatial variation; random variables not located in the same block have a covariance of zero.

In spatial effect models, **R** takes the form $\mathbf{R} = \mathbf{I}_n\sigma^2 + \sigma^2{}_s\mathbf{F}$, where $\sigma^2{}_s$ is the covariance parameter of spatial structure variation, **F** is a square matrix with a dimension reflecting the number of observations, whose ij^{th} element is $f(d_{ij})$, in which d_{ij} is the Euclidian distance between spatial observation points *i* and *j*. Suppose (x_i,y_i) and (x_j ,y_j) describe the coordinates of the median points of plots for observations *i* and *j*, respectively, then their distance is:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \qquad (6.2),$$

where x and y denote horizontal and vertical directions. The variable $f(d_{ij})$ is generally a function of d_{ij} and its form is dependent on the spatial model used,

which is dependent on the characteristics of spatial variation. The spatial covariance structures available for analysing field trials are listed in Table 1.

In Table 1 *c-list* contains the names of the numeric variables used as coordinates of the location of the observation in space, and d_{ij} is the Euclidean distance between the i^{th} and j^{th} vectors of these coordinates, which correspond to the i^{th} and j^{th} observations in the input data set. For SP(POWA) and SP(EXPA), c is the number of coordinates, and $d(i,j,k)$ is the absolute distance between the k^{th} coordinate, $k = 1 \ldots, c$, of the i^{th} and j^{th} observations in the input data set. For the geometrically anisotropic structures SP(EXPGA), SP(GAUGA), and SP(SPHGA), exactly two spatial coordinate variables must be specified as c_1 and c_2 . Geometric anisotropy is corrected by applying a rotation θ and λ scaling to the coordinate system, $d_{ij}(\theta, \lambda)$which represents the Euclidean distance between two points in the transformed space. SP(MATERN) and SP(MATHSW) represent covariance structures in a class defined by Matérn see (Handcock, 1994, Handcock and Stein, 1993, Matérn, 1986). The function $K_v$is the modified Bessel function of the second kind of (real) order $v > 0$; the parameter governs the smoothness of the process (for further detail see SAS 9.3 help and documentation). The five spatial-variance covariance structures presented above belong to isotropic models, *i.e.*, the variation properties are the same in both directions x and y; the other models, as their names show, belong to anisotropic covariance structures, *i.e.*, the variation properties can be different in directions x and y.

Table 6.1 Spatial Covariance Structures

Structure	Description	Parameters	(i ,j) elements
SP(EXP)(c-list)	Exponential	2	$\sigma^2\exp\{-d_{ij}/\theta\}$
SP(EXPA)(c-list)	Anisotropic Exponential	2c + 1	$\sigma^2\prod_{k=1}^{c}\exp\{-\theta_k d(i,j,k)^{pk}\}$
SP(EXPGA)(c1c2)	2DExponential,Geometrically Anisotropic	4	$\sigma^2\exp\{-d_{ij}(\theta,\lambda)/\rho\}$
SP(GAU)(c-list)	Gaussian	2	$\sigma^2\exp\{-{d_{ij}}^2/\rho^2\}$
SP(GAUGA) (c1c2)	2DGaussian,Geometrically Anisotropic	4	$\sigma^2\exp\{-d_{ij}(\theta,\lambda)^2/\rho^2\}$
SP(LIN)(c-list)	Linear	2	$\sigma^2(1-\rho d_{ij})1((\rho d_{ij}\le 1)$
SP(LINL)(c-list)	Linear Log	2	$\sigma^2(1-\rho\log(d_{ij}))\times 1((\rho\log d_{ij}\le 1)$
SP(MATERN)(c-list)	Matérn	3	$\sigma^2\frac{1}{\Gamma(v)}\left(\frac{d_{ij}}{2\rho}\right)^{v}2K_v\left(\frac{d_{ij}}{\rho}\right)$
SP(MATHSW)(c-list)	Matérn (Handcock-Stein-Wallis)	3	$\sigma^2\frac{1}{\Gamma(v)}\left(\frac{d_{ij}\sqrt{v}}{2\rho}\right)^{v}2K_v\left(\frac{2d_{ij}\sqrt{v}}{\rho}\right)$
SP(POW)(c-list)	Power	2	$\sigma^2\rho^{d_{ij}}$
SP(POWA)(c-list)	Anisotropic Power	C + 1	$\sigma^2{\rho_1}^{d(i,j,1)}{\rho_2}^{d(i,j,2)}\dots{\rho_c}^{d(i,j,c)}$
SP(SPH)(c-list)	Spherical	2	$\sigma^2\left[1-\left(\frac{3d_{ij}}{2\rho}\right)+\left(\frac{{d^3}_{ij}}{2\rho^3}\right)\right]1((\rho d_{ij}\le\rho)$
SP(SPHGA)(c1c2)	2D Spherical, Geometrically Anisotropic	4	$\sigma^2\left[1-\left(\frac{3d_{ij}(\theta,\lambda)}{2\rho}\right)+\left(\frac{d_{ij}(\theta,\lambda)^3}{2\rho^3}\right)\right]\times 1(d_{ij}(\theta,\lambda)\le\rho)$

Estimation and statistical test of varietal effects for the classical analysis of block designs uses ANOVA, which is, equating the observed mean squares to the expected mean squares with the assumption of independence, normality and homogeneity of the variances of the residuals. Spatial models analyses use REML for estimating variance components. Estimable functions $\boldsymbol{L\beta}$ of linear contrast of fixed effects (variety) are estimated based on $\boldsymbol{L\hat{\beta}} = \boldsymbol{L}(\boldsymbol{X}'V^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'V^{-1}\boldsymbol{y}$ with V being replaced by a REML estimate $\hat{V}$. The

variance of $\boldsymbol{L\hat{\beta}}$ is determined based on $var(\boldsymbol{L\hat{\beta}}) = \boldsymbol{L}(\boldsymbol{X}'\hat{\boldsymbol{V}}^{-1}\boldsymbol{X})^{-}\boldsymbol{L}'$ (Hartley and Rao, 1967; Harville, 1977). Null hypotheses of the form of H_o: $\boldsymbol{L\hat{\beta}} = 0$ are tested using the statistic

$$t = \frac{\boldsymbol{L\hat{\beta}}}{\sqrt{var(\boldsymbol{L\hat{\beta}})}} \sim t(d.f.) \qquad (6.3)$$

In general, the test statistic in (6.3) is only approximately t-distributed and its degrees of freedom must be estimated. The approximate degrees of freedom in this research were determined using the Kenward-Roger method (Kenward and Roger, 1997). This approximation also uses the basic idea of Satterthwaite (1941). Its extension relative to the Satterthwaite method of Giesbrecht and Burns (1985) and Hrong-Tai Fai and Cornelius (1996) is an asymptotic correction of the estimated standard error of fixed effects due to Kackar and Harville (1984) in small and/ or unbalanced data structures.

6.3 Statistical tools for model selection and test of consistency

Two questions in the analysis of practical trials are whether there is significant spatial variability and whether spatial models should be used (and if so, which models are most appropriate for data analysis). To answer these questions, statistical tools include likelihood-based methods (Oman, 1991, Wolfinger, 1993).The likelihood-ratio test (LRT) allows the comparison of the model's fit, provided that one of the models is hierarchically subordinated to the other or similarly the smaller model is nested with the larger one. This is the case if one model can be seen as a special case of a more general model due to certain model restrictions. The LRT then results from

$$LRT = -2\left(lnLL_g - lnLL_s\right) \sim \chi^2(d.f) \qquad (6.4),$$

where lnLLg and $\ln LL_s$ denote the log likelihood of the general model g and the special models, respectively. Given certain regularity conditions, the LRT testing statistic asymptotically follows a χ^2 distribution, with the degrees of freedom (d.f.) resulting from the number of restrictions that are necessary to transform the general model g into the special model (Fahrmeir L and Hamerle A, 1984, Greene WH, 2003). The general model fit, when compared to the special model, is considered better if LRT $>\chi^2$ (1-α, d.f.) with a significant level of a. If the model comparison focuses on the covariance structure of a constant expectation structure, the likelihoods are employed via the REML method (Wolfinger, 1993). This can be used for the first question. In this case, g corresponds to the model with spatial correlations among observations, and corresponds to the model without spatial correlation among observations. The LRT based on formula (4) can also be used for testing the difference between the block design ANOVA model (block effects as random) and the model without correlations among observations, because the latter is also a special model variation of the former. Thus, it can be used for testing the difference between the spatial models with and without block effects.

As mentioned above, the LRT is only applicable when comparing two nested models. For model comparisons that do not require hierarchical models, there are a number of analytical criteria. These are so-called "Information Criteria" based on likelihood estimations. In the current work Akaike's Information Criterion (AIC) is used for comparing the covariance structures for an identical expectation structure using the REML estimation methods and is generally given by:

$$AIC = -2\, lnLL + 2q \qquad (6.5),$$

where lnLL is the log-likelihood same as in formula (4) and q is the number of the parameters of the variance-covariance structure. Thus, the formula of the information criteria is given in such a way that the model with the smaller

value for the AIC is preferred (Burnham and Anderson, 2002, Bozdogan, 1987). For un-nested model we prefer to use the AIC but we note that there are other available information criteria, such as the Schwarz Bayesian Information Criterion (BIC) (Schwarz, 1978). Guerin L and Stroup WW (2000) compared the performance of AIC and BIC on covariance model selection for repeated measures and stated that AIC tends to select a more complex model but with a better control of type I error than the BIC. To assess consistency (or inconsistency) in the statistical tests on varietal effects between two models one can use the test consistency ratio, which is computed as follows:

$$test\ consistency\ ratio = \frac{the\ number\ of\ significant\ varietial\ differencess\ tested\ simultaneously\ in\ two\ models}{\max(the\ number\ of\ significan\ varietial\ differencess\ tested\ under\ the\ two\ models\ considered)}$$

6.4 Data set and analysis

The data sets used in this study are taken from the Ethiopian Agricultural Research Institute National Variety Trials for Bread Wheat (BW00RVTI data) and Barley Trial (BW01RVII data) of 2006-2008. Some 20 bread wheat (*Triticum aestivum*) varieties were tested in at six locations (environments) on the first year (2006/7) and five locations (environments) among the six of the first used on the second year (2007/8). Similarly 25 barley (*Hordeum vulgare*) varieties were tested in five locations (environments) in 2007/8. All the trials in each location were laid out as a randomized complete block (RCB) design with four replicates. There are two approaches to analysing MET data using mixed model, the so-called one- and two-stage approaches (Welham et al., 2010b). In a one-stage analysis, individual plot data from all trials are combined in a single analysis (Cullis et al., 1998a). In a two-stage analysis, variety means are first obtained from the separate analysis of individual trials (Stage I), and are then combined in an overall mixed model analysis (Stage II). The two-stage analysis can be unweighted (*e.g.* Patterson & Silvey, 1980) or weighted to

reflect the relative precision of variety means from each trial (*e.g.*, Smith et al., 2001a). A one-stage approach provides the most accurate predictions of variety performance, but it can be computationally difficult to use when the variance models involved are complex. With the steady improvements in computing power, single-stage analyses are becoming feasible. Apart from computational speed, the main advantages of the two-stage approach are that one can carefully analyse each trial individually, taking into account any specifics of the design or field trends.

In this study we used two approaches for analysis; the first one was a separate individual analysis of each location of the BW00RVTI data set of wheat and BW01RVII data set for barley. The second one was a one-stage analysis, individual plot where data from all trials (locations)are combined in a single analysis of a two year BW00RVTI data set of wheat and a one year BW01RVII data set of six location. Each data set was separately fitted per location and per year using the mixed model with fourteen variance-covariance structures. The mixed model with compound symmetry (CS) variance-covariance structures was identical to the ANOVA model. The optimally fitted spatial model and the ANOVA model are used for further varietal effect assessment and statistical tests (or inference). The single-stage analysis was applied to each of the data sets by fitting one spatial-variance covariance structure at a time for all location. Putting location as random group factor on SAS (proc mixed) analysis gave a different random parameter estimate for each location. All the analyses ware conducted using standard SAS software version 9.3. The results from the two models were compared and used to assess consistency (or inconsistency) in statistical tests on varietal effects between the two models, using consistency ratio defined earlier.

6.5 Results and discussion

The results of the fitted ANOVA and the mixed model with various variance-covariance structures are summarised in Tables 6.2, 6.3 and 6.4. Note that "-" denotes the failure of a model to converge. This occurred with the sp(lin) and sp(linlog) structures in any of the locations, which shows that these models may not suitable for that trial data (Schabenberger and Pierce, 2002). The smallest AIC value (bold in Tables 6.2, 6.3 and 6.4) indicates that for BW00RVTI trial data set year 1 and 2 support the Anisotropic Power [spa (powa)] and Exponential [spa (exp)] variance-covariance structures as the best compared to the ANOVA model for seven trials (locations) out of eleven. Similarly for the BW00RVTI trial five different spatial variance-covariance structures [sp(pow), sp(expga), sp(mathsw), sp(expga) and sp(powa)] models were selected as the best compared to the ANOVA model for the five location BW01RVII trial data set.

A model comparison between a block effect without spatial structure (ANOVA) and a model with a block and spatial effect using the LRT χ^2-test for the trials for the two (BW00RVTI and BW01RVII) data sets suggested that the selected spatial variance-covariance structure fitted the data significantly better than the ANOVA model. However the optimally-fitted spatial variance-covariance structures were not the same from one location to the other. Optimally fitted spatial variance-covariance structure was spatial power [sp(powa)] for most of the locations. These results showed that assuming a homogeneous variance-covariance structure in the ANOVA model is generally not realistic, and therefore using a linear mixed model with spatial variance-covariance is necessary to improve the efficiency of the data analysis and accommodation of local stationary trend of MET data.

It appears the year to year effect on variance-covariance of varieties is greatly exhibited in the BW00RVTI data set. This is shown through the variance-covariance structures being mostly consistent for different locations in the

same year, but obviously not consistent between years as shown in Table 6.2 and 6.3. This result is easily understood by realising that within a year we expect only between location differences, but between years there could be differences in environments (years). The failure of some spatial variance-covariance structures to converge may indicate that they are not suitable or compatible with the structure of the current MET data but could work with other data sets.

To examine the impact of the spatial variance-covariance structures on estimates on test of varieties, the number of significant varietal means yield differences by the t-test are given in Table 6.5. Using the ANOVA model and mixed model with the optimally-fitted spatial variance-covariance for each location, we assessed the consistency between these two models. The number of significant varietal differences by t-test is not the same between the ANOVA model and the mixed model with optimally fitted spatial variance-covariance structures. The consistence ratio test between the two models falls in the range of 33-84%. From the average of all trials (locations), the test consistency ratio of two models is approximately 64%, which means that approximately 36% of the pairwise varietal yield differences being tested as significant in one model cannot be tested as significant or very significant by the other model. This indicates the need for consideration of spatial correlation between plots during the analysis which can lead to better understanding of the performance of genotypes.

Table 6.2 Related fitting statistics of ANOVA model and linear mixed model with spatial variance-covariance structures for the first year BW00RVTI data set

	location- 1			location-2			location-3			location-4			location-5			location-6		
Model	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2
RCBD	385.1	387.1	-	344.2	346.2	-	398.8	400.8	-	426.4	**428.4**	-	421.7	**423.7**	-	377.3	379.3	-
sp(sph)	408.5	412.5	1	360.8	364.8	1	407.5	411.5	1	442.1	446.1	1	441.4	445.4	1	364.4	368.4	0.0003
sp(exp)	385.1	387.1	1	343.7	347.7	0.4869	394.2	398.2	0.0303	426.4	**428.4**	1	421.7	**423.7**	1	361.8	365.8	<.0001
sp(gau)	384.9	388.9	0.6537	343.8	347.8	0.5743	394.7	398.7	0.0425	426.3	430.3	0.746	421.7	**423.7**	1	365.5	369.5	0.0006
sp(lin)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sp(linlog)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sp(pow)	385.1	389.1	0.865	343.7	347.7	0.4869	394.2	398.2	0.0303	426.2	430.2	0.63	421.7	425.7	0.9213	361.8	365.8	<.0001
sp(mat)	385.1	387.1	1	343.7	349.7	0.7806	394.1	400.1	0.0953	426.1	432.1	0.8603	421.6	427.6	0.9648	359.7	365.7	0.0001
sp(EXPA)	-	-	-	-	-	-	386.4	396.4	0.0144	-	-	-	-	-	-	-	-	-
sp(EXPGA)	385.1	391.1	0.0608	334.5	342.5	0.0179	386.1	394.1	0.0747	426.4	432.4	0.2721	421.7	427.7	0.1855	353	361	0.0034
sp(GAUGA)	379.3	387.3	0.0339	344.2	352.2	1	398.8	404.8	1	426.4	432.4	0.3575	421.7	427.7	0.2295	363.9	371.9	0.1199
sp(MATHSW)	385.1	389.1	1	343.7	349.7	0.7806	394.1	400.1	0.0953	426.4	430.4	1	421.7	425.7	1	359.7	365.7	0.0001
sp(POWA)	372.7	**378.7**	0.002	332.5	**338.5**	0.003	386.4	**392.4**	0.002	424.8	430.8	0.4545	420.1	426.1	0.4665	349	**355**	<.0001
sp(SPHGA)	393.8	399.8	1	-	-	-	-	-	-	-	-	-	439.2	447.2	1	355.6	363.6	<.0001

Table 6.3 Related fitting statistics of ANOVA model and linear mixed model with spatial variance-covariance structures for the second year BW00RVTI data set.

	location- 1			location-2			location-3			location-4			location-5		
Model	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2
RCBD	370	372	-	334.2	336.2	-	345.6	**347.6**	-	395.2	397.2	-	284.1	**286.1**	-
sp(sph)	382.4	386.4	1	351.9	355.9	1	366.4	370.4	1	394.2	398.2	0.317	303	307	1
sp(exp)	367.1	**371.1**	0.091	333.5	337.5	0.428	345.6	**347.6**	1	387.3	**391.3**	0.005	284.1	**286.1**	1
sp(gau)	368.3	372.3	0.195	332.4	336.4	0.189	345.6	**347.6**	1	388.7	392.7	0.011	284.1	**286.1**	1
sp(lin)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sp(linlog)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sp(pow)	367.1	**371.1**	0.091	333.5	337.5	0.428	345.5	349.5	0.659	387.3	**391.3**	0.005	284	288	0.762
sp(mat)	365.8	371.8	0.124							386.9	392.9	0.016			
sp(EXPA)	366.2	374.2	0.29	322.9	**330.9**	0.01	343.9	353.9	0.79	384.9	394.9	0.036			
sp(EXPGA)	363.8	371.8	0.339	323.1	331.1	0.003	340.5	348.5	0.108	384.6	392.6	0.232	284.1	290.1	0.234
sp(GAUGA)	370	378	1	325.6	333.6	0.032	344	352	0.976	395.2	403.2	1	284.1	290.1	0.594
sp(MATHSW)	365.8	371.8	0.124	-	-	-	-	-	-	386.9	392.9	0.016	284.1	288.1	1
sp(POWA)	367.9	373.9	0.367	322.7	328.7	0.003	344.1	350.1	0.464	386.3	392.3	0.011	283.6	289.6	0.773
sp(SPHGA)	382.4	390.4	1	-	-	-	356.1	364.1	1	391.3	399.3	0.273	-	-	-

Table 6.4 Related fitting statistics of ANOVA model and linear mixed model with spatial variance-covariance structures for the one year BW01RVII data set

	location- 1			location-2			location-3			location-4			location-5		
Model	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2	LL	AIC	Pr > x2
RCBD	475.5	477.5	-	454.5	456.5	-	421.6	423.6	-	506.9	508.9	-	524	526	-
sp(sph)	475.5	477.5	1	454.5	456.5	1	427.7	431.7	1	506.9	508.9	1	551.3	555.3	1
sp(exp)	468.5	**472.5**	0.0083	452.3	456.3	0.142	410.9	414.9	0.001	503.7	507.7	0.0712	524	526	1
sp(gau)	471	475	0.0341	452.4	456.4	0.145	414.1	418.1	0.006	504.8	508.8	0.1409	523.9	527.9	0.689
sp(lin)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sp(linlog)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
sp(pow)	468.5	**472.5**	0.0083	452.3	456.3	0.142	410.9	414.9	0.001	503.7	507.7	0.0712	524	528	0.8947
sp(mat)	-	-	-	452	458	0.283	407.7	**413.7**	0.001	502.9	508.9	0.1344	524	526	1
sp(EXPA)	-	-	-	452.4	462.4	0.711	-	-	-	-	-	-	-	-	-
sp(EXPGA)	466.7	474.7	0.347	447.7	**455.7**	0.07	421.6	429.6	1	493.2	**501.2**	0.0141	524	530	0.0323
sp(GAUGA)	469	477	0.3427	449.6	457.6	0.157	410.1	418.1	0.275	498.6	506.6	0.0619	529.9	530.5	0.1424
sp(MATHSW)	-	-	-	452	458	0.283	407.7	**413.7**	0.001	502.9	508.9	0.1344	524	528	1
sp(POWA)	467.4	473.4	0.0179	458.8	459.2	0.435	413.7	419.7	0.02	497.1	503.1	0.0071	519.2	**525.2**	0.0873
sp(SPHGA)	475.5	481.5	1	458.7	466.7	1	420.4	428.4	0.742	495	503	0.0077	-	-	-

Table 6.5 The number of significant and very significant variety contrasts of t-test for trials of the BW00RVTI and BW01RVII data set trials and the test of consistency ratio between the ANOVA model and the spatial linear mixed model with optimally fitting spatial variance –covariance structure(SLMM).

Data set BW00RVTI									Data set BW01RVII				
Year-1					Year-2				Year-1				
	ANOVA	SLMM	Consistency		ANOVA	SLMM	Consistency			ANOVA	SLMM	Consistency	
			No	Ratio (%)			No	Ratio (%)				No	Ratio (%)
location-1	60	78	53	67.94	44	40	37	84.09	location-1	46	65	43	66.15
location-2	53	50	37	69.81	33	35	26	74.29	location-2	67	79	60	75.95
location-3	10	18	6	33.33	*	*			location-3	97	157	94	59.87
location-4	*	*			45	50	41	82	location-4	47	36	22	46.8
location-5	*	*			*	*			location-5	36	45	26	57.78
location-6	45	20	15	33.33									
Average	42	41.5	22.75	51.11	40.67	41.66	34.67	80.13	Average	58.6	76.4	49	61.31

Note * the optimally fitting model is ANOVA

6.6 Varietal ranking

Apart from contrasts between new varieties, the ranking of varietal productivity and a comparison of new varieties with standard variety is also important for variety trials. We consider the trial from the five locations of BW01RVII data to compare variety estimated means ranking between the ANOVA model and the optimal spatial variance-covariance model. A trial corresponds to a single experiment at a single location. Table 6.6 shows the ranking for the first eight entries from the optimal spatial variance-covariance mixed model compared to the ranking from the ANOVA model across the locations. The model with spatial structure is relatively more consistent in its top eight ranking than the ANOVA model. The ranking are different for different locations and differ between the spatially structured model and ANOVA. A rank difference of genotype between the locations is showing the presences of genotype by environment interaction. This also indicates the advantage of single stage spatial models on the handling of the spatial trend and variation of the trials.

Table 6.6 The first eight genotype ranking comparison between ANOVA and optimally fitting spatial variance –covariance structure (SLMM) of five trials of data set BW01RVII location by location and a single- stage analysis

Rank	Location -1		Location-2		Location-3		Location- 4		Location- 5		ALL
	ANOVA	SLMM	ANOVA	SLMM	ANOVA	SLMM	ANOVA	SLMM	ANOVA	SLMM	SLMM
1	G23	G23	G21	G23	G23	G23	G23	G23	G23	G23	G23
2	G13	G13	G23	G21	G4	G4	G11	G2	G2	G2	G21
3	G5	G21	G3	G3	G17	G17	G19	G1	G4	G1	G4
4	G17	G3	G17	G2	G2	G21	G15	G5	G8	G19	G13
5	G21	G17	G7	G17	G21	G10	G2	G19	G13	G10	G2
6	G4	G4	G11	G13	G10	G2	G5	G8	G14	G4	G17
7	G3	G5	G9	G16	G8	G8	G8	G15	G1	G3	G8
8	G19	G15	G13	G9	G6	G15	G13	G11	G19	G13	G3

The simple homogenous variance–covariance structures implied by ANOVA models, which assume that the interaction effects of varieties are independent, is mostly not appropriate for data analyses of MET. The fact that the goodness of fit of one variance–covariance structure was different for various trial data sets, and that none fitted all trial data sets optimally throughout, indicates that the heterogeneous characteristics of variance–covariance are not identical across the trials. Therefore, the arbitrary use of a homogeneous variance–covariance structure (*e.g.* ANOVA model) to analyse the MET cannot ensure a high degree of accuracy. In this study, the ANOVA model, as a special case form of mixed models, showed obvious inconsistency in estimates and tests of varietal effects compared to the linear mixed model with the optimally-fitted spatial variance–covariance structures.

Both effective experimental results and spatial analyses can have an important role in improving the reliability and precision of experimental results. The importance of spatial variability to be expected from a logical assumption of the presence of external local or global variation and subjective-related perspective is confirmed in a variety of experiments. As presented in much of the literature, spatial analysis may lead to higher efficiency with regard to standard error of estimation of fixed effects than a non-spatial analysis, provided that spatial variability is present. Based on this work, the commonly used ANOVA mixed model is not an appropriate model for data analysis of MET trials. The spatial variance–covariance models are more useful in a practical sense, given that they can describe actual existing variance–covariance characteristics more accurately than the ANOVA model. Of course, with one-stage analyses, the proposed spatial variance–covariance models are expected to yield identical mean yields for balanced data, and differences are expected only for unbalanced data. Even so, a selection of variance–covariance structures based on the mixed model framework is important since the standard error of varietal effect estimates (*i.e.* the accuracy of varietal effect estimates) is different under the various models, and unbalanced data are common in MET (Möhring and

Piepho, 2009). The advantage and validity of using spatial variance-covariance structure depends on the present spatial variability. Most of the investigated spatial models showed better data fitting and smaller standard error for variety contrasts than the ANOVA model.

The main purposes of the present chapter was to show the importance of variance–covariance structure selection and to illustrate that the classical ANOVA model is inferior to more elaborate mixed models in the analysis of MET data. This does not imply that the models considered in this paper are appropriate for any situation. For example, in some locations (trials) the ANOVA model still optimally fitted the data better than the spatial models.

6.7 Single-stage analysis

The analysis of the resulting MET data are often done by mixed models, which may be complex when different types of models are needed to characterise within-environment variation in the various environments. Single-stage analysis is regarded as preferable from theoretical considerations, because it provides best linear unbiased estimators (BLUE) of all fixed effects and best linear unbiased predictors (BLUPs) of all random effects under the assumed single-stage model. Optimal performance of this approach has been substantiated based on simulation evidence (Welham et al., 2010b).

Thus, single-stage analysis is usually considered the gold standard Smith et al., (2001). Consider a series of trials (BW00RVTI) laid out as randomised complete block designs in different locations (environments). Factors used for analysis are location (environment), complete replicate (nested within location) and genotype. All model effects will be defined in terms of these factors. A simple linear mixed model for analysing of this kind of experiment is

$$y_{ijk} = \mu_{b(i)} + t_{b(j)} + v_{b(ij)} + r_{w(jk)} + \varepsilon_{w(ijk)}$$

where y_{ijk} (i = 1, . . . , p; j = 1, . . . q, k = 1, . . . , s) is the response of the i-th genotype in the k-th replicate of the j-th location (environment), $\mu_{b(i)}$ is the i-th genotype mean, $t_{b(j)}$ is the j-th location(environment) main effect, $v_{b(ij)}$ is the ij-th genotype–environment interaction effect, $r_{w(jk)}$ is the effect of the k-th replicate in the j-th environment, and $\varepsilon_{w(ijk)}$ is the within-environment plot error associated with y_{ijk}. The expected genotype means $\mu_{b(i)}$ are taken as fixed. For the random effects we assume $t_{b(j)}) \sim N(0, \sigma^2{}_t)$ and $v_{b(ij)} \sim N(0, \sigma^2{}_v)$, $r_{w(jk)} \sim N(0, \sigma^2{}_{r(t)})$ and $\sim N(0, \sigma^2{}_{e(j)})$.

A location by location analysis in the above discussion indicates the need of spatial-variance covariance and the presence of a variation on estimates from on location to the other. To accommodate these differences a mixed model with spatial variance covariance structure with a group factor of location on the random model was fitted for the one stage analysis using the MIXED procedure of SAS. A model with spatial-variance covariance of sp(exp), sp(gau), sp(linl), sp(pow) and sp(sph) have the smallest AIC value compare to the rest of structures. As long as these models have equal AIC value, we can select a model with a smaller number of parameters and has a simple structure.

Single stage analysis estimates of parameters as it shown in Table 6.7 indicate the presence of variation from location to location and its advantage on handling of additional sources of variation that is location and genotype by location (environment) interaction variation. Most variation of replicate within a location, between locations variations and a genotype (entry) by location (environment) interaction variation are very significant and significant at 5% level of significance.

A significant variation of estimated values of a parameter for power covariance structure is also observed on this single stage analysis. These estimated values

indicate the presence of plot to plot correlation. The inclusions of these variations in single stage analysis will also improve the real estimation of genotype effect and their ranking.

Table 6.7 Variance parameter estimates for single-stage analysis of BW00RVTI data sets.

Covariance Parameter Estimates					
Cov Parm	Group	Estimate	Standard Error	Z Value	Pr Z
rep(loc)	loc 1	12.1452	3.3177	3.66	<.0001
rep(loc)	loc 2	0	.	.	.
rep(loc)	loc 3	8.3059	3.5686	2.306	0.0104
rep(loc)	loc 4	4.7991	2.8205	1.702	0.0455
rep(loc)	loc 5	0	.	.	.
loc	loc 1	90.3913	30.81	2.9338	<.0001
loc	loc 2	32.3509	18.3752	1.76	0.0392
loc	loc 3	115.43	157.78	0.73	0.2322
loc	loc 4	48.425	23.7017	1.436	0.0207
loc	loc 5	0	.	.	.
entry*loc	loc 1	0	.	.	.
entry*loc	loc 2	5.0156	3.023	1.659	0.0475
entry*loc	loc 3	4.5663	2.0231	2.257	<.0001
entry*loc	loc 4	12.2776	4.8818	2.515	0.0061
entry*loc	loc 5	10.6646	5.4038	1.973	0.0244
Variance	loc 1	24.8749	3.7637	6.61	<.0001
SP(POW)	loc 1	0.9	0.35	2.571	<.0001
Variance	loc 2	17.597	2.8736	6.12	<.0001
SP(POW)	loc 2	0.9	0.35	2.571	<.0001
Variance	loc 3	11.8619	1.9521	6.08	<.0001
SP(POW)	loc 3	0.9	0.35	2.571	<.0001
Variance	loc 4	45.7003	7.4987	6.09	<.0001
SP(POW)	loc 4	0.9	0.35	2.571	<.0001
Variance	loc 5	55.5086	9.0645	6.12	<.0001
SP(POW)	loc 5	0.9	0.35	2.571	<.0001

Model diagnostic plots in fig 6.1, plot of residuals do not show any systematic pattern.

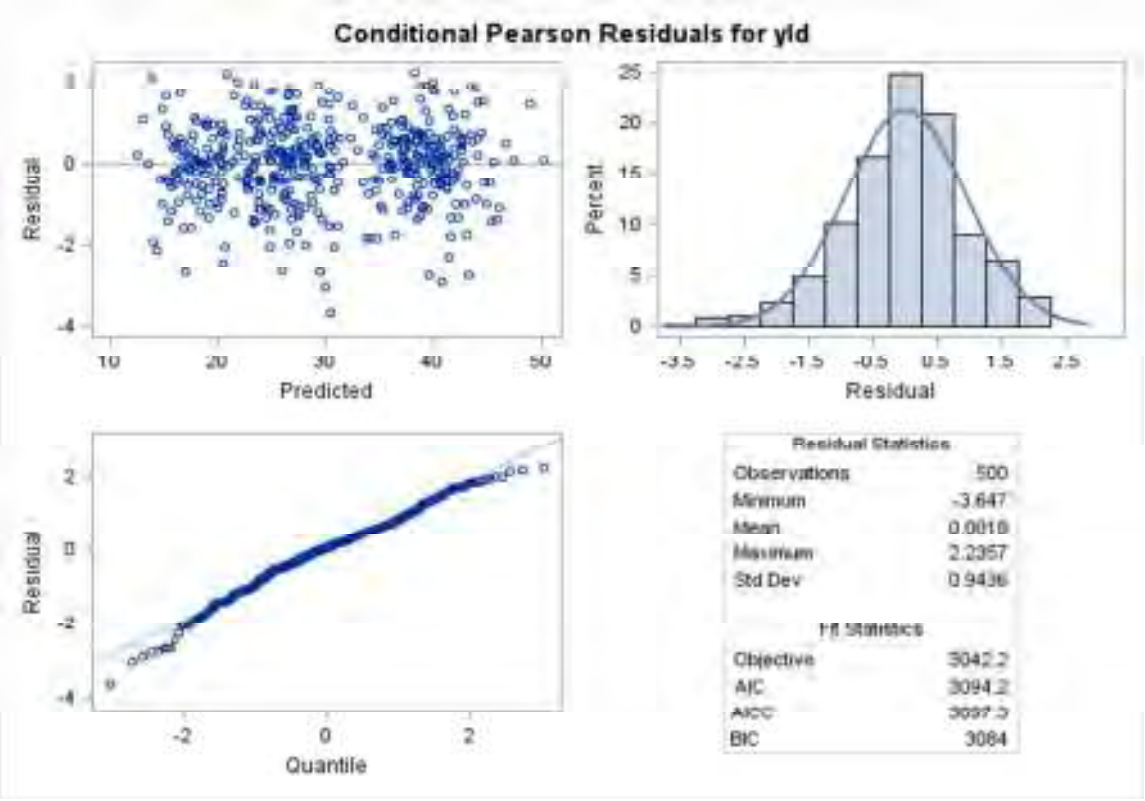


Figure 6. 1 Model diagnostic plots

Chapter 7
Conclusions and Future Research

In Ethiopia there is national and regional system for variety evaluation. The national agricultural research organization (EARO) with collaborations of regional agricultural research institutes has helped to promote consistency in trial design, methods of analysis and reporting results. The recommendation of new plant varieties for small holder farmers or commercial use requires reliable and accurate predictions of the average yield of each variety across a range of target environments and knowledge of important interactions with environment. This information is obtained from a series of plant variety trials, also known as multi-environment trials (MET). Each year, a huge amount of money is spent world-wide on the acquisition of such data. Unfortunately the value of this effort may be discounted with the use of inappropriate or inefficient statistical analysis.

The challenge of controlling field variability has been felt since the beginning of the 20th century. A design-based approach was widely used as a method of controlling variability. But, presence of correlation between adjacent plots led to a search for more efficient methods of controlling variation. Although several spatial models were implemented to account for spatial variations over years, problem of confounding between local and global variations hindered efficiency of these methods. In this thesis we cover some problems on the analysis of multi-environmental data analysis starting with the problem of handling missing data to accommodating of spatial variations on the large experimental fields. In the case of missing data, the loss of information produces unbalanced designs that lose their symmetry and, therefore, hypothesis tests of interest such as those for the difference between the treatments may need special theoretical development. We cover methods of handling missing data for multi-environment trial such as a complete-case, available case, single imputation,

multiple imputation and missing data patterns. A multiple data imputation with cross validation and principal component analysis methods are an options for missing data handling at plot level. Missing data at a plot level is common on multi environment agricultural trial. A multiple data imputation is necessary for AMMI and GGE analysis which is only applicable for a balanced data case. All the five methods (PCA, BPCA, SVDI, Nipls and NLPCA) show nearest imputation values of missing data observation when we compare each other based on eigenvalue structure and plot for fitted and observed values. A cross validation or PCA methods of imputation have great advantages of easy computationally and non-dependency on the pattern or mechanisms of missing data in experiments which covered on Chapter 2.

We covered the most commonly applied statistical method, the additive main effects and multiplicative interactions model (AMMI) and genotype main effect and genotype by environment interaction (GGE) biplot analysis on chapter 3. Here we see the advantages of this method on classification of the environment in to mega environment classification and which-won-where pattern of genotypes. The problem of repeatability on the performance and stability of genotype on two years repeated over location data analysis leads to a decision making on genotype recommendations and mega–environment delineation should be done with a caution. A selection of the number of multiple interactions and testing of a multiplicative term on the AMMI and GGE model using a parametric bootstrap method for the two data sets on chapter 4 is also one example for a way of handling problems in the selection and test for these models. The different statistical tests results comparisons also supporting the advantage of these methods.

A common procedure for analyzing multi-environmental trials is based on the assumption that the residual error variance is homogenous across all locations considered. However, this may often be unrealistic, and therefore limit the accuracy of variety evaluation or the reliability of variety recommendations.In

chapter 5 and 6 we try to see the advantages of applying some linear trends on the row and column and spatial variance covariance structures to improve the analysis results on the varietal comparisons using a mixed model. A smooth spatial local and global variation are accommodated in the model by including linear terms on the row and column terms to the design and a spatial variance-covariance structure respectively. A correlated mixed model with power, AR1 or AR2 with a random linear trend model result showed an improvement and need for more appropriate and strong spatial variance covariance mixed model to accommodate local trend and its influence. The advantages of mixed models with spatial variance–covariance structures, and direct implications of model choice on the inference of varietal performance, ranking and testing based on two multi-environmental data sets from realistic national trials was shown in chapter 6.

A model comparison with a χ^2-test for the trials in the two data sets (wheat dataset BW00RVTI and barley dataset BW01RVII) suggested that selected spatial variance-covariance structures fitted the data significantly better than the ANOVA model. The forms of optimally-fitted spatial variance-covariance, ranking and test consistence ratio were not the same from one trial (location) to the other. Linear mixed models with single stage analysis including spatial variance-covariance structure with a group factor for location on the random model also improved the real estimation i.e estimation which considers spatial variation, of genotype effect on yield and their ranking. The model also improved varietal performance estimation because of its capacity to handle additional sources of variation, location and genotype by location (environment) interaction variation and accommodating of local stationary trend. A relationship between plot size and covariance structure was developed as a justification to use spatial modelling. It is necessary for national agricultural research system to develop optimal plot size, block size and orientation using the relationship established to improve results from data analysis.

The PCA method and spatial-variance covariance structures mixed model analysis provided some evidence regarding the missing data imputation and handling of spatial variation on the experimental fields and variety ranking. A potential future research area are combination of a design and modelling approach of applying different spatial variance covariance structure for each location on a single stage analysis to obtain a maximum benefit in improving the multi-environment data analysis. Further research can be focussed a multivariate or a covariate approach of AMM or GGE biplot analysis.

References

ALLISON, P. D. 2001. *Missing data*, Sage publications.

ARCINIEGAS-ALARC´ON S. 2008. *Data Imputation in Trials with Genotype by Environment Interaction: An Application on Cotton Data.* University of S˜ao Paulo, .

ARCINIEGAS-ALARCÓN, S. & DIAS, C. T. D. S. 2009. AMMI analysis with imputed data in genotype x environment interaction experiments in cotton. *Pesquisa Agropecuária Brasileira,* 44**,** 1391-1397.

ARCINIEGAS-ALARCÓN, S., GARCÍA-PEÑA, M., DIAS, C. & KRZANOWSKI, W. J. 2010. An alternative methodology for imputing missing data in trials with genotype-by-environment interaction. *Biometrical Letters,* 47**,** 1-14.

ARCINIEGAS-ALARCÓN, S., GARCIA-PENA, M. & DIAS, C. D. S. 2011. Data imputation in trials with genotype × environment interaction. *Interciencia,* 36**,** 444-449.

ATKINSON, A. C., DONEV, A. N. & TOBIAS, R. D. 2007. Optimum experimental designs, with SAS, UK, Oxford Univ. Press.

AWEKE, G. T. 2005. *Using spatial modelling techniques to improve data analysis from agricultural field trials.* PhD, University of Kwa-Zulu Natal.

BARTLETT, M. S. 1978. Nearest Neighbour Models in the Analysis of Field Experiments. *Journal of the Royal Statistical Society. Series B (Methodological),* 40**,** 147-174.

BELLO, A. 1994. A bootstrap method for using imputation techniques for data with missing values. *Biometrical journal,* 36**,** 453-464.

BELLO, A. L. 1993. Choosing among imputation techniques for incomplete multivariate data: a simulation study. *Communications in Statistics-Theory and Methods,* 22**,** 853-877.

BERGAMO, G. C., DIAS, C. T. D. S. & KRZANOWSKI, W. J. 2008. Distribution-free multiple imputation in an interaction matrix through singular value decomposition. *Scientia Agricola,* 65**,** 422-427.

BOZDOGAN, H. 1987. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika,* 52**,** 345-370.

BRADLEY, J. V. 1978. Robustness? *British Journal of Mathematical and Statistical Psychology,* 31**,** 144-152.

BRO, R., KJELDAHL, K., SMILDE, A. & KIERS, H. 2008. Cross-validation of component models: a critical look at current methods. *Analytical and bioanalytical chemistry,* 390**,** 1241-1251.

BROWNIE, C. & GUMPERTZ, M. L. 1997. Validity of Spatial Analyses for Large Field Trials. *Journal of Agricultural, Biological, and Environmental Statistics,* 2**,** 1-23.

BURNHAM, K. P. & ANDERSON, D. R. 2002. *Model selection and multi-model inference: a practical information-theoretic approach*, Springer.

CALINSKI, T., CZAJKA, S., DENIS, J. & KACZMAREK, Z. 1992. EM and ALS algorithms applied to estimation of missing data in series of variety trials. *Biuletyn Oceny Odmian,* 24**,** 25.

CALIŃSKI, T., CZAJKA, S., DENIS, J. & KACZMAREK, Z. 1999. Further study on estimating missing values in series of variety trials. *Biuletyn Oceny Odmian,* 30**,** 7-38.

CALIŃSKI, T., CZAJKA, S., KACZMAREK, Z., KRAJEWSKI, P. & PILARCZYK, W. 2009. A mixed model analysis of variance for multi-environment variety trials. *Statistical Papers,* 50**,** 735-759.

CASANOVES, F., MACCHIAVELLI, R. & BALZARINI, M. 2005. Error Variation in Multienvironment Peanut Trials. *Crop Sci.,* 45**,** 1927-1933.

CASANOVES, F., MACCHIAVELLI, R. & BALZARINI, M. 2013. Models for multi-environment yield trials with fixed and random block effects and homogeneous and heterogeneous residual variances. *Journal of Agriculture of the University of Puerto Rico,* 91.

CHOUKAN, R. 2010. Genotype, environment and genotype by environment interaction effects on the performance of maize (zea mays l.) inbred lines. *Crop Breeding Journal,* **,** 97-103.

CLARKE, F. R. & BAKER, R. J. 1996. Spatial Analysis Improves Precision of Seed Lot Comparisons. *Crop Sci.,* 36**,** 1180-1184.

COCHRAN, W. G. & COX, G. M. 1957. *Experimental Designs,* New York, Wiley Press.

COCKERHAM, C. C. 1963. Estimation of genetic variances. *Statistical genetics and plant breeding,* 982**,** 53-94.

COMSTOCK, R. & MOLL, R. H. 1963. Genotype-environment interactions. *Statistical genetics and plant breeding***,** 164-196.

COOPER, M., DELACY, I. & BASFORD, K. 1996. Relationships among analytical methods used to analyse genotypic adaptation in multi-environment trials. *Plant adaptation and crop improvement***,** 193-224.

CORNELIUS, P. L. 1993. Statistical Tests and Retention of Terms in the Additive Main Effects and Multiplicative Interaction Model for Cultivar Trials. *Crop Sci.,* 33**,** 1186-1193.

CORNELIUS, P. L., CROSSA, J. & SEYEDSADR, M. S. 1996. *Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In Genotype-by-Environment Interaction,* Boca Raton, Florida, CRC Press.

CROSSA, J. 1990. Statistical analyses of multilocation trials. *Advances in agronomy,* 44**,** 55-85.

CSA 2013/14. Report on area and crop production forecast for major grain crops (For private peasant holding, Meher Season. Statistical Bulletins 534. Addis Ababa, Ethiopia.

CULLIS, B. & GLEESON, A. 1991a. Spatial analysis of field experiments-an extension to two dimensions. *Biometrics***,** 1449-1460.

CULLIS, B., GOGEL, B., VERBYLA, A. & THOMPSON, R. 1998a. Spatial Analysis of Multi-Environment Early Generation Variety Trials. *Biometrics,* 54**,** 1-18.

CULLIS, B., THOMSON, F., FISHER, J., GILMOUR, A. & THOMPSON, R. 1996a. The analysis of the NSW wheat variety database. I. Modelling trial error variance. *Theoretical and applied genetics,* 92**,** 21-27.

CULLIS, B., THOMSON, F., FISHER, J., GILMOUR, A. & THOMPSON, R. 1996b. The analysis of the NSW wheat variety database. II. Variance component estimation. *Theoretical and Applied Genetics,* 92**,** 28-39.

CULLIS, B. R. & GLEESON, A. C. 1991b. Spatial Analysis of Field Experiments-An Extension to Two Dimensions. *Biometrics,* 47**,** 1449-1460.

DAVIDOFF, B. & SELIM, H. M. 1988. Correlation between spatially variable soil moisture content and soil temperature. *Soil Science,* 145**,** 1-10.

DEHGHANI, H., EBADI, A. & YOUSEL, A. 2006. Biplot analysis of genotype by envi-ronment interaction for barley yield in iran. *Agronomy Journal,* 98**,** 388-393.

DEHGHANI, H., SABAGHNIA, N., MOGHADDAM, M. & 2009. Interpretation of genotype-by-environment interaction for late maize hybrids' grain yield using a biplot method. *Turk. J. Agric. Fores,* 33**,** 139–148.

DEHGHANI, H., SABAGHPOUR, S. H., EBADI, A. & 2010. Study of genotype × environment interaction for chickpea yield in Iran. *Agron. J. ,* 102.

DEMIRTAS, H. & SCHAFER, J. L. 2003. On the performance of random-coefficient pattern-mixture models for non-ignorable drop-out. *Statistics in medicine,* 22**,** 2553-2575.

DEMPSTER, A. P., LAIRD, N. M. & RUBIN, D. B. 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological),* 39**,** 1-38.

DENIS, J.-B., PIEPHO, H.-P. & EEUWIJK, F. A. V. 1997. Modelling expectation and variance for genotype by environment data. *Heredity,* 79**,** 162-171.

DENIS, J. 1991. Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Revue de statistique appliquée,* 39**,** 5-24.

DENIS, J. & BARIL, C. 1992. Sophisticated models with numerous missing values: the multiplicative interaction model as an example. *Biuletyn Oceny Odmian,* 24**,** 33-45.

DENIS, J. B. & GOWER, J. C. 1994. Biadditive modeles. *Biometrics,* 50**,** 310-311.

DIAS, C. & KRZANOWSKI, W. 2003. Model selection and cross validation in additive main effect and multiplicative interaction models. *Crop Sci,* 43**,** 865-873.

DIAS, C. T. D. S. & KRZANOWSKI, W. J. 2006. Choosing components in the additive main effect and multiplicative interaction (AMMI) models. *Scientia Agricola,* 63**,** 169-175.

DODGE, Y. 1985. Analysis of experiments with missing data. *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1985,* 1.

EARO 1996. Biomerics report. Addis Ababa, Ethiopia.

EASTMENT, H. & KRZANOWSKI, W. 1982a. Cross-validatory choice of the number of components from a principal component analysis. *Technometrics,* 24**,** 73-77.

ECEA 2008. Underestanding Wheat: A review of supply and marketing issues. Research report, Addis Ababa, Ethiopia.

EEUWIJK, F. A. V. & KROONENBERG, P. M. 1998. Multiplicative Models for Interaction in Three-Way ANOVA, with Applications to Plant Breeding. *Biometrics,* 54**,** 1315-1333.

EFRON, B. & TIBSHIRANI, R. 1993. *An Introduction to the Bootstrap,* Boca Raton, Florida, Chapman & Hall/CRC.

FAHRMEIR L & HAMERLE A 1984. *Multivariate Statistische Verfahren,* Berlin, NY., Walter de Gruyter.

FARSHADFAR, E., ZALI, H. & MOHAMMADI, R. 2011. Evaluation of phenotypic stabil-ity in chickpea genotypes using gge-biplot. *Annals of Biological Research,* 2(6)**,** 282-292.

FARSHADFAR E., MOHAMMADI, R., AGHAEE, M. & ZAHRA VAISI, Z. 2012. GGE biplot analysis of genotype × environment interaction in wheat-barley disomic addition lines. *Australian journal of crop science,* 6(6)**,** 1074-1079.

FERRAUDO, G. M. & PERECIN, D. 2014. Mixed Model, AMMI and Eberhart-Russel Comparison via Simulation on Genotype× Environment Interaction Study in Sugarcane. *Applied Mathematics,* 2014.

FINLAY, K. W. & WILKINSON, G. N. 1963. Analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research,* 14**,** 742-&.

FISHER, R. A. 1992. The Arrangement of Field Experiments. *In:* KOTZ, S. & JOHNSON, N. (eds.) *Breakthroughs in Statistics.* Springer New York.

FITZMAURICE, G. M., LAIRD, N. M. & WARE, J. H. 2012. *Applied longitudinal analysis*, John Wiley & Sons.

FORKMAN, J. & PIEPHO, H. P. 2014. Parametric bootstrap methods for testing multiplicative terms in GGE and AMMI models. *Biometrics,* 70**,** 639-647.

FREEMAN, G. H. & PERKINS, J. M. 1971. Environmental and genotype-environmental components of variability. VIII. Relationships between genotypes grown in different environments and measures of these environments. Heredity, 27, 15-23.

FRENSHAM, A., CULLIS, B. & VERBYLA, A. 1997. Genotype by environment variance heterogeneity in a two-stage analysis. *Biometrics***,** 1373-1383.

FRENSHAM, A. B., BARR, A. R., CULLIS, B. R. & PELHAM, S. D. 1998. A mixed model analysis of 10 years of oat evaluation data: use of agronomic information to explain genotype by environment interaction. *Euphytica,* 99**,** 43-56.

GAIN 2012. Ethiopian grain and feed annual report.Grain report number Et-1201.

GARCÍA-PEÑA, M. & DIAS, C. 2009. Analysis of bivariate additive models with multiplicative interaction (AMMI). *Biometric Brazilian Journal,* 27**,** 586-602.

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

GAUCH, H. G. 2006. Statistical Analysis of Yield Trials by AMMI and GGE. *Crop Sci.,* 46**,** 1488-1500.

GAUCH, H. G., JR. 1988. Model Selection and Validation for Yield Trials with Interaction. *Biometrics,* 44**,** 705-715.

GAUCH, H. G., PIEPHO, H.-P. & ANNICCHIARICO, P. 2008a. Statistical Analysis of Yield Trials by AMMI and GGE: Further Considerations. *Crop Sci.,* 48**,** 866-889.

GAUCH, H. G., PIEPHO, H. P. & ANNICCHIARICO, P. 2008b. Statistical analysis of yield trials by AMMI and GGE: Further considerations. *Crop Sci. ,* 48**,** 866–889.

GAUCH, H. G. & ZOBEL, R. W. 1996. *AMMI analysis of yield trials. In M.S. Kang & Gauch, H.G. eds. Genotype-by-environment interaction,* Boca Raton, FL, CRC Press. .

GAUCH JR, H. G. & ZOBEL, R. W. 1990. Imputing missing yield trial data. *Theoretical and Applied Genetics,* 79**,** 753-761.

GIESBRECHT, F. G. & BURNS, J. C. 1985. Two-Stage Analysis Based on a Mixed Model: Large-Sample Asymptotic Theory and Small-Sample Simulation Results. *Biometrics,* 41**,** 477-486.

GILMOUR, A. R., CULLIS, B. R. & VERBYLA, A. P. 1997a. Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological, and Environmental Statistics***,** 269-293.

GLEESON, A. C. 1997. Spatial analysis. *In:* KEMPTON, R. A., FOX, P. N. & CEREZO, M. (eds.) *Statistical Methods for Plant Variety Evaluation.* Springer Netherlands.

GLEESON, A. C. & CULLIS, B. R. 1987. Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. *Biometrics***,** 277-287.

GODFREY, A. 2004. *Dealing with sparsity in genotype x environment analysis.* Dissertation, Massey University.

GODFREY, A., WOOD, G., GANESALINGAM, S., NICHOLS, M. & QIAO, C. 2002. Two-stage clustering in genotype-by-environment analyses with missing data. *The Journal of Agricultural Science,* 139**,** 67-77.

GOLLOB, H. 1968. A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika,* 33**,** 73-115.

GOOD, I. J. 1969. Applications of the singular decomposition of a matrix,". *Technometrics, ,* 11**,** 823–831.

Gower, J.C, Gardner-Lubbe.,and Le Roux, N.J. (2011). *Understanding Biplot.* John Wiley and Sons, Sussex

GREENE WH 2003. *Econometric analysis,* University Prentice Hall, NY, Pearson Educ Int.

GUERIN L & STROUP WW. A simulation study to evaluate PROC MIXED analysis of repeated measures data. . Proc 12th Ann Conf on Applied Statistics in Agriculture. , 2000. Kansas State Univ (Manhattan KS, ed.), 170-203.

HANDCOCK, M. S. 1994. An Approach to Statistical Spatial-Temporal Modeling of Meteorological Fields: Rejoinder. *Journal of the American Statistical Association,* 89**,** 388-390.

HANDCOCK, M. S. & STEIN, M. L. 1993. A Bayesian Analysis of Kriging. *Technometrics,* 35**,** 403-410.

HARTLEY, H. O. & RAO, J. N. 1967. Maximum-likelihood estimation for the mixed analysis of variance model. Biometrika, 54, 93-108.

HARVILLE, D. A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. Journal of the American Statistical Association, 72, 320-338.

HASTIE, T., TIBSHIRANI, R., SHERLOCK, G., EISEN, M., BROWN, P. & BOTSTEIN, D. 1999. Imputing missing data for gene expression arrays. Stanford University Statistics Department Technical report.

HEDDERLEY, D. 1995. A comparison of imputation techniques for internal preference mapping, using Monte Carlo simulation. *Food quality and preference,* 6**,** 281-297.

HEDEKER, D. & GIBBONS, R. D. 1997. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological methods,* 2**,** 64.

HEDEKER, D., MERMELSTEIN, R. J. & DEMIRTAS, H. 2007. Analysis of binary outcomes with missing data: missing= smoking, last observation carried forward, and a little multiple imputation. *Addiction,* 102**,** 1564-1573.

HONG, N., WHITE, J. G., GUMPERTZ, M. L. & WEISZ, R. 2005. Spatial Analysis of Precision Agriculture Treatments in Randomized Complete Blocks This project was funded in part by USDA Initiative for Future Agricultural and Food Systems (IFAFS) grant no. 00-52103-9644. *Agron. J.,* 97**,** 1082-1096.

HRONG-TAI FAI, A. & CORNELIUS, P. L. 1996. Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. *Journal of Statistical Computation and Simulation,* 54**,** 363-378.

JOHNSON, D. E. & GRAYBILL, F. A. 1972. An Analysis of a Two-Way Model with Interaction and No Replication. *Journal of the American Statistical Association,* 67**,** 862-868.

JOSSE, J., PAGÈS, J. & HUSSON, F. 2011. Multiple imputation in principal component analysis. *Advances in data analysis and classification,* 5**,** 231-246.

KACKAR, R. N. & HARVILLE, D. A. 1984. Approximations for Standard Errors of Estimators of Fixed and Random Effect in Mixed Linear Models. *Journal of the American Statistical Association,* 79**,** 853-862.

KAYA, Y., AKCURA, M. & TANER, S. 2006. GGE-biplot analysis of multi-environment yield trials in bread wheat. *Turk J. Agric,* 30**,** 325-337.

KELLY, A. M., SMITH, A. B., ECCLESTON, J. A. & CULLIS, B. R. 2007a. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Science,* 47**,** 1063-1070.

KEMPTON, R. 1984. The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science,* 103**,** 123-135.

KENWARD, M. G. & ROGER, J. H. 1997. Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics,* 53**,** 983-997.

KNIGHT, R. 1970. The measurement and interpretation of genotype-environment interactions. *Euphytica,* 19**,** 225-235.

KRZANOWSKI, W. 1988. Missing value imputation in multivariate data using the singular value decomposition of a matrix. *Biometrical letters,* 25**,** 31-39.

LITTELL, R. C., MILLIKEN, G. A., STROUP, W. W., WOLFINGER, R. D. & SCHABENBERGER, O. 2006. *SAS for Mixed Models,* Cary North Carolina SAS Institute Inc.

LITTLE, R. J. & RUBIN, D. B. 2014. *Statistical analysis with missing data*, John Wiley & Sons.

LIU, G. Z. & CORNELIUS, P. L. 2001. Simulations and derived approximations for the means and standard deviations of the characteristic roots of a

Wishart matrix. *Communications in Statistics-Simulation and Computation,* 30**,** 963-989.

M. S. KANG, M. G. BALZARINI & J. L. L. GUERRA 2004. *"Genotype-byenvironment interaction," in Genetic Analysis of Complex Traits* Cary, NC, USA,, SAS Institute nc.

MANDEL, J. 1961. Non-Additivity in Two-Way Analysis of Variance. *Journal of the American Statistical Association,* 56**,** 878-888.

MANDEL, J. 1971. A New Analysis of Variance Model for Non-additive Data. *Technometrics,* 13**,** 1-18.

MARASINGHE, M. G. 1985. Asymptotic tests and monte carlo studies associated with the multiplicative interaction model. *Communications in Statistics - Theory and Methods,* 14**,** 2219-2231.

MARTIN, R. 1990. The use of time-series models and methods in the analysis of agricultural field trials. *Communications in Statistics-Theory and Methods,* 19**,** 55-81.

MASJKUR, M. 2012. Model Comparison for Multienvironment Trials.

MATÉRN, B. 1986. Spatial Variation. *In:* SECOND (ed.) *Lecture Notes in Statistics.* New York: Springer-Verlag.

MO, H. & SI, Y. 1986. Trend variation and it's control in field experiment *Acta Agronomica Sinnica,* 12(04)**,** 233-240.

MÖHRING, J. & PIEPHO, H.-P. 2009. Comparison of Weighting in Two-Stage Analysis of Plant Breeding Trials. *Crop Sci.,* 49**,** 1977-1988.

MUIRHEAD, R. J. 1978. Latent Roots and Matrix Variates: A Review of Some Asymptotic Results. *The Annals of Statistics***,** 5-33.

MUTHÉN, B. O. & KHOO, S.-T. 1998. Longitudinal studies of achievement growth using latent variable modeling. *Learning and individual differences,* 10**,** 73-101.

NG, M. P. & GRUNWALD, G. K. 1997. Nonlinear Regression Analysis of the Joint-Regression Model. *Biometrics,* 53**,** 1366-1372.

OBA, S., SATO, M.-A., TAKEMASA, I., MONDEN, M., MATSUBARA, K.-I. & ISHII, S. 2003. A Bayesian missing value estimation method for gene expression profile data. *Bioinformatics,* 19**,** 2088-2096.

OMAN, S. D. 1991. Multiplicative Effects in Mixed Model Analysis of Variance. PATTERSON, H. N. & ABUGOOMU, F. REML and the analysis of series of crop variety trials. In Proceedings from the 16th International Biometric Conference, 1992 Hamilton, New Zealand. 77–93.*Biometrika,* 78**,** 729-739.

PATTERSON, H. N. & ABUGOOMU, F. REML and the analysis of series of crop variety trials. In Proceedings from the 16th International Biometric Conference, 1992 Hamilton, New Zealand. 77–93.

PATTERSON, H. & SILVEY, V. 1980. Statutory and recommended list trials of crop varieties in the United Kingdom. *Journal of the Royal Statistical Society. Series A (General)***,** 219-252.

PATTERSON, H., SILVEY, V., TALBOT, M. & WEATHERUP, S. 1977. Variability of yields of cereal varieties in UK trials. *The Journal of Agricultural Science,* 89**,** 239-245.

PATTERSON, H. D. & THOMPSON, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika,* 58**,** 545-554.

PEREIRA, D. G., MEXIA, J. T. & RODRIGUES, P. C. 2007. Robustness of joint regression analysis. *Biometrical Letters,* 44**,** 105-128.

PEUGH, J. L. & ENDERS, C. K. 2004. Missing data in educational research: A review of reporting practices and suggestions for improvement. Review of educational research, 74, 525-556.

PIEPHO, H.-P. 1996. A Simplified Procedure for Comparing the Stability of Cropping Systems. *Biometrics,* 52**,** 315-320.

PIEPHO, H.-P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. *Theoretical and Applied Genetics,* 97**,** 195-201.

PIEPHO, H.-P. 1999. Fitting A Regression Model for Genotype-By-Environment Data on Heading Dates in Grasses by Methods for Nonlinear Mixed Models. *Biometrics,* 55**,** 1120-1128.

PIEPHO, H.-P. & MÖHRING, J. 2006. Selection in cultivar trials—is it ignorable? *Crop Science,* 46**,** 192-201.

PIEPHO, H.-P., RICHTER, C. & WILLIAMS, E. 2008. Nearest Neighbour Adjustment and Linear Variance Models in Plant Breeding Trials. *Biometrical Journal,* 50**,** 164-189.

PIEPHO, H. P. 1995. Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials. *Theoretical and Applied Genetics,* 90**,** 438-443.

PIEPHO, H. P. 1997. Analyzing Genotype-Environment Data by Mixed Models with Multiplicative Terms. *Biometrics,* 53**,** 761-766.

PIEPHO, H. P. 1998a. Methods for Comparing the Yield Stability of Cropping Systems. *Journal of Agronomy and Crop Science,* 180**,** 193-213.

PIEPHO, H. P. 1999a. Cultivar comparisons in three-way data based on mixed models with flexible variance-covariance structure. *Biuletyn Oceny Odmian,* 30**,** 31-48.

PIEPHO, H. P. & MÖHRING, J. 2010. Generation Means Analysis Using Mixed Models. *Crop Sci.,* 50**,** 1674-1680.

RAGHUNATHAN, T. E., LEPKOWSKI, J. M., VAN HOEWYK, J. & SOLENBERGER, P. 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey methodology, 27, 85-96.

RAJU, B. 2002. A study on AMMI model and its biplots. *J. Ind. Soc. Agril. Statist,* 55**,** 297-322.

RAJU, B. & BHATIA, V. 2003. Bias in the estimates of sensitivity from incomplete GxE tables. *J. Ind. Soc. Agril. Statist,* 56**,** 294-30.

RAJU, B., BHATIA, V. & BHAR, L. 2009. Assessing stability of crop varieties with incomplete data. *Journal of the Indian Society of Agricultural Statistics,* 63**,** 139-149.

RAJU, B., BHATIA, V. & KUMAR, V. 2006. Assessment of sensitivity with incomplete data. *Journal of the Indian Society of Agricultural Statistics,* 60**,** 118-125.

RAO, A. R. & POLIGNANO PRABHAKARAN, V. T. 2005. Use of AMMI in simultaneous selection of genotypes for yield and stability. . *Ind Soc Agril Statist,* 59**,** 76–82.

ROBERTSON, A. 1959. The sampling variance of the genetic correlation coefficient. *Biometrics,* 15**,** 469-485.

RODRIGUES, P. C., PEREIRA, D. G. S. & MEXIA, J. T. 2011. A comparison between joint regression analysis and the additive main and multiplicative interaction model: the robustness with increasing amounts of missing data. *Scientia Agricola,* 68**,** 679-686.

RODRIGUES, P. J. C. 2012. *New Strategies to Detect and Understand Genotype-by-Environment Interactions and QTL-by-Environment Interactions. PhD* Universidade Nova de Lisboa.

ROMAGOSA, I., VOLTAS, J., MALOSETTI, M. & VAN EEUWIJK, F. 2008. Interacción genótipo por ambiente. *La adaptación al ambiente y los estreses abióticos en la mejora vegetal.* Junta de Andalucia.

RUBIN, D. B. 1987. Multiple Imputation for Nonresponse in Surveys, New York, J. Wiley & Sons.

SABAGHNIA, N., MOHAMMADI, M. & KARIMIZADEH, R. 2012b. The evaluation of genotype × environment interactions of durum wheat's yield using of the AMMI model. *Agric. Fores. ,* 55**,** 5–21.

SATTERTHWAITE, F. 1941. Synthesis of variance. *Psychometrika,* 6**,** 309-316.

SCHAALJE, G. B., MCBRIDE, J. & FELLINGHAM, G. 2002. Adequacy of approximations to distributions of test statistics in complex mixed linear models. *Journal of Agricultural, Biological, and Environmental Statistics,* 7**,** 512-524.

SCHABENBERGER & PIERCE 2002. *Contemporary statistical models for the plant and soil sciences. ,* Boca Raton, CRC Press.

SCHARF, P. C. & ALLEY, M. M. 1993. Accounting for Spatial Yield Variability in Field Experiments Increases Statistical Power. *Agronomy Journal,* 85**,** 1254-1256.

SCHOTT, J. R. 1986. A note on the critical values used in stepwise tests for multiplicative components of interaction. *Communications in Statistics - Theory and Methods,* 15**,** 1561-1570.

SCHWARZBACH, E. 1984. A new approach in the evaluation of field trials. The determination of the most likely genetic ranking of varieties. Vortrage Pflanzen, 249-259.

SCHWARZ, G. 1978. Estimating the Dimension of a Model. 461-464.

SEARLE, S. R., CASELLA, G. & MCCULLOCH, C. E. 2008. Frontmatter. *Variance Components.* John Wiley & Sons, Inc.

SEARLE, S. R., CASELLA, G., MCCULLOCH, C. E. & 1992. *Variance Components,* Hoboken, New Jersey, Wiley.

SHIFERAW, B., KASSIE, M., JALETA, M. & YIRGA, C. 2014. Adoption of improved wheat varieties and impacts on household food security in Ethiopia. *Food Policy,* 44**,** 272-284.

SMILDE A., BRO R., P. & GELADI P 2004. *Multi-Way Analysis with Applications in the Chemical Sciences*, John Wiley and Sons, Chichester, UK,.

SMITH, A., CULLIS, B. & GILMOUR, A. 2001a. Applications: The Analysis of Crop Variety Evaluation Data in Australia. *Australian & New Zealand Journal of Statistics,* 43**,** 129-145.

SMITH, A., CULLIS, B. & THOMPSON, R. 2001b. Analyzing Variety by Environment Data Using Multiplicative Mixed Models and Adjustments for Spatial Field Trend. *Biometrics,* 57**,** 1138-1147.

SMITH, A., CULLIS, B. R. & THOMPSON, R. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science,* 143**,** 449-462.

SO, Y.-S. & EDWARDS, J. 2009. A comparison of mixed-model analyses of the Iowa Crop Performance Test for Corn. *Crop science,* 49**,** 1593-1601.

SPILKE, J., PIEPHO, H.-P. & HU, X. 2005. A Simulation Study on Tests of Hypotheses and Confidence Intervals for Fixed Effects in Mixed Models for Blocked Experiments with Missing Data. *Journal of Agricultural, Biological, and Environmental Statistics,* 10**,** 374-389.

SPILKE, J., PIEPHO, H. P. & MEYER, U. 2004. Approximating the degrees of freedom for contrasts of genotypes laid out as subplots in an alpha-SPRENT, P. & SMEETON, N. C. 2001. Applied Nonparametric Statistical Methods, Boca Raton, Florida, CRC Press LLC.design in a split-plot experiment. *Plant Breeding,* 123**,** 193-197.

STACKLIES, W., REDESTIG, H., SCHOLZ, M., WALTHER, D. & SELBIG, J. 2007. pcaMethods—a bioconductor package providing PCA methods for incomplete data. *Bioinformatics,* 23**,** 1164-1167.

STEFANOVA, K. T. & BUIRCHELL, B. 2010. Multiplicative Mixed Models for Genetic Gain Assessment in Lupin Breeding. *Crop Sci.,* 50**,** 880-891.

STROUP, W. W. 2002. Power analysis based on spatial effects mixed models: A tool for comparing design and analysis strategies in the presence of spatial variability. *Journal of Agricultural, Biological, and Environmental Statistics,* 7**,** 491-511.

TALBOT, M. 1984. Yield variability of crop varieties in the UK. *The Journal of Agricultural Science,* 102**,** 315-321.

TAYE, G., GETACHEW, T. & BEJIGA, G. 2000. AMMI adjustment for yield estimate and classification of genotypes and environments in field pea (Pisum sativum L.). *Journal of Genetics & Breeding,* 54**,** 183-191.

TROYANSKAYA, O., CANTOR, M., SHERLOCK, G., BROWN, P., HASTIE, T., TIBSHIRANI, R., BOTSTEIN, D. & ALTMAN, R. B. 2012. Missing value estimation methods for DNA microarrays. *Bioinformatics,* 17**,** 520-525.

TUKAMUHABWA, P., ASIIMWE, M., NABASIRYE, M., KABAYL, P. & MAPHOSA, M. 2012. Genotype by environment interaction of advanced generation soybean lines for grain yield in Uganda. *African Crop Science Journal,* 20(2)**,** 107-115.

TUKEY, J. W. 1949. One Degree of Freedom for Non-Additivity. *Biometrics,* 5**,** 232-242.

VAN BUUREN, S. & OUDSHOORN, C. 2000. Multivariate imputation by chained equations. MICE V1. 0 user's manual. Leiden: TNO Preventie en Gezondheid.

VAN EEUWIJK, F. A. & KROONENBERG, P. M. 1998. Multiplicative models for interaction in three-way ANOVA, with applications to plant breeding. *Biometrics***,** 1315-1333.

VAN EEUWIJK, F. A., MALOSETTI, M. & BOER, M. P. 2007. Modelling the genetic basis of response curves underlying genotype x environment interaction. *Frontis,* 21**,** 113-124.

VAN EEUWIJK, F. A., MALOSETTI, M., YIN, X., STRUIK, P. C. & STAM, P. 2005. Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. *Crop and Pasture Science,* 56**,** 883-894.

WARREN, J. & MENDEZ, I. 1982. Methods for estimating background variation in field experiments. *Agronomy Journal,* 74**,** 1004-1009.

WASITO, I. 2003. *Least Squares Algorithms with Nearest Neighbour Techniques for Imputing Missing Data Values.* Birkbeck College.

WELHAM, S. J., GOGEL, B. J., SMITH, A. B., THOMPSON, R. & CULLIS, B. R. 2010b. A comparison of anlysis methods for late-stage variety evaluation trials. *Australian & New Zealand Journal of Statistics,* 52**,** 125-149.

WILKINSON, G., ECKERT, S., HANCOCK, T. & MAYO, O. 1983a. Nearest neighbour (NN) analysis of field experiments. *Journal of the Royal Statistical Society. Series B (Methodological)***,** 151-211.

WILLIAMS, E. R. 1986. A neighbour model for field experiments. *Biometrika,* 73**,** 279-287.

WOLD, S. 1978. Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models. *Technometrics,* 20**,** 397-405.

WOLFINGER, R. 1993. Covariance structure selection in general mixed models. *Communications in Statistics - Simulation and Computation,* 22**,** 1079-1106.

WU, T. & DUTILLEUL, P. 1999. Validity and Efficiency of Neighbor Analyses in Comparison with Classical Complete and Incomplete Block Analyses of Field Experiments. *Agronomy Journal,* 91**,** 721-731.

YAN, W. 1999. *Methodology of cultivar evaluation based on yield trial data-with special reference to winter wheat in Ontario.* PhD thesis. .

YAN, W. 2002. Singular-value partitioning in biplot analysis of multienvironment trial data. *Agronomy Journal,* 94**,** 990-996.

YAN, W., CORNELIUS, P. L., CROSSA, J. & HUNT, L. A. C. S. 2001. Two types of gge biplots for analyzing multi-environment trial data. . *40***,** 656-663.

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. & KANG, M. S. 2002. *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists.* Boca Raton,Florida: , CRC Press.

YAN, W. & KANG, M. S. 2003. *GGE biplot analysis : a graphical tool for breeders, geneticists, and agronomists.* Boca Raton, Flaorida, CRC Press.

YAN, W., KANG, M. S., MA, B., WOODS, S. & CORNELIUS, P. L. 2007. GGE Biplot vs. AMMI Analysis of Genotype-by-Environment Data. *Crop Sci.,* 47**,** 643-653.

YAN, W. & RAJCAN, I. 2002. Biplot analysis of test sites and trait relations of soybean in ontario. *Crop Science,* 42.

YAN, W. & TINKER, N. A. 2006a. Biplot analysis of multi-environment trial data: principles and application. *Candian Jornal of plant science,,* 86**,** 623-645.

YANG, R.-C., CROSSA, J., CORNELIUS, P. L. & BURGUEÑO, J. 2009a. Biplot Analysis of Genotype × Environment Interaction: Proceed with Caution. *Crop Sci.,* 49**,** 1564-1576.

YANG, R.-C., YE, T. Z., BLADE, S. F. & BANDARA, M. 2004. Efficiency of Spatial Analyses of Field Pea Variety Trials. *Crop Science,* 44**,** 49-55.

YATES, F. & COCHRAN, W. G. 1938. The analysis of groups of experiments. *The Journal of Agricultural Science,* 28**,** 556-580.

YOCHMOWITZ, M. G. & CORNELL, R. G. 1978. Stepwise Tests for Multiplicative Components of Interaction. *Technometrics,* 20**,** 79-84.

YUAN, Y. C. Multiple Imputation for Missing Data: Concepts and New Development. Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference, 2000. SAS Institute, Cary NC, 267-25.

ZIMMERMAN, D. L. & HARVILLE, D. A. 1991a. A Random Field Approach to the Analysis of Field-Plot Experiments and Other Spatial Experiments. *Biometrics,* 47**,** 223-239.

Appendix A

Table A 1 Results from the application of the linear mixed correlation models with four correlation structures to the BW01RVII data set in Location 2.

Model for location2	Correlation structures	AIC	Fixed term	df	F-value	P-value
Response yield Fixed model: Constant +entry Random model: row + column +row.column	Identity	330.60	Entry(genotypes)	24	2.83	<0.001
	Power	333.30	Entry(genotypes)	24	2.84	<0.001
	auto-regressive order 1 (AR1)	333.30	Entry(genotypes)	24	2.84	<0.001
	auto-regressive order 2 (AR2)	331.46	Entry(genotypes)	24	5.14	<0.001
Response yield Fixed model: Constant + entry + lin_row + lin_col Random model: row.column	Identity	331.71	Entry(genotypes)	24	2.83	<0.001
			lin_row	1	0.89	0.348
			lin_col	1	1.52	0.222
	Power	334.52	Entry(genotypes)	24	2.83	<0.001
			lin_row	1	0.71	0.414
			lin_col	1	0.95	0.343
	auto-regressive order 1 (AR1)	334.52	Entry(genotypes)	24	2.84	<0.001
			lin_row	1	0.71	0.414
			lin_col	1	0.95	0.343
	auto-regressive order 2 (AR2)	335.48	Entry(genotypes)	24	3.18	<0.001
			lin_row	1	2.38	0.150
			lin_col	1	2.27	0.147
Response :yield Fixed model: Constant + entry + lin_row + lin_col Random model: row + column + row.column	identity	335.70	Entry(genotypes)	24	2.81	<.001
			lin_row	1	0.90	0.346
			lin_col	1	1.41	0.248
	Power	338.52	Entry(genotypes)	24	2.83	<0.001
			lin_row	1	0.71	0.414
			lin_col	1	0.95	0.343
	auto-regressive order 1 (AR1)	338.52	Entry(genotypes)	24	2.83	<0.001
			lin_row	1	0.71	0.414
			lin_col	1	0.95	0.343
	auto-regressive order 2 (AR2)	334.75	Entry(genotypes)	24	5.29	<0.001
			lin_row	1	2.45	0.232
			lin_col	1	2.20	0.155

Table A 2 Results from the application of the linear mixed correlation models with four correlation structures to the BW01RVII data set in Location 3.

Model for location3	Correlation structures	AIC	Fixed term	df	F-value	P-value
Response yield Fixed model: Constant +entry Random model: row + column +row.column	identity	297.85	Entry(genotypes)	24	4.20	<0.001
	Power	296.95	Entry(genotypes)	24	4.62	<0.001
	auto-regressive order 1 (AR1)	296.50	Entry(genotypes)	24	4.62	<0.001
	auto-regressive order 2 (AR2)	296.60	Entry(genotypes)	24	5.11	<0.001
Response yield Fixed model: Constant + entry + lin_row + lin_col Random model: row.column	identity	304.24	Entry(genotypes)	24	3.48	<0.001
			lin_row	1	4.47	0.038
			lin_col	1	0.00	0.984
	Power	298.88	Entry(genotypes)	24	4.35	<0.001
			lin_row	1	2.21	0.153
			lin_col	1	0.03	0.865
	auto-regressive order 1 (AR1)	298.88	Entry(genotypes)	24	4.35	<0.001
			lin_row	1	2.21	0.153
			lin_col	1	0.03	0.865
	auto-regressive order 2 (AR2)	298.15	Entry(genotypes)	24	5.10	<0.001
			lin_row	1	0.82	0.380
			lin_col	1	0.06	0.809
Response :yield Fixed model: Constant + entry + lin_row + lin_col Random model: row + column + row.column	identity	303.04	Entry(genotypes)	24	4.18	<0.001
			lin_row	1	1.63	0.330
			lin_col	1	0.00	0.967
	Power	302.04	Entry(genotypes)	24	4.60	<0.001
			lin_row	1	1.40	0.325
			lin_col	1	0.02	0.877
	auto-regressive order 1 (AR1)	302.04	Entry(genotypes)	24	4.60	<0.001
			lin_row	1	1.40	0.325
			lin_col	1	0.02	0.877
	auto-regressive order 2 (AR2)	302.16	Entry(genotypes)	24	5.02	<0.001
			lin_row	1	0.81	0.416
			lin_col	1	0.08	0.787

*(ANOVA model Entry(genotypes) F-value 3.73 <.001)

Table A 3 Results from the application of the linear mixed correlation models with four correlation structures to the BW01RVII data set in Location 4.

Model for location4	Correlation structures	AIC	Fixed term	df	F-value	P-value
Response yield Fixed model: Constant +entry Random model: row + column +row.column	identity	392.66	Entry(genotypes)	24	2.24	0.005
	Power	387.75	Entry(genotypes)	24	1.53	0.102
	auto-regressive order 1 (AR1)	387.47	Entry(genotypes)	24	1.50	0.121
	auto-regressive order 2 (AR2)	391.23	Entry(genotypes)	24	1.51	0.157
Response yield Fixed model: Constant + entry + lin_row + lin_col Random model: row.column	identity	401.36	Entry(genotypes)	24	1.76	0.035
			lin_row	1	5.01	0.028
			lin_col	1	1.58	0.212
	Power	384.92	Entry(genotypes)	24	1.52	0.099
			lin_row	1	1.01	0.343
			lin_col	1	1.73	0.211
	auto-regressive order 1 (AR1)	384.13	Entry(genotypes)	24	1.50	0.119
			lin_row	1	0.95	0.359
			lin_col	1	2.22	0.165
	auto-regressive order 2 (AR2)	388.03	Entry(genotypes)	24	1.43	0.182
			lin_row	1	0.88	0.383
			lin_col	1	2.32	0.167
Response :yield Fixed model: Constant + entry + lin_row + lin_col Random model: row + column + row.column	identity	393.70	Entry(genotypes)	24	2.27	0.004
			lin_row	1	0.62	0.514
			lin_col	1	1.98	0.164
	Power	387.73	Entry(genotypes)	24	1.52	0.104
			lin_row	1	0.50	0.553
			lin_col	1	1.87	0.198
	auto-regressive order 1 (AR1)	387.20	Entry(genotypes)	24	1.49	0.126
			lin_row	1	0.50	0.560
			lin_col	1	2.29	0.161
	auto-regressive order 2 (AR2)		Entry(genotypes)	24	1.45	0.176
			lin_row	1	0.49	0.564
			lin_col	1	2.20	0.177

Table A 4 Results from the application of the linear mixed correlation models with four correlation structures to the BW01RVII data set in Location 5.

Model for location5	Correlation structures	AIC	Fixed term	df	F-value	P-value
Response yield Fixed model: Constant +entry Random model: row + column +row.column	identity	409.20	Entry(genotypes)	24	1.74	0.038
	Power	409.93	Entry(genotypes)	24	1.95	0.020
	auto-regressive order 1 (AR1)	408.41	Entry(genotypes)	24	2.20	0.011
	auto-regressive order 2 (AR2)	409.08	Entry(genotypes)	24	2.42	0.006
Response yield Fixed model: Constant + entry + lin_row + lin_col Random model: row.column	identity	417.45	Entry(genotypes)	24	1.43	0.123
			lin_row	1	0.41	0.522
			lin_col	1	0.09	0.769
	Power	411.02	Entry(genotypes)	24	2.14	0.010
			lin_row	1	0.25	0.622
			lin_col	1	0.37	0.547
	auto-regressive order 1 (AR1)	409.95	Entry(genotypes)	24	2.29	0.008
			lin_row	1	0.34	0.570
			lin_col	1	0.36	0.559
	auto-regressive order 2 (AR2)	408.63	Entry(genotypes)	24	2.48	0.005
			lin_row	1	0.28	0.616
			lin_col	1	0.11	0.754
Response :yield Fixed model: Constant + entry + lin_row + lin_col Random model: row + column + row.column	identity	412.52	Entry(genotypes)	24	1.71	0.042
			lin_row	1	0.06	0.829
			lin_col	1	0.10	0.748
	Power	412.52	Entry(genotypes)	24	2.00	0.017
			lin_row	1	0.07	0.815
			lin_col	1	0.33	0.573
	auto-regressive order 1 (AR1)	411.21	Entry(genotypes)	24	2.30	0.008
			lin_row	1	0.08	0.804
			lin_col	1	0.31	0.583
	auto-regressive order 2 (AR2)	411.85	Entry(genotypes)	24	2.45	0.005
			lin_row	1	0.11	0.775
			lin_col	1	0.12	0.742

*(ANOVA model Entry(genotypes) F-value 1.74, P-value 0.038)

Appendix B Codes used in the study

B.1 Data Imputation using the cross-validation by Eigenvectors method

```
#########################################################################
#####################################

#
#

#       Data Imputation Using the Cross-Validation by Eigenvector
Method

#
#

#
            #

#########################################################################
#####################################


######Importing a matrix with missing values

X<-matrix(scan('C:/SAA/Eigenvector(2008)_Pesquisa/MissingValues.txt'),
ncol=5, byrow=T)

X # Data Matrix

is.matrix(X)

nfilasX<-nrow(X)

ncolX<-ncol(X)

nfilasX

ncolX

totalelementos<-nfilasX*ncolX
```

```
#######Creating a vector of missing positions and replacing the
missing values by the mean of each column


XMISSING=X

indicamissing<-is.na(XMISSING)

indicadora<-indicamissing*1

totalfaltantes<-sum(indicadora)

medcol<-t(matrix(colMeans(XMISSING, na.rm=TRUE)))

incompleta<-XMISSING

XMISSING

posicfaltantes<-matrix(0,totalfaltantes,2)

indi3<-0

for (indi1 in 1:nfilasX){

   for(indi2 in 1:ncolX){

      if (indicadora[indi1,indi2]==1){

      indi3<-indi3+1

      posicfaltantes[indi3,1]=indi1

      posicfaltantes[indi3,2]=indi2

      XMISSING[indi1,indi2]=medcol[1,indi2]

      } #FIN DEL if (indicadora[indi1,indi2]

   }# FIN DEL for(indi2 in 1:ncolX)

} # FIN DEL for (indi1 in 1:nfilasX)


posicfaltantes

XMISSING
```

```
completed_matrix<-XMISSING
maxiter<-50
iter<-0
desvest<-sqrt(t(diag(var(XMISSING))))
#print(desvest)
Xmissingestand<-scale(XMISSING)
epsilon<-1*(10**(-6))
stabilitycrit<-1
stabilityiter<-0

######Imputation by eigenvector

while (stabilitycrit>epsilon & stabilityiter<=500){


for (i in 1:totalfaltantes){
```

B.2 AMMI and GEE codes

```
OPTIONS PS = 5000 LS=78 NODATE;
FILENAME BIPLOT 'EXAMPLE1.CGM';
GOPTIONS DEVICE=CGMMWWC GSFNAME=BIPLOT GSFMODE=REPLACE;
DATA RAW;
   INFILE 'c:\mydata\amii\EXAMPLE1.DAT';
   INPUT ENV $ GEN $ YIELD;
   YLD=YIELD/1000;
PROC GLM DATA=RAW OUTSTAT=STATS ;
```

```
   CLASS ENV GEN;
   MODEL YLD = ENV GEN ENV*GEN/SS4;
DATA STATS2;
SET STATS ;
DROP _NAME_ _TYPE_;
   IF _SOURCE_ = 'ERROR' THEN DELETE;
MSE=0.1580245;
DFE=94;
NREP=3;
SS=SS*NREP;
MS=SS/DF;
F=MS/MSE;
PROB=1-PROBF(F,DF,DFE);
PROC PRINT DATA=STATS2 NOOBS;
   VAR _SOURCE_ DF SS MS F PROB;
PROC GLM DATA=RAW NOPRINT;
   CLASS ENV GEN;
   MODEL YLD = ENV GEN / SS4 ;
   OUTPUT OUT=OUTRES R=RESID;
PROC SORT DATA=OUTRES;
   BY GEN ENV;
PROC TRANSPOSE DATA=OUTRES OUT=OUTRES2;
   BY GEN;
   ID ENV;
   VAR RESID;
PROC IML;
```

```
USE OUTRES2;
READ ALL INTO RESID;
NGEN=NROW(RESID);
NENV=NCOL(RESID);
USE STATS2;
READ VAR {MSE} INTO MSEM;
READ VAR {DFE} INTO DFEM;
READ VAR {NREP} INTO NREP;
CALL SVD (U,L,V,RESID);
MINIMO=MIN(NGEN,NENV);
L=L[1:MINIMO,];
SS=(L##2)*NREP;
SUMA=SUM(SS);
PORCENT=((1/SUMA)#SS)*100;
MINIMO=MIN(NGEN,NENV);
PORCENTA=0;
   DO I = 1 TO MINIMO;
      DF=(NGEN-1)+(NENV-1)-(2*I-1);
      DFA=DFA//DF;
      PORCEACU=PORCENT[I,];
      PORCENTA=PORCENTA+PORCEACU;
      PORCENAC=PORCENAC//PORCENTA;
   END;
DFE=J(MINIMO,1,DFEM);
MSE=J(MINIMO,1,MSEM);
SSDF=SS||PORCENT||PORCENAC||DFA||DFE||MSE;
```

```
L12=L##0.5;
SCOREG1=U[,1]#L12[1,];
SCOREG2=U[,2]#L12[2,];
SCOREG3=U[,3]#L12[3,];
SCOREE1=V[,1]#L12[1,];
SCOREE2=V[,2]#L12[2,];
SCOREE3=V[,3]#L12[3,];
SCOREG=SCOREG1||SCOREG2||SCOREG3;
SCOREE=SCOREE1||SCOREE2||SCOREE3;
SCORES=SCOREG//SCOREE;
CREATE SUMAS FROM SSDF;
APPEND FROM SSDF;
CLOSE SUMAS;
CREATE SCORES FROM SCORES;
APPEND FROM SCORES ;
CLOSE SCORES;
/* obtaining the polygon and its perpendiculars */
d1=scoreg[,1:2][cvexhull(scoreg[,1:2])[loc(cvexhull(scoreg[,1:2])>0),],];
d=d1//d1[1,];
xxx=J(nrow(d)-1,1,0);
yyy=J(nrow(d)-1,1,0);
ppp={0 1,1 0};
   do i=1 to nrow(d)-1 ;
      dd=d[i:i+1,];
         if dd[1,1]>dd[2,1] then ddd=ppp*dd;
         else ddd=dd;
```

```
      p=(ddd[2,2]-ddd[1,2])/(ddd[2,1]-ddd[1,1]) ;
         if p<0 then ss=1 ;
         else ss=-1 ;
      r=tan((180-90-abs(atan(p)*180/3.14156))*3.14156/180)*ss ;
      aa=(ddd[1,2]+ddd[2,2])/2-p*(ddd[1,1]+ddd[2,1])/2;
      xx=aa/(r-p) ;
         if abs(r)<1 then xxx[i,]=1;
         else xxx[i,]=1/abs(r);
            if xx<0 then xxx[i,]=-xxx[i,] ;
            else xxx[i,]=xxx[i,];
      yyy[i,]=xxx[i,]*r;
   end;
kk=xxx||yyy;
xx1={V1 V2};
create pol from d[colNAME=xx1];
append from d ;
close pol;
xx2={V3 V4};
create perp from kk[colNAME=xx2];
append from kk ;
close perp;
data pol; set pol; TYPE="pol";
data perp; set perp; TYPE="per";

DATA SSAMMI;
SET SUMAS;
```

```
SSAMMI =COL1;
PORCENT =COL2;
PORCENAC=COL3;
DFAMMI =COL4;
DFE =COL5;
MSE =COL6;
DROP COL1 - COL6;
MSAMMI=SSAMMI/DFAMMI;
F_AMMI=MSAMMI/MSE;
PROBF=1-PROBF(F_AMMI,DFAMMI,DFE);
PROC PRINT DATA=SSAMMI NOOBS;
   VAR SSAMMI PORCENT PORCENAC DFAMMI MSAMMI F_AMMI PROBF;
PROC SORT DATA=RAW;
   BY GEN;
PROC MEANS DATA = RAW NOPRINT;
   BY GEN ;
   VAR YLD;
   OUTPUT OUT = MEDIAG MEAN=YLD;
DATA NAMEG;
   SET MEDIAG;
   TYPE = 'GEN';
   NAME = GEN;
   KEEP TYPE NAME YLD;
PROC SORT DATA=RAW;
   BY ENV;
PROC MEANS DATA = RAW NOPRINT;
```

```
   BY ENV ;
   VAR YLD;
   OUTPUT OUT = MEDIAE MEAN=YLD;
DATA NAMEE;
   SET MEDIAE;
   TYPE = 'ENV';
   NAME1 = 'S'||ENV;
   NAME = COMPRESS(NAME1);
   KEEP TYPE NAME YLD;
DATA NAMETYPE;
   SET NAMEG NAMEE;
DATA BIPLOT0 ;
   MERGE NAMETYPE SCORES;
   DIM1=COL1;
   DIM2=COL2;
   DIM3=COL3;
   DROP COL1-COL3;
data biplot ;
   set biplot0 pol perp;
PROC PRINT DATA=BIPLOT NOOBS;
   VAR TYPE NAME YLD DIM1 DIM2 DIM3;
Data labels;
   set biplot ;
   retain xsys '2' ysys '2' ;
   length function text $8 ;
   text = name ;
```

```
if type = 'GEN' then do;
   color='black ';
   size = 0.6;
   style = 'hwcgm001';
   x = dim1;
   y = dim2;
      if dim1 >=0
         then position='5';
      else position='5';
   function = 'LABEL';
   output;
end;
if type = 'ENV' then DO;
   color='black ';
   size = 0.6;
   style = 'hwcgm001';
   x = 0.0;
   y = 0.0;
   function='MOVE';
   output;
   x = dim1;
   y = dim2;
   function='DRAW' ;
   output;
      if dim1 >=0
         then position='5';
```

```
            else position='5';
         function='LABEL';
         output;
         end;
      if type = "per" then do;
         color='red';
         line=2;
         size = 0.6;
         style = 'hwcgm001';
         x=0.0;
         y=0.0;
         function='MOVE';
         output;
         x=v3;
         y=v4;
         function='DRAW';
         output;
      end;
Proc gplot data=biplot;
Plot dim2*dim1 v2*v1 / overlay Annotate=labels frame
   Vref=0.0 Href = 0.0
   cvref=black chref=black
   lvref=3 lhref=3
   vaxis=axis2 haxis=axis1
   vminor=1 hminor=1 nolegend;
   symbol1 v=none c=black h=0.7 ;
```

```
    symbol2 v=none c=blue i=j line=3 ;
    axis2
       length = 6.0 in
       order = (-1.0 to 1.0 by 0.2)
       label=(f=hwcgm001 h=1.2 a=90 r=0 'Factor 2')
       value=(h=0.8)
       minor=none;
    axis1
       length = 6.0 in
       order = (-1.0 to 1.0 by 0.2)
       label=(f=hwcgm001 h=1.2 'Factor 1')
       value=(h=0.8)
       minor=none;
Title1 f=hwcgm001 h=1.0 'AMMI biplot for Example 1 using adjusted
means';
run;
```

Appendix C Published Papers

Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)
Available online at www.inia.es/sjar
http://dx.doi.org/10.5424/sjar/2014121-4926

Spanish Journal of Agricultural Research 2014 12(1): 195-205
ISSN: 1695-971-X
eISSN: 2171-9292

Mixed model with spatial variance-covariance structure for accommodating of local stationary trend and its influence on multi-environmental crop variety trial assessment

Asnake Worku Negash[1]*, Henry Mwambi[1], Temesgen Zewotir[1] and Girma Aweke[2]

[1] School of Mathematics, Statistics and Computer Science. University of KwaZulu-Natal. Pietermaritzburg, South Africa. [2] CIMMYT. Addis Ababa. Ethiopia

Abstract

The most common procedure for analyzing multi-environmental trials is based on the assumption that the residual error variance is homogenous across all locations considered. However, this may often be unrealistic, and therefore limit the accuracy of variety evaluation or the reliability of variety recommendations. The objectives of this study were to show the advantages of mixed models with spatial variance-covariance structures, and direct implications of model choice on the inference of varietal performance, ranking and testing based on two multi-environmental data sets from realistic national trials. A model comparison with a χ^2-test for the trials in the two data sets (wheat data set BW00RVTI and barley data set BW01RVII) suggested that selected spatial variance-covariance structures fitted the data significantly better than the ANOVA model. The forms of optimally-fitted spatial variance-covariance, ranking and consistency ratio test were not the same from one trial (location) to the other. Linear mixed models with single stage analysis including spatial variance-covariance structure with a group factor of location on the random model also improved the real estimation of genotype effect and their ranking. The model also improved varietal performance estimation because of its capacity to handle additional sources of variation, location and genotype by location (environment) interaction variation and accommodating of local stationary trend.

Additional key words: multi-environmental trials; chi-squared test; spatial variance-covariance; consistency ratio test; wheat; barley.

Introduction

National multi-environmental yield trials (MET), allow assessment of the potential yield performance of different varieties across a range of environments (locations and possibly over years, as well as combination of the two). These trials play an important part in crop variety evaluation in breeding programs and varietal recommendations for plant production. It is therefore vital that the statistical methods used to design the studies and analyse data from national yield trial evaluation programs are as accurate, efficient and informative as possible. Although the development of statistical methods for analysing variety trial data has a long history, due to the complexity of varietal and environmental interactions there is no specific model that is generally suitable for analysing combined data sets from national trials. Spatial variability often exists in field experiments due to factors such as moisture, fertility, pH and structure of the soil, as well as the pressure of diseases and pests (Davidoff & Selim, 1988; Scharf & Alley, 1993; Wu & Dutilleul, 1999; Stroup, 2002). Multi-environment crop variety trials and field evaluations are a particularly well-known example of this. Failure to effectively control for spatial variability greatly increases the risk of misleading interpretations or erroneous inferences (Mo & Si, 1986; Stroup, 2002; Yang *et al*., 2004).

* Corresponding author: 208517201@stu.ukzn.ac.za
Received: 30-08-13. Accepted: 19-02-14

Abbreviations used: AIC (Akaike's information criteria); AR (autoregressive); BIC (Bayesian information criterion); CS (compound symmetry); d.f. (degrees of freedom); LL (log-likelihood); LRT (likelihood-ratio test); MET (multi-environmental trial); MVN (multivariate normal); NNA (nearest neighbour adjustment); RCBD (randomized complete block design); REML (restricted maximum likelihood); SLMM (spatial liner mixed model).

Historically, the analysis of variance (ANOVA), along with randomised block designs (including complete, incomplete blocks), has been used to deal with the spatial variability of these trials. Numerous studies have shown that such design-based control of the spatial variation of field trials are often not optimal and results in poor analysis efficiency (Yang *et al*., 2004). Statistical procedures that account for spatial variation between plots within trials have been proposed to address the topic of modelling spatial variation in crop evaluation trials using polynomial trend analysis, nearest neighbour analysis and a model with correlated errors.

The problem with the ANOVA method as a means to analyse multi-environmental crop variety trials is that it requires the assumption of homogenous variance-covariance structures across locations or environments. This homogeneity of variance and covariance may be unrealistic in many circumstances (Kempton, 1984; Piepho, 1999a). As a result, a range of more complex and informative models that can account for variance or/and covariance heterogeneity have been proposed for analysing MET data. While other models are available, the problem of how the models should be assessed and which model is more suitable for a given trial's data has not been solved. This restricts the applicability of the models and model choice. Therefore, a linear mixed model approach with flexible spatial variance-covariance structures is proposed. Correspondingly, model-based approaches for analysing field trials that focus on the need to control spatial variation have been put forward. These approaches include nearest neighbour adjustment (NNA) analysis and its modifications (Bartlett, 1978; Cullis & Gleeson, 1991; Clarke & Baker, 1996; Yang *et al*., 2004). Other options include linear mixed models with spatial covariance structures such as those used in geostatistics (Zimmerman & Harville, 1991; Gilmour *et al*., 1997; Stroup, 2002). The efficiency of spatial approaches has been compared with the no spatial analyses found in the literature (Brownie & Gumpertz, 1997; Wu & Dutilleul, 1999; Smith *et al*., 2001; Yang *et al*., 2004; Hong *et al*., 2005).

However, most comparisons of efficiency in the literature appear to focus on the nearest neighbour adjustment (including its modification or extensions) and/or the linear mixed model with one special covariance structure (usually the first order autoregressive model, AR(1)) against the analysis of variance of block designs. There have been few comparisons of mixed models with different spatial covariance structures. Now a migration seems to be taking place from the NNA to a fully-fledged mixed model analysis with different spatial components for spatial variability because of the flexibility, simplicity of use and other advantages of mixed model analysis (Piepho *et al*., 2008). Recently, linear mixed models have become well developed, and range from simple variance component models that provide information similar to ANOVA, to models with complex variance-covariance structures that aim to explore complex sources of variability and better accommodate interactions. Specifically, different analytical models can be cast in a unified mixed modelling framework (Denis *et al*., 1997; Piepho, 1998, 1999b).Within such a framework, different models can be handled as mixed models with different variance-covariance structures. Thus candidate models can be assessed and selected for MET data analyses, which can result in high accuracy when estimating and testing varietal effects.

Within advanced experimental designs, many spatial methods were proposed for adjusting the spatial trend (Bartlett, 1978; Wilkinson *et al*., 1983; Schwarzbach, 1984; Williams, 1986; Gilmour *et al*., 1997; Gleeson, 1997; Piepho, 1999a). A common feature of these methods is that plots that are closer together are assumed to have a higher correlation than plots farther apart. Via such models the precision of genotypic value estimates can be improved through both blocking and the adjustment of spatial trend in one or two dimensions.

With regard to the practical application of the linear mixed model with a spatial component, various unsolved problems must be dealt with. Among other issues, these are concerned with the selection of a suitable covariance model, *i.e*., a model with criteria that form the basis for a user's choice of whether or not to use a spatial model at all. Another point in this regard is the fact that the covariance parameters are unknown in practice and the estimated values based on observed data have to be used. In this case the statistical tests about the fixed effects of linear mixed models are generally not exact and their degrees of freedom must be determined by approximation. For some types of mixed models, the available methods for approximating degrees of freedom have been well examined (Schaalje *et al*., 2002; Spilke *et al*., 2004, 2005). For mixed models with spatial covariance structures, however, the use of the approximation methods has to be undertaken with care. In addition of the approxi-

mation, further consideration has to be given to the question of what influence the various spatial models have on the statistical tests used for, ranking and selection of lines in cultivar trial evaluations, apart from on efficiency vis a vis standard errors for line effect estimations. In MET, the local spatial tendency within trials and the residual heterogeneity between trials can be jointly modelled in the context of linear mixed models. By using a two-dimensional coordinate system at each trial, it is possible to define the plot location in a field, for example by specifying the latitude and longitude of plot centres (Casanoves *et al*., 2005, 2013).

The main objectives and contribution of this paper were (1) to highlight the advantages of mixed effect models in the data analysis of a national MET; (2) to show the importance of several main spatial variance-covariance structures, and direct implications of model choice for the inference of varietal performance, ranking and testing based on two data sets from real national trials by comparing blocking without spatial effect (ANOVA) model and a model with a block and spatial effect; the mixed models with spatial variance-covariance structure models were fitted using restricted maximum likelihood (REML) approach; and finally (3) to compare parameter estimates, ranking the varieties and ranking order and tests of varietal effects between the ANOVA model with only block effects and the mixed effects model with a block effect with selected spatial variance-covariance structure.

Material and methods

Linear mixed models have become well developed, and range from simple variance component models that provide information similar to ANOVA, to models with complex variance-covariance structures that aim to explore or better accommodate interactions. Specifically, different analytical models can be cast in a unified mixed modelling framework (Denis *et al*., 1997; Piepho, 1998, 1999b). Within such a framework, different models with specific variance-covariance structures can be formulated. Thus candidate models can be assessed and selected for MET data analyses, which result in high accuracy when estimating and testing varietal effects. Although there are already some general reviews of crop breeding analysis and variety evaluation trials (Davidoff & Selim, 1988; Smith *et al*., 2001, 2005), as well as studies on the analysis of MET data using the mixed models (Bartlett, 1978; Piepho, 1997; Kelly *et al*., 2007; Piepho & Möhring, 2010; Stefanova & Buirchell, 2010), most references just contain some examples for demonstration, or contain just one specific type of mixed model in data analysis.

Both traditional block design ANOVA models and spatial effect models can take the general form of the linear mixed model:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}u + e \qquad [1]$$

where **y** stands for the vector of observations, **X** is a matrix of constants associated with the fixed effects contained in the vector β, β is a vector of unknown fixed effects, **Z** is a matrix of constants associated with the random effects, *u* is a vector of random effects, and *e* is a vector of random residual errors. The random effects are assumed to be distributed as multivariate normal (MVN) or more precisely $\mathbf{u} \sim$ MVN (0, **G**) and the residual errors (*e*) distributed as MVN (0, **R**). It follows that the vector of observations is distributed as $\mathbf{y} \sim$ MVN ($\mathbf{X}\beta$, **V**) where $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$, The matrix **G** is the covariance matrix among random effects, **R** is the covariance matrix among the random residual errors, and **V** is the covariance matrix of **y**. For block designs, block effects may be regarded as fixed or random effects. A random block analysis makes additional use of the so-called inter-block information and is generally the preferred approach (Littell *et al*., 2006). In this article, block effects will be considered random in a combined analysis of data from different location. In this situation, **u** is the vector of block effects, and **Z** corresponds to the block effect design.

For analysis of variance models for block designs, block effects are assumed to be iid $\sim N(0,\sigma^2_b)$, and residual errors are assumed to be iid $N \sim (0,\sigma^2_b)$, where iid denotes independent and identically distribution, and and are variance components of blocks and residual errors, respectively. Hence, $\mathbf{G} = \mathbf{I}_b\sigma^2_b$ and $\mathbf{R} = \mathbf{I}_n\sigma^2_b$, where I_b is an identity matrix whose dimension equals the number of blocks, I_n is an identity matrix whose rank equals the number of observations. The main feature of analysis of variance models for block designs is that random variables located in the same block have the same covariance regardless of the extent of spatial variation; random variables not located in the same block have a covariance of zero.

In spatial effect models, **R** takes the form $\mathbf{R} = \mathbf{I}_n\sigma^2 + \sigma^2_s\mathbf{F}$, where σ^2_s is the covariance parameter of spatial structure variation, **F** is a square matrix with

Table 1. Spatial covariance structures

Structures	Description	Parameters	(i,j) elements
SP(EXP)(*c-list*)	Exponential	2	$\sigma^2 exp\{-d_{ij}/\theta\}$
SP(EXPA)(*c-list*)	Anisotropic Exponential	2c + 1	$\sigma^2\prod_{k=1}^{c} exp\{-\theta_k d(i,j,k)^{pk}\}$
SP(EXPGA)(c_1c_2)	2D Exponential, Geometrically Anisotropic	4	$\sigma^2 exp\{-d_{ij}(\theta,\lambda)/\rho\}$
SP(GAU)(*c-list*)	Gaussian	2	$\sigma^2 exp\{-d_{ij}^{\ 2}/\rho^2\}$
SP(GAUGA) (c_1c_2)	2D Gaussian, Geometrically Anisotropic	4	$\sigma^2 exp\{-d_{ij}(\theta,\lambda)^2/p^2\}$
SP(LIN)(*c-list*)	Linear	2	$\sigma^2(1-\rho d_{ij})1((\rho d_{ij}\le 1)$
SP(LINL)(*c-list*)	Linear Log	2	$\sigma^2(1-\rho log(d_{ij}))\times 1((\rho log d_{ij}\le 1)$
SP(MATERN)(*c-list*)	Matérn	3	$\sigma^2\frac{1}{\Gamma(v)}\left(\frac{d_{ij}}{2\rho}\right)^v 2K_v\left(\frac{d_{ij}}{\rho}\right)$
SP(MATHSW)(*c-list*)	Matérn (Handcock-Stein-Wallis)	3	$\sigma^2\frac{1}{\Gamma(v)}\left(\frac{d_{ij}\sqrt{v}}{2\rho}\right)^v 2K_v\left(\frac{2d_{ij}\sqrt{v}}{\rho}\right)$
SP(POW)(*c-list*)	Power	2	$\sigma^2\rho^{d_{ij}}$
SP(POWA)(*c-list*)	Anisotropic Power	c + 1	$\sigma^2\rho_1^{\ d(i,j,1)}\rho_2^{\ d(i,j,2)}\ldots\rho_c^{\ d(i,j,c)}$
SP(SPH)(*c-list*)	Spherical	2	$\sigma^2\left[1-\left(\frac{3d_{ij}}{2\rho}\right)+\left(\frac{d^3_{\ ij}}{2\rho^3}\right)\right]1((\rho d_{ij}\le\rho)$
SP(SPHGA)(c_1c_2)	2D Spherical, Geometrically Anisotropic	4	$\sigma^2\left[1-\left(\frac{3d_{ij}(\theta,\lambda)}{2\rho}\right)+\left(\frac{d_{ij}(\theta,\lambda)^3}{2\rho^3}\right)\right]\times 1(d_{ij}(\theta,\lambda)\le\rho)$

a dimension reflecting the number of observations, whose ijth element is $f(d_{ij})$, in which d_{ij} is the Euclidian distance between spatial observation points *i* and *j*. Suppose (x_i,y_i) and (x_j, y_j) describe the coordinates of the median points of plots for observations *i* and *j*, respectively, then their distance is:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \qquad [2]$$

where *x* and *y* denote horizontal and vertical directions. The variable $f(d_{ij})$ is generally a function of d_{ij} and its form is dependent on the spatial model used, which is dependent on the characteristics of spatial variation. The spatial covariance structures available for analysing field trials are listed in Table 1. In Table 1 *c-list* contains the names of the numeric variables used as coordinates of the location of the observation in space, and is the Euclidean distance between the *i*th and *j*th vectors of these coordinates, which correspond to the *i*th and *j*th observations in the input data set. For SP(POWA) and SP(EXPA), *c* is the number of coordinates, and $d(i, j, k)$ is the absolute distance between the *k*th coordinate, $k = 1 \ldots, c$, of the *i*th and *j*th observations in the input data set. For the geometrically anisotropic structures SP(EXPGA), SP(GAUGA), and SP(SPHGA), exactly two spatial coordinate variables must be specified as c_1 and c_2 . Geometric anisotropy is corrected by applying a rotation ı and Ï scaling to the coordinate system, $d_{ij}(\theta, \lambda)$ which represents the Euclidean distance between two points in the

transformed space. SP(MATERN) and SP(MATHSW) represent covariance structures in a class defined by Matérn (see Matérn, 1986; Handcock & Stein, 1993; Handcock, 1994). The function K_v is the modified Bessel function of the second kind of (real) order $v > 0$; the parameter governs the smoothness of the process (for further detail see SAS 9.3 help and documentation). The five spatial-variance covariance structures presented above belong to isotropic models, *i.e.*, the variation properties are the same in both directions x and y; the other models, as their names show, belong to anisotropic covariance structures, *i.e.*, the variation properties can be different in directions x and y.

Estimation and statistical test of varietal effects for the classical analysis of block designs uses ANOVA, which is, equating the observed mean squares to the expected mean squares with the assumption of independence, normality and homogeneity of the variances of the residuals. While spatial models analyses use REML for estimating variance components. Estimable functions $\boldsymbol{L}\beta$ of linear contrast of fixed effects (variety) are estimated based on $\boldsymbol{L}\hat{\beta} = \boldsymbol{L}(\boldsymbol{X}'V^{-1}\boldsymbol{X})^{-}X'V^{-1}\boldsymbol{y}$ with V being replaced by a REML estimate $\hat{V}$. The variance of $\boldsymbol{L}\hat{\beta}$ is determined based on $var(\boldsymbol{L}\hat{\beta}) = \boldsymbol{L}(\boldsymbol{X}'\hat{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{L}'$ (Hartley & Rao, 1967; Harville, 1977). Null hypotheses of the form of H_0: $\boldsymbol{L}\hat{\beta} = 0$ are tested using the statistic

$$t = \frac{\mathbf{L}\hat{\beta}}{\sqrt{\mathrm{var}\left(\mathbf{L}\hat{\beta}\right)}} \sim t\left(d.f.\right) \qquad [3]$$

In general, the test statistic in [3] is only approximately t-distributed and its degrees of freedom must be estimated. The approximate degrees of freedom in this research were determined using the Kenward-Roger method (Kenward & Roger, 1997). This approximation also uses the basic idea of Satterthwaite (1941). Its extension relative to the Satterthwaite method of Giesbrecht & Burns (1985) and Hrong-Tai Fai & Cornelius (1996) is an asymptotic correction of the estimated standard error of fixed effects due to Kackar & Harville (1984) in small and/or unbalanced data structures.

Statistical tools for model selection and test of consistency

Two questions in the analysis of practical trials are whether there is significant spatial variability and whether spatial models should be used (and if so, which models are most appropriate for data analysis). To answer these questions, statistical tools include likelihood-based methods (Oman, 1991; Wolfinger, 1993).The likelihood-ratio test (LRT) allows the comparison of the model's fit, provided that one of the models is hierarchically subordinated to the other or similarly the smaller model is nested with the larger one. This is the case if one model can be seen as a special case of a more general model due to certain model restrictions. The LRT then results from

$$LRT = -2\left(\ln LL_g - \ln LL_s\right) \sim \chi^2\left(d.f.\right) \qquad [4]$$

where $\ln LL_g$ and $\ln LL_s$ denote the log likelihood of the general model g and the special models, respectively. Given certain regularity conditions, the LRT testing statistic asymptotically follows a χ^2 distribution, with the degrees of freedom (*d.f.*) resulting from the number of restrictions that are necessary to transform the general model g into the special model (Fahrmeier & Hamerle, 1984; Greene, 2003). The general model fit, when compared to the special model, is considered better if $LRT > \chi^2(1\text{-}\alpha, d.f.)$ with a significant level of α. If the model comparison focuses on the covariance structure of a constant expectation structure, the likelihoods are employed via the REML method (Wolfinger, 1993). This can be used for the first question. In this case, g corresponds to the model with spatial correlations among observations, and corresponds to the model without spatial correlation among observations. The LRT based on formula [4] can also be used for testing the difference between the block design ANOVA model (block effects as random) and the model without correlations among observations, because the latter is also a special model variation of the former. Thus, it can be used for testing the difference between the spatial models with and without block effects.

As mentioned above, the LRT is only applicable when comparing two nested models. For model comparisons that do not require hierarchical models, there are a number of analytical criteria. These are so-called «Information Criteria» based on likelihood estimations. In the current work Akaike's Information Criterion (AIC) is used for comparing the covariance structures for an identical expectation structure using the REML estimation methods and is generally given by:

$$AIC = -2\ln LL + 2q \qquad [5]$$

where $\ln LL$ is the log-likelihood same as in formula [4] and q is the number of the parameters of the variance-

covariance structure. Thus, the formula of the information criteria is given in such a way that the model with the smaller value for the AIC is preferred (Bozdogan, 1987; Burnham & Anderson, 2002). For un-nested model we prefer to use the AIC but we note that there are other available information criteria, such as the Schwarz Bayesian Information Criterion (BIC) (Schwarz, 1978). Guerin & Stroup (2000) compared the performance of AIC and BIC on covariance model selection for repeated measures and stated that AIC tends to select a more complex model but with a better control of type I error than the BIC. To assess consistency (or inconsistency) in the statistical tests on varietal effects between two models one can use the test consistency ratio, which is computed as follows:

$$\text{text consistency ratio} = \frac{\text{number of significant varietal differences tested simultaneously in two models}}{\max\ (\text{number of significant varietal differences tested under the two models considerated})} \quad [6]$$

Data set and analysis

The data sets used in this study are taken from the Ethiopian Agricultural Research Institute National Variety Trials for Bread Wheat (BW00RVTI data) and Barley Trial (BW01RVII data) of 2006-2008. Some 20 bread wheat (*Triticum aestivum* L.) varieties were tested in at six locations (environments) on the first year (2006/7) and five locations (environments) among the six of the first used on the second year (2007/8). Similarly 25 barley (*Hordeum vulgare* L.) varieties were tested in five locations (environments) in 2007/8. All the trials in each location were laid out as a randomized complete block (RCB) design with four replicates. There are two approaches to analysing MET data using mixed model, the so-called one- and two-stage approaches (Welham *et al*., 2010). In a one-stage analysis, individual plot data from all trials are combined in a single analysis (Cullis *et al*., 1998). In a two-stage analysis, variety means are first obtained from the separate analysis of individual trials (Stage I), and are then combined in an overall mixed model analysis (Stage II). The two-stage analysis can be unweighted (*e.g*., Patterson & Silvey, 1980) or weighted to reflect the relative precision of variety means from each trial (*e.g*., Smith *et al*., 2001). A one-stage approach provides the most accurate predictions of variety performance, but it can be computationally difficult to use when the variance models involved are complex. With the steady improvements in computing power, single-stage analyses are becoming feasible. Apart from computational speed, the main advantages of the two-stage approach are that one can carefully analyse each trial individually, taking into account any specifics of the design or field trends.

In this study we used two approaches for analysis; the first one was a separate individual analysis of each location of the BW00RVTI data set of wheat and BW01RVII data set for barley. The second one was a one-stage analysis, individual plot where data from all trials (locations)are combined in a single analysis of a two year BW00RVTI data set of wheat and a one year BW01RVII data set of six location. Each data set was separately fitted per location and per year using the mixed model with fourteen variance-covariance structures. The mixed model with compound symmetry (CS) variance-covariance structures was identical to the ANOVA model. The optimally fitted spatial model and the ANOVA model are used for further varietal effect assessment and statistical tests (or inference). The single-stage analysis was applied to each of the data sets by fitting one spatial-variance covariance structure at a time for all location. Putting location as random group factor on SAS (proc mixed) analysis gave a different random parameter estimate for each location. All the analyses ware conducted using standard SAS software version 9.3. The results from the two models were compared and used to assess consistency (or inconsistency) in statistical tests on varietal effects between the two models, using consistency ratio defined earlier.

Results and discussion

Model fit statistics from ANOVA and the mixed model with various spatial variance-covariance structures and results of possible LRT and AIC for all models are summarised in Tables 2, 3 and 4. Note that "—" denotes the failure of a model to converge. This occurred with the sp(lin) and sp(linlog) structures in any of the locations, which shows that these models are not suitable for that trial data (Schabenberger & Pierce, 2002). The smallest AIC value (bold in Tables 2, 3 and 4) indicates that for BW00RVTI trial data set year 1 and 2 support the anisotropic power [spa (powa)] and exponential [spa (exp)] variance-covariance structures as the best compared to the ANOVA model for seven trials (locations) out of eleven. Similarly for the BW00RVTI trial five different spatial variance-covariance structures [sp(pow),

Table 2. Related fitting statistics for the ANOVA model and the linear mixed model with spatial variance-covariance structures for the first year BW00RVTI data set

Model	Location-1			Location-2			Location-3			Location-4			Location-5			Location-6		
	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2
RCBD	385.1	387.1	—	344.2	346.2	—	398.8	400.8	—	426.4	**428.4**	—	421.7	**423.7**	—	377.3	379.3	—
sp(sph)	408.5	412.5	1	360.8	364.8	1	407.5	411.5	1	442.1	446.1	1	441.4	445.4	1	364.4	368.4	0.0003
sp(exp)	385.1	387.1	1	343.7	347.7	0.4869	394.2	398.2	0.0303	426.4	**428.4**	1	421.7	**423.7**	1	361.8	365.8	<.0001
sp(gau)	384.9	388.9	0.6537	343.8	347.8	0.5743	394.7	398.7	0.0425	426.3	430.3	0.746	421.7	**423.7**	1	365.5	369.5	0.0006
sp(pow)	385.1	389.1	0.865	343.7	347.7	0.4869	394.2	398.2	0.0303	426.2	430.2	0.63	421.7	425.7	0.9213	361.8	365.8	<.0001
sp(mat)	385.1	387.1	1	343.7	349.7	0.7806	394.1	400.1	0.0953	426.1	432.1	0.8603	421.6	427.6	0.9648	359.7	365.7	0.0001
sp(EXPA)	—	—	—	—	—	—	386.4	396.4	0.0144	—	—	—	—	—	—	—	—	—
sp(EXPGA)	385.1	391.1	0.0608	334.5	342.5	0.0179	386.1	394.1	0.0747	426.4	432.4	0.2721	421.7	427.7	0.1855	353	361	0.0034
sp(GAUGA)	379.3	387.3	0.0339	344.2	352.2	1	398.8	404.8	1	426.4	432.4	0.3575	421.7	427.7	0.2295	363.9	371.9	0.1199
sp(MATHSW)	385.1	389.1	1	343.7	349.7	0.7806	394.1	400.1	0.0953	426.4	430.4	1	421.7	425.7	1	359.7	365.7	0.0001
sp(POWA)	372.7	**378.7**	0.002	332.5	**338.5**	0.003	386.4	**392.4**	0.002	424.8	430.8	0.4545	420.1	426.1	0.4665	349	**355**	<.0001
sp(SPHGA)	393.8	399.8	1	—	—	—	—	—	—	—	—	—	439.2	447.2	1	355.6	363.6	<.0001

—: denotes the failure of a model to converge. Bold values indicate smallest AIC (Akaike's information criteria). LL: log-likelihood. Locations 1, 2, 3, 4, 5 and 6 are Kulumsa, Adet, Bekoji, Sinana, Holeta, and DeberZeit (Ethiopia), respectively.

Table 3. Related fitting statistics for the ANOVA model and the linear mixed model with spatial variance-covariance structures for the second year BW00RVTI data set

Model	Location-1			Location-2			Location-3			Location-4			Location-5		
	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2
RCBD	370	372	—	334.2	336.2	—	345.6	**347.6**	—	395.2	397.2	—	284.1	**286.1**	—
sp(sph)	382.4	386.4	1	351.9	355.9	1	366.4	370.4	1	394.2	398.2	0.317	303	307	1
sp(exp)	367.1	**371.1**	0.091	333.5	337.5	0.428	345.6	**347.6**	1	387.3	**391.3**	0.005	284.1	**286.1**	1
sp(gau)	368.3	372.3	0.195	332.4	336.4	0.189	345.6	**347.6**	1	388.7	392.7	0.011	284.1	**286.1**	1
sp(pow)	367.1	**371.1**	0.091	333.5	337.5	0.428	345.5	349.5	0.659	387.3	**391.3**	0.005	284	288	0.762
sp(mat)	365.8	371.8	0.124	—	—	—	—	—	—	386.9	392.9	0.016	—	—	—
sp(EXPA)	366.2	374.2	0.29	322.9	**330.9**	0.01	343.9	353.9	0.79	384.9	394.9	0.036	—	—	—
sp(EXPGA)	363.8	371.8	0.339	323.1	331.1	0.003	340.5	348.5	0.108	384.6	392.6	0.232	284.1	290.1	0.234
sp(GAUGA)	370	378	1	325.6	333.6	0.032	344	352	0.976	395.2	403.2	1	284.1	290.1	0.594
sp(MATHSW)	365.8	371.8	0.124	—	—	—	—	—	—	386.9	392.9	0.016	284.1	288.1	1
sp(POWA)	367.9	373.9	0.367	322.7	328.7	0.003	344.1	350.1	0.464	386.3	392.3	0.011	283.6	289.6	0.773
sp(SPHGA)	382.4	390.4	1	—	—	—	356.1	364.1	1	391.3	399.3	0.273	—	—	—

—: denotes the failure of a model to converge. Bold values indicate smallest AIC (Akaike's information criteria). LL: log-likelihood.

sp(expga), sp(mathsw), sp(expga) and sp(powa)] models were selected as the best compared to the ANOVA model for the five location BW01RVII trial data set.

A model comparison between a block effect without spatial structure (ANOVA) and a model with a block and spatial effect using the LRT χ^2-test for the trials for the two (BW00RVTI and BW01RVII) data sets suggested that the selected spatial variance-covariance structure fitted the data significantly better than the ANOVA model. However the optimally-fitted spatial variance-covariance structures were not the same from one location to the other. The optimally fitting spatial variance-covariance structure was spatial power [sp(powa)] for most of the locations. These results showed that assuming a homogeneous variance-covariance structure in the ANOVA model is generally not realistic, and therefore using a linear mixed model with spatial variance-covariance is necessary to improve the efficiency of the data analysis and accommodation of local stationary trend of MET data.

It appears the year to year effect on variance-

Table 4. Related fitting statistics of ANOVA model and linear mixed model with spatial variance-covariance structures for the one year BW01RVII data set

Model	Location-1			Location-2			Location-3			Location-4			Location-5		
	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2	LL	AIC	Pr>χ^2
RCBD	475.5	477.5	—	454.5	456.5	—	421.6	423.6	—	506.9	508.9	—	524	526	—
sp(sph)	475.5	477.5	1	454.5	456.5	1	427.7	431.7	1	506.9	508.9	1	551.3	555.3	1
sp(exp)	468.5	**472.5**	0.0083	452.3	456.3	0.142	410.9	414.9	0.001	503.7	507.7	0.0712	524	526	1
sp(gau)	471	475	0.0341	452.4	456.4	0.145	414.1	418.1	0.006	504.8	508.8	0.1409	523.9	527.9	0.689
sp(pow)	468.5	**472.5**	0.0083	452.3	456.3	0.142	410.9	414.9	0.001	503.7	507.7	0.0712	524	528	0.8947
sp(mat)	—	—	—	452	458	0.283	407.7	**413.7**	0.001	502.9	508.9	0.1344	524	526	1
sp(EXPA)	—	—	—	452.4	462.4	0.711	—	—	—	—	—	—	—	—	—
sp(EXPGA)	466.7	474.7	0.347	447.7	**455.7**	0.07	421.6	429.6	1	493.2	**501.2**	0.0141	524	530	0.0323
sp(GAUGA)	469	477	0.3427	449.6	457.6	0.157	410.1	418.1	0.275	498.6	506.6	0.0619	529.9	530.5	0.1424
sp(MATHSW)	—	—	—	452	458	0.283	407.7	**413.7**	0.001	502.9	508.9	0.1344	524	528	1
sp(POWA)	467.4	473.4	0.0179	458.8	459.2	0.435	413.7	419.7	0.02	497.1	503.1	0.0071	519.2	**525.2**	0.0873
sp(SPHGA)	475.5	481.5	1	458.7	466.7	1	420.4	428.4	0.742	495	503	0.0077	—	—	—

—: denotes the failure of a model to converge. Bold values indicate smallest AIC (Akaike's information criteria). LL: log-likelihood.

Table 5. The number of significant and highly significant variety contrasts of t-test for trials of the BW00RVTI and BW01RVII data sets and the consistency ratio test between the ANOVA model and the spatial linear mixed model with optimally fitting spatial variance-covariance structure (SLMM)

	Data set BW00RVTI									Data set BW01RVII			
	Year-1				Year-2					Year-1			
	ANOVA	SLMM	Consistency		ANOVA	SLMM	Consistency			ANOVA	SLMM	Consistency	
			No.	Ratio (%)			No.	Ratio (%)				No.	Ratio (%)
Location-1	60	78	53	67.94	44	40	37	84.09	Location-1	46	65	43	66.15
Location-2	53	50	37	69.81	33	35	26	74.29	Location-2	67	79	60	75.95
Location-3	10	18	6	33.33	*	*			Location-3	97	157	94	59.87
Location-4	*	*			45	50	41	82	Location-4	47	36	22	46.8
Location-5	*	*			*	*			Location-5	36	45	26	57.78
Location-6	45	20	15	33.33									
Average	42	41.5	22.75	51.11	40.67	41.66	34.67	80.13	Average	58.6	76.4	49	61.31

*: the optimally fitting model is ANOVA.

covariance of varieties is greatly exhibited in the BW00RVTI data set. This is shown through the variance-covariance structures being mostly consistent for different locations in the same year, but obviously not consistent between years as shown in Table 2 and 3. This result is easily understood by realising that within a year we expect only between location differences, but between years there could be differences in environments (years). The failure of some spatial variance-covariance structures to converge may indicate that they are not suitable or compatible with the structure of the current MET data but could work with other data sets.

To examine the impact of the spatial variance-covariance structures on estimates on test of varieties, the number of significant (at $\alpha = 0.05$) varietal differences by the t-test are given in Table 5. Using the ANOVA model and mixed model with the optimally-fitted spatial variance-covariance for each location, we assessed the consistency between these two models. The number of significant varietal differences by t-test is not the same between the ANOVA model and the

Table 6. The first eight genotype ranking comparison between the ANOVA model and the optimally fitting spatial variance-covariance structure (SLMM) of five trials of data set BW01RVII location by location and a single-stage analysis

Rank	Location-1		Location-2		Location-3		Location-4		Location-5		All
	ANOVA	SLMM	ANOVA	SLMM	ANOVA	SLMM	ANOVA	SLMM	ANOVA	SLMM	SLMM
1	G23	G23	G21	G23	G23	G23	G23	G23	G23	G23	G23
2	G13	G13	G23	G21	G4	G4	G11	G2	G2	G2	G21
3	G5	G21	G3	G3	G17	G17	G19	G1	G4	G1	G4
4	G17	G3	G17	G2	G2	G21	G15	G5	G8	G19	G13
5	G21	G17	G7	G17	G21	G10	G2	G19	G13	G10	G2
6	G4	G4	G11	G13	G10	G2	G5	G8	G14	G4	G17
7	G3	G5	G9	G16	G8	G8	G8	G15	G1	G3	G8
8	G19	G15	G13	G9	G6	G15	G13	G11	G19	G13	G3

mixed model with optimally fitted spatial variance-covariance structures. The consistency ratio test between the two models falls in the range of 33-84%. From the average of all trials (locations), the test consistency ratio of two models is approximately 64%, which means that approximately 36% of the varietal differences being tested as significant or very significant in one model cannot be tested as significant or very significant by the other model.

Varietal ranking

Apart from contrasts between new varieties, the ranking of varietal productivity and a comparison of new varieties with standard variety is also important for variety trials. We consider the trial from the five locations of BW01RVII data to compare variety mean ranking between the ANOVA model and the optimal spatial variance-covariance model. A trial corresponds to a single experiment at a single location. Table 6 shows the ranking for the first eight entries from the optimal spatial variance-covariance mixed model compared to the ranking from the ANOVA model across the locations. The model with spatial structure is relatively more consistent in its top eight ranking than the ANOVA model. The ranking are different for different locations and differ between the spatially structured model and ANOVA. A rank difference of genotype between the locations is showing the presences of genotype by environment interaction. This also indicates the advantage of single stage spatial models on the handling of the spatial trend and variation of the trials.

The simple homogenous variance-covariance structures implied by ANOVA models, which assume that the interaction effects of varieties are independent, is mostly not appropriate for data analyses of MET. The fact that the goodness of fit of one variance-covariance structure was different for various trial data sets, and that none fitted all trial data sets optimally throughout, indicates that the heterogeneous characteristics of variance-covariance are not identical across the trials. Therefore, the arbitrary use of a homogeneous variance-covariance structure (*e.g.* ANOVA model) to analyse the MET cannot ensure a high degree of accuracy. In this study, the ANOVA model, as a special case form of mixed models, showed obvious inconsistency in estimates and tests of varietal effects compared to the linear mixed model with the optimally-fitted spatial variance-covariance structures.

Both effective experimental designs and spatial analyses can have an important role in improving the reliability and precision of experiment results. The importance of spatial variability to be expected from a logical and subjective-related perspective is confirmed in a variety of experiments. As presented in much of the literature, spatial analysis may lead to higher efficiency with regard to standard error of estimation of fixed effects than a non-spatial analysis, provided that spatial variability is present. Based on this work, the commonly used ANOVA mixed model is not an appropriate model for data analysis of MET trials. The spatial variance-covariance models are more useful in a practical sense, given that they can describe actual existing variance-covariance characteristics more accurately than the ANOVA model. Of course, with one-stage analyses, the proposed spatial variance-covariance models are expected to yield identical mean yields for balanced data, and differences are expected only for unbalanced data. Even so, a selection of

variance-covariance structures based on the mixed model framework is important since the standard error of varietal effect estimates (*i.e*. the accuracy of varietal effect estimates) is different under the various models, and unbalanced data is common in MET (Möhring & Piepho, 2009). The advantage and validity of using spatial variance-covariance structure depends on the present spatial variability. Most of the investigated spatial models showed better data fitting and smaller standard error for variety contrasts than the ANOVA model.

The main purposes of the present paper was to show the importance of variance-covariance structure selection and to illustrate that the classical ANOVA model is inferior to more elaborate mixed models in the analysis of MET data. This does not imply that the models considered in this paper are appropriate for any situation. For example, in some locations (trials) the ANOVA model still optimally fitted the data better than the spatial models.

Acknowledgment

We gratefully acknowledge Institute of Agricultural Research (EIAR, Addis Ababa. Ethiopia) for supplying data sets used in this study. The authors also thank the UKZN for funding the PhD study for the first author. We also sincerely thank the anonymous reviewers for their comments which helped to substantially improve the writing of the paper.

References

Bartlett MS, 1978. nearest neighbour models in the analysis of field experiments. J R Stat Soc B Method 40: 147-174.

Bozdogan H, 1987. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika 52: 345-370.

Brownie C, Gumpertz ML, 1997. Validity of spatial analyses for large field trials. J Agr Biol Environ Stat 2: 1-23.

Burnham KP, Anderson DR, 2002. Model selection and multi-model inference: a practical information-theoretic approach, Springer.

Casanoves F, Macchiavelli R, Balzarini M, 2005. Error variation in multienvironment peanut trials. Crop Sci 45: 1927-1933.

Casanoves F, Macchiavelli R, Balzarini M, 2013. Models for multi-environment yield trials with fixed and random block effects and homogeneous and heterogeneous residual variances. Journal of Agriculture of the University of Puerto Rico, 91: 117-131.

Clarke FR, Baker RJ, 1996. Spatial analysis improves precision of seed lot comparisons. Crop Sci 36: 1180-1184.

Cullis BR, Gleeson AC, 1991. Spatial analysis of field experiments-an extension to two dimensions. Biometrics 47: 1449-1460.

Cullis B, Gogel B, Verbyla A, Thompson R, 1998. Spatial analysis of multi-environment early generation variety trials. Biometrics 54: 1-18.

Davidoff B, Selim HM, 1988. Correlation between spatially variable soil moisture content and soil temperature. Soil Sci 145: 1-10.

Denis JB, Piepho HP, Eeuwijk FA, 1997. Modelling expectation and variance for genotype by environment data. Heredity 79: 162-171.

Fahrmeir L, Hamerle A, 1984. Multivariate Statistische Verfahren. Walter de Gruyter, Berlin, NY.

Giesbrecht FG, Burns JC, 1985. Two-stage analysis based on a mixed model: large-sample asymptotic theory and small-sample simulation results. Biometrics 41: 477-486.

Gilmour AR, Cullis BR, Verbyla AP, 1997. Accounting for natural and extraneous variation in the analysis of field experiments. J Agr Biol Environ Stat 2: 269-293.

Gleeson AC, 1997. Statistical methods for plant variety evaluation (Kempton RA, Fox PN, eds). Chapman & Hall, London, Chapter 5, pp: 68-85.

Greene WH, 2003. Econometric analysis, 5th edition. Pearson Educ Int, University Prentice Hall, NY

Guerin L, Stroup WW, 2000. A simulation study to evaluate PROC MIXED analysis of repeated measures data. Proc 12th Ann Conf on Applied Statistics in Agriculture. Kansas State Univ (Manhattan KS, ed.). pp: 170-203.

Handcock MS, 1994. An approach to statistical spatial-temporal modeling of meteorological fields: rejoinder. J Am Stat Assoc 89: 388-390.

Handcock MS, Stein ML, 1993. A Bayesian analysis of kriging. Technometrics 35: 403-410.

Hartley HO, Rao CR, 1967. Maximum likelihood estimation for the mixed analysis of variance model. Biometrica 54: 93-108.

Harville DA, 1977. Maximum likelihood approaches to variance component estimation and to related problems. J Am Stat Assoc 72: 320-338.

Hong N, White JG, Gumpertz ML, Weisz R, 2005. Spatial analysis of precision agriculture treatments in randomized complete blocks. Agron J 97: 1082-1096.

Hrong-Tai Fai A, Cornelius PL, 1996. Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. J Stat Comput Sim 54: 363-378.

Kackar RN, Harville DA, 1984. Approximations for standard errors of estimators of fixed and random effect in mixed linear models. J Am Stat Assoc 79: 853-862.

Kelly AM, Smith AB, Eccleston JA, Cullis BR, 2007. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. Crop Sci 47: 1063-1070.

Kempton R, 1984. The use of biplots in interpreting variety by environment interactions. J Agric Sci 103: 123-135.

Kenward MG, Roger JH, 1997. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53: 983-997.

Littell RC, Milliken GA, Stroup WW, Wolfinger RD, Schabenberger O, 2006. SAS for mixed models. SAS Institute Inc., Cary, NC, USA.

Matérn B, 1986. Spatial variation, 2nd edition. In: Lecture notes in statistics. Springer-Verlag, NY. pp: 705-711.

Mo H, Si Y, 1986. Trend variation and it's control in field experiment. Acta Agr Sinnica 12(4): 233-240.

Möhring, J, Piepho HP, 2009. Comparison of weighting in two-stage analysis of plant breeding trials. Crop Sci 49: 1977-1988.

Oman SD, 1991. Multiplicative effects in mixed model analysis of variance. Biometrika 78: 729-739.

Patterson HD, Silvey V, 1980. Statutory and recommended list trials of crop varieties in the United Kingdom. J R Stat Soc A 143: 219-252.

Piepho HP, 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. Biometrics 53: 761-766.

Piepho HP, 1998. Methods for comparing the yield stability of cropping systems. J Agr Crop Sci 180: 193-213.

Piepho HP, 1999a. Cultivar comparisons in three-way data based on mixed models with flexible variance-covariance structure. Biuletyn Oceny Odmian 30: 31-48.

Piepho HP, 1999b. Stability analysis using the SAS System. Agron J 91: 154-160.

Piepho HP, Möhring J, 2010. Generation means analysis using mixed models. Crop Sci 50: 1674-1680.

Piepho HP, Richter C, Williams E, 2008. Nearest Neighbour adjustment and linear variance models in plant breeding trials. Biometrical J 50: 164-189.

Satterthwaite FE, 1941. Synthesis of variance. Psychometrika 6(5): 309-316.

Schaalje GB, McBride J, Fellingham G, 2002. Adequacy of approximations to distributions of test statistics in complex mixed linear models. J Agr Biol Environ Stat 7: 512-524.

Schabenberger O, Pierce FJ, 2002. Contemporary statistical models for the plant and soil sciences. CRC Press, Boca Raton. 738 pp.

Scharf PC, Alley MM, 1993. Accounting for spatial yield variability in field experiments increases statistical power. Agron J 85: 1254-1256.

Schwarz G, 1978. Estimating the dimension of a model. Ann Stat 6: 461-464.

Schwarzbach E, 1984. A new approach in the evaluation of field trials. The determination of the most likely genetic ranking of varieties. Vortrage Pflanzen 6: 249-259.

Smith A, Cullis B, Gilmour A, 2001. Applications: the analysis of crop variety evaluation data in Australia. Aust New Zeal J Stat 43: 129-145.

Smith A, Cullis BR, Thompson R, 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. J Agr Sci 143: 449-462.

Spilke J, Piepho HP, Meyer U, 2004. Approximating the degrees of freedom for contrasts of genotypes laid out as subplots in an alpha-design in a split-plot experiment. Plant Breeding 123: 193-197.

Spilke J, Piepho HP, Hu X, 2005. A simulation study on tests of hypotheses and confidence intervals for fixed effects in mixed models for blocked experiments with missing data. J Agr Biol Environ Stat 10: 374-389.

Stefanova KT, Buirchell B, 2010. Multiplicative mixed models for genetic gain assessment in lupin breeding. Crop Sci 50: 880-891.

Stroup WW, 2002. Power analysis based on spatial effects mixed models: A tool for comparing design and analysis strategies in the presence of spatial variability. J Agr Biol Environ Stat 7: 491-511.

Welham SJ, Gogel BJ, Smith AB, Thompson R, Cullis BR, 2010. A comparison of anlysis methods for late-stage variety evaluation trials. Aust New Zeal J Stat 52: 125-149.

Wilkinson GN, Eckert SR, Hancock TW, Mayo O, 1983. Nearest neighbour (NN) analysis of field experiments. J R Stat Soc B 45: 151-211.

Williams ER, 1986. A neighbour model for field experiments. Biometrika 73: 279-287.

Wolfinger R, 1993. Covariance structure selection in general mixed models. Commun Stat-Simul C 22: 1079-1106.

Wu T, Dutilleul P, 1999. Validity and efficiency of neighbor analyses in comparison with classical complete and incomplete block analyses of field experiments. Agron J 91: 721-731.

Yang RC, Ye TZ, Blade SF, Bandara M, 2004. Efficiency of spatial analyses of field pea variety trials. Crop Sci 44: 49-55.

Zimmerman DL, Harville DA, 1991. A random field approach to the analysis of field-plot experiments and other spatial experiments. Biometrics 47: 223-239.

academicJournals

Vol. 8(12), pp. 1033-1040, 4 April, 2013
DOI: 10.5897/AJAR2012.6648
ISSN 1991-637X ©2013 Academic Journals
http://www.academicjournals.org/AJAR

African Journal of Agricultural Research

Full Length Research Paper

Additive main effects and multiplicative interactions model (AMMI) and genotype main effect and genotype by environment interaction (GGE) biplot analysis of multi-environmental wheat variety trials

Asnake Worku Negash[1], Henry Mwambi[1], Temesgen Zewotir[1] and Girma Taye[2]

[1]School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal Pietermaritzburg, South Africa.
[2]CIMMIT, Ethiopia.

Accepted 26 March, 2013

Genotype by environment (G×E) interaction is associated with the differential performance of genotypes tested at different locations and in different years, and influences selection and recommendation of cultivars. Wheat genotypes were evaluated in six environments to determine the G×E interactions and stability of the genotypes. Additive main effects and multiplicative interactions (AMMI) was conducted for grain yield of both year and it showed that grain yield variation due to environments, genotypes and (G×E) were highly significant (p <0.01). Stability for grain yield was determined using genotype plus genotype by environment interaction (GGE) biplot analysis. The first two principal components (PC1 and PC2) were used to create a 2-dimensional GGE biplot. Which-won-where pattern was based on six locations in the first and five locations in the second year for all the 20 genotypes. The resulting pattern is one realization among many possible outcomes, and its repeatability in the second was different and a future year is quite unknown. A repeatability of which-won-where pattern over years is the necessary and sufficient condition for mega-environment delineations and genotype recommendation.

Key words: Additive main effects and multiplicative interactions (AMMI), genotype×environment (G×E) interactions, wheat, stability.

INTRODUCTION

The increase in population and the subsequent rise in demand for agricultural produce are expected to be greater in regions where production is already insufficient, in particular in Sub-Saharan Africa. The necessity and demand to increase agricultural production represents a huge challenge to local farming systems given it must come mainly from increased yield per unit area in addition to the limited extension of cultivated land in the country. To meet this requirement various crop improvement programmes have been initiated by the Ethiopian Institute of Agricultural Research (EIAR). Under any crop improvement programme a sample of promising genotypes are performance tested each year at a number of sites, representing major crop growing areas with the a view to identify genotypes which possess the dual qualities of high yield capacity and low sensitivity to

*Corresponding author. E-mail: 208517201@stu.ukzn.ac.za.

adverse change in environmental condition. One of the important focuses in the current paper is to assess the performance of improved genotypes in multi environment (multi-location, multi-year or both) trials. Multi-Environment Yield Trials (MEYT) are conducted for different crops throughout the world (Yan and Rajcan, 2002; Dehghani et al., 2006) not only to identify high yielding cultivars but also to identify sites that best represent the target environment (Yan, 1999; Yan et al., 2000, 2001). As usual in MEYT, a number of genotypes are tested over a number of sites and years to see adaptation of the crop. But, it is often difficult to determine the pattern of genotypic responses across environments without the use of appropriate analytical and statistical tools such as additive main effects and multiplicative interactions (AMMI) and Genotype main effect and Genotype×Environment interaction (GGE) biplot (Gauch, 1992; Gauch and Zobbel, 1996; Yan et al., 2000; Yan, Tinker, 2006) for graphical display of data.

The measured yield of each cultivar in each test environment is a result of genotype main effect (G), and environment main effect (E) and genotype by environment (G×E) interaction (Yan and Kang, 2003). Though, E mostly accounts for about 80% of the total yield variation; it is only G and G×E interaction that are relevant to cultivar evaluation and mega environment classification (Rao et al., 2005; Yan et al., 2000; Yan, 2002; Yan and Rajcan 2002; Kaya et al., 2006). AMMI and GGE models are singular value decomposition (SVD) based statistical methods often applied to yield trial studies for visualizing the data. The methods helps in understanding complex genotype by environment(G×E) interactions, determining which genotype has been in which environments, and also helping in grouping environments with the same winner (or similar winners) into mega-environments.

Wheat is the most important cereal crop in Ethiopia and represents nearly 14% of grain crop production. It covers 71,786.86 ha of cropped land area with average productivity of 9.86 qut/ha but it is less than half of the world average yield (ECSA, 2011). Understanding genotype by environment interaction (GEI) helps plant breeders to design better breeding strategies. Therefore, the objectives of this study are to evaluate the yield performance and stability of genotypes in relation to environment (location) on year to year basis. Secondly the study will examine the possible existence of different mega environments and the wining genotype for each mega environments.

MATERIALS AND METHODS

Description of the data

The data used in the current paper are from a study carried out between 2004 and 2005 in six different research stations in Ethiopia. The locations consist of loc1 (Kulumsa),loc2(Adet), loc3 (Bekoji), loc4 (Sinana), loc5 (Holeta) and loc6 (DeberZeit). Twenty bread wheat genotypes were evaluated in each of the above locations (environments) in a randomized complete block design with four replications. These Twenty genotypes are coded from G1-G20.

The model

In terms of effects, the basic model for a multi-environment trial can be written as

$$Y_{ijl} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijl} \quad (1)$$

Where Y_{ijl} is the measured yield value of the i^{th} genotype in the j^{th} environment and l^{th} replicate, μ is the grand mean, α_i is the main effect of the i^{th} genotype, β_j is the main effect of j^{th}environment, γ_{ij} is interaction between i^{th} genotype and j^{th} environment and ϵ_{ijl} is random error. Were we assume that $\epsilon_{ijl} \sim$indep $N(0,\delta_j^2)$. The ranges of indices are i=1, 2,...,20j =1, 2,..., 6 l =1, 2, 3, 4. Thus the cell mean for the model is

$$\mu_{ij} = E(Y_{ijl}) = \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijl} \quad (2)$$

In GGE biplots genotype plus genotype × environment (*G* + *GE*) interaction are studied together and to achieve this G+GE effect is separated out from the observed mean and eventually the model becomes (omitting the random error)

$$\mu_{ij} - (\mu + \beta_j) = \alpha_i + \gamma_{ij} \quad (3)$$

However in the case of the AMMI model, the effect of genotypes is also separated out only genotype × environment (GE) interaction is studied for biplot, and eventually the model becomes

$$\mu_{ij} - (\mu + \beta_j + \alpha_i) = \gamma_{ij} \quad (4)$$

The mathematical expressions for partitioning of G+GE for GGE biplots and GE for AMMI models are similar except a difference in model formulation. The G+GE for GGE and GE for AMMI effects are partitioned into multiplicative terms by using the singular value decomposition (SVD) as

$$\mu_{ij} - \mu - \beta_j = \lambda_1 \xi'_{i1} \eta_{j1} + \lambda_2 \xi'_{i2} \eta_{j2} + \gamma_{ij}$$

and

$$\mu_{ij} - \mu - \beta_j - \alpha_i = \lambda^*_1 \xi'^*_{i1} \eta^*_{j1} + \lambda^*_2 \xi'^*_{i2} \eta^*_{j2} + \gamma^*_{ij} \quad (7)$$

respectively, where λ_1 (λ^*_1) and λ_2 (λ^*_2) are the singular values (SV) for the first and second principal component (PC1 and PC2), ξ'_{i1}(ξ'^*_{i1}) andξ'_{i2} (ξ'^*_{i2}) are eigenvectors of genotype i for PC1 and PC2, η_{j1} (η^*_{j1}) and η_{j2} (η^*_{j2}) are eigenvectors of environment j for PC1 and PC2 and γ_{ij} (γ^*_{ij}) is the residual not explained by PC1 and PC2 for genotype i in environment j. The PC1 and PC2 eigenvectors cannot be plotted directly to construct a meaningful biplot before the singular values are partitioned into the genotype and environment eigenvectors. To generate a biplot that can be used in visual analysis of MEYT data, the SVs have to be partitioned into the genotype and environment eigenvectors so that Equation (5) can be written in the form of

$$\mu_{ij} - \mu - \beta_j = \sum_{l=1}^{k} g'_{il} e'_{lj} + \gamma_{ij} \text{ and } \mu_{ij} - \mu - \beta_j - \alpha_i = \sum_{l=1}^{k} g'_{il} e'_{lj} + \gamma^*_{ij} \quad (6)$$

Table 1. ANOVA table for AMMI model.

Source	Year 2004					Year 2005				
	df	SS	MS	F	F-prob	df	SS	MS	F	F-prob
Total	479	54590	114			399	27188	68.1		
Treatments	119	41599	349.6	10.2	0	99	19806	200.1	9.93	0
Genotypes	19	1187	62.5	1.82	0.01944	19	2779	146.3	7.26	0
Environments	5	35212	7042.4	99.8	0	4	13988	3497.1	31.97	0
Block	18	1270	70.6	2.06	0.00706	15	1641	109.4	5.43	0
Interactions	95	5200	54.7	1.6	0.00134	76	3038	40	1.98	0.00003
IPCA	23	2035	88.5	2.58	0.00012	22	1459	66.3	3.29	0
IPCA	21	1588	75.6	2.21	0.00193	20	897	44.9	2.23	0.00227
Residuals	51	1577	30.9	0.9	0.66493	34	682	20.1	1	0.47979
Error	342	11721	34.3			285	5742	20.1		

The block source of variation refers to blocks within environments.

Where g'_{il} and e'_{lj} are called PCI scores for genotype i and environment j, respectively. In a biplot, genotype i is displayed as a point defined by all g'_{il} values, and environment j is displayed as a point defined by all e'_{lj} values (l = 1 and 2 for a two-dimensional biplot). Singular-value partitioning is implemented by

$$g'_{il} = \lambda_l^{f_i}\xi_{il} \text{ and } e'_{lj} = \lambda_l^{1-f_i}\eta_{lj} \quad (7)$$

where f_i is the partition factor for PCI. Theoretically, f_i can be anything between 0 and 1 although 0.5 is so far the most commonly used partition factor (Yan, 2002). In this paper we have use a value of 0.5 to give equal importance to both genotype and environment.

RESULTS AND DISCUSSION

The AMMI analysis of variance of grain yield (Table 1) showed significant effects of genotype, environment (location) and genotype by environment interaction. Location explained 84.65% of the total (G + E + GE) variation of year 2004 and 70.63% for year 2005, whereas the genotype by environment interaction and genotype captured 12.5 and 0.0029% of year 2004 and 15.34 and 14.03% for year 2005, respectively. The magnitude of genotype by environment interaction as compared to genotype suggested a possible existence of different mega environments in year 2004. The partitioning of GGE sum of squares through the GGE biplot analysis showed that PC1 and PC2 accounted 43.21 and 26.43% of GGE sum of squares of year 2004 and 58.01 and 22.14% for year 2005, respectively. The two principal components explained a total of 69.6 and 80.16% variation in the two years respectively. Nonetheless agricultural biplot literature provides no guidance concerning how much of the total variability accounted for by the first two principal components are considered adequate (Sabaghnia et al., 2012b; Yang et al., 2009). This result revealed that there was a differential yield performance among wheat genotypes across testing environment (location) due to the presence of genotype by environment interaction.

Graphical statistical methods based on GGE biplot analysis

Relationship among test environments

GGE biplot, which was based on environment focussed scaling, was used to estimate the pattern of environments (locations) as shown in Figure 1. Environment PC1 score had both negative and positive scores indicating that there was a difference in rankings of yield performance among genotypes across environments leading to cross-over G ×E interactions.

Like PC1, the environment PC2 scores had both positive and negative values. This gave rise to crossover, leading to inconsistent genotype yield performance across environment (locations). To visualize the relationship between environments, lines are drawn to connect the test environments to the biplot origin known as environment vectors. The cosine of the angle between two environments is used to approximate the correlation between them as described and used in Dehghani et al. (2009, 2010), Kaya et al. (2006), Yan and Tinker (2006).b For example locations 2,3 and 6 were positively correlated (an acute angle), location 1 and 5 were negatively correlated (an obtuse angle), and location 1 and 4 were not correlated (a right angle) in year 2004.The presence of wide obtuse angle (that is, strong negative correlations) among test environments is an indication of high cross over GEI (Yan and Tinker, 2006).

The distance between two environments measures their dissimilarity in discriminating the genotype, thus the six locations in (Figure 1a) fell into 4 apparent groups where locations 2,3 and 6 form the first group while lactations 1,4 and 5 each of them separately form their own group. The presence of close associations among some test locations in year 2004, suggest that the same

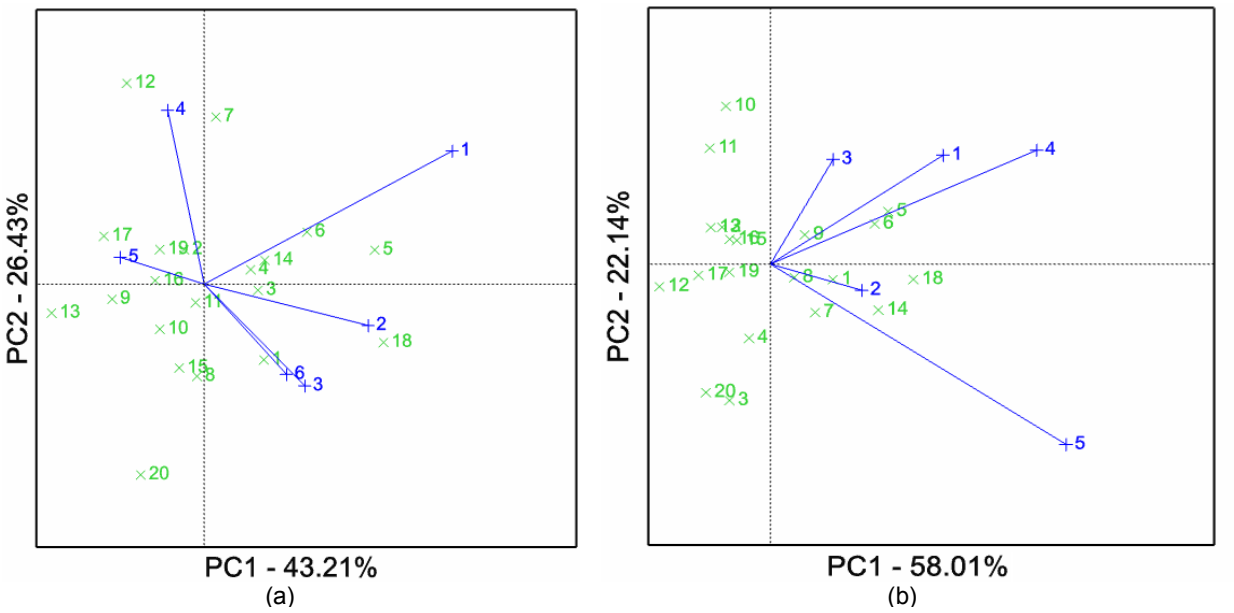


Figure 1. Scatter plot of environments (a) year 1 (b) year 2.

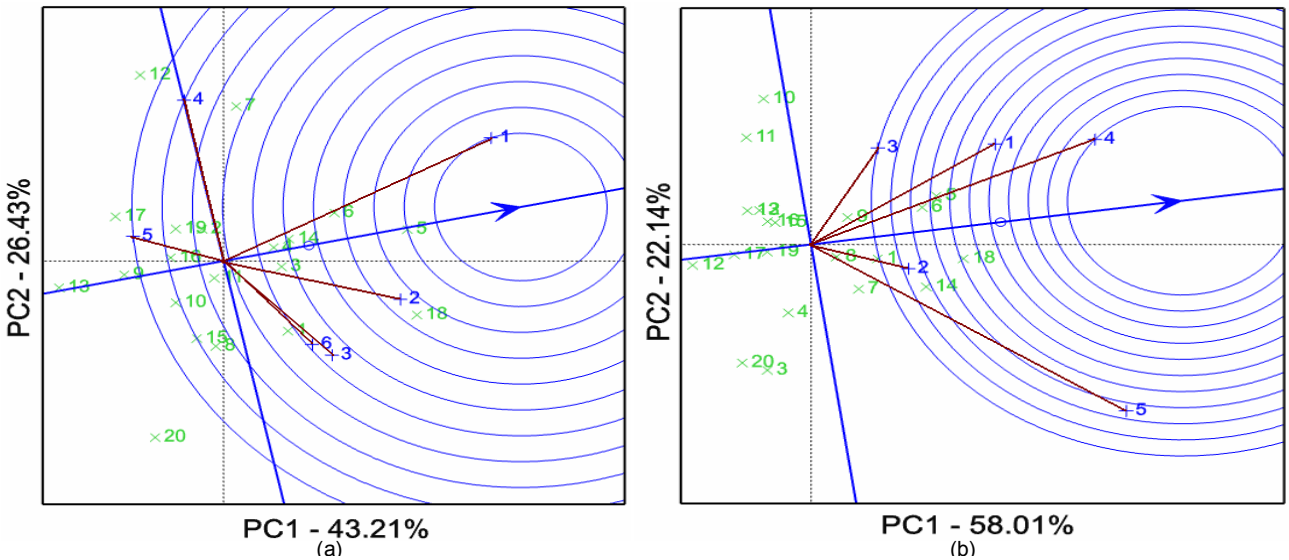


Figure 2. GGE biplot based on environment-focused scaling for comparison of the environment with ideal environment (a) year 1 (b) year 2.

information about genotypes could be obtained from fewer test locations, and hence the potential to reduce test cost (Choukan, 2010; Tukamuhabwa et al., 2012). If two test locations are closely correlated consistently across years, one of them can be drooped without loss of much in-formation about the genotypes. However, in reality the correlation consistency between the test locations vary from year to year as it shown in Figure 1. Clearly Figure 1a and Figure 2b show differing genotype and environment structure. However it should be noted that data in 2005 had only five of the location in 2004.

Discriminating ability and representativeness of the test environment

GGE biplot discriminating ability and representativeness is an important measure of the testing environments. The concentric circles on the biplot as shown in Figure 2 help to visualize the length of the environment vectors, which is proportional to the standard deviation within the respective environments and is a measure of the discriminatory ability of the environments. Therefore, among the six environments, E1 and E4 were most

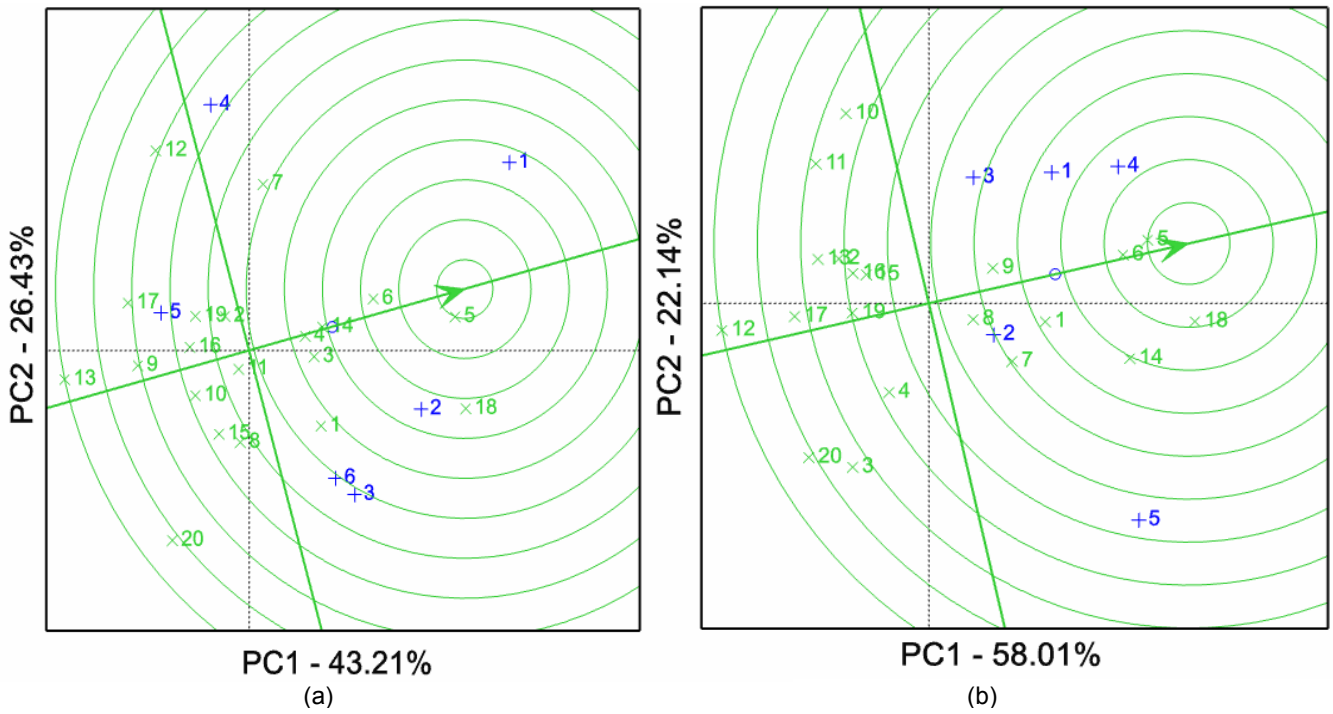


Figure 3. GGE biplot based on genotype-focused scaling for comparison of the genotype with ideal genotype (a) year 1 (b) year 2.

discriminating (informative) and E5 least discriminating in year 1; whereas in year 2 (Figure 2) E5 and E4 are most discriminating and E2 was least-discriminating. Test environments that are consistently non-discriminating (non-informative) provide little information on the genotypes and, therefore, should not be used as test environments. The average environment (represented by the small circle at the end of the arrow) has the average coordinates of all test environments, and Average-Environment Axis (AEA) or Average-Tester-Axis (ATA) (Yan, 2002) is the line that passes through the average environment and the biplot origin. A test environment that has a smaller angle with the AEA is more representative of other test environments. Thus, E1 and E4 are most representative whereas E5 and E3 least representative in their respective year. Test environments (locations) that are both discriminating and representative (e.g., E1) are good test environments for selecting generally adaptable genotypes. Discriminating but non-representative test environments like E3 are useful for selecting specifically adapt-able genotypes if the target environments can be divided into mega-environments or they are useful for culling unstable genotypes if the target environment is a single mega-environment.

Ranking genotypes relative to the ideal genotype

An ideal genotype should have the highest mean performance and be absolutely stable (that is, performs the best in all environments). Such an ideal genotype is defined by having the greatest vector length of the high yielding genotypes and with zero GEI, as represented by an arrow pointing to it (Figure 3). Although such an ideal genotype may not exist in reality, it can be used as a reference for genotype evaluation (Yan and Tinker, 2006). A genotype is more desirable if it is located closer to the ideal genotype. Thus, using the ideal genotype as the centre, concentric circles were drawn to help visualize the distance between each genotype and the ideal genotype. Because the units of both PC1 and PC2 for the genotypes are the original unit of yield in the genotype-focused scaling (Figure 3), the units of the AEC abscissa (mean yield) and ordinate (stability) should also be in the original unit of yield. The unit of the distance between genotypes and the ideal genotype, in turn, will be in the original unit of yield as well. Therefore, the ranking based on the genotype-focused scaling assumes that stability and mean yield are equally important (Farshadfar et al., 2012; Yan, 2002). Figure 3 revealed that G5, which fell into the centre of concentric circles, was the ideal genotype in terms of higher yielding ability and stability, compared with the rest of the genotypes. In addition, G6 and G14, located on the next consecutive concentric circle, may be regarded as desirable genotypes.

Mean performance and stability of the genotypes

Yield performance and stability of genotypes were

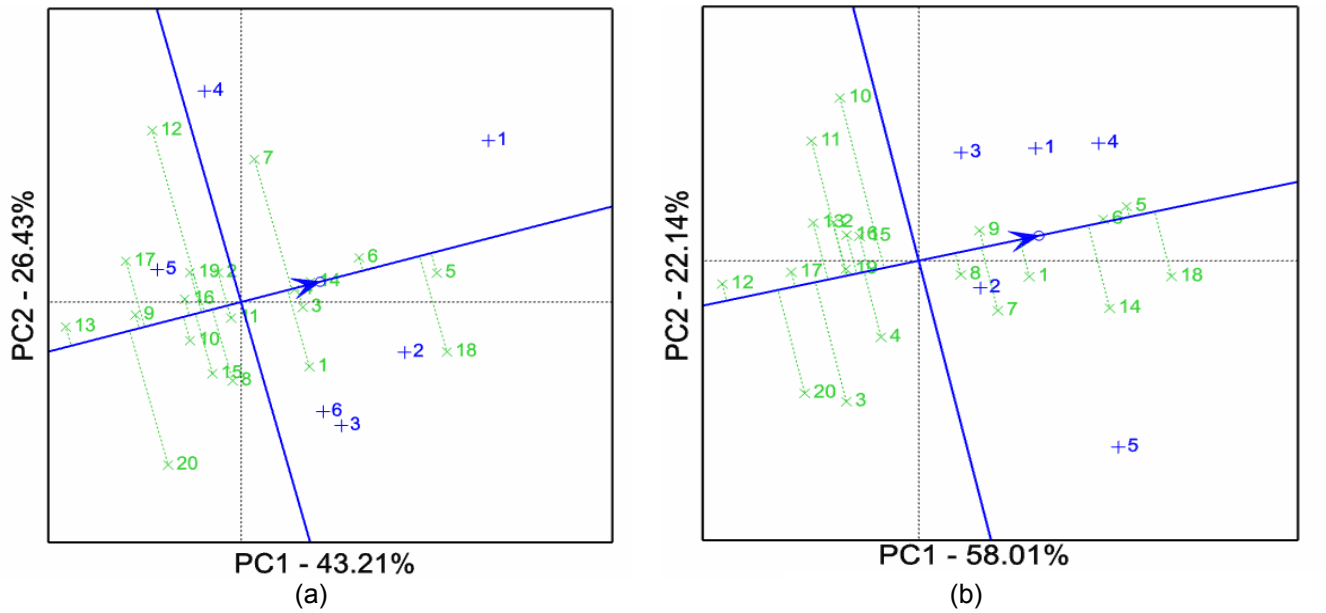


Figure 4. GGE biplot based on environment-focused scaling for mean performance and stability of the genotypes (a) year 1 (b) year 2.

evaluated by an average environment coordination (AEC) method in Farshadfar et al. (2011). Within a single mega-environment, genotypes should be evaluated on both mean performance and stability across environments. Figure 4a gives the average environment coordination (AEC) view of the GGE biplot. The single-arrowed line is the AEC abscissa, it points to higher mean yield across environments. Thus, G5, G18, G6 and G14 had the highest mean yield. The non-arrowed line is the AEC ordinate; it points to greater variability (poorer stability) in either direction. Thus, G12 and G20 ware highly unstable and below average yield, whereas G4 and G14 highly stable, were followed by G5, G6 and G3 with above average yield in the first year.

The mean performance and stability of these 20 genotypes in five locations (environment) in the second year of the trial shows some variation from the first year as it shown in Figure 4b. However G6, G5, G4 and G18 were relatively high yielding and stabile genotypes in both trial years.

Which genotype won where and mega environments with GGE bi-plot

One of the most attractive features of a GGE biplot is its ability to show the which-won-where pattern of a genotype by environment data set (Figure 5). Many researchers find this use of a biplot intriguing, as it graphically addresses important concepts such as crossover GE, mega environment differentiation, specific adaptation, etc as discussed in Yan and Tinker (2006). The polygon is formed by connecting the markers of the genotypes that are further away from the biplot origin such that all other genotypes are contained in the polygon. Genotypes located on the vertices of the polygon performed either the best or the poorest in one or more locations since they had the longest distance from the origin of biplot. The perpendicular lines are equality lines between adjacent genotypes on the polygon, which facilitate visual comparison of them. For example, the equality line between G5 and G18 in 2004 indicates that G5 was better in E1, whereas G18 was better in E2, E3 and E6. An interesting feature of this view of a GGE biplot is that the vertex genotype(s) for each sector has higher (some times the highest) yield than the others in all environments that fall in the sector (Gauch et al., 2008; Yan, 2002). These six equality lines divide the biplot into six sectors, and the environments fall into four of them (Figure 5).This pattern suggest that the target environment may consist of four different mega-environments and that different cultivars should be selected and deployed for each.

In which-win-where GGE biplot for the second year (Figure 5b), eight equality lines divide the biplot into eight sectors and the five locations fell into three of them. The mega-environment classification of these five trial location is different from the first year. This difference leads to a different wining genotype in different locations (environment) across a year.

Conclusions

The GGE biplotsof MEYT data allow visualizing the inter-relationship among genotypes including the ranking of genotypes based on both mean performance and stability, inter-relationship among environments, and

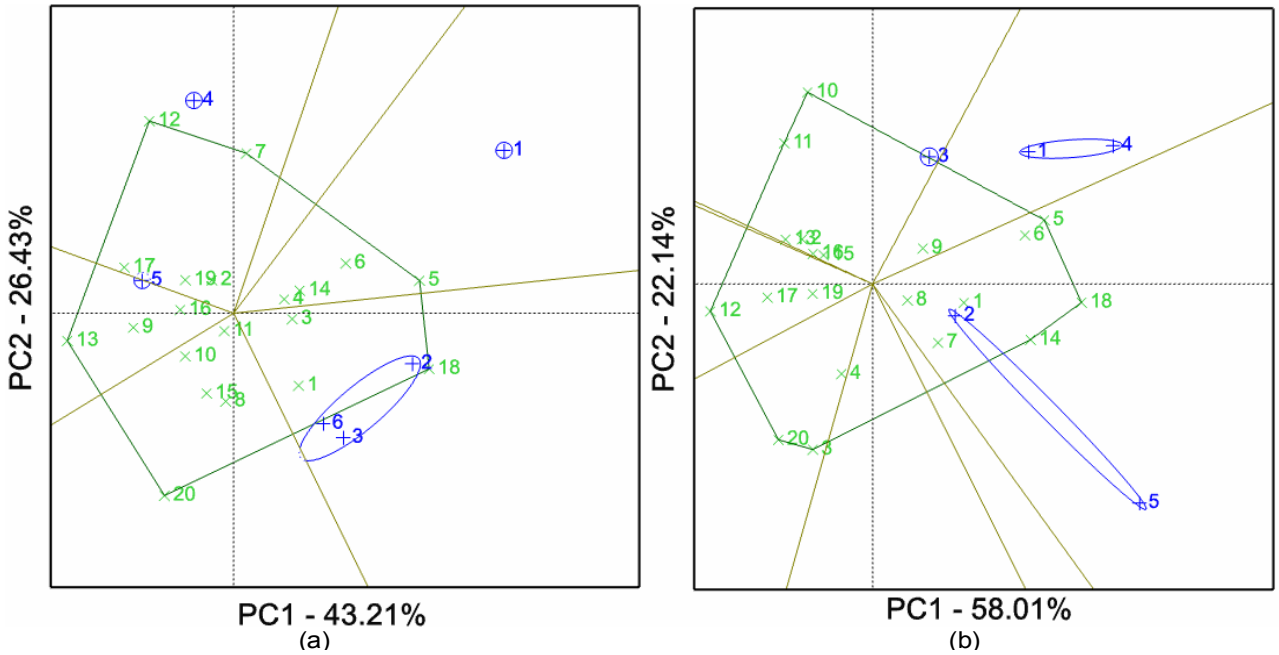


Figure 5. The which-won-where view of the GGE biplot to show which genotypes performed best in which environment (a) year 1 (b) year 2.

interaction between genotypes and environments including the which-won-where pattern. The result of this study indicated that wheat yield performance was highly influenced by the environment effect followed by the magnitude of GEI and genotype. Total yield variation by the genotype increased from 0.0029% in first year to 14.03% in the second year which had almost equal effect with the G×E interactions. These two years repeated over location data analysis result; which-win-where pattern, yield performance and stability of genotype indicate that repeatability pattern over years is the necessary and sufficient condition for mega-environment delineation and genotype recommendation. Decision making recommendation based on one year data should be done with caution.

ACKNOWLEDGMENT

The authors gratefully acknowledge Institute of Agricultural Research (EIAR) for supplying data sets used in this study.

REFERENCES

ECSA (2010 / 2011). Area and production of major crops (private peasant holdings, meher season).Technical report.

Choukan R (2010). Genotype, environment and genotype by environment interaction effects on the performance of maize (zea mays l.) inbred lines. Crop Breed. J. (1(2)):97-103.

Dehghani H, Ebadi A, YouseL A(2006). Biplot analysis of genotype by environment interaction for barley yield in iran. Agron.J., 98:388-393.

Dehghani H, Sabaghnia N, Moghaddam M (2009). Interpretation of genotype-by-environment interaction for late maize hybrids' grain yield using a biplot method. Turk. J. Agric. Fores. 33:139–148.

Dehghani H, Sabaghpour SH andEbadi A (2010). Study of genotype × environment interaction for chickpea yield in Iran. Agron. J. 102:1–8

Farshadfar E, Zali, H, Mohammadi R (2011). Evaluation of phenotypic stabil-ity in chickpea genotypes using gge-biplot. Ann. Biol. Res. (2 (6)282-292):282-292.

Farshadfar E, Mohammadi R, Aghaee M, Zahra Vaisi Z. (2012).GGE biplot analysis of genotype × environment interaction in wheat-barley disomic addition lines. Australian Crop Sci. 6(6):1074-1079

Gauch HG, Zobel RW (1996): AMMI analysis of yield trials. In M.S. Kang & Gauch H.G. eds. Genotype-by-environment interaction, pp 85-122. Boca Raton, FL, CRC Press.

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

Gauch HG, Piepho HP, Annicchiarico P (2008). Statistical analysis of yieldtrials by AMMI and GGE: Further considerations. Crop Sci. 48:866–889.

Kaya Y, Akcura M, Taner S (2006). GGE-biplot analysis of multi-environment yield trials in bread wheat. Turk J. Agric. 30:325-337.

Rao AR, PolignanoPrabhakaran VT (2005). Use of AMMI in simultaneous selection of genotypes for yield and stability. Ind. Soc. Agric. Statist. 59:76–82.

Sabaghnia N, Mohammadi M. Karimizadeh R (2012b). The evaluation of genotype × environment interactions of durum wheat's yield using of the AMMI model. Agric. Fores. 55:5–21.

Tukamuhabwa P, Asiimwe M, Nabasirye M, Kabayl P, Maphosa M (2012). Genotype by environment interaction of advanced generation soybean lines for grain yield in Uganda Africa. Crop Sci. J. 20(2):107-115.

Yan W (1999). Methodology of cultivar evaluation based on yield trial data-with special reference to winter wheat in Ontario. PhD thesis.

Yan W (2002). Singular-value partitioning in biplot analysis of multienviron-ment trial data. Agron. J. 94:990-996.

Yan W. Cornelius PL, Crossa, J, Hunt LA (2001). Two types of ggebiplots for analyzing multi-environment trial data. Crop Sci. 40:656-663.

Yan W, Hunt LA, Sheng Q, Szlavnics Z (2000). Cultiver evaluation and mega-environment investigation based on the ggebiplot. Crop Sci, 40:597-605.

Yan W, Kang MS (2003). GGE biplotanalysis : a graphical tool for breeders, geneticists, and agronomists. Boca Raton, Florida: CRC Press, Florida.
Yan W, Rajcan I (2002). Biplot analysis of test sites and trait relations of soybean in ontario. Crop Sci. 42:11-20.
Yan W, Tinker NA (2006). Biplot analysis of multi-environment trial data: principles and application. Candian J. Plant Sci. 86:623-645.
Yang R.-C, Crossa J, Cornelius PL, Burgueo J (2009). Biplot analysis of geno-type environment interaction: Proceed with caution. Crop Sci. Soc. Am. 49:1564-1576.