ESTIMATION OF GENOTYPE X ENVIRONMENT INTERACTION FOR YIELD IN GREEN BEANS(*Phaseolus vulgaris*)

BY

KIRUBASHIN NADARAJH PILLAY

Thesis submitted in accordance with the requirements for the Master of Science degree in the Faculty of Agriculture, Department of Plant Breeding at the University of the Orange Free State.

University of the Orange Free State

BLOEMFONTEIN

November

2000

Supervisor: Prof. C.S. van Deventer

Contents

Chapter 1

Introduction

The green bean(*Phaseolus vulgaris)* is consumed in the fresh or processed state as compared to the dry bean which is used dry. The green bean is referred to by many names across the world viz. snap bean, garden bean, french bean, amongst others. Although the indeterminate growth habit is dominant and generally more adaptable than determinate type, the determinate type is primarily used for domestic production. The green bean is believed to originate from the South and Central Americas.

In South Africa, green beans can be produced throughout the country. The main commercial production areas are the North and Northwest Province, Kwazulu-Natal, Mpumulanga Province, Eastern Cape and Western Cape. There is approximately 4000ha of green bean production annually, with the average yields of between 8-14 tons/ha. The average national consumption of green beans is estimated to be at 40 000 tons annually, with 40% of this consumed as frozen or canned products.

There are two main production limitations viz., water availability and disease pressure. Green beans are high consumers of water thus water needs to be available for commercial productions. Our climate is highly favourable for disease development especially rust(*Uromyces appendiculatis*) thus the disease pressure is high on green beans. Farmers use both genetic resistance and chemicals in the control of diseases. High temperature is also an adverse factor which is dealt with by appropriate planting dates.

The green bean is considered a luxury food with its main contribution being a source of vitamins. Relatively little work is done on genotype x environment interaction (GxE) and minimal information is available on the response of green bean varieties to different environments. The objectives of this study are therefore:

1) To detect the presence of GxE in green beans.
2) The determine the appropriate analytical method to use in the study of GxE in green beans.
3) The optimum allocation of resources (locations, years and replications) in a green bean evaluation program.

Chapter 2

Literature Study

Nature of Genotype x Environment Interaction

Biological or statistical concepts can be used to define genotype x environment interactions(GxE). The GxE system can be defined as a combination of a population of genotypes and a population of environments that are relevant to the objectives of a breeding program. The genotype is the genetic factors that influence the expression of the trait under investigation and the environment is all things non-genetic that influence the expression of the trait. GxE can be detected statistically by a significant difference in the relative performance of the same group of genotypes in the different environments. The biological basis for this interaction is that the contribution of the different genes that control the trait, varies in the different environments. The statistical tests employed in the study of GxE are attempting to detect this biological basis for the interaction (Basford and Cooper, 1998).

GxE complicates analysis of data by confounding the genotypes observed performance with its true value (Crossa, 1990; Freeman, 1973). Should there not be any such interaction, then selection is made much easier since a high performing variety in one location will result in a similar performance in the other locations (Basford and Cooper, 1998). The degree of uncertainty introduced by GxE when deciding on the superiority of a variety will complicate selection for broad adaptability. Genotype inferences, with the presence of GxE, can only be made in the particular set of environments in which the experiment was done. Assessing the performance of a variety without looking at its

interaction with the environment will result in an analysis that is incomplete and may result in inaccurate estimates of yield. Thus the application of resources to study these interactions through multi-environment trials (MET) is essential (Crossa, 1990).

GxE that cause rank changes of genotypes amongst environments are the components of the interaction that will potentially complicate selection (Basford and Cooper, 1998). Not all changes in rank will complicate the selection process. The impact of crossover interactions should be assessed on their impact on selection response. Allard and Bradshaw(1964) found that there are 10^{145} possible interactions in a trial with 10 varieties at 10 locations and only a small number of these may be of importance to the breeder. The estimates of the magnitude of GxE is only a rough approximation since a very small number of possible types of interaction are used in the determination. In the multitude of possible interactions, there is only one where a single variety will be the best in all environments.

GxE for yield or yield stability is very complex (Kang and Magari, 1996). Yield stability is genetically controlled. The amount of genetic variation is estimated by the statistics available and the amount of environmental variation. The interaction with the environment can be in many forms, for example, disease and pest resistance; in an environment with a disease problem, a variety with resistance to the disease will exhibit a more stable performance than a variety without the resistance. The contribution of individual factors to GxE needs to be determined. Dealing with the more important factors can lead to improved yield stability. A misconception amongst some researchers is that selecting for both yield and stability simultaneously leads to a reduction in yield.

The main objective of yield trials is to predict future performance using past data. The mean yield of varieties selected using methods combining yield and yield stability would be lower than if varieties were selected on yield alone. This lower yield is based on previous data and does not necessarily translate into lower yields in production areas in the future.

There are two broad categories of GxE namely, defined causes and undefined causes (Basford and Cooper, 1998). The two categories can be distinguished on the basis of the level of understanding of the environmental and genetic factors that are influencing the system. Defined causes can include such items as specific disease, soil constraints and experimental error. Undefined causes can include all factors which influence the system and are unexplained.

There are usually three stages in the testing of a new cultivar. They are first tested at a single location where the analysis involves comparisons of means, selected cultivars are then tested in other locations (comparisons are made over locations) and the third stage of testing is over years where reproducibility of performance is being tested. The second and third stage differ from the first stage in that the analysis not only includes the comparisons of means, but also includes a GxE analysis (Basford and Cooper, 1998). Data for GxE studies is usually gathered by testing varieties in different environments. A large component of GxE can be attributed to interactions with years and location, it is therefore important to test the varieties across both years and locations. There are three fundamental aspects to the data collected in multi-environment trials: a) structural patterns; b)nonstructural noise and c) relationships

among genotypes, environments and genotype and environments together (Crossa, 1990). Cooper *et al* (1993b), developed a model for the evaluation of germplasm in a multi-environment system. There were three major recurring steps: a) the sampling of the genotypes and environments, b) analysis of the results of the trials and c) selection and prediction of response in the target environment. Multi-environment trials have three major objectives in agriculture: a) to accurately predict the yield of a variety based on a limited set of data, b) to determine the stability of yield and the response pattern of genotypes or agronomic traits across environments and c) for the reliable selection of genotypes that will perform well in future years and other locations (Crossa, 1990).

The environments that are used in GxE studies are sampled from a population of environments that are assumed to represent some target environment being investigated (Comstock, 1977). The target environment is sampled by testing the genotypes at a number of locations over a number of years (Basford and Cooper, 1998). Environments are often defined as site-year combinations and in doing so, there is no indication of the biophysical factors that affect the traits measured. A question often asked is: do the environments sampled for MET represent the target environment? Due to the large number of different environments possible, it is most likely that sampling is inadequate and will result in inaccurate estimates of the genetic merit of the individuals. One of the objectives in selection in current MET is to identify environments that will maximize indirect selection response for future years (Cooper *et al*, 1993a). There are various methods that may assist in the reduction of the effect of sampling such as pattern analysis and the use of crop models (Basford and Cooper, 1998). Specific screening environments can be set up to aid in screening for example,

disease resistance. However, there must be useful genetic variation and the ability to manage reliably the environment that is required to manipulate the genetic variation. The relevance of the environment conditions to the target population of environments is very important for the establishment of specific screening environments. The relevance of multi-environment trials to performances in farmers lands is still being debated. Each location can constitute a different environment and thus selection can be considered to be indirect. Pederson and Rathjen (1981) observed that sites other than research stations were always more preferable to carry out selection for target environments. The environments selected for testing must be as similar as possible to the target environments for maximum selection gains and probability of successful production of the variety. Cooper *et al* (1997) investigated how well yield performance of wheat in the target environments were predicted by well-irrigated (low stress) breeding nurseries. They found that the low stress nurseries were able to predict performance in similar target environments. The predictive value decreased as stress increased.

The repeatability of the interactions in the GxE system will have a major influence in determining the level of understanding of the interaction and the associated adaptation. A high proportion of the interaction is not repeatable. Tests need to be developed that would be able to detect the repeatable component of the interaction and those that are consistently observed need to be incorporated into a genetic model that is used to improve a quantitative trait (Basford and Cooper, 1998). The proportion of the contribution of the variance components for genotype relative to the total variation gives an indication of the repeatability of the trait (Odendaal and van Deventer, 1987). If yield

was the trait under consideration, the higher the variance component for genotype, the more predictable the variety performance for yield will be.

Strategies to Deal with Genotype by Environment Interaction

When GxE is present, it must be taken into account when selecting for performance at other locations (Basford and Cooper, 1998).There are many different opinions on the appropriate strategies to deal with GxE, for example, agroclimatic analysis (Nix, 1975), more critical investigation of plant breeding trials (Rathjen, 1994), statistical genetical methods (Comstock and Moll, 1963) and molecular markers (Beavis and Keim, 1996). The breeding material needs to be tested under as similar conditions as the environment in which it is to be produced in the future (Allard and Bradshaw, 1964). The major successes in dealing with GxE by plant breeders was largely as a result of careful observation and interpretation of experimental results obtained in farmers' fields (Basford and Cooper, 1998). The breeding methods that reduce the effect of GxE when selecting superior varieties would result in more rapid genetic improvement in a breeding program. Although GxE may be partitioned statistically, the biology underlying the performances must be understood before one can determine the appropriate combination of genes. There are a number of statistical methods that can be used to investigate GxE, but the investigation should go beyond just this aspect and consider aspects such as connections between the experiments and the biophysical components of the environments (Basford and Cooper, 1998; Crossa, 1990). More adequate models will have to be developed to explain GxE and these models must enable us to identify and manipulate positive components of these interactions (Basford and Cooper, 1998). A breeding program has the ability to utilize GxE, whereas optimal exploitation of

specific adaptational requirements is the aim of cultivar trials (Marais, 1986).

Lin and Binns (1994) suggested that the study of GxE can be classified into three groups based on goals: 1) finding a suitable model to explain the structure of the interaction or to predict it, 2) estimation of the size of the interaction so that accurate estimates of other genetic parameters(for example, heritability) can be obtained and 3) selection of superior cultivars and finding the most suitable locations for testing so that good cultivar recommendations can be made.

Significant genotype x location interaction (GxL) suggests that there are different environments within the region. The breeding program should therefore be directed towards producing a number of varieties for the region, each adapted to one of the environments (Basford and Cooper, 1998; Allard and Bradshaw, 1964). Genotype x year (GxY) and genotype x year x location (GxLxY) interaction is almost always larger than GxL. The year forms an important part in the unpredictable component of the environment. If GxY is significant, varieties need to be tested over years. By testing over a number of years and locations, the chances of finding an adaptable variety are greater (Allard and Bradshaw, 1964).

Adaptability over the entire test region is referred to as general adaptability and specific adaptability is where the cultivar is adapted to a specific sub-region. The application of these terms needs some caution, because it is dependent on the size of the area in the analysis (Lin and Binns, 1994). Using the random model, selection is directed towards broad adaptation (Basford and Cooper, 1998). Not all GxE are completely random,

there exists some repeatable and heritable patterns within the complex that can be used to exploit specific adaptation during selection. The fixed model can be used to determine specific adaptation. There are many cases where broad adaptation was used successfully for example, extensive use of semi-dwarf wheat cultivars. There are also examples of specific adaptability being used successfully. The larger the size of the interaction component, the more difficult it is to identify broadly adapted genotypes. Broad and specific adaptability are in most cases treated as being mutually exclusive although there is no clear definition of the genetic or environmental basis of either. A clearer definition of the breeding objectives, that is, broad or specific adaptability is required for better advancement in the program. If the yield of a cultivar is high relative to the other cultivars, it is considered to be well adapted to the region (Lin and Binns, 1994). Stability on the other hand deals with the variability of the yield, thus a stable cultivar has low variability.

Whether broad or specific adaptability is required in a variety, is dependent on the diversity of environments encountered in the target area and the resources available for the breeding program (Shorter *et al*, 1991). Broad adaptability programs cost less than when breeding for specific adaptability. A potential cost of using broad adaptability is developing a variety that is broadly adaptable but would not perform as well as a specifically adapted variety. This is more likely to occur if there is a large diversity in environments in the sub-set of environments. By exposing a number of genotypes to different environments, widely adapted genotypes (high average yield, low GxE) and specifically adapted genotypes (high average yields, high GxE and low average yields, high GxE) can be identified (Ceccarelli, 1989). Genotypes with low average yields and

low GxE are undesirable. The varieties that yield the highest in the high yielding environments are generally poor in stress environments, and those varieties that perform well in stress environments cannot be selected if grown only in optimal environments. Stress environments are environments where conditions are below optimum. The perceived adaptability of a variety is dependent on the definition of the stress environment. Breeding specifically for stress environments, that is, specific adaptation has been less accepted because it is believed that the environmental effects in stress environments are larger than the genetic effects, thus the heritability and response to selection will be less than in favourable environments. Broadly adapted varieties do exist but only within a given range of environments. Specific adaptation makes use of the GxE that is encountered whereas broad adaptation looks for low GxE and high yield. Selection for broad adaptation results in the selection of varieties that are adapted to environments were the yield potential of the genotypes can be expressed. In high stress environments, it is essential to develop varieties with high yield stability than with high yield potentials. Wright(1976)(cited by Bramel-Cox, 1996) concluded that breeding for broad adaptability is equal to selection for specific adaptability but never greater. If the interaction with environments is very large, breeding for broad adaptation is less effective.

Indirect selection has the potential to be used to identify varieties with broad or specific adaptation from international trials (Cooper *et al,*1993c). It was found that there were positive correlated responses between locations in Australia and Mexico. It is believed that these correlations can be determined for different locations in different countries, allowing different breeders to exploit this response for sourcing germplasm for their

breeding programs. Using this approach, breeding lines from foreign countries can be selected for adaptability prior to testing in the home country. This allows for more efficient selection of germplasm.

A number of breeders believe that specific adaptation can be exploited (Basford and Cooper, 1998). For effective exploitation of the specific adaptability, the nature and causes must be established (Allard and Bradshaw,1964; Kang, 1990). It would seem that in general, a detailed understanding of the causes is not attained.

Another strategy that will aid in the understanding of GxE is the characterization of the environment (Basford and Cooper, 1998). The environment can be separated into predictable and unpredictable components (Allard and Bradshaw, 1964). The predictable component includes characters such as soil type, whereas the unpredictable component consists of factors such as rainfall and temperature. The crop is usually the best indicator of the environment. Westcott (1986) pointed out that there are large variations in the environment but environmental measurements are often not available. There is a need to define the environment with improved data collection and appropriate analysis of the environmental variables. Often reports of results of trials fail to include agency factors, that is, factors under the control of the farmer such as a soil analysis and land preparation. Most of the effort in the study of GxE is being put into understanding the genotype responses and relatively minimal time is spent on defining the environment, this may be the cause for the lack of more detailed understanding of GxE.

As mentioned earlier, environments are often defined as site-year combinations, the environment effect can be defined further by considering specific factors that may influence the expression of the traits under investigation. Eisemann *et al* (1990) argued that greater attention needs to be given to the definition of the key environmental factors if significant progress is to be made in the understanding of complex undefined GxE. They also noted that there was a difference in the quality of the definition of the environment when it came to biotic or abiotic factors. Where biotic factors were involved in the interaction, a lot more work is done to define the cause of the interaction as compared to when abiotic factors are involved. In the latter, terms such as "high stress" or "low stress" environments are usually used to define the environment. The definition of the environment for which breeding is being done, is essential for an effective crop improvement program (Shorter *et al*, 1991).

The effect of different environmental variables on GxE was investigated by Gorman *et al* (1989). They found that their environmental index and rainfall were the main contributors to the GxE in sorghum. The environmental index was a measure of differences in fertility and/or cultural practices. It was also said that there are factors other than their environmental index and weather variables that contributed to the GxE. Grafius (1958) proposed a biological measurement of the environment. A standard set of oat varieties (with known adaptation and responses to disease) were to be grown and their performance evaluated. The season was then to be classified according to these performances. Selection criteria are then adapted depending on the classification of the environment. In this way, the probability of discarding good material can be reduced, especially early in the testing program where seed resources limit trialing to a single trial

in most cases.

The use of exact physical measures of the environment in the assessment of the environment is not always possible since the exact nature of the variable is not always known (Fripp, 1972). The use of exact measures also becomes difficult because more than one environmental variable can be changing at any one time and the contribution, in kind and magnitude, to the biological response of the genotype is not always known. The different methods of estimating the environment, need to estimate the environmental values precisely else this may lead to different values for adaptability and stability of the genotypes. Although it is preferable to have an independent environmental index in the joint linear regression in the study of GxE, the use of a non-independent environmental index as proposed by Finlay and Wilkinson (1963) is adequate, although the proposed index induces a bias on the regression coefficient. The interpretation of the data was minimally influenced by the choice of methods used in assessing the environment.

The response to selection can be predicted from estimates of components of variance, heritability and genetic correlations among environments when a completely random model is been adopted (Basford and Cooper, 1998). Response to selection can be direct or indirect genetic improvement. Direct is where the environments are from the target population of environments and indirect is where the environment is from other target populations. Direct selection is the most effective method (Simmonds, 1991). Selection to exploit the positive components of GxE is based on the ability to distinguish between repeatable and non-repeatable GxE (Basford and Cooper, 1998). Non-

repeatable interactions can be treated as a source of error and can be used for selection of broad adaptation, whilst repeatable components can be used for specific adaptations. A better definition of the environment in GxE can be used in selection to accommodate repeatable components of the interaction.

A selection strategy in which the number of environments used in the test would increase as the number of lines is decreased is suggested by Brennan *et al*, 1981. Large breeding populations would be tested at a small number of environments, the environments chosen to allow for maximum prediction of performance of the lines in a single year. The best performers (which will be a few lines) will be tested over a larger number of locations to provide estimates of performance across locations. The testing over the larger number of locations would be repeated to provide information on specific or general adaptability of the lines.

The primary objective of any multi-environment trial program is to optimize selection amongst genotypes (Cooper *et al*, 1995). This selection is an indirect selection for performance in environments that may be encountered. Response to selection can be improved by choosing the test environments so that they are as similar to the target environments. Selection for broad adaptability involves selection for a target population of environments. If the test environments can be repeated, this will increase confidence in the predictions made when carrying out variety selection. Where little is known about the target environment, a random sample of environments can be used. Where information on the environment is available, a managed or controlled environment approach can be used where the managed environment is representative of the target

environment. All environmental influences in the target environment will have to be taken into account in the setting up of the managed environment, at this stage it is unrealistic that this will be achieved. The managed environments can be made to resemble the target environment as close as possible. Where the nature of the target environment can be clearly defined, the managed environment approach may be useful.

A selection index that will be able to cope with GxE was proposed by Louw (1990). With this index, greater weight was given to the performance in the target environment. It was found that the index for the local ranking was more efficient in ranking the genotypes than ranking without the index. The joint index proposed allows for GxE.

It has been found that alternating selection between environments in early generations allows for the selection of varieties with wider adaptability than selecting in a single environment (St-Pierre *et al*, 1967). It was also found that selection in certain environments results in more adaptability than others. Stress environments differentiate the varieties more than non-stress environments. Selection for widely adapted varieties will come from environments that allow for the best expression of wide adaptation genes.

GxE effects can be reduced by subdividing a region into smaller sub-regions (Horner and Frey,1957). The differences in performance within the sub-region will then more likely be due to varietal differences as opposed to the interaction of varieties with the environment. The potential gain in selection and recommendation of a variety within a

sub region will be larger than the potential gain achieved in selecting for the region as a whole. Various methods have been proposed to determine the appropriate locations to use in a breeding program to maximize selection response. Some of these methods include reduction of the GxE variance by varying combinations of locations (Horner and Frey, 1957), heritability of correlated response (Pederson and Rathjen, 1981), correlation of varietal performance at one location to the performance over a large area (Hamblin *et al*, 1980), the grouping of environments with similar GxE using the cluster method (Alagarswamy and Chandra, 1998; Lin and Morrison,1992) and the additive main effects multiplicative interaction (AMMI) analysis (Gauch and Zobel, 1996). Alagarswamy and Chandra (1998) believe that Pattern Analysis is more suitable for environmental classification than AMMI. When locations are grouped, it is implicit that selection in one environment would result in a greater correlated response in another environment that is within the same group (Cooper, *et al.,* 1993a), that is, indirect selection will be effective. The similarity between environments can be based on the pattern of discrimination among the genotypes expressed in the environments, which in turn can be used to reduce the number of testing sites required. Classification of the environment is an important prerequisite for the effective targeting of multi-environment trials to the target environments and also for the characterization of pattern of adaptability, that is, broad or specific (Alagarswamy and Chandra, 1998). Saindon and Schaalje(1993) used AMMI, cluster analysis and genotype-environment mean square decomposition to establish the number of locations required for efficient testing of dry beans. The cluster analysis accounted for more of the GxE sum of squares than the first IPCA axis in the AMMI analysis. The first two IPCA axes in the AMMI analysis accounted for most of the GxE sum of squares indicating that these two IPCA axes

were the essential elements of the GxE. The amount of GxE sum of squares accounted for by the first two PCA axes were more than that of the cluster analysis. The interpretations from the AMMI analysis were consistent with the interpretations from cluster analysis.

New tools are being developed to aid in the study of GxE. Molecular markers are being developed for the better understanding of genes that contribute to the GxE. The methodologies such as RFLP, AFLP and QTL analysis (Asins *et al,* 1994) have been used with varying degrees of success (Basford and Cooper, 1998). The use of QTL analysis in GxE studies has been reviewed by Beavis and Keim (1996). Consistent results on QTL by environment interactions could not be found using the Interval Mapping approach. The Multiple-QTL Model (Jansen *et al*, 1995) may be more effective in the analysis QTL by environment interactions than the Interval Mapping method.

Crop modelling provides a means for the assessment of quantitative traits. With models, hypothetical genotypes can be set up that can be used in the investigation of the effects of different traits on adaptation (Shorter *et al*, 1991). With more development, crop simulation models (Hammer and Vanderlip, 1989) may be used in determining response to a range of environments. These models, however, will not replace variety trials completely, but rather improve the efficiency of conducting these trials.

Analytical Methods

There are five broad categories of analytical methods that have been used in the study

of GxE, 1) correlation; 2)stability and regression analysis; 3) heritability and variance components; 4) general combining ability (GCA), specific combining ability (SCA), additive and dominance models; 5)pattern analysis (Basford and Cooper, 1998).

Statistical analysis of multi-environment trials and experimental design are used to eliminate as much of the unexplainable and irrelevant variation (noise) present in the data (Crossa, 1990). If data is the sum of pattern and noise, the analysis should be able to separate out as much of the pattern as possible while eliminating maximum noise (Freeman, 1973). The more detailed information will reveal more of the structure underlying the data which would be more beneficial in the study of GxE.

The majority of research in the study of GxE so far has been in the development of statistical methodologies to quantify the magnitude of the interaction, characterize the nature of the interaction and to develop selection strategies using these statistical procedures (Cooper and DeLacy, 1994). The breeder needs to use whichever method is suitable to his needs (Hohls, 1995). The effective use of these analytical methods is however reduced by the lack of understanding of the biophysical basis for differences detected by these analyses (Basford and Cooper, 1998).

In statistical terms, the guarantee that a variety would perform consistently as expected is based on Type I and Type II error rates for the selection criterion (Kang and Magari, 1996). If the null hypothesis is $h_0: \mu_1 \geq \mu_0$ where μ_1 is the variety performance and μ_0 is the mean yield of all genotypes, a Type I error would occur if the hypothesis is rejected when it is true. A Type II error would occur if the hypothesis is accepted when it is false.

If the Type I error is committed, the farmer may not lose because he is not using the best variety. The economic loss to this farmer is dependent on the alternative variety chosen. If a Type II error is committed, an inferior variety can be recommended. The farmer would definitely suffer an economic loss if the inferior variety is used. Thus a Type II error are more harmful to growers than a Type I error.

Phenotypic performance of genotypes in a combination of environments can be use to calculate the amount of variation attributed to genotypic effects, environmental effects, GxE effects and experimental error (Basford and Cooper, 1998). The partitioning of GxE into those that cause rank changes and those that do not cause rank changes may be useful since it is this component that causes rank changes that can complicate the selection process. As the target populations are a sample of environments, multi-environment trials are subject to sampling errors. This sampling variance and alternating directional selection that is inherent in multi-environment trials can cause varieties that were selected previously as being superior, to be discarded.

Analysis of Variance

Total yield variation in a Genotype-Environment system having G genotypes and E environments can be partitioned into a) additive main effects for genotypes and environments and b) non-additive effects due to GxE (Crossa, 1990). This system can be represented by the following model:

$$Y_{ij} = \mu + G_i + E_j + (GE)_{ij} + \epsilon_{ij} \qquad (1)$$

where μ is the general mean, G_i, E_j and G_{ij} represent the effect of the i^{th} genotype , j^{th} environment and the interaction of the i^{th} genotype with the j^{th} environment respectively. ϵ_{ij} is the random error. This model suggests that the performance of the variety is not only dependent on the additive levels of G and E separately, but also on the specific combinations of the levels of G and E.

The main criticism against the use of analysis of variance in GxE analysis is the lack of a suitable unbiased F-test for testing of the null hypothesis due to the homogeneous or heterogeneous nature of the error mean square (Crossa, 1990). Also, analysis of variance does not elaborate on the underlying structure of the GxE. The pattern of response of genotypes and environments is also neglected in the analysis.

Variance components due to different sources of variation, including genotype and GxE components, can be estimated from the analysis of variance of multi-environment trial. Variance components are important in measuring yield performance, as a large proportion of the error is due to GxE. The estimate of the size of the interaction is important in efficiently estimating the genotype effects and for the determining of the optimal allocation of resources. The variance components can be determined by solving simultaneous linear equations (called "estimated mean squares") that are estimated by the observed mean squares in the analysis of variance. REML (restricted maximum likelihood) analysis can also be used in estimating the variance components. REML is capable of handling both balanced and unbalanced data efficiently (Crossa, 1990). The relative size of the variance components indicate the relative importance of the corresponding sources of variation ie. the higher the component, the more influence the

corresponding source has on variation (Miller *et al*, 1962).

Using variance components, Miller *et al*, (1959) found that in North Carolina, the GxL and GxY were not significant for lint yield in cotton. This meant that the performance of the varieties were basically the same in all three years and nine locations of testing. This is true if the environments and years were an adequate sample of the years and locations. The lack of a significant GxL, suggests that dividing the testing area into smaller sub-areas will not improve the estimates of performance and is therefore not necessary. The second order interaction ie. the GxLxY interaction was highly significant. The significance of the second order interaction lead them to conclude that cotton testing needs to be done over a number of different environments (years and locations) in the breeding area.

Linear Regression

Joint linear regression has been extensively used in plant breeding to determine the yield stability of genotypes (Crossa, 1990). The linear regression approach is the most widely used method for selecting high yielding and stable genotypes (Hernandez *et al*, 1993). The GxE component in Equation 1 is partitioned into components due to the linear regression (b_i) of the i^{th} genotype on environmental mean and a deviation from regression (d_{ij}), thus the model becomes:

$$Y_{ij} = \mu + G_i + E_j + (b_iE + d_{ij}) + \epsilon_{ij} \qquad (2)$$

The first limitation of the application of linear regression to the study of GxE is that the

genotype means are being regressed on a non-independent variable viz. environmental marginal means (Crossa, 1990). This violates one of the assumptions for a valid linear regression ie. both sets of values in the regression must be independent. This interdependence may have less of an impact if the number of entries is large. Another limitation is that the error is not statistically independent as the sum of squares for deviation cannot be divided orthogonally among the genotypes. The third criticism against this method is the assumption that the relationship between the interaction and the environmental means is linear. The regression can also be influenced by the locations and genotypes included in the regression (Westcott, 1986), the magnitude of the influence is variable and this could lead to erroneous conclusions.

Finlay and Wilkinson (1963) developed a model for the study of GxE based on linear regression. For each variety, a linear regression of each variety's yield on the mean yield of all the varieties at each location in each season was calculated. The mean of all the varieties at a location in a season (site mean) was used as an index of the environment. A logarithmic transformation was used to induce more linearity for the regression of variety means on site means. Varieties that had a regression coefficient close to or equal to 1 (b=1), have average stability over all environments. If the same variety had above average yields in all environments, this indicates that it has general adaptability; if the yields was below average in all environments then it is poorly adapted to all environments. If a variety has a regression coefficient significantly greater than 1 (b>1), the stability is below average and it is specifically adapted to high-potential environments. A variety with a regression coefficient significantly less than 1 (b<1), has above average stability. This variety is less sensitive to environmental change and will

produce above average yields in the low environments, but the yields will not improve as the potential of the environment increases. This kind of variety is adapted to low-potential environments. Finlay and Wilkinson(1963) used the variety mean yield as a second index of variety performance and plotted the regression coefficient of each variety against the variety mean yield as given in Figure 2.a. They noted that the distribution of varieties with regard to mean yield was dependent on the sample of seasons and years included in the analysis. The ideal variety with general adaptability is described as one with maximum yield in high potential environments with maximum phenotypic stability. It was also noted that testing over seasons is essential for more relevant recommendations because seasonal fluctuations are inherent in the genotype-environment system.

The method of regression analysis as proposed by Yates and Cochran (1938) for the study of GxE was unused in the study of GxE until 1963 when it was developed further and applied by Finlay and Wilkinson (1963) (Crossa, 1990). The major contribution to the study of GxE by Finlay and Wilkinson (1963) was the definition of the index that quantified the environmental effect (Lin and Binns, 1994; Knight, 1970). This violated the assumption of the independent regressor variable used in regression analysis (Freeman and Perkins, 1971) but Freeman (1973) withdrew his criticism on the basis that this did not matter if the dataset was large. Finlay and Wilkinson's (1963) model is a descriptive model rather than a predictive model (Lin and Binns, 1994). As long as the R^2 value is large, the regression coefficient is a useful indicator of response characteristics.

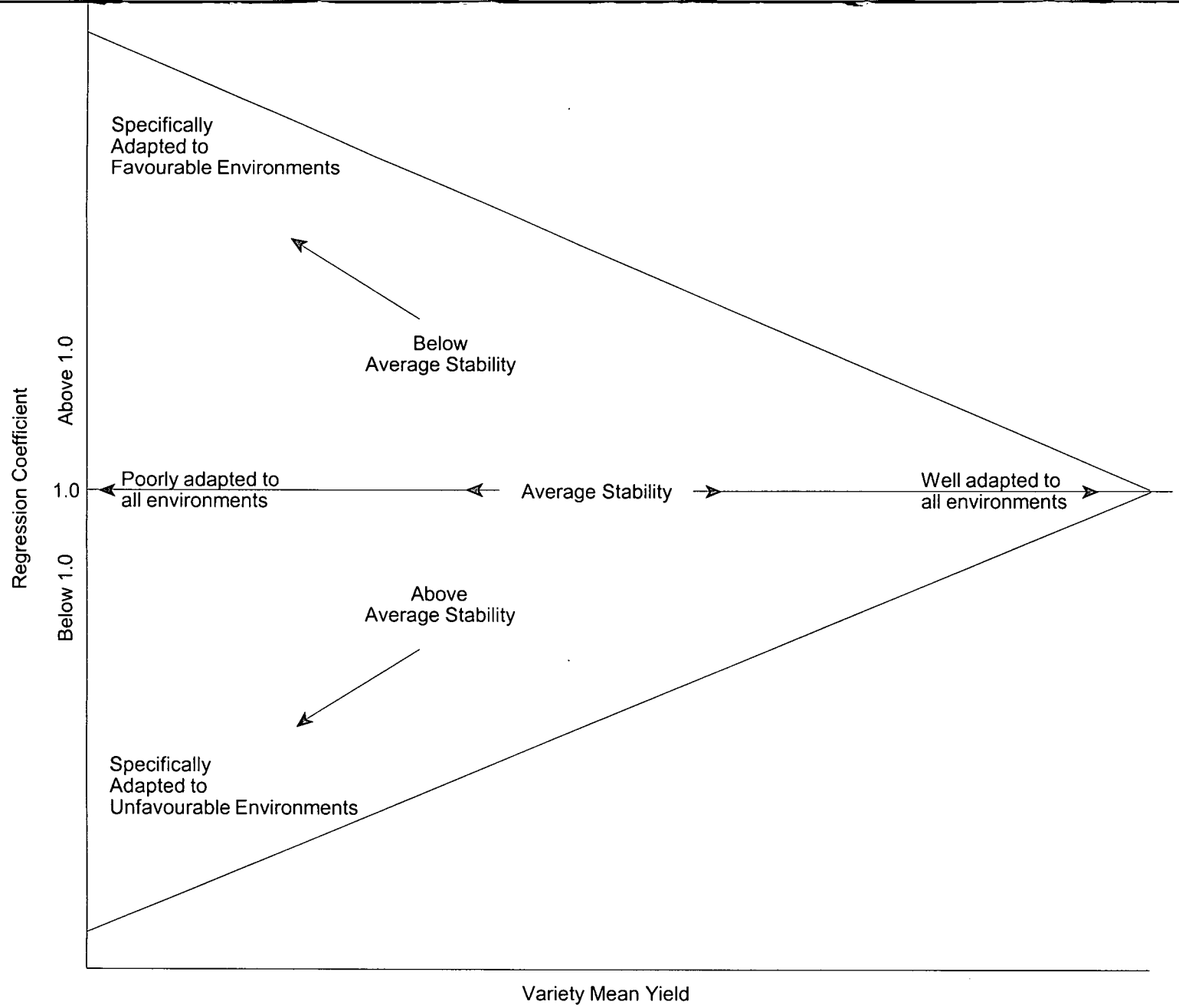


Figure 2.a.Generalized interpretation of regression coefficients plotted against variety yield (Finlay and Wilkinson, 1963)

Knight (1970) noted that a few of the genotypes may influence the analysis of Finlay and Wilkinson (1963). The interpretations are therefore limited to the set of varieties in the study. Any genotype that differs from the majority will show a larger deviation from its regression line. Finlay and Wilkinson's (1963) terminology in interpreting the regression coefficient is more indicative of the response of the variety than Eberhart and Russell's (1966). The effectiveness of the logarithmic transformation to increase linearity is dependent on the data. If the data is already linear, the transformation may induce curvature. Thus, it must be tested whether transformation of the data is necessary or not. Although the definition of the environment using the environmental means is acceptable, a better definition of the environment is required. In Finlay and Wilkinson's (1963) analysis, biological interpretation of the results would be difficult. The Finlay and Wilkinson (1963) analysis will be a valuable tool to plant breeders as long as the limitations of the method are taken into account when interpreting the results.

Many breeders have applied regression techniques for the analysis of adaptability and stability with success. Hardwick and Wood (1972) said that most of this work lacks a theoretical foundation. As a working hypothesis, the assumption that the response of a variety to environment is linear, seems reasonable. This assumption will only be seriously wrong if the performance is a discontinuous function of some environmental variable. Multiple regression has the advantage of making the regression coefficients for each genotype independent of the number of genotypes included in the analysis.

Freeman (1973) felt that the linear regression method was suitable for the study of GxE when the response was linear. Problems could be encountered when the data moved

away from linearity. He also predicted more use of multivariate techniques in the study of GxE as computational power increased.

Becker and Leon (1988) said that the linear regression technique is of little use if the heterogeneity of variance due to regression is non-significant. The regression coefficient is of little use if included in the definition of stability. The regression coefficient merely gives more information on the average response of the genotype to the environment. They also pointed out that there is no reason to reject the regression approach for the study of GxE. The multivariate techniques available may be able to give a more detailed analysis of the GxE, but these techniques will not replace the regression approach because of its simplicity and biological relevance. The most severe limitation of the regression approach is the low repeatability of the regression coefficient and the large number of environments required for a reliable estimate of this statistic.

AMMI

The AMMI model integrates both analysis of variance and principal components analysis (PCA) into a single analysis that can be used to analyse multi-environment trials (Crossa *et al*, 1990; Zobel *et al*, 1988). The additive genotypic and environmental main effects are described by the usual analysis of variance and the non-additive GxE is described by the principal components analysis (Crossa, 1990).The three main uses of AMMI are a) model testing when used as the base model, b) clarification of GxE through pattern relationships between genotypes and environments and c) improved accuracy of yield estimates, which in turn will lead to reduced trialing costs by increasing efficiency of trials.

Analysis of variance, linear regression and principal component analysis are subcases of the AMMI model (Zobel *et al*, 1988). The AMMI model can account for more of the sum of squares attributable to treatments than the other models. The analysis results in improved partitioning of sum of squares, which in turn will result in increased sensitivity of the F-test. The AMMI analysis should be the first analysis done on multi-environment trials as it includes both the multiplicative and additive main effects in the analysis which will result in more information being obtained from the analysis. Biplots (Kempton, 1984) are a useful tool to graphically display cultivar responses when using the AMMI Model.

The AMMI model has been extensively reviewed by Gauch and Zobel(1996). They claimed that increases in trialing efficiency of between 200-400% is commonly achieved where this analysis is used. The increase in efficiency is a result of the reduction of test sites and replications required. AMMI achieves the gain in precision by removing more of the noise that is inherent in the dataset, resulting in estimates of yields that are more precise and predictively accurate than when estimated using treatment means (Nachit *et al*, 1992). AMMI uses all GER observations (ie. G Genotypes, E Environments and R replications) in the estimates of genotype yields, whereas treatment means is concerned only with the R replicates which results in more precision with AMMI estimates (Gauch and Zobel, 1996). The adjustment that is carried out is done by using information from other locations to refine the estimates of the yields within a given location (Crossa *et al*, 1991).

The AMMI model had the ability to account for more of the variation in the interaction

component than the regression model (Nachit *et al*, 1992). In the regression model, a larger proportion of the sum of squares of the regression was included in the residual sum of squares for regression. This study showed that AMMI is more effective in capturing and partitioning the sum of squares attributable to the GxE than the linear regression analysis.

Piepho (1994a) found that BLUP (Best Linear Unbiased Prediction) may be a worthwhile alternative to AMMI in obtaining reliable estimates of yields especially if the dataset is large enough for good estimates of variance components to be determined. BLUP falls short of AMMI in that the former cannot investigate the interaction term.

Crossa *et al*,(1991) reported that a variety that had a good adaptability had almost been discarded when the analysis was done not using AMMI. The good adaptability response of the variety was clouded by the noise that was present in the data. Conversely, poorly adapted lines may show up as highly adaptable.

In the AMMI Biplot of IPCA1(first principal component) *versus* mean yield, displacement along the abscissa showed differences in main effects, whereas displacement in the ordinate shows differences in the GxE (Crossa *et al*, 1991). Genotypes that have IPCA scores near zero have little interaction across environments, similarly for environments with IPCA scores close to zero. Environments with IPCA scores close to zero also had low discrimination amongst genotypes ie. they were less able to separate out varieties on differences of response. Rankings of genotypes with IPCA scores close to zero was more reliable than ranking of those genotypes with large (positive or negative) IPCA

scores. Combinations of genotype and environment interaction with IPCA scores of the same sign produced positive specific interaction effects, whereas combinations with opposite signs produced negative specific interaction effects.

Methods to determine the number of multiplicative terms that are relevant to the dataset has been reviewed by Cornelius (1993). Both tests of significance and cross validation can be used for determining the number of terms. The use of the statistical tests and cross-validation were also discussed by Piepho (1995) and Annicchiarico (1997). The number of IPCAs to include in the analysis can also be determined by using the postdictive or predictive approach. In the predictive approach, data is split into two, one part is used to determine the number of IPCA axes required, the remaining data is used to validate the models. The postdictive approach uses the amount of variation explained by the IPCA axes as a guide to the number of axes to include (Crossa, 1990).

The three major physiological components of yield (namely, net accumulated biomass, harvest index and the time needed to harvest maturity) are determined by GxE (Wallace *et al*, 1993). AMMI can quantify the deviations caused by GxE for each of the major components and sub-components of yield due to each genotype and each environment. Integration of physiological analysis and AMMI analysis was attempted and it was found that integrating some of the plant processes that result in higher yield (measuring these physiologically) with statistical analysis, such as AMMI, has some value (Romagosa *et al*, 1993).

Yau (1995) showed that regression analysis becomes less effective as the data set gets larger whereas with AMMI this is not the case. The effectiveness of regression analysis

was inversely proportional to the number of sites that were sampled. He concluded that AMMI was at least as effective as regression analysis and often better, but AMMI is not a general replacement for regression analysis. For detailed analysis of GxE, AMMI should be used, but if knowledge of responsiveness of entries to the environment is required, regression analysis is preferred. Regression analysis is effective only when a heterogeneity of regression accounts for a significant proportion of the GxE interaction.

Other Methods

Lin and Binns (1988b) proposed the Superiority Measure in investigating adaptability. The superiority measure gave an indication of general adaptability, but cultivars with specific adaptability would be discarded (Lin and Binns, 1994). They included another test for the detection of specific adaptability. The superiority measure combines the genotype performance and GxE into one parameter for each cultivar.

There are various multiplicative models that can be used in the study of GxE. Crossa and Cornelius (1993) reviewed a number of multiplicative models that can be used for the study of GxE. He discusses the clustering and fusion methods amongst others in the grouping of genotypes and environments.

Various other methods have been proposed for the study of adaptation: Pairwise GxE with checks (Lin and Binns, 1985), PCA (Freeman and Dowker, 1973), GEAR (Genotype, Environment and Attribute in Regression) (Moreno-Gonzalez and Crossa, 1998), Desirability Index (Hernandez *et al*, 1993), AMMI using Macro and Micro-

environments (Wu and O'Malley, 1998), Redundancy Analysis (van Eeuwijk, 1992), Alternative partitioning of GxE (Muir *et al*, 1992), Cluster or pattern analysis (Bull *et al*, 1992).

Stability

Adaptability is one of the major concerns when selecting cultivars, another concern is the stability of the performance of the cultivar across different environments (Lin and Binns, 1994). Stability is the consistency of genotype performance across environments (Basford and Cooper, 1998). There is a lack of a globally accepted definition of what a stable genotype is.

Lin *et al*(1986) said that most stability models are ineffective in capturing the contribution of stability because the response of genotypes across environments is multivariate and these measures try to simplify the response to a univariate response.

Lin *et al* (1986) suggested three types of stability statistics (Type 1 to Type 3). An additional stability statistic (Type 4) was defined by Lin and Binns (1988a). A cultivar is considered to be stable if:

Type 1: its variation over the entire range of environments is small. Examples of this type are Finlay and Wilkinson's (1963) regression (where b_i=0 for stable variety), variance of cultivar across environments (s^2), and coefficient of variation (CV_i) (Francis and Kannenberg,1978).

Type 2: its performance across environments is parallel to the mean of all the cultivars in the trial, eg. Eberhart and Russell (1966) regression coefficient

(where b=1 is stable) and Shukla's stability variance (σ_i^2) (Shukla,1972).

Type 3: its deviation mean square from the regression is small eg. Eberhart and Russell's (1966) δ_i^2 parameter.

Type 4: the year means square is small (Linn and Binns 1988a)

Lin and Binns (1991b) have shown that Type 1 and Type 4 stability have a genetic basis and Type 3 and Type 4 have a statistical basis ie. they are not biological measurements of stability. As such, Type 1 and Type 4 are heritable measurements of biological characteristics and are thus more suitable for variety selection. Lin and Binns (1994) believe that the unsuccessful attempts at developing a high-yielding, high-stability variety using stability statistics is the use, by most plant breeders, of Type 2 and Type 3 statistics that do not have a genetic basis.

Federer and Scully (1993) said that the current definitions of stability are based on statistical considerations when stability should be defined from the perspective of the grower. Precise definitions of poor and optimal environments are also required for proper recommendations to be made. The range of environments used in the stability analysis should be those that will be encountered in the production areas.

There are two concepts of stability, static and dynamic, depending on the goal of the breeder (Becker and Leon, 1988). With the static concept, a variety's performance is unchanged regardless of the environmental variations encountered. The dynamic concept on the other hand allows for a predictable response of the genotypes to the environments and a stable variety does not deviate from this response to environments. The variance of genotypes across environments (s_i^2) is an example of the static concept

of stability. This concept of stability can be used for traits, such as quality traits and disease resistance, where the levels of the traits have to be maintained across the environments. The dynamic concept is recommended for use in the investigation of yield stability. The main limitation on any stability analysis is the influence of the cultivars included in the analysis. Large deviations from regression does not necessarily mean that the variety is unstable, it could be that this specific variety reacts differently from the rest of the varieties included the analysis. The influence of the genotypes included in the analysis is not only found in the regression method but other methods, such as multivariate methods, also show similar influences.

The poor repeatability of stability statistics is a major problem in selecting for yield stability (Becker and Leon, 1988). The ranking of genotypes varies from year to year, therefore testing needs to take place over years. The heritability of stability measures is relatively low. The creation of artificial environments to increase the number of environments in the analysis has hardly resulted in the effect hoped for. Location, years and cultural practices may be able to replace each other in the analysis, but this is not the rule and is dependent on the material used in the experiment. It is not possible to calculate useful stability measures from a few environments only.

Pham and Kang (1988) concluded that stability statistics are useful to the breeder only for a particular set of environments as the repeatability of these statistics were low in alternate environments. They also investigated relationships amongst various stability statistics. Peltonen-Saino *et al* (1993) investigated the use of five different stability statistics on phenotypic stability of oats, and it was concluded that these different stability statistics ranked the varieties similarly.

Other proposed methods of determining stability are: rank sum method (Kang *et al*, 1991), crossover interactions (Baker, 1988; Hühn *et al*, 1993, Hühn, 1996), combination of regression and Type 4 stability (Lin and Binns, 1988a), combination of various stability methods (Brandle and Brule-Babel, 1991; Lin and Binns, 1991a), Westcott (1987)) and Safety-First Rule (Eskridge, 1990). The analysis of stability using unbalanced data-sets was investigated by Piepho (1994b).

GxE and stability analysis have been applied to a diversity of crops including amongst others(not cited previously) wheat (van Deventer, 1986; Robert and Denis, 1996, Purchase, 1997), quinoa (Jacobsen *et al*, 1996; Risi and Galwey, 1991), Perennial ryegrass (Charmet *et al*, 1993), lobolly pine (McKeand *et al*,1997), soya bean (Schutz and Bernard, 1967), Bermuda grass (Chakroun *et al*, 1990), cocksfoot (Breese, 1969), potato (Steyn *et al*, 1993). All these studies detected GxE and agreed that this interaction needs to be studied further and must be included in the decision making in breeding programs and variety selection.

When conducting the analysis of multi-environment trials, the most appropriate analytical method to meet the breeding objective must be used and not the best analytical method available (Basford and Cooper, 1998). The choice of the method is dependent on the objectives of the breeder (DeLacey *et al*, 1996).

Cooper and DeLacey(1994) investigated the relationships between analysis of variance, indirect selection and pattern analysis in the study of GxE. They found that by using the three methods in conjunction with each other, they were able to get more clarity on the

nature of the GxE. With this improved understanding of the GxE, the breeder would be in a better position to develop strategies for selection.

Optimum Allocation of Resources

The optimal breeding strategy for highly variable range of target environments involves the careful characterization of the target area so that the selection criteria and selection environments can be identified (Bramel-Cox, 1996). This will allow for the best allocation of resources for maximum gain in varietal performance.

The optimum number of sites, years and replications required for multi-environment trials can be determined by manipulating these factors to minimise the variance of the genotype means (Basford and Cooper, 1998; Crossa, 1990). The genotype mean is a function of the components of variance for GxL, GxY, GxLxY interactions and error. The AMMI model can also be used to reduce the amount of "noise" in the data (Basford and Cooper, 1998). With this increase in the efficiency of the data, the number of replications required to maintain the desired precision can be reduced.

The conducting of multi-environment trials are costly and time consuming (Sprague and Federer, 1951). The data gathered in one year or at one location is not sufficient to make general recommendations. Variance components were used to determine the optimal number of replications, locations and years that needed to be included in the trials for maximum genetic gain. They also went further, attempting to establish what the cost per unit of genetic gain was when conducting multi-environment trials. It was found that as the number of plots per location increased, the cost per unit decreased. Disregarding costs, where the GxE is large, testing must be done using fewer replicates

and more locations and years.

Miller *et al*, (1962) used the theoretical variance of the mean to evaluate the relationship between the number of testing environments used and the precision with which the evaluation of the variety could be made. The theoretical variance of a variety mean($V_{\bar{x}}$) from replicated tests over locations and years may be given by:

$$V_{\bar{x}} = \frac{\sigma^2_e}{rpy} + \frac{\sigma^2_{vpy}}{py} + \frac{\sigma^2_{vp}}{p} + \frac{\sigma^2_{vy}}{y}$$

where the numerators are the variance components and r,p and y are the number of replications, locations and years respectively in which the varieties are to be tested. The numerators can be substituted by the estimates of the variance components estimated from the analysis of variance. The variance of a mean can then be predicted for any combination or combinations of r, p and y. The smaller the variance of the mean with the different combinations of r, p and y, the more precise the estimates of variety performance would be. It was noted that there was a gain in precision with each addition of a test environment, but this increase was less than the previous increase in precision attained ie. diminishing returns. There is a point where increasing the number of environments results in very little increase in precision, the number of environments should not be increased beyond this point. It is more efficient to increase testing environments than to increase replications.

The influence of particular test environments cannot be determined from variance components (Brennan *et al*, 1981). This method also does not discriminate amongst the cultivars on the form of the cultivars response. It is unknown whether the level of

sampling of the environments is adequate when using variance components. A combination of methods (including cluster analysis) were used to determine the optimum number of environments required for variety testing. An increase in the precision specified and a decrease in the probability of error, required a significant increase in the number of trials conducted. When determining superiority of a variety's general adaptability, fewer trials are required than when testing for specific adaptability to a region. Restricted testing will increase the risk of error in determining the performance of a cultivar.

The use of cost per unit information obtained in the optimal allocation of resources to variety testing was investigated by Swallow and Wehner (1989). The traditional use of variance components (as by Sprague and Federer (1951)) does not allow for compromises especially when more than one trait is under consideration. Sometimes, the optimal allocation of resources as determined by the traditional analysis using variance components is not necessarily the best. Vermeer (1990) used the coefficient of variation of the different sources of variation in a multi-environment trial to determine the optimal allocation of resources.

Chapter 3

Materials and Methods

3.1 Materials

Nine green bean varieties (Table 3.1a) were included in multi-environment trials over a three year period (Y1 to Y3) between 1998 and 2000 at four locations coded A to D (Table 3.1b). All the varieties were white flowered. The varieties included both commercially available lines and experimental lines, and all are known to have high yield potentials.

Table 3.1a Snap bean varieties included in the multi-environment trial

Code	Variety	Origin	Type
G1	Variety 1	USA	Bobby, Determinate
G2	Variety 2	USA	Bobby, Determinate
G3	Variety 3	RSA	Bobby, Determinate
G4	Variety 4	USA	Bobby, Determinate
G5	Variety 5	RSA	Fine, Determinate
G6	Variety 6	USA	Bobby, Determinate
G7	Variety 7	USA	Fine, Determinate
G8	Variety 8	USA	Bobby, Determinate
G9	Variety 9	USA	Bobby, Determinate

The locations were selected in an attempt to include a diversity of environments. Location A is included as it is where cultivar selection is currently being carried out. Locations B and C are representative of the major green bean producing areas in South Africa. Location D was included as a stress environment.

To allow for timeous harvesting, the trials were sown approximately two weeks apart. The sowing dates were similar for all three years. In Year3, the floods experienced had washed away the trial at Location C; this trial had to be resown. The trial at Location B in Year 3 was severely stressed by the excessive rainfall, this data is included as part of the stress environment.

Table 3.1b Locations included in the multi-environment trial

Location	Climate	Sowing Date
A	Coastal/Mistbelt	31 December
B	Highveld	18 January
C	Middleveld	3 February
D	Desert	20 February

The land was prepared according to the standard procedures used by bean farmers in the area. The land preparation was kept as similar as possible amongst the locations. Fertilizer 2:3:4(40) was applied at preplanting at a rate of 600kg/ha (the equivalent was used if 2:3:4(40) was not available). Thereafter, LAN (28) was used as a topdressing at a rate of 300kg/ha at 14 and 35 days after emergence. A Dithane/Kocide mixture was sprayed every 10 days from 21days after emergence to flowering to control rust. In Location C, Lannate was used for the control of bollworm. The total precipitation (rainfall and irrigation) per week was 30mm. Irrigation was applied using a dragline system. Weed control was carried out by hand.

A randomized blocks design with three replications was used at each location and each trial was independently randomized. Sowing was done by hand at a between row

spacing of 0.5m and within row spacing of 0.07m. The plot dimensions were 4m x 3m. The plot consisted of six rows, each 4m in length. The net plot had an area of $8m^2$. The middle four rows were used for the pod yield determinations with the aid of an electronic scale.

3.2 Character measured

An important characteristic in the production of green beans is pod yield. Picking of the pods was conducted in two stages, the first pick was at 20 days after flowering and the second pick six days after the first pick. Picking was conducted as close as possible to these guidelines. The combined pod masses of the first pick and second pick was used to determine the total yield.

3.3 Statistical Analysis

3.3.1 Analysis of Variance

The analysis of variance (ANOVA) was conducted on the pooled data across years and locations. The ANOVA was carried out for yield. The expected form of the ANOVA and expected mean squares is given in Table 3.3a. The estimated variance components for the treatments were then computed using the methods described in Table 3.3b from the ANOVA .

3.3.2. Finlay and Wilkinson

A regression analysis was done for each genotype by regressing the genotype performance on the environmental index to investigate adaptability as carried out by Finlay and Wilkinson (1963). The mean yield of all the varieties at a location within a single year was used as an index of the environment. The regression coefficients were

Table 3.3a Expected Anova and Expected Mean Squares(adapted from Rasmusson and Lambert, 1961)

Source	Df	Mean Square	Expected Mean Square
Year(Y)	(y-1)		
Location(L)	(l-1)		
Y x L	(y-1)(l-1)		
Replication(R)	ly(r-1)		
Variety(V)	(v-1)	M_5	$\sigma^2_e + r\sigma^2_{vyl} + ry\sigma^2_{vl} + rl\sigma^2_{vy} + rly\sigma^2_v$
V x L	(v-1)(l-1)	M_4	$\sigma^2_e + r\sigma^2_{vyl} + ry\sigma^2_{vl}$
V x Y	(v-1)(y-1)	M_3	$\sigma^2_e + r\sigma^2_{vyl} + rl\sigma^2_{vy}$
V x Y x L	(v-1)(y-1)(l-1)	M_2	$\sigma^2_e + r\sigma^2_{vyl}$
Error	ly(v-1)(r-1)	M_1	σ^2_e

y,l,v and r are the number of years, locations, varieties, and replications respectively; σ^2_e and σ^2_v are components of variance for error and varieties respectively. The components of variance for the interactions are identified by the combinations of the subscript. M_1 to M_5 are the observed values of the various mean squares.

Table 3.3b Method of Determination of Variance Components (adapted from Rasmusson and Lambert, 1961)

Variance Component	Method of Determination
Varieties(σ^2_v)	$\frac{M_5 + M_2 \cdot M_3 - M_4}{r\,l\,y}$
Variety x Location(σ^2_{vl})	$\frac{M_4 \cdot M_2}{r\,y}$
Variety x Year(σ^2_{vy})	$\frac{M_3 \cdot M_2}{r\,l}$
Variety x Year x Location(σ^2_{vyl})	$\frac{M_2 \cdot M_1}{r}$
Plot Error(σ^2_e)	M_1

then plotted against the variety mean yields to create a scatter diagram for further analysis.

According to Finlay and Wilkinson (1963), a genotype that has a regression coefficient approximately equal to one ($b \approx 1$) displays average stability over all the environments; when this is associated with above average yields, the genotype has good general adaptability and low adaptability when this b is associated with below average yields. A genotype with $b>1$ has below average stability and is specifically adapted to high yielding environments, whereas a genotype with $b<1$ has above average stability as is adapted to low yielding environments.

3.3.3 AMMI analysis

The AMMI analysis was applied to the data for yield; this model applies the additive main effects using the usual ANOVA with the non-additive interaction effect being determined using PCA. The analysis of the data included genotype x location within years and genotype x environment across years where the environment is made up of location-year combinations (Zobel *et al*, 1988). Each environment is represented by a single letter for the location and a single digit for the year. For example, Location A in Year3 would be denoted A3 and Location C in Year1 would be C1, etc. Thus to look at the influence of the Location C across years, all the environments with C in its combination will be considered. The location-year combinations are used as the environment in the combined analysis because AMMI requires a two-way table, and the addition of years to the table will result in a three way table. For year data to be included in the analysis, the use of location-year combinations are necessary. The requirement of a two way table is a limitation of AMMI.

In considering the interaction component in the AMMI analysis, a genotype is considered stable if the value of the IPCA score is zero. Genotypes with IPCA scores closer to zero are more stable for yield performance than genotypes with IPCA scores further away from zero, the distance away from zero can therefore be used as a measure of stability. In considering stability of genotypes, the absolute value of the IPCA score needs to be used.

The IPCA scores for yield can be plotted against the yield performance of the genotypes and the environments on a single set of axes to produce a biplot. The similarity between environments or genotypes is a function of the distance between the points of the specific environments or genotypes on the biplot. The smaller the distance between the two points, the more similar the environments or genotypes are. The biplot can also be used to determine the AMMI estimates of yield for the genotypes at different locations.

3.3.4. Optimum Allocation of Resources

The variance components calculated in 3.1.1 above were then used to determine the optimum allocation of resources for efficient variety testing. The effect of the number of replications (r), locations (l) and years (y) on the theoretical variance of a mean ($V_{\bar{x}}$) was determined using the following equation (Rasmusson and Lambert, 1961):

$$V_{\bar{x}} = \frac{\sigma^2_{vl}}{l} + \frac{\sigma^2_{vy}}{y} + \frac{\sigma^2_{vly}}{ly} + \frac{\sigma^2_{e}}{rly}$$

Different levels of l, y and r are substituted into the formula. The lower the $V_{\bar{x}}$ the more precise the estimate of the mean. Thus the combination of l, y and r that results in the lowest $V_{\bar{x}}$ will produce the best estimate of the mean.

The least significant difference (LSD) can be used as a measure of precision where the smaller the LSD, the higher the precision. $V_{\bar{x}}$ can be used to determine the different LSD for the various combinations of l, y, and r. This LSD can then be expressed as a percentage of the grand mean for standardization and comparison purposes. The LSD percentage (LSD%) can be calculated using the following formula:

$$LSD\% = \frac{(\text{t-value}).(\sqrt{2.V_x}).(100)}{\text{Grand Mean}}$$

Chapter 4

Results and Discussion

4.1 Analysis of Variance

The ANOVA for the nine genotypes evaluated in trials conducted over three years at four locations is given in Table 4.1a. As can be seen, both the genotype and environment main effects and GxE interaction are highly significant. There is definitely an interaction between genotype and the environment in snap beans.

The YxL effect in Table 4.1b is also highly significant, this indicates that the effect of the locations on the genotypes is different from year to year. In the ANOVA given in Table 4.1b, the location and year effect are treated as fixed effects and the genotypes as random effects. From here on, the discussion will concentrate on the random effects.

The significant genotype effect suggests that there are consistent differences in performance amongst the genotypes. Locations and Years also have an influence on the performance of the genotypes as can be seen by the significant GxL, GxY and GxYxL interaction in Table 4.1b. The status of the interaction of genotypes with locations and years implies that the genotypes perform differently at different locations in different years. Therefore, testing of the genotypes needs to be conducted at different locations over different years for a more reliable evaluation of genotype performance.

Variance components for the random effects can be estimated from the ANOVA in Table 4.1b. The estimates for these variance components are given in Table 4.1c. The higher

Table 4.1a Factorial Anova for nine snap bean genotypes evaluated at four locations over three years.

Source	d.f.	Sum of Squares	Mean Square	F Prob.
Total	323	4749		
Replication	2	42	20.9	
Treatment	107	4442	41.5	
E	11	2700	245.5	<.001
G	8	657	82.2	<.001
G x E	88	1085	12.3	<.001
Error	214	266	1.2	
Grand Mean = 10.54			Coefficient of Variation = 10.6%	

Table 4.1b Combined Anova for nine snap bean genotypes evaluated at four locations over three years with interactions expanded.

Source	d.f.	Sum of Squares	Mean Square	F Prob.
Total	323	4749		
Replication	2	42	20.9	
Treatment	107	4442	41.5	
Y	2	770	385.2	<.001
L	3	748	249.2	<.001
Y x L	6	1182	197.0	<.001
G	8	657	82.2	<.001
G x L	24	481	20.1	<.001
G x Y	16	199	12.5	<.001
G X L x Y	48	405	8.4	<.001
Error	214	266	1.2	
Grand Mean = 10.54			Coefficient of Variation = 10.6%	

Table 4.1c Estimated variance components for yield and there relative contributions to total variation for nine snap bean genotypes evaluated at four locations over three years.

Variance Component	Estimated Value	Percentage of Total
σ^2_g	1.613	23.46
σ^2_{gl}	1.294	18.82
σ^2_{gy}	0.336	4.89
σ^2_{gly}	2.392	34.79
σ^2_e	1.241	18.05
Total	6.876	100.00

the variance component, the more influence that effect has on the performance of the variety. The heritability of yield is estimated to be 23.46%. This is also an indication of the repeatability of the genotype performance thus the chance of repeating the yield performance of the genotypes is 23.46%. Genotype has 25% and 380% more influence on yield performance than location and year respectively, but the combination of location and year has the most influence on genotype performance (48% more than that of the genotype effect). The year effect has the least influence on genotype performance which suggests that testing of genotypes across years may not be as important as the testing of genotypes across the locations but the high σ^2_{gly} requires testing across years and locations because the locations are different in different years.

Table 4.1d ranks, in descending order, the mean genotype performance of the nine genotypes evaluated at four locations over three years. The LSD was used to test the difference between the genotypes. The results of the LSD test are given in the third column of Table 4.1d, genotypes with the same letter are not significantly different from each other. Three groups of genotypes can be identified:

Group 1	G3, G5, G9
Group 2	G8, G9
Group 3	G1, G2, G4, G6, G7

Genotype G9 falls within Group 1 and Group 2. Group 1 includes the high yielding genotypes and Group 3 the lower yielding genotypes. It is known that Genotype G8 is also high yielding, but there seems to be something different about the response exhibited, which puts it into Group 2 and not Group 1. Genotypes from Group 1 and Group 2 would give the best yield performance in the production areas.

Table 4.1d Mean Yield(kg/plot) and Least Significant Difference pairwise comparison of the genotype effect of nine snap bean varieties evaluated at four locations over three years.

Rank	Genotype	Mean Yield (kg/plot)	Comparison of Means
1	G3	12.51	a
2	G5	12.32	a
3	G9	11.94	ab
4	G8	11.67	b
5	G6	9.57	c
6	G4	9.36	c
7	G7	9.20	c
8	G2	9.17	c
9	G1	9.15	c
$LSD_{t(0.05)}$ for Genotype = 0.52 kg/plot			

The ANOVA does not allow for a more detailed analysis of the GxE that has been detected. Information on the underlying pattern of response for the interaction cannot be elucidated further using this method. This has resulted in the limited use of ANOVA directly in the study of GxE. However, ANOVA is part of other analyses that are used in the study of GxE such as the Finlay and Wilkinson (1963) regression analysis and the AMMI analysis.

4.2 Finlay and Wilkinson Analysis

The regression analysis was applied as described by Finlay and Wilkinson (1963). The logarithmic transformation of the data did not result in an increase in linearity, thus it was decided to use the untransformed data.

Becker and Leon's (1988) term "dynamic stability" can be applied to Finlay and Wilkinson's (1963) definition of average stability (where b=1). A good breeding strategy would be to look for genotypes with average stability and above average yields. This will allow for an increase in production when the conditions are favourable but the potential to still produce above average yields under unfavourable conditions is greater. Specific adaptation can also be exploited by using genotypes with b>1.

The regression of genotype mean yield on the environmental index for the nine varieties evaluated at four locations over the three years is given in Figure 4.2a. Genotypes G3, G5 and G9 can be described as having average stability and good adaptability, so these genotypes will respond to environmental changes and will perform well in most environments. Genotype G8 has b<1 therefore exhibits above average stability, it can be seen that this genotype performs well at low potential environments, but the relative performance decreases as the potential of the environment increases. This means that there are environments where G8 will not perform above average, in this case it will be the high potential environments. Genotypes G1, G2, G4, G6 and G7 have $b \approx 1$ but have below average yields, thus they have average stability with poor adaptability. None of the genotypes showed specific adaptability to high potential environments. A composite of all the genotypes in all environments is shown in Figure 4.2b, it can be seen that

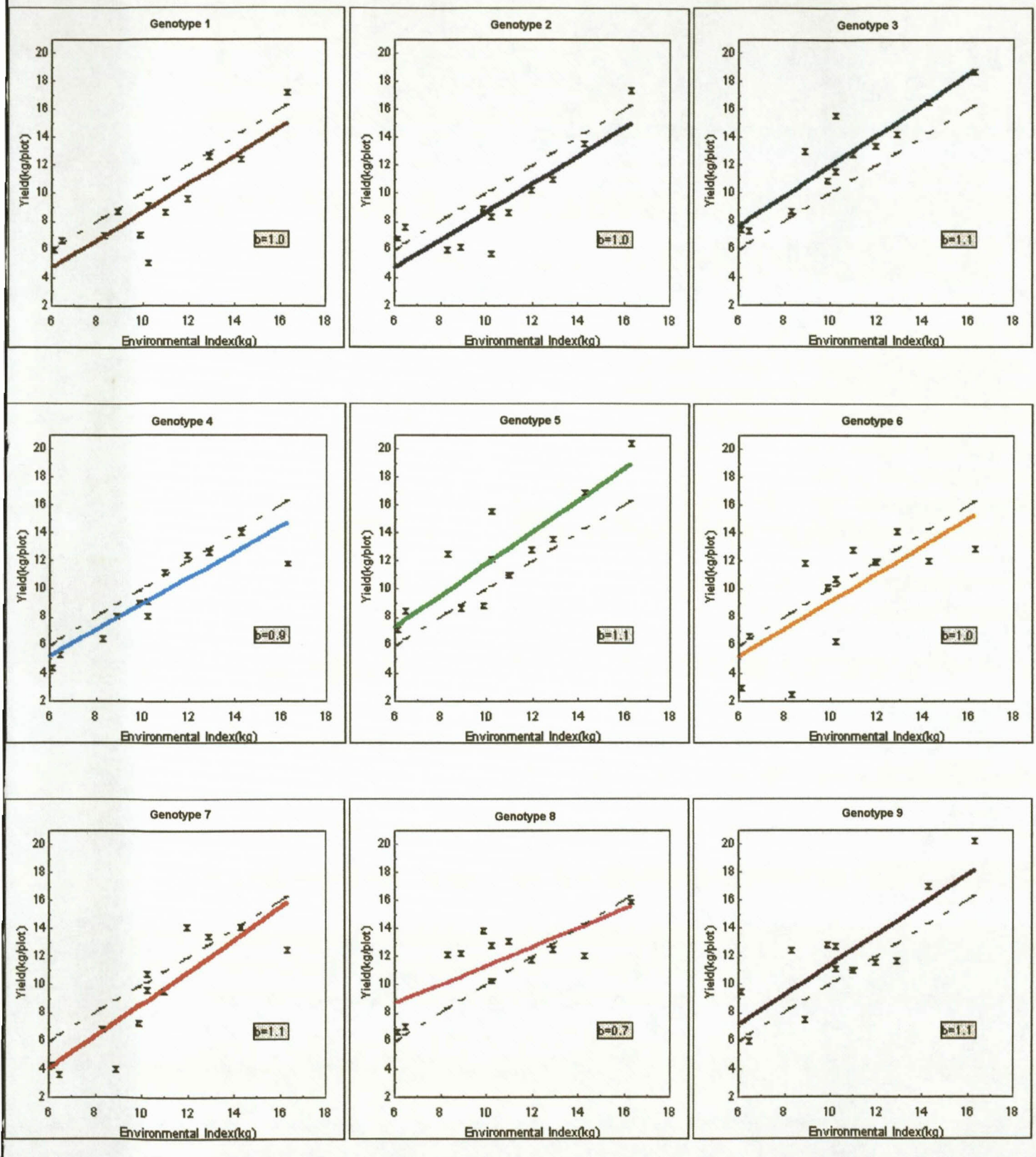


Figure 4.2a Regression of genotype yield on environmental index across four locations and three years. The hourglasses are the actual yield, the dashed line indicates the mean yield of all genotypes, and the regression coefficient(b) is also given.

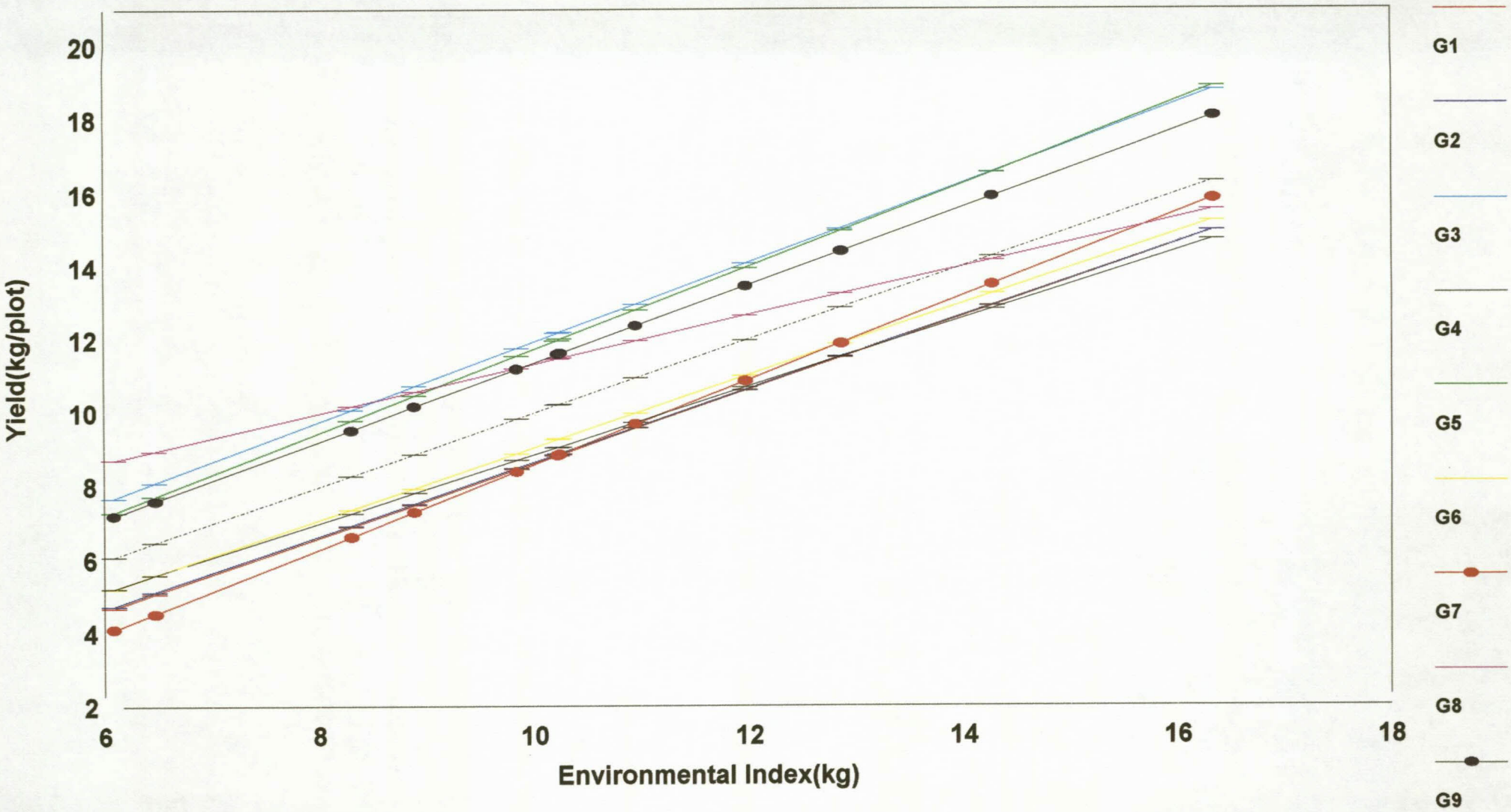


Figure 4.2b Composite of regressions of genotype performance on environmental index over three years.

there are three distinct groups. Group 1 consists of genotypes that have average stability with above average yields thus good adaptability, Group 2 consists of genotypes with above average stability and Group 3 genotypes have average stability with below average yields, therefore poor adaptability.

The regression coefficients for the genotypes were plotted against the mean yield of the genotypes as given in Figure 4.2c. The three groups as identified in Figure 4.2b can also be seen here. The groups are:

Group 1	G3, G5, G9
Group 2	G8
Group 3	G1, G2, G4, G6, G7

Genotype G9 is more similar to G8 in yield, but is more similar to G5 and G3 in terms of the stability coefficient. Genotype G9 could also be classified in Group 2 together with G8 if the classification is based on yield alone. The link between G8 and G9 detected earlier in the ANOVA conducted in 4.1 has been elucidated in this regression analysis. G8 and G9 may be classified together on yield but separate if classified on regression coefficient. This may explain the groupings obtained in the ANOVA in 4.1

From the ANOVA given in Table 4.2a, it can be seen that the GxE interaction is further partitioned into SS due to the regression and residual SS (deviations from regression). Both the SS for regression and deviations from regression were significant. This indicates that there are differences in the regressions of the response of genotypes and also the amount of variation accounted for by the regression is not consistent from genotype to genotype. The Finlay and Wilkinson (1963) model was able to explain 77%

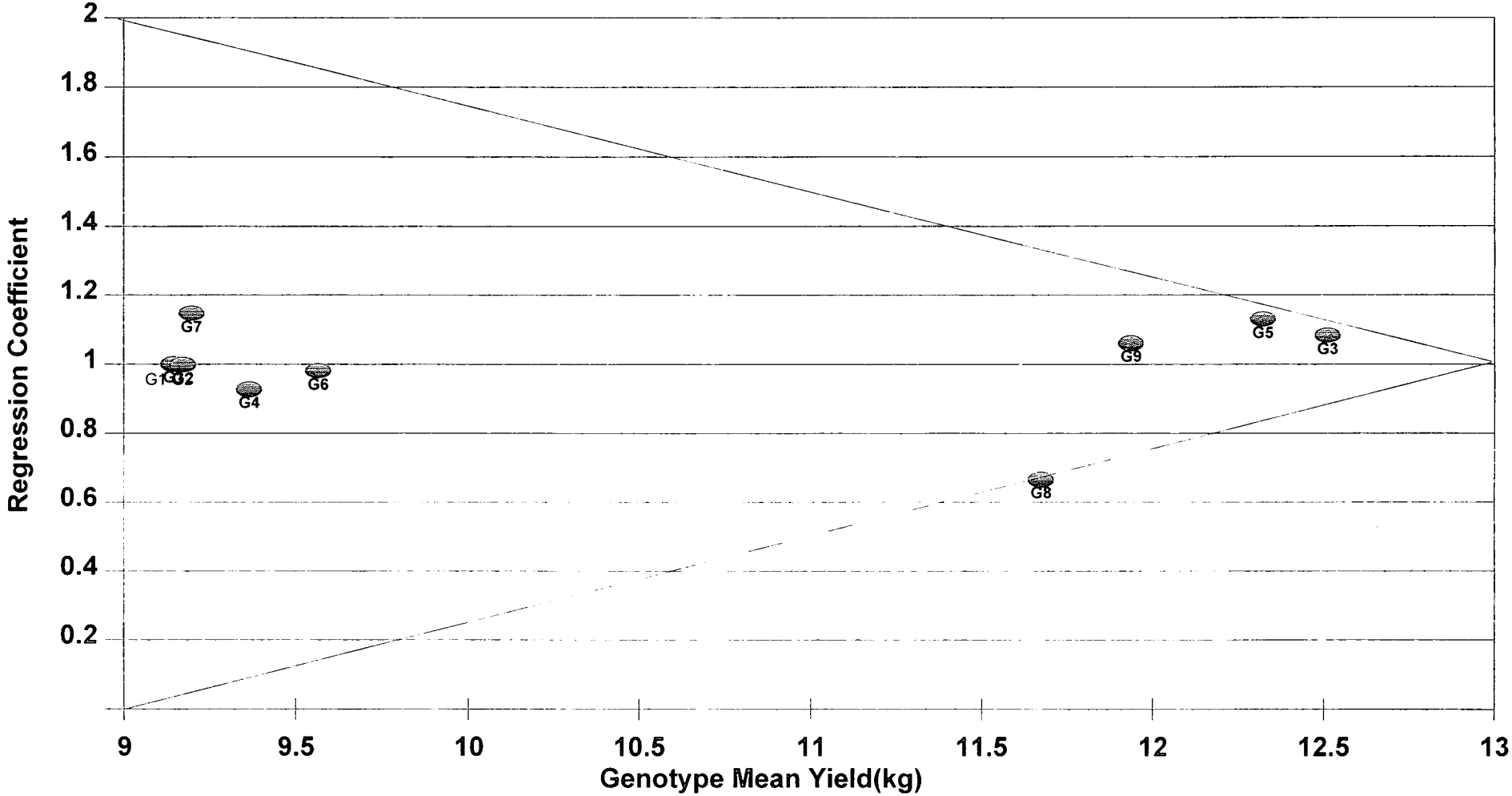


Figure 4.2c Regression coefficients plotted against genotype mean yield for nine snap bean genotypes evaluated at four locations over three years.

Table 4.2a Combined Anova for nine snap bean genotypes evaluated at four locations over three years using the Finlay and Wilkinson(1963) regression analysis.

Source	d.f.	Sum of Squares	Mean Square	F Prob.
Total	323	4749		
Replication	2	42	21.0	
Treatment	107	4442	41.5	
E	11	2700	245.5	
G	8	657	82.2	<.001
G x E	88	1085	12.3	<.001
Regression	8	50	6.3	<.001
Residual	80	1035	12.9	<.001
Error	214	266	1.2	<.001

Table 4.2b Regression coefficients of variety performance on the environmental index for nine snap bean varieties evaluated at four locations over three years.

Genotype	Regression Coefficient(b)			
	Y1	Y2	Y3	3 Year Composite
G1	1.21	0.93	0.29	1.00
G2	1.40	0.67	0.05	1.00
G3	1.26	0.94	1.45	1.08
G4	0.75	1.22	1.11	0.93
G5	0.98	0.94	0.97	1.13
G6	1.28	1.57	0.97	0.98
G7	0.75	1.42	1.54	1.15
G8	0.38	0.87	1.30	0.67
G9	1.01	0.43	1.32	1.06

of the SS for treatments. The partitioning of the GxE interaction using the regression analysis resulted in an extraction of 4.6% of the SS of the interaction leaving 95.4% of the SS of the interaction being unexplained.

Figures 4.2d, 4.2e and 4.2f (Appendix A) represent the regression of the genotype yield on the environmental index within each of the three years for the nine genotypes across four locations. It can be seen that the genotype yield performance varied from year to year, thus repeatability of yield performance is low. From Table 4.2b, it can be seen that the regression coefficients also varied from year to year for the genotypes except for Genotype G5 which had the regression coefficient with the highest repeatability, the response of this genotype is probably highly linear. There were very few points available to carry out the regression within years, thus slight changes in any of the points may result in a large influence on the slope of the lines. In all likelihood, the regression analysis as carried out within each year will be less reliable since there may be too much influence on the end result by changes in individual performances. The regression as carried out in Figure 4.2a is more reliable because of the larger number of points available for the regression. As the number of data points increases, the regression is less influenced by individual points and also allows for more accurate estimation of the regression coefficient. It must be remembered though that for an increase in the number of data points, there is also an increase in the number of environments (locations and/or years) which will result in a greater demand on the resources available.

4.3 AMMI Analysis

From the initial analysis of variance, it was determined that the first IPCA axis was able to remove 14% of the "noise" in the data, and the second IPCA axis was able to remove a further 3-5%. It was subsequently decided to use only the first IPCA axis in the analysis. The ANOVA for the AMMI analysis is given in Table 4.3a. It can be seen that the GxE interaction is partitioned using principal component analysis, and that the first PCA axis and the residual is highly significant, this indicates that the IPCA1 has accounted for a high proportion of the GxE interaction. The significance of the residual indicates that there is confounded information still not extracted from the GxE interaction. The first IPCA score was plotted against the mean yield of the varieties and environments to create the biplot as given in Figure 4.3a, this figure therefore represents 86% of the SS for treatments (Table 4.3a). The remaining 14% of SS can be considered to be noise in the data and is not of interest.

The AMMI estimate of the yield for any of the genotypes in any environment (ie. genotype-environment combination) can be determined from Figure 4.3a. The main effect is equal to the sum of the genotype mean and environment mean subtract the grand mean, and the interaction effect is the genotype IPCA score multiplied by the environment IPCA score. For example, Genotype G9 grown in environment D2 has a main effect of 11.94 + 8.89 - 10.54 = 10.29 and the interaction effect is -1.70 x 1.53 = -2.6, therefore the AMMI estimate of the yield of Genotype G9 in environment D2 is 10.29 + (-2.6) = 7.69kg. This compares well with the observed yield of 7.45kg, whereas the ANOVA estimate would have been 11.94kg.

Table 4.3a Combined Anova for nine snap bean genotypes evaluated at four locations over three years using the AMMI model.

Source	d.f.	Sum of Squares	Mean Square	F Prob.
Total	323	4749		
Replication	2	42	21.0	
Treatment	107	4442	41.5	
E	11	2700	245.5	
G	8	657	82.2	<.001
G x E	88	1085	12.3	<.001
PCA1	18	465	25.8	<.001
Residual	70	620	8.9	<.001
Error	214	266	1.2	<.001

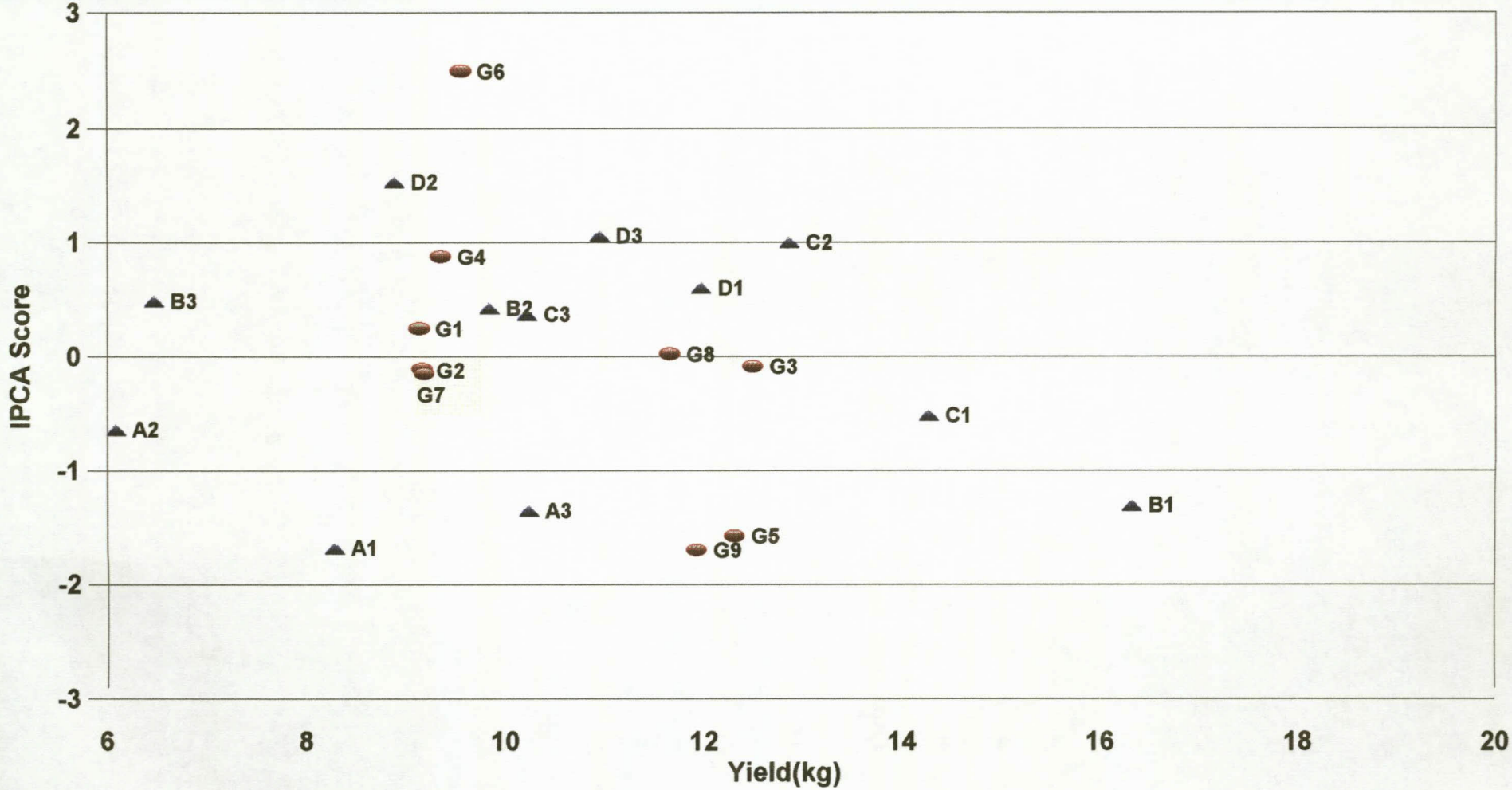


Figure 4.3a IPCA1 score plotted against yield for Genotypes(ellipse) and Environments(triangles) over three years.

From Figure 4.3a, four groups of genotypes can be identified. Group 1 consists of genotypes with high yields and low stability, Group 2 contains genotypes with high yields and high stability, genotypes with low yields and high stability are found in Group 3 and genotypes with low yield and low stability are found in Group 4. The genotypes can be grouped:

Group 1	G5, G9	High yields, Low stability
Group 2	G3, G8	High yields, High stability
Group 3	G1, G2, G7, G4	Low yields , High stability
Group 4	G6	Low yields , Low stability

In considering the environments, the further away from zero the IPCA score for the environment is, the more interaction the environment has with the genotypes, thus it will be more difficult to chose genotypes for these environments. Location C seems to have the lowest interaction with the genotypes followed by Location B, this is probably because the locations are generally low stress and optimum for snap bean production. Location A and D seem to have similar magnitudes of interactions with the genotypes but the effects of the interaction on the genotypes are different. Locations A and D are very effective in discriminating between genotypes whereas locations B and C are less so.

It can be seen from Figure 4.3a that the effect of the locations is not consistent from year to year but, the environments at a single location can be broadly grouped together (Figure 4.3b) as being similar except for Location B; at this location the response in Year1 was very different from the response in Year2 and Year3. The locations can be

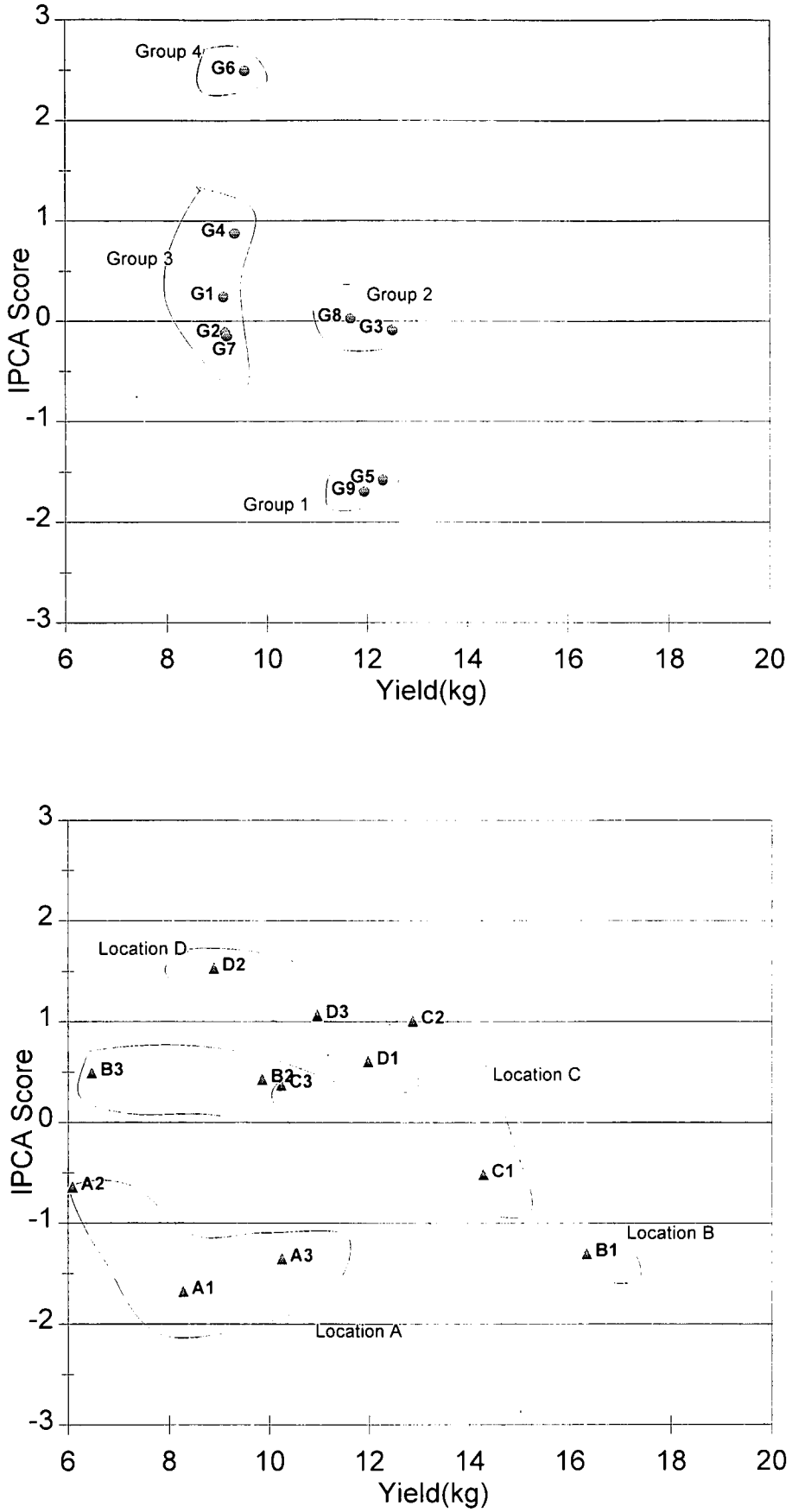


Figure 4.3b Grouping of genotypes with similar responses(above) and the grouping of the environments experienced at specific locations(below).

classified into four broadly defined groups:

Group 1	Location A
Group 2	Location B
Group 3	Location C
Group 4	Location D

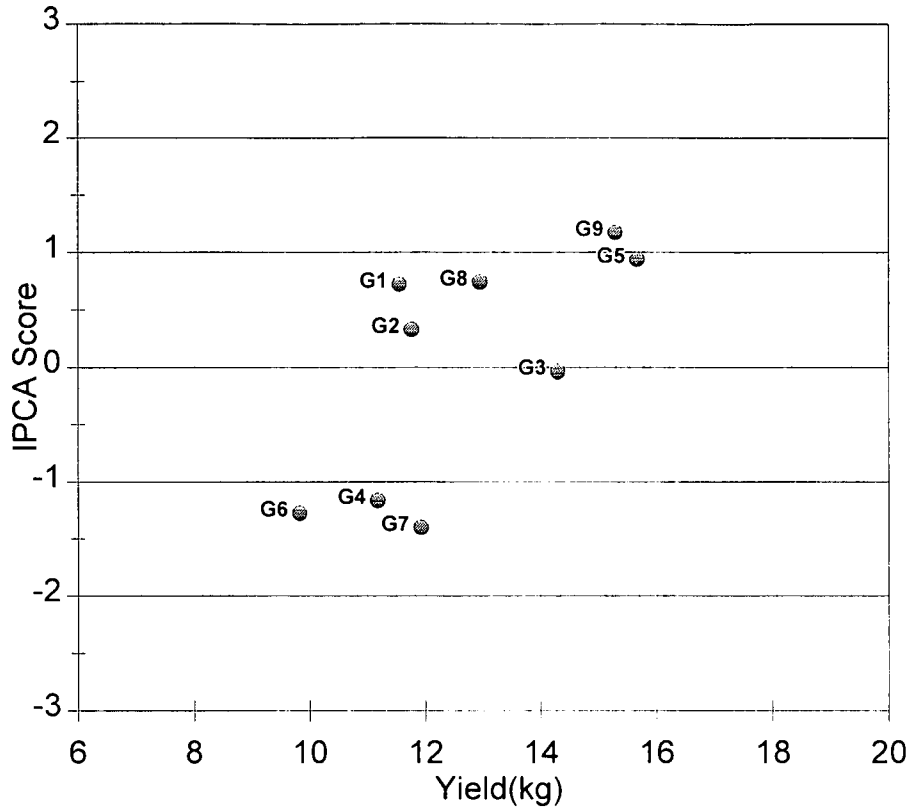
Group 1 locations have environments that result in moderate to low yields with high discrimination amongst genotypes whereas Group 2 locations have high variation in yield potentials but low discrimination among genotypes, Group 3 locations have more consistent, moderate-high yielding environments with low discrimination amongst genotypes. The group 4 locations produce environments conducive to moderate yields, but have high discrimination amongst the varieties.

Location A (Figure 4.3b) is generally a moderate to low potential environment with a yield potential of between 10-12.5 tons/ha. Location B seems to be the environment with the largest variation in yield potential. Taking into account that B3 was a very unusual year, it was decided to exclude Year3 for the estimation of the yield potential of Location B. The yield potential of location B is thus determined to be 12.5-20.0 tons/ha. Location C has a yield potential of 12.5-17.5 tons/ha and 10-15 tons/ha is the estimated yield potential of Location D.

The selection of the environments is important in a variety testing program. The careful selection of the environments may result in a reduction in the resources required. Although the environments have been allocated to different groups, there is one important characteristic of the environments that has been mentioned, this is the environments ability to discriminate between the genotypes. Locations A and D can

discriminate amongst the genotypes the most, so at least one of these locations needs to be included in the testing program. Inclusion of only the discriminating environment in the program is not sufficient to test the superiority of the variety, a reliable test location is required. The location can be either Location B or C, but Location C seems to be more reliable on consistency than Location B. From this discussion, a test program should include Location A or D and Location C.

As can be seen from Figures 4.3c to 4.3e, the relative performance amongst the genotypes were similar. Genotypes G1, G2, G4, G6 and G7 consistently produced the lowest yield with the remaining genotypes generally producing high yields across years. It must be noted though, that absolute yield of genotype performance across years was inconsistent. The repeatability of the absolute genotype performance and IPCA score from year to year is very low. It is thus necessary to test the genotypes over years for more precise estimates of performance to be determined. It can also be seen that the responses of the locations across years is not consistent, this requires that the year effect to be included in the analysis. The relative performance of Location A and D seem to be consistent across years, and Location C seems to be the most repeatable environment.



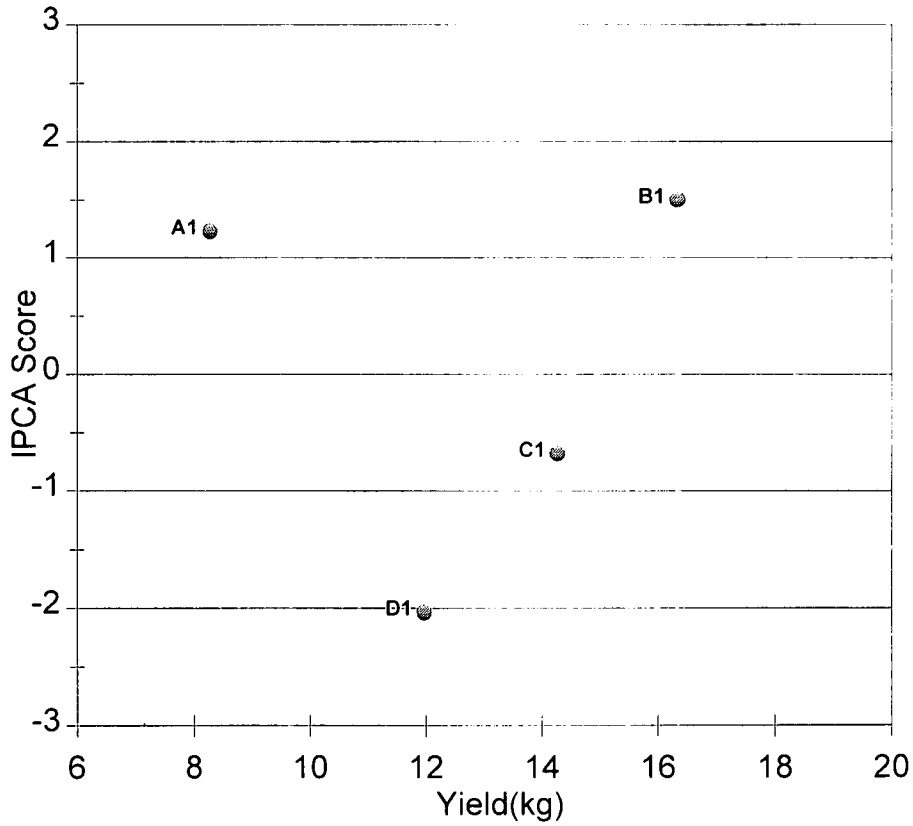
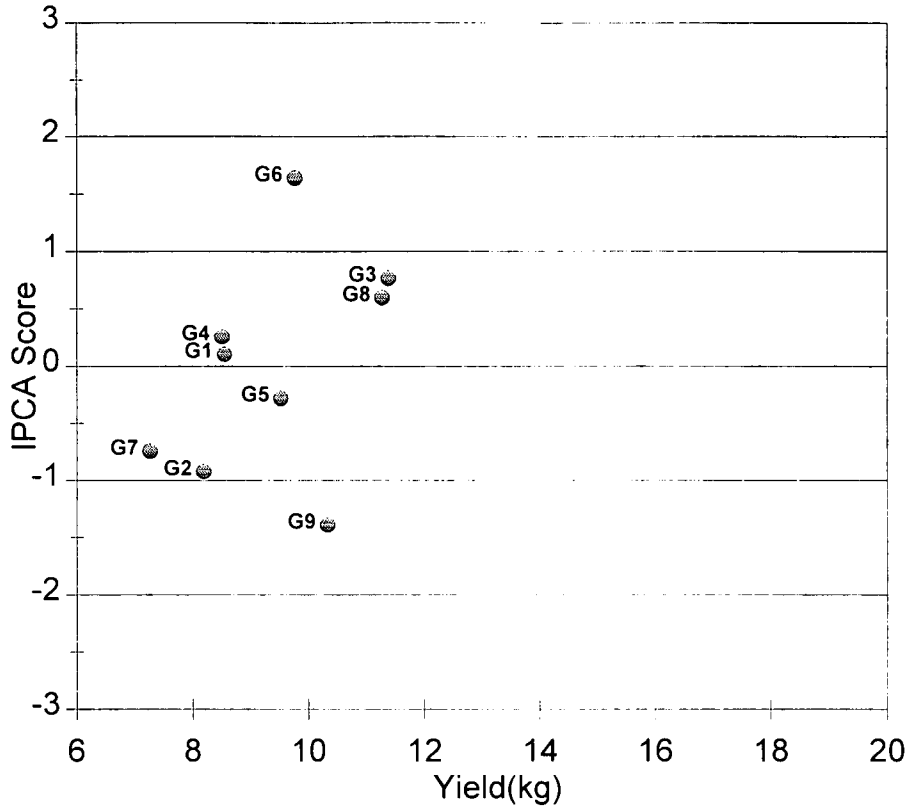


Figure 4.3c IPCA1 score plotted against yield for Genotypes(above) and Environments(below) for Year 1.



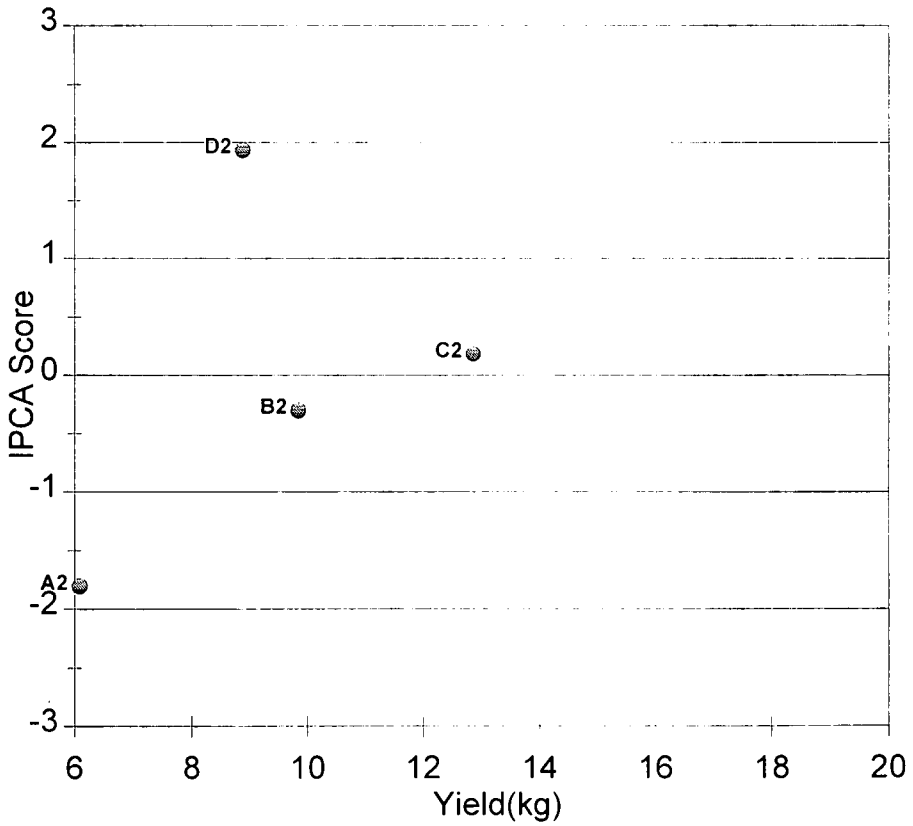
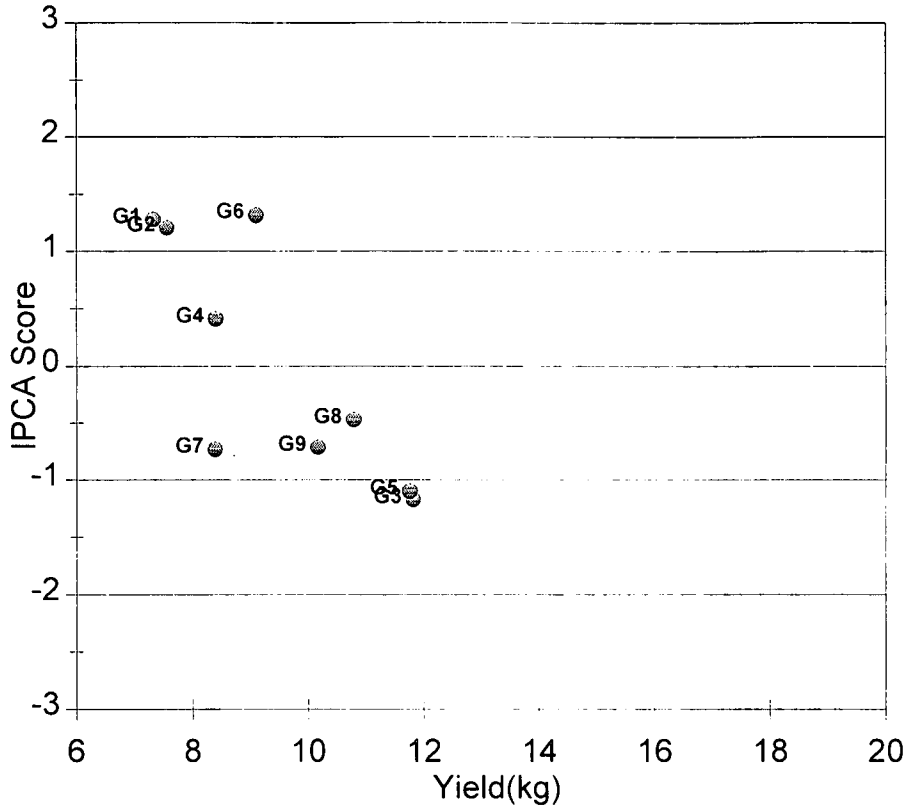


Figure 4.3d IPCA1 score plotted against yield for Genotypes(above) and Environments(below) for Year 2.



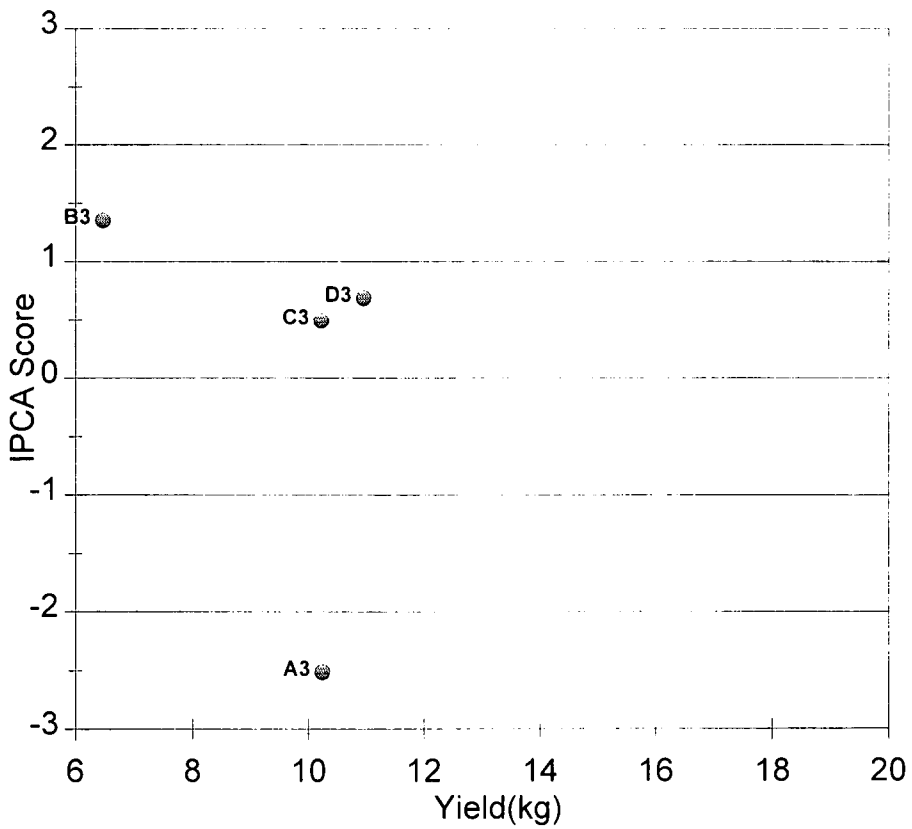


Figure 4.3e IPCA1 score plotted against yield for Genotypes(above) and Environments(below) for Year 3.

4.4 Comparison of the Analytical Methods

The genotypes included in the trials were ranked on the basis of stability as given in Table 4.4a. There seems to be some degree of similarity in rankings between each of the methods for some of the genotypes but the similarity is not consistent for all the methods. It appears that generally where AMMI rankings are similar to Finlay and Wilkinson, the ANOVA is very different for the same genotype, vice versa. Using the Spearman rank correlation coefficient as tabulated in Table 4.4b, it can be seen that the correlation between the different methods is very low, thus ranking of genotypes using the different methods will more than likely result in different rank orders dependent on method used. The low correlation could also be related to the closeness of the values of the stability parameters such that small differences result in a change in rank. The almost complementary results of the Finlay and Wilkinson model and the ANOVA to produce similar results to the AMMI model seems to suggest that the two models are submodels of AMMI.

In Table 4.4a, the ANOVA ranks Genotype G3 as being most stable whereas Genotype G8 is identified as being most stable when using the Finlay and Wilkinson and AMMI model. The genotypes with the lowest stability are G1, G7 and G6 according to the ANOVA, Finlay and Wilkinson and the AMMI model respectively.

One of the uses for yield trials is to accurately predict yield performance. The ANOVA uses the mean yield across all locations as an indication of yield potential. With the AMMI model, the yield is estimated by using both the genotypic performance and the environment in which it is to be planted. As given in the example in section 4.3, the

Table 4.4a Ranking of the different genotypes on the basis of stability using different stability parameters.

Genotype	G X E Analysis Method					
	Anova		Finlay and Wilkinson		AMMI	
	Yield	Rank	b	Rank	IPCA1	Rank
G1	9.15	9	1.002	5	0.242	5
G2	9.17	8	0.999	4	-0.110	3
G3	12.51	1	1.083	7	-0.089	2
G4	9.36	6	0.928	2	0.874	6
G5	12.32	2	1.131	8	-1.578	7
G6	9.57	5	0.980	3	2.491	9
G7	9.20	7	1.147	9	-0.154	4
G8	11.67	4	0.667	1	0.023	1
G9	11.94	3	1.063	6	-1.699	8

Table 4.4b Spearman rank correlation coefficient for three stability analysis

Method	Anova	Finlay and Wilkinson
Finlay and Wilkinson	-0.217	-
AMMI	-0.050	0.100

AMMI yield estimate was much closer to the actual genotype yield than the estimate given by ANOVA. The yield estimate given by the Finlay and Wilkinson model will be dependent on the linearity of the response for yield, the more linear the response, the more accurate the yield prediction. AMMI uses more of the data in the calculation of prediction of yield than the Finlay and Wilkinson or ANOVA model.

As a result of the ability of AMMI to correct for environmental influences, the model can create rankings of genotypes specific to an environment, in this way genotypes with specific and broad adaptability can be identified. The Finlay and Wilkinson model uses the regression coefficient and yield to identify specific and broad adaptability. The ANOVA can be used for the identification of broad adaptability only.

The different models account for different amounts of variation in the data (Table 4.1a, 4.2a and 4.3a). The ANOVA model resulted in 24.4% of the treatment SS being included in the GxE interaction term. By the nature of the calculation of the GxE term, residual from treatment SS, it includes all variation not accounted for by the main effects. In this case, all the SS included on the GxE term can be treated as noise since the variation due to different sources are confounded and cannot be separated out. The Finlay and Wilkinson model (Table 4.2a) reduced the amount of unaccounted treatment SS to 23.3% as only a very small proportion of the GxE SS was attributed to the regressions. The AMMI model (Table 4.3a) was able to reduce the Treatment SS due to unknown sources to 14.0%. Thus, the AMMI model was able to extract the most information from the data.

The nature of the analyses results in the use of past data to predict future performance

of varieties. As a result of this, the analyses are dependent on the data available. A shortcoming in all the methods is the dependence on past data, even the modern multiplicative models such as AMMI cannot escape this limitation. However, the effect of this limitation is greater in the Finlay and Wilkinson model when compared to the AMMI model because of the extent of the use of the data available.

The Finlay and Wilkinson model concentrates on the variation of the genotypes whereas the AMMI model has similar emphasis on both the genotypes and the environments. The genotype groupings are interestingly very similar for all the analyses. This is an indication that the different analyses are identifying similar trends in the data. The analysis of the environments is useful in identifying those that are similar and in so doing redundant environments can be removed. The ANOVA cannot elucidate the environment effect further than the estimation of the variance component.

4.5 Optimum Allocation of Resources

An investigation on the effect of manipulating the number of years, locations and replications on the theoretical variance of the mean ($V_{\bar{x}}$) was used to determine the optimum number of these factors that need to be included in the trial program. The combination of the number of years, locations and replications that results in the lowest $V_{\bar{x}}$ will be the best combination to use; noting that the lower the $V_{\bar{x}}$, the more precise the yield estimates. It must be remembered though, an increase in any of these factors leads to an exponential increase in the amount of total resources required. The effectiveness of these three different factors on increasing precision was also considered, the thinking being that some of the factors being considered may be more effective than others in reducing the $V_{\bar{x}}$. The measure of precision is given in the methods section.

The three dimensional graph given in Figure 4.5a represents the LSD% plotted for six replications and twenty locations. As can be seen, both locations and replications have the ability to increase the precision of trial means. The increase in precision by increasing the number of locations, increases the gain in precision far more effectively than an increase in the number of replications. The increase in precision by increasing the replication is minimal when compared to the increase gained by an increase in the number of locations. Thus, for an increase in precision in a trial program, an increase in the number of locations used in the testing of varieties is far more preferable than an increase in the number of replications. From Figure 4.5b, it can clearly be seen that the rate of gain in precision decreases as the number of locations increases, the law of diminishing returns applies; consider the instance of two replications, increasing the

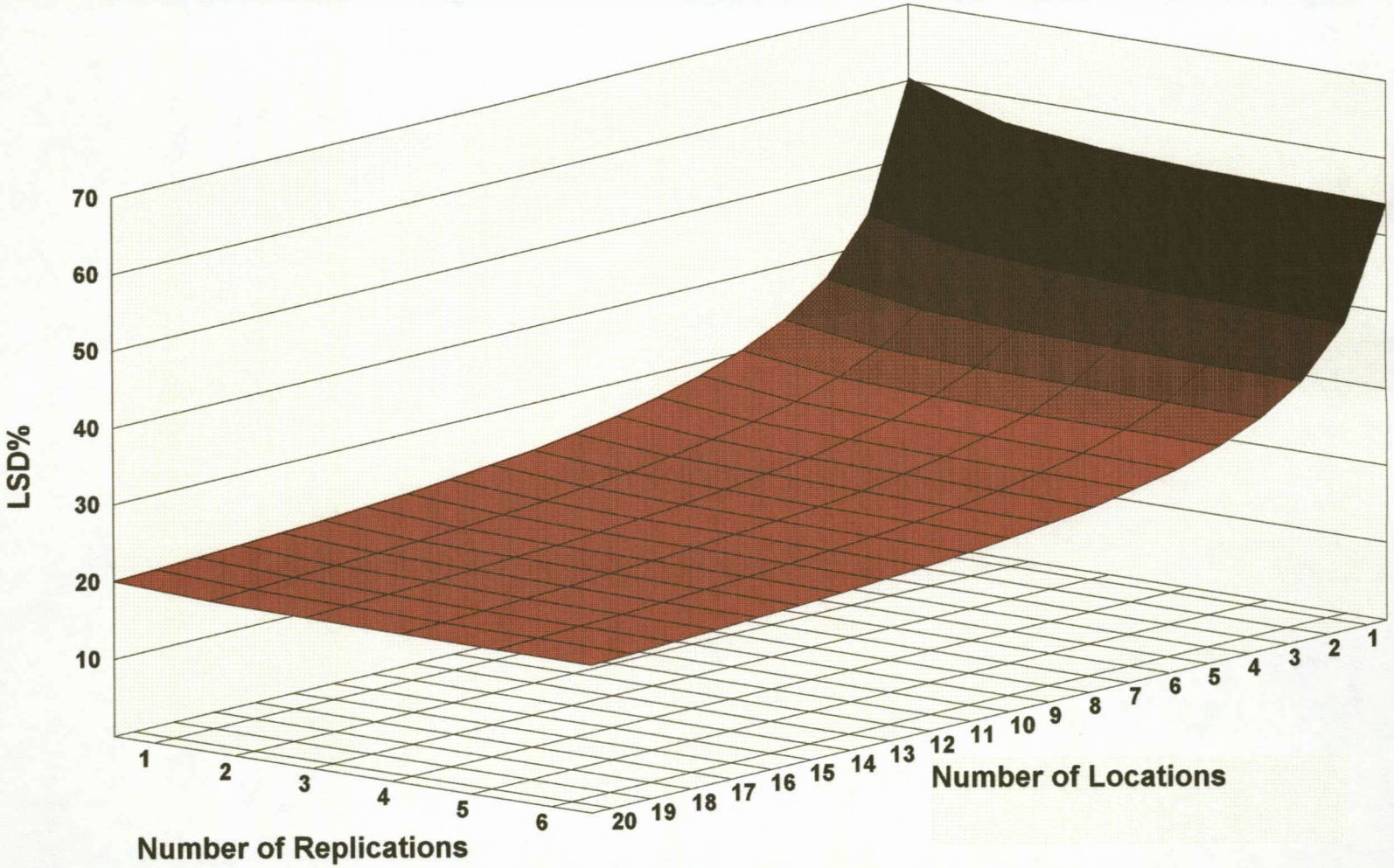


Figure 4.5a Effect of number of locations and replications on LSD%

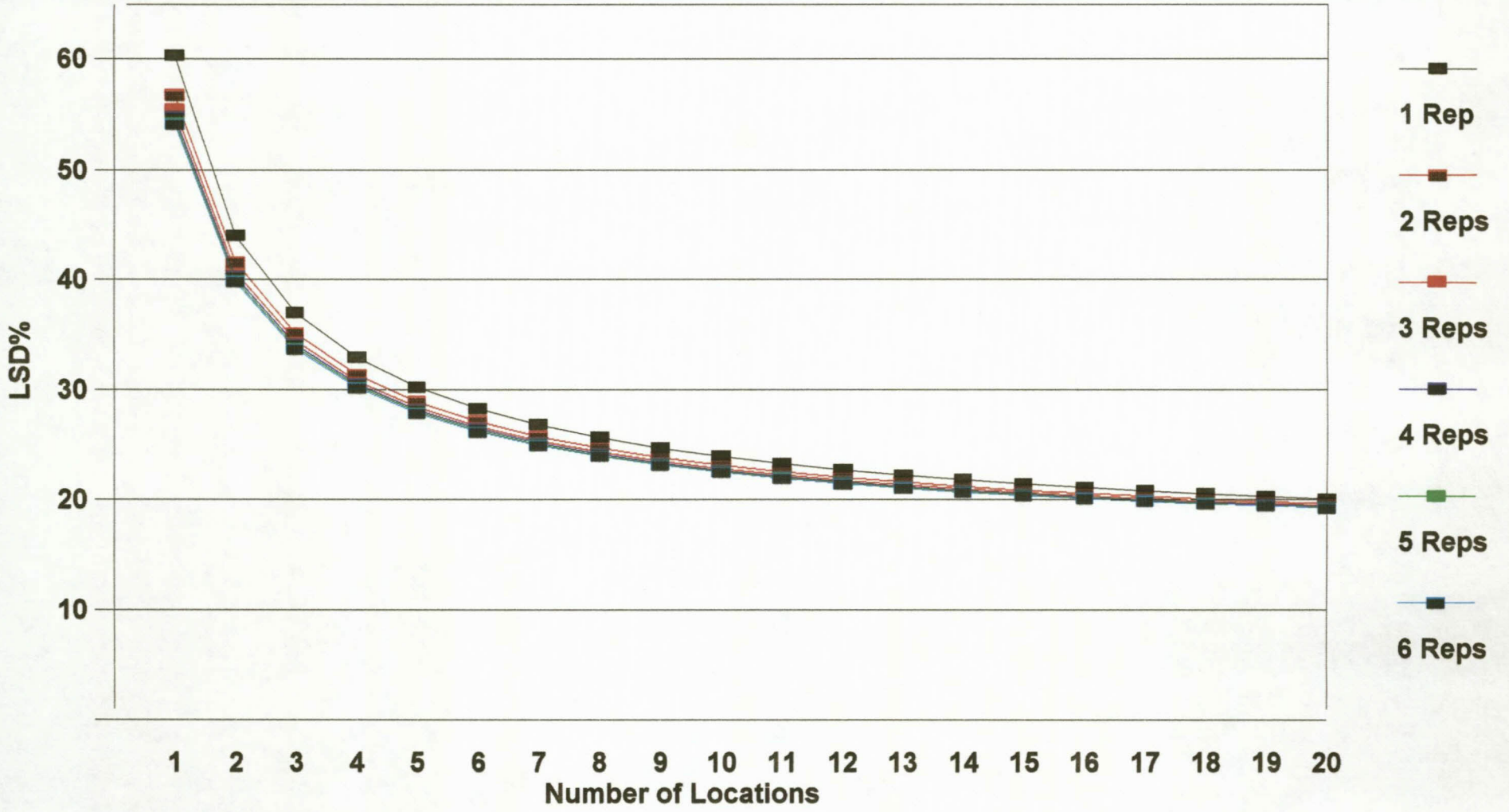


Figure 4.5b LSD% for number of locations using different numbers of replication

number of locations from one to two results in a 15.18% gain in precision, but the addition of another location only results in a 6.49% gain in precision. With the addition of a fourth location only a 3.74% gain in precision is achieved and so on. The optimum number of locations that need to be used in a variety testing program is between four and six. Increasing the number of locations to beyond six will result in very little increase in precision. The optimum number of replications to use is two.

Figure 4.5c represents the effect of both number of years and locations on gain in precision. An increase in the number of years in the testing program will also result in a gain in precision. However, this rate of gain in precision is lower than that of the gain achieved by increasing the number of locations. The gain in precision obtained by increasing the number of years included in the test is substantially more than that achieved by increasing the number of replications. From Figure 4.5d, it can be seen that the law of diminishing returns also applies to the gain in precision by increasing the number of years included in the testing. Assuming that trials at four locations is optimum, then this testing must be carried out over three years for the most efficient maximum gain in precision. If for instance only two years of trials are possible, then the same precision as three years of trials can be achieved in two years of testing by increasing the number of locations to six per year. The number of locations and years can thus be manipulated to achieve the gain in precision desired within a specific time frame.

A snap bean variety evaluation program should include between 4-6locations, and 2-3 years of testing. The actual number of these factors that are to be used depends on the

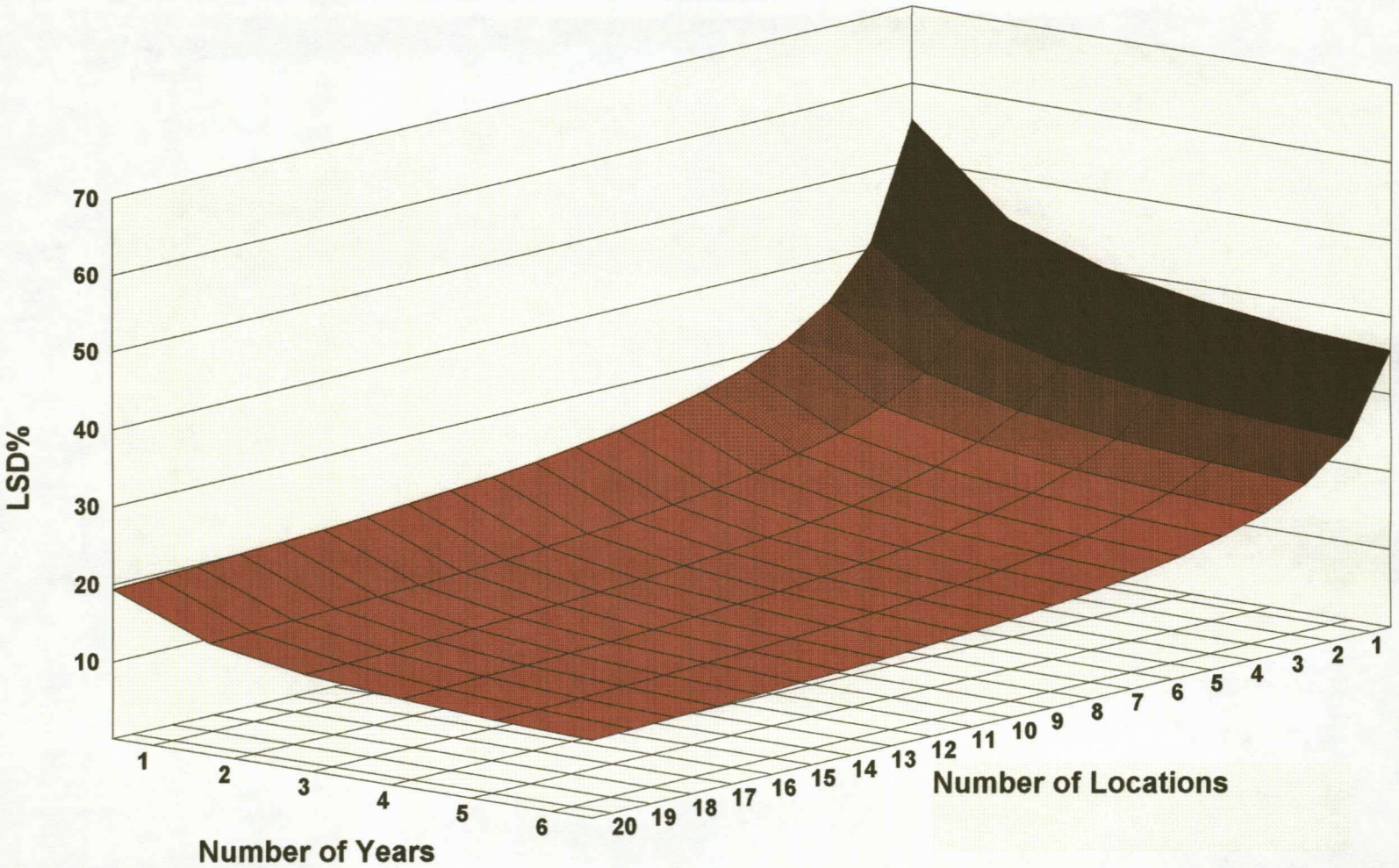


Figure 4.5c Effect of number of locations and replications on LSD%

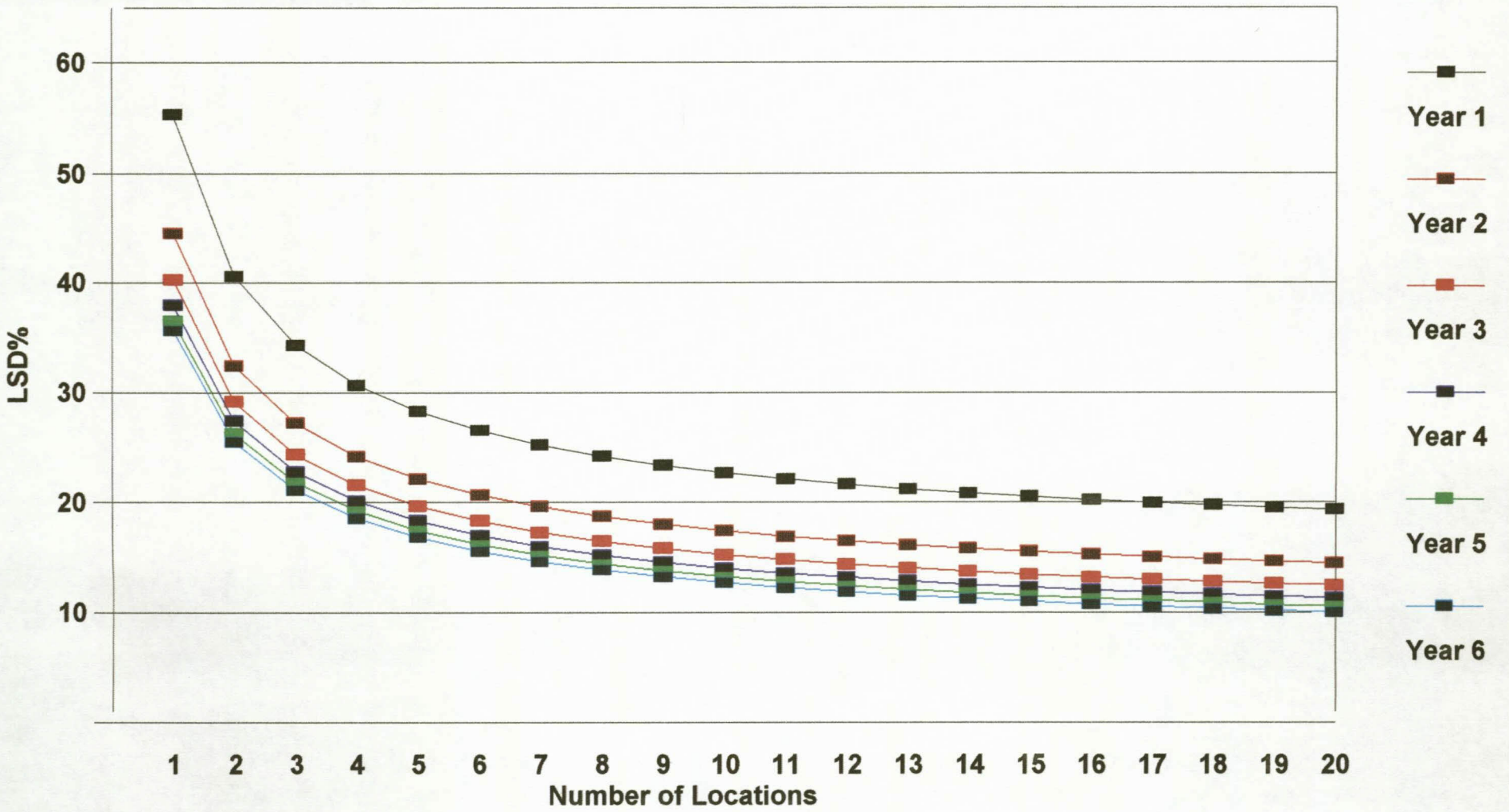


Figure 4.5d LSD% for number of locations using different numbers of year

researchers needs and the availability of other resources such as time. The number of replications should not exceed two, as this is the least efficient in reducing the variance of the mean.

Chapter 5

Summary

1. The objective of this study was to establish the significance of GxE interaction on yield in snap beans, secondly to determine and compare the appropriate analytical method to describe the GxE interaction. Thirdly, to determine the optimal number of locations, years and replications required for efficient testing.

2. Nine snap bean varieties were included in trials at four locations, with the trials being repeated over three years (1998-2000).The pod yield was measured. The data was analysed applying the ANOVA, Finlay and Wilkinson, and the AMMI models. Variance components were also calculated to determine the influence of the different effects on the theoretical variance of the mean thus allowing for the optimal allocation of locations, years and replications.

3. The GxL, GxY and GxYxL interactions were significant for pod yield. The estimated variance components for yield shows that σ^2_g, σ^2_{gl}, σ^2_{gy} and σ^2_{gly} contribute 23.5, 18.8, 4.9 and 34.8 percent respectively to the total phenotypic variance. The contribution of the error variance in these trials was 18.0 percent.

4. According to Finlay and Wilkinson, genotypes G3, G5 and G9 were classified as having average stability and high yields therefore good adaptability. Genotype G8 has above average stability while G1, G2, G4, G6 and G7 were classified as genotypes having average stability with low yields therefore poor adaptability.

5. The AMMI ANOVA subdivides the GxE effect further using principal component analysis. Genotypes G5 and G9 are classified as having high yields but low stability whereas genotypes G3 and G8 possessed both high yields and high stability. Genotypes G1, G2, G4 and G7 have low yields but are highly stable. Genotype G6 has low yields and low stability. The AMMI model was also used to classify the different locations. Location A and D can differentiate more amongst the genotypes whereas Location B and C have low discrimination amongst the genotypes.

6. According to the ANOVA, genotype G3 was classified as being most stable whereas the Finlay and Wilkinson, and the AMMI models identified genotype G8 as being most stable. The ANOVA ranked genotype G8 fourth for stability. Both AMMI and Finlay and Wilkinson models ranked genotype G1 fifth for stability but according to the ANOVA it was most unstable. Finlay and Wilkinson identified genotype G7 as being the most unstable while the most unstable genotype in the AMMI analysis was G6. The correlations between the stability rankings of the genotypes for the different methods were relatively low.

7. According to the estimated variance of a variety mean, two years, six locations and two replications is the optimum number of years, locations and replications for efficient testing of snap beans for yield in South Africa.

Opsomming

1. Die doel van die studie is om die betekenisvolheid van GxE vir opbrengs by groenbone te ondersoek. Tweedens om die verskillende analitiese metodes vir

die beskrywing van genotipe x omgewingsinteraksie met mekaar te vergelyk. Derdens om die optimale aantal lokaliteite, jare en herhalings vir doeltreffende evaluering te bepaal.

2. Nege groenboon variëteite is ingesluit in die proewe wat op vier lokaliteite oor 'n tydperk van drie jaar (1998-2000) aangeplant is. Die perseelopbrengste is bepaal. Die data is met behulp van ANOVA, Finlay and Wilkinson en AMMI modelle, geannaliseer. Variansiekomponente is beraam om die invloed van die verskillende effekte op die teoretiese variansie van 'n gemiddeld te bereken, ten einde die optimale allokasie van lokaliteite, jare en herhalings te bepaal.

3. Die GxL, GxY en GxYxL interaksies was betekenisvol vir opbrengs. Die beraamde variansiekomponente vir opbrengs toon dat $\sigma^2_{g,}$, σ^2_{g1}, σ^2_{gy} en σ^2_{gly} onderskeidelik 23.5, 18.8, 4.9, en 34.8 persent van die totale fenotipiese variansie verklaar. Die bydrae van die fout variansie was 18.0 persent.

4. Volgens Finlay en Wilkinson beskik genotipes G3, G5 en G9 oor gemiddelde stabiliteit met 'n hoë opbrengs met ander woorde oor goeie aanpasbaarheid. Genotipe G8 beskik oor bogemiddelde stabiliteit, terwyl G1, G2, G4, G6 en G7 geklassifiseer kan word as genotipes met gemiddelde stabiliteit en lae opbrengs met ander woorde lae aanpasbaarheid.

5. Die AMMI ANOVA onderverdeel die G X E effekte by wyse van hoofkomponent ontledings. Genotipes G5 en G9 word volgens die metode geklassifiseer as genotipes met lae stabiliteit en hoë opbrengste. Genotipes G3 en G8 beskik oor

hoë stabiliteit tesame met hoë opbrengste. Genotipes G1, G2, G4 en G7 is hoogs stabiel met lae opbrengste. Genotipe G6 beskik oor lae stabiliteit en lae opbrengs. Die AMMI model is ook gebruik om die verskillende lokaliteite te klassifiseer. Lokaliteite A en D differensieer tot 'n groter mate tussen genotipes, terwyl lokaliteite B en C tot 'n mindere mate tussen genotipes diskrimineer.

6. Volgens die variansie-ontleding word genotipe G3 as meer stabiel geklassifiseer, terwyl die Finlay & Wilkinson en AMMI modelle genotipe G8 as die mees stabiele identifiseer. Die variansie-ontleding het genotipe G8 vierde beste vir stabiliteit gegroepeer. Beide die AMMI en Finlay and Wilkinson modelle het genotipe G1 vyfde vir stabiliteit gegroepeer, maar volgens die variansie-ontleding is die genotipe hoogs onstabiel. Finlay & Wilkinson het G7 as die mees, onstabiel genotipe geïdentifiseer, terwyl die AMMI ontleding genotipe G6 as die mees onstabiele genotipe geklassifiseer het. Die korrelasies tussen die stabiliteits rangordes van die genotipes vir die verskillende modelle is relatief laag.

7. Volgens die beraamde variansies van 'n cultivar gemiddelde is twee jaar, ses lokaliteite en twee replikasies die optimale aantal jare, lokaliteite en herhalings om groenbone vir opbrengs in Suid-Afrika doeltreffend te evalueer.

Chapter 6

Conclusion and Recommendations

The interaction of genotype with the environment is significant in snap beans. This GxE results in different genotype rankings in different environments, thus testing of the genotypes needs to be conducted at different locations. The genotype performance at the different locations were not consistent from year to year therefore testing must also take place across years. The GxE has complicated variety selection by requiring testing over locations and years for reliable estimates of yields to be obtained. This increased testing procedure required in turn places a bigger demand on the resources available. Thus analytical methods that effectively take account of the GxE and the efficient use of the resources available are essential for a successful variety evaluation program.

Among the analytical methods investigated, the simplest model was the ANOVA followed by the Finlay and Wilkinson model. The AMMI model was the most complex model to fit. With the advent of computers, the complexity of the models is not important. AMMI was the most effective in reducing the amount of noise present in the data, thus allowing for the maximum elucidation of the underlying structure of the GxE. The AMMI estimates for yield were representative of field performance. Although, the genotype groupings obtained were generally similar for all three analyses, the AMMI model seems to encompass the Finlay and Wilkinson and ANOVA model. Analysis of the environment can also be conducted using AMMI. The AMMI biplot is convenient and effective in displaying the analysis of the results. For the investigation of adaptability, the Finlay and Wilkinson model is effective whereas the AMMI model is suitable for investigations into static stability. It is therefore recommended that the AMMI analysis be used for the study

of GxE and the Finlay and Wilkinson model for the investigation of adaptability.

The most efficient gain in precision can be achieved using 4-6 locations over 2-3 years of trials using two replications. From the AMMI analysis, it was found that Location A and Location D have the same discriminating value for genotypes. Thus only one of these locations is necessary. It was also felt that one other location, either Location B or Location C, needs to be included as an alternate test environment; the use of two locations is recommended by AMMI. A compromise needs to be reached between the most gain in precision and maximum gain in environmental variation. It is recommended that the variety trials be conducted at three locations using either Location A or Location D, and Location B and Location C over 2-3 years. The number of replications to use is two.

Acknowledgements

I am grateful to:

1. Pannar (Pty) Ltd for the funding of this study.

2. Dr Bryan Featonby-Smith for his assistance in initiating this project.

3. My family and Nirvana for their support and assistance.

Appendix A

Figure 4.2d

Figure 4.2e

Figure 4.2f

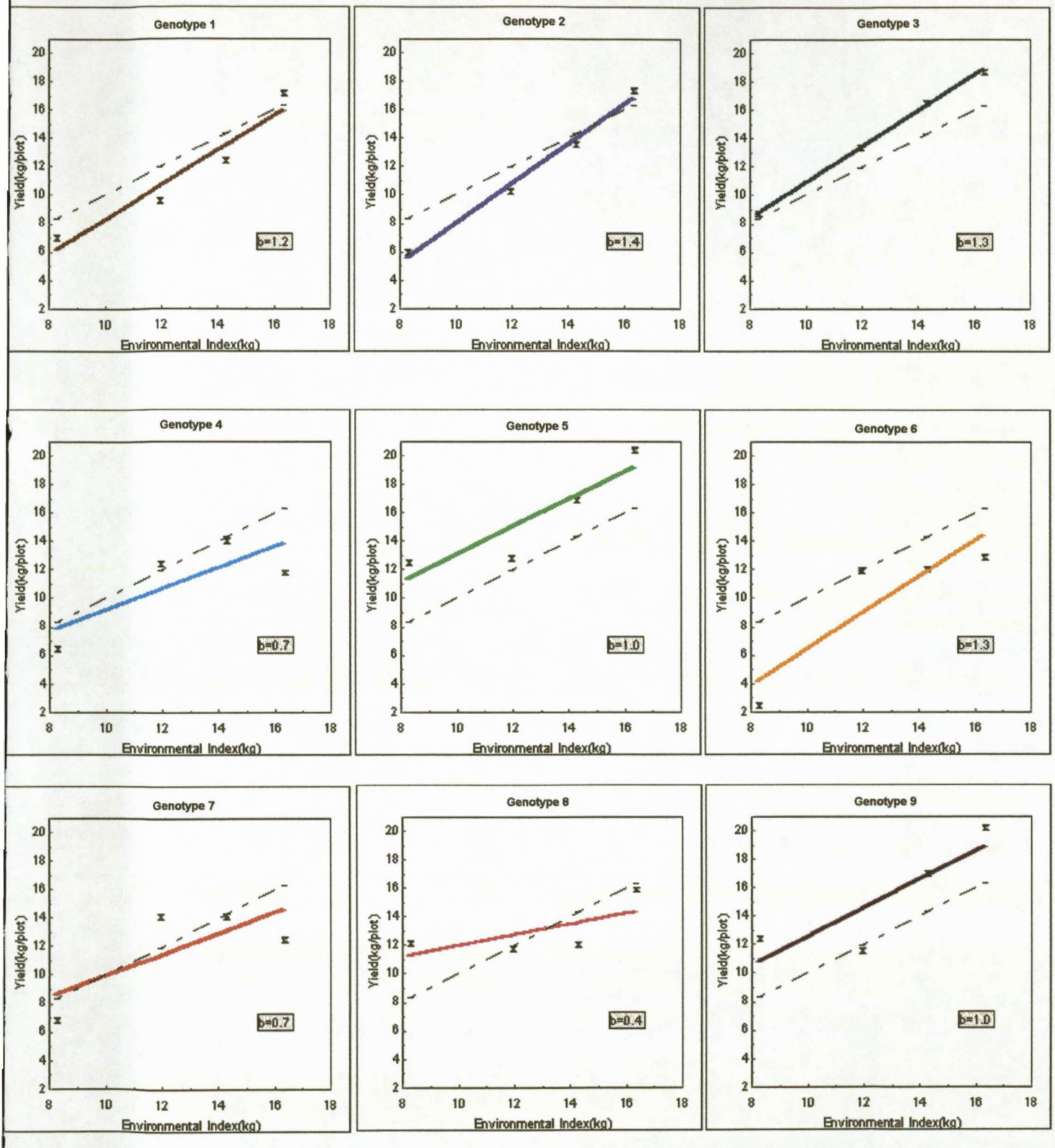


Figure 4.2d Regression of genotype yield on environmental index across four locations within Year1. The hourglasses are the actual yield, the dashed line indicates the mean yield of all genotypes, and the regression coefficient(b) is also given.

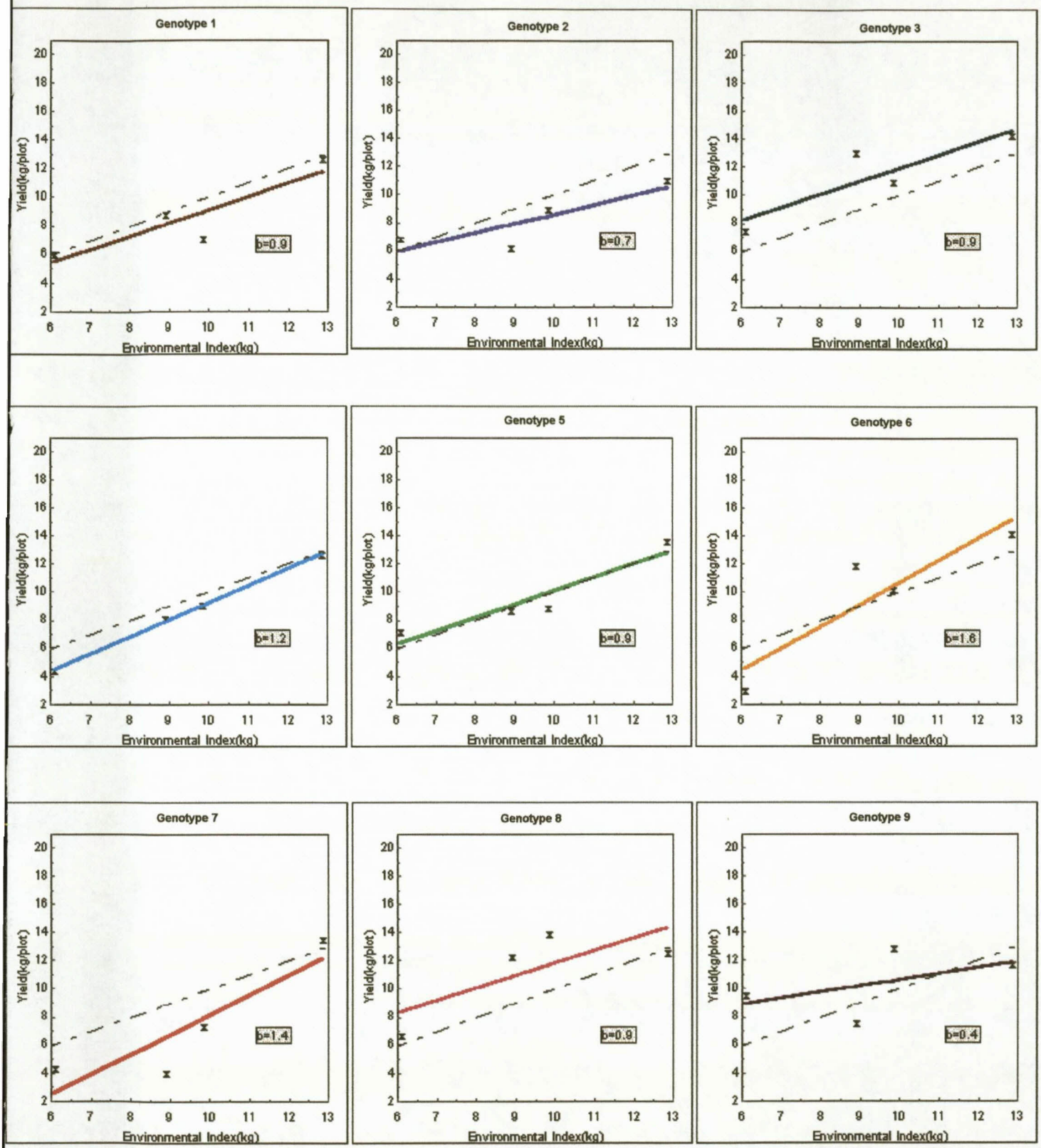


Figure 4.2e Regression of genotype yield on environmental index across four locations within Year2. The hourglasses are the actual yield, the dashed line indicates the mean yield of all genotypes, and the regression coefficient(b) is also given.

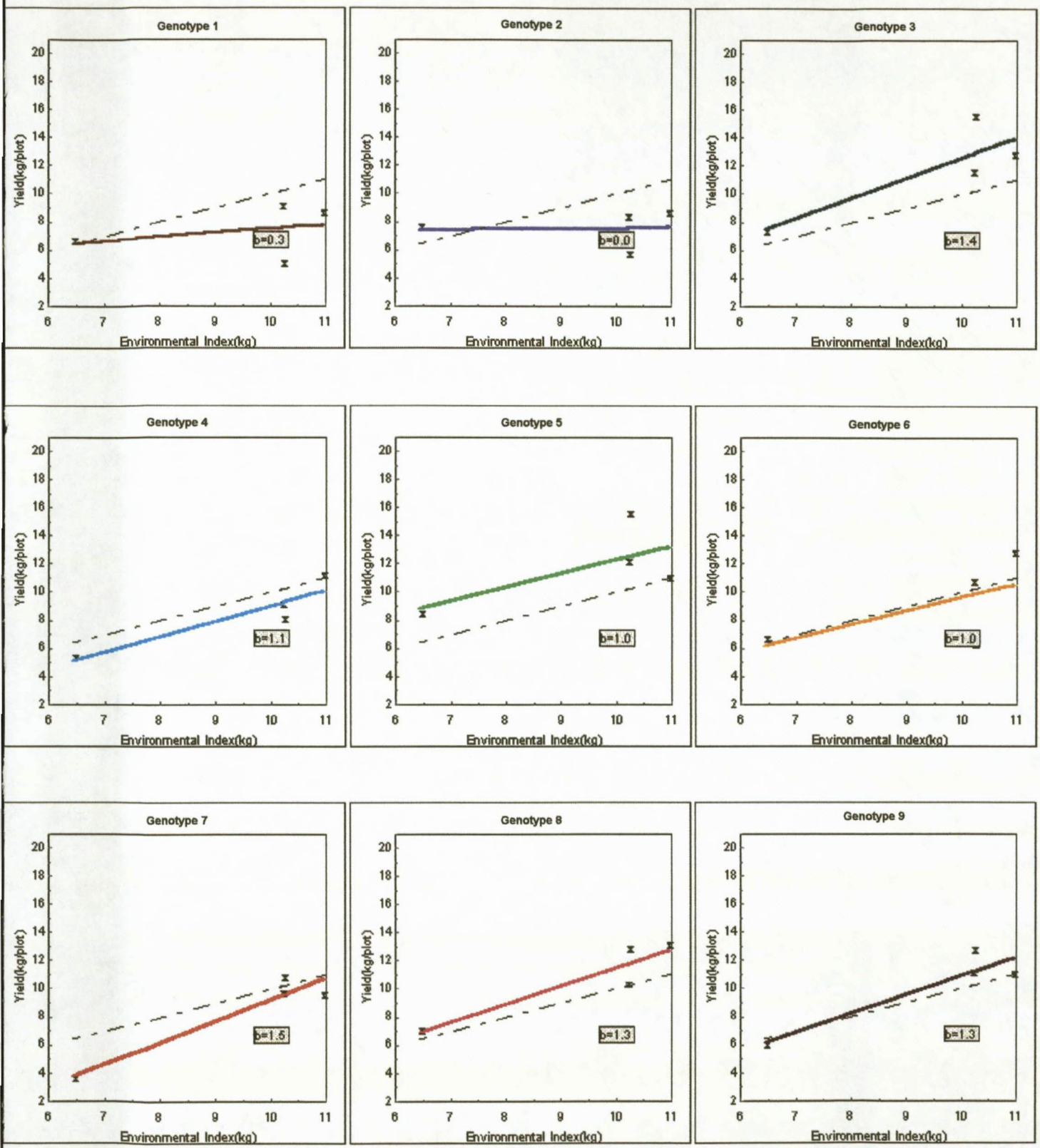


Figure 4.2f Regression of genotype yield on environmental index across four locations within Year3. The hourglasses are the actual yield, the dashed line indicates the mean yield of all genotypes, and the regression coefficient(b) is also given.

References

Alagarswamy, G., Chandra, S., 1998. Pattern analysis of international sorghum multi-environment trials for grain-yield adaptation. *Theor. Appl. Genetics* 96, 397-405.

Allard, R.W., Bradshaw, A.D., 1964. Implications of genotype x environmental interactions in applied plant breeding. *Crop Sci.* 4, 503-508.

Annicchiarico, P., 1997. Additive main effects and multiplicative interaction (AMMI) analysis of genotype x environment interaction in variety trials repeated over years. *Theor. Appl. Genetics* 94, 1072-1077.

Asins, M.J., Mestre, P., Garcia, J.E., Dicenta, F., Carbonell, E.A., 1994. Genotype x environment interaction in QTL analysis of an intervarietal almond cross by means of genetic markers. *Theor. Appl. Genetics* 89, 358-364.

Baker, R.J., 1988. Tests for crossover genotype x environmental interactions. *Canadian J. Plant Sci.* 68, 405-410.

Basford, K.E., Cooper, M., 1998. Genotype x environment interactions and some considerations of their implications for wheat breeding in Australia. *Aust. J. Agric. Res.* 49, 153-174.

Beavis, W.D., Keim, P., 1996. Identification of quantitative trait loci that are affected by environment. In: *Genotype-by-Environment Interaction*. Eds. Kang and Gauch, Jr. CRC Press, Boca Raton, Florida , 123-149.

Becker, H.C., Leon, J., 1988. Stability analysis in plant breeding. *Plant Breeding* 101, 1-23.

Bramel-Cox, P.J., 1996. Breeding for reliability of performance across unpredictable environments. In: *Genotype-by-Environment Interaction*. Eds. Kang and Gauch, Jr. CRC Press, Boca Raton, Florida , 309-340.

Brandle, J.E., Brule-Babel, A.L., 1991. An integrated approach to oilseed rape cultivar selection using phenotypic stability. *Theor. Appl. Genetics* 81, 679-684.

Breese, E.L., 1969. The measurement and significance of genotype x environment interactions in grasses. *Heredity* 24, 27-44.

Brennan, P.S., Byth, D.E., Drake, D.W., De Lacy, I.H., Butler, D.G., 1981. Determination of the location and number of test environments for a wheat cultivar evaluation program. *Aust. J. Agric. Res.* 32, 189-201.

Bull, J.K., Basford, K.E., DeLacy, I.H., Cooper, M., 1992. Classifying genotypic data from plant breeding trials: a preliminary investigation using repeated checks. *Theor. Appl. Genetics* 85, 461-469.

Ceccarelli, S., 1989. Wide adaptation:How wide? *Euphytica* 40, 197-205.

Chakroun, M., Takiaferro, C.M., McNew, R.W., 1990. Genotype-environment interactions of bermudagrass forage yields. *Crop Sci.* 30, 49-53.

Charmet, G., Balfourier, F., Ravel, C., Denis, J.B., 1993. Genotype x environment interactions in a core collection of french perennial ryegrass populations. *Theor. Appl. Genetics* 86, 731-736.

Comstock, R.E., 1977. Quantitative genetic and design of breeding programs. In: *Proceedings of the International Conference on Quantitative Genetics*. Eds. Pollak, Kempthorn and Bailey. The Iowa State University Press: Ames, IA. 705-718.

Comstock, R.E., Moll, R.H., 1963. Genotype-environment interactions. In: *Statistical Genetics and Plant Breeding*. Eds. Hanson and Robinson. 982:164-196.

Cooper, M., Byth, D.E., DeLacy, I.H., 1993a. A procedure to assess the relative merit of classification strategies for grouping environments to assist selection in plant breeding regional evaluation trials. *Field Crops Res.* 35, 63-74.

Cooper, M., DeLacy, I.H., and Eisemann, R.L., 1993b. Recent advances in the study of genotype by environment interaction and their application to plant breeding. In: *Focused Plant Improvement: Towards Responsible and Sustainable Agriculture. Proceedings of the Tenth Australian Plant Breeding Conference*. Eds. Imrie and Hacker. 116-131.

Cooper, M., Byth, D.E., DeLacy, I.H., Woodruff, D.R., 1993c. Predicting grain yield in Australian environments using data from CIMMYT international wheat performance trials. 1. Potential for exploiting correlated response to selection. *Field Crops Res*. 32, 305-322.

Cooper, M., DeLacy, I.H., 1994. Relationships amongst analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theor. Appl. Genetics* 88, 561-572.

Cooper, M., Stucker, R.E., DeLacy, I.H., Harch, B.D., 1997. Wheat breeding nurseries, target environments, and indirect selection for grain yield. *Crop Sci.* 37, 1168-1176.

Cooper, M., Woodruff, D.R., Eisemann, R.L., Brennan, P.S., DeLacy, I.H., 1995. A selection strategy to accommodate genotype-environment interaction for grain yield of wheat: managed-environments for selection among genotypes. *Theor. Appl. Genetics* 90, 492-502.

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.

Crossa, J., 1990. Statistical analyses of multilocation trials. *Adv. in Agron*. 4, 55-85.

Crossa, J., Cornelius, P.L., 1993. Recent developments in multiplicative models for cultivar trials. In: *Crop Science I.* 1, 571-577.

Crossa, J., Fox, P.N., Pfeiffer, W.H., Rajaram, S., Gauch Jr, H.G., 1991. AMMI adjustment for statistical analysis of an international wheat yield trial. *Theor. Appl. Genetics* 81, 27-37.

Crossa, J., Gauch Jr, H.G., Zobel, R.W., 1990. Additive main effects and multiplicative interaction analysis of two international maize yield trials. *Crop Sci.* 30, 493-500.

DeLacy, I.H., Cooper, M., Basford, K.E., 1996. Relationships among analytical methods used to study genotype-by-environment interactions and evaluation of their impact on response to selection. In: *Genotype-by-Environment Interaction*. Eds. Kang and Gauch, Jr. CRC Press, Boca Raton, Florida , 51-84.

Eberhart, S.A., Russell, W.A., 1966. Stability parameters for comparing varieties. *Crop Sci.* 6, 36-40.

Eisemann, R.L., Cooper, M., Woodruff, D.R., 1990. Beyond the analytical methodology-better interpretation and exploitation of genotype-by-environment interaction in breeding. In: *Genotype-By-Environment Interaction and Plant Breeding*. Ed. Kang. Louisiana State University: Baton Rouge, LA. 108-117.

Eskridge, K.M., 1990. Selection of stable cultivars using a safety-first rule. *Crop Sci.* 30, 369-374.

Federer, W.T., Scully, B.T., 1993. A parsimonious statistical design and breeding procedure for evaluating and selecting desirable characteristics over environments. *Theor. Appl. Genetics* 86, 612-620.

Finlay, K.W., Wilkinson, G.N., 1963. The analysis of adaptation in a plant-breeding programme. *Aust. J. Agric. Res.* 14, 742-754.

Francis, T.R., and Kannenberg, L.W., 1978. Yield stability studies in short-term maize. 1. A descriptive method for grouping genotypes. *Canadian J. of Plant Sci.* 58,1029-1034.

Freeman, G.H., 1973. Statistical methods for the analysis of genotype-environment interactions. *Heredity* 31, 339-354.

Freeman, G.H., Dowker, B.D., 1973. The analysis of variation between and within genotypes and environments. *Heredity* 30, 97-109.

Freeman, G.H., Perkins, J.M., 1971. Environmental and genotype-environmental components of variability. *Heredity* 27, 15-23.

Fripp, Y.J., 1972. Genotype-environmental interactions in *Schizophyllum commune*. *Heredity* 28, 223-238.

Gauch, H.G., Zobel, R.W., 1996. AMMI analysis of yield trials. In: *Genotype-by-Environment Interaction*. Eds. Kang and Gauch, Jr. CRC Press, Boca Raton, Florida , 85-122.

Gorman, D.P., Kang, M.S., Milam, M.R., 1989. Contribution of weather variables to genotype x environment interaction in grain sorghum. *Plant Breeding* 103, 299-303.

Grafius, J.E., 1958. Biologic measurement of environment as an aid in selection. *Agron. J.* 50, 506-508.

Hamblin, J., Fisher, H.M., Ridings, H.I., 1980. The choice of locality for plant breeding when selecting for high yield and general adaptation. *Euphytica* 29, 161-168.

Hammer, G.L., Vanderlip, R.L., 1989. Genotype x environment interaction in grain sorghum. III. Modelling the impact in field environments. *Crop Sci.* 29, 385-391.

Hardwick, R.C., Wood, J.T., 1972. Regression methods for studying genotype-environment interactions. *Heredity* 28, 209-222.

Hernandez, C.M., Crossa, J., Castillo, A., 1993. The area under the function: an index for selecting desirable genotypes. *Theor. Appl. Genetics* 87, 409-415.

Hohls, T., 1995. Analysis of genotype-environment interactions. *South African J. Sci.* 91, 121-124.

Horner, T.W., Frey, K.J., 1957. Methods for determining natural areas for oat varietal recommendations. *Agron. J.* 49, 313-315.

Hühn, M., 1996. Nonparametric analysis of genotype x environment interactions by ranks. In: *Genotype-by-Environment Interaction*. Eds. Kang and Gauch, Jr. CRC Press, Boca Raton, Florida , 235-272.

Hühn, M., Lotito, S., Piepho, H.P., 1993. Relationships between genotype x environment interactions and rank orders for a set of genotypes tested in different environments. *Theor. Appl. Genetics* 86, 943-950.

Jacobsen, S.E., Hill, J., Stolen, O., 1996. Stability of quantitative traits in quinoa(*Chenopodium quinoa*). *Theor. Appl. Genetics* 93, 110-116.

Jansen, R.C., Van Ooijen, J.W., Stam, P., Lister, C., Dean, C., 1995. Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. *Theor. Appl. Genetics* 91, 33-37.

Kang, M.S., 1990. Understanding and utilization of genotype-by-environment interaction in plant breeding. In: *Genotype-By-Environment Interaction and Plant Breeding.* Ed. Kang. Louisiana State University: Baton Rouge, LA. , 52-68.

Kang, M.S., Gorman, D.P., Pham, H.N., 1991. Application of a stability statistic to international maize yield trials. *Theor. Appl. Genetics* 81, 162-165.

Kang, M.S., Magari, R., 1996. New developments in selecting for phenotypic stability in crop breeding. In: *Genotype-by-Environment Interaction.* Eds. Kang and Gauch, Jr. CRC Press, Boca Raton, Florida , 1-14.

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

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

Lin, C.S., Binns, M.R., 1988a. A method of analysing cultivar x location x year experiments: a new stability parameter. *Theor. Appl. Genetics* 76, 425-430.

Lin, C.S., Binns, M.R., 1988b. A superiority measure of cultivar performance for cultivar x location data. *Canadian J. Plant Sci* 68, 193-198.

Lin, C.S., Binns, M.R., 1991a. Assessment of a method for cultivar selection based on regional trial data. *Theor. Appl. Genetics* 82, 379-388.

Lin, C.S., Binns, M.R., 1991b. Genetic properties of four types of stability parameter. *Theor. Appl. Genetics* 82, 505-509.

Lin, C.S., Binns, M.R., 1994. Concepts and methods for analysing regional trial data for cultivar and location selection. *Plant Breeding Reviews* 12, 271-297.

Lin, C.S., Binns, M.R., 1985. Procedural approach for assessing cultivar-location data: pairwise genotype x environment interactions of test cultivars with checks. *Canadian J. Plant Sci.* 65, 1065-1071.

Lin, C.S., Binns, M.R., Lefkovitch, L.P., 1986. Stability analysis: Where do we stand?. *Crop Sci.* 26, 894-900.

Lin, C.S., Morrison, M.J., 1992. Selection of test locations for regional trials of barley. *Theor. Appl. Genetics* 83, 968-972.

Louw, J.H., 1990. A selection index to cope with genotype x environment interaction with an application to wheat breeding. *Plant Breeding* 104, 346-352.

Marais, G.F., 1986. The nature of genotype x environment interactions for yield in wheat extension trials in the winter rainfall region. *South African J. of Plant and Soil* 3, 123-129.

McKeand, S.E., Erikson, G., Roberds, J.H., 1997. Genotype by environment interaction for index traits that combine growth and wood density in lobolly pine. *Theor. Appl. Genetics* 94, 1015-1022.

Miller, P.A., Robinson, H.F., Pope, O.A., 1962. Cotton variety testing: additional information on variety x environment interactions. *Crop Sci.* 2, 349-352.

Miller, P.A., Williams, J.C., Robinson, H.F., 1959. Variety x environment interactions in cotton variety tests and their implications on testing methods. *Agron. J.* 51, 132-134.

Moreno-Gonzalez, J., Crossa, J., 1998. Combining genotype, environment and attribute variables in regression models for predicting the cell-means of multi-environment cultivar trials. *Theor. Appl. Genetics* 96, 803-811.

Muir, W., Nyquist, W.E., Xu, S., 1992. Alternative partitioning of the genotype-environment interaction. *Theor. Appl. Genetics* 84, 193-200.

Nachit, M.M., Nachit, G., Ketata, H., Gauch Jr, H.G., Zobel, R.W., 1992. Use of AMMI and linear regression models to analyse genotype-environment interaction in durum wheat. *Theor. Appl. Genetics* 83, 597-601.

Nix, H.A., 1975. The Australian climate and its effects on grain yield and quality. In: *Australian Field Crops 1: Wheat and Other Temperate Cereals*. Eds. Lazenby and Matheson. Angus and Robertson Publishers: Australia, 183-226.

Odendaal, T.E.M., van Deventer, C.S., 1987. Genotipe-omgewingsinteraksie by sojabone. *South African J. of Plant and Soil* 4, 65-69.

Pederson, D.G., Rathjen, A.J., 1981. Choosing trial sites to maximize selection response for grain yield in spring wheat. *Aust. J. Agric. Res.* 32, 411-424.

Peltonen-Sainio, P., Moore, K., Pehu, E., 1993. Phenotypic stability of oats measured with different stability analyses. *J. of Agric. Sci.* 121, 13-19.

Pham, H.N., Kang, M.S., 1988. Interrelationships among and repeatability of several stability statistics estimated from international maize trials. *Crop Sci.* 28, 925-928.

Phiepo, H.P., 1994a. Best Linear Unbiased Prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interaction (AMMI) analysis. *Theor. Appl. Genetics* 89, 647-654.

Piepho, H.P., 1994b. Missing observations in the analysis of stability. *Heredity* 72, 141-145.

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

Purchase, J.L., 1997. Parametric analysis to describe genotype x environment interaction and yield stability in winter wheat. PhD Thesis. University of the Orange Free State, Bloemfontein, Republic of South Africa.

Rasmusson, D.C., Lambert, J.W., 1961. Variety x environment interactions in barley variety tests. *Crop Sci.* 1, 261-262.

Rathjen, A.J., 1994. The biological basis of genotype by environment interaction-its definition and management. In: *Proceedings of the Seventh Assembly: Wheat Breeding-Into the Century*. Eds. Paull, Dundas, Shepherd and Hollamby. Wheat Breeding Society of Australia, 13-17.

Risi, J., Galwey, N.W., 1991. Genotype x environment interaction in the Andean grain crop Quinoa(*Chenopodium quinoa*) in temperate environments. *Plant Breeding* 107, 141-147.

Robert, N., Denis, J.B., 1996. Stability of baking quality in bread wheat using several stability statistics. *Theor. Appl. Genetics* 93, 172-178.

Romagosa, I., Fox, P.N., Garcia del Moral, L.F., Ramos, J.M., Garcia del Moral, B., Roca de Togores, F., Molina-Cano, J.L., 1993. Integration of Statistical and Physiological analyses of Adaptation of near-isogenic barley lines. *Theor. Appl. Genetics* 86, 822-826.

Saindon, G., Schaalje, G.B., 1993. Evaluation of locations for testing dry bean cultivars in western Canada using statistical procedures, biological interpretation and multiple trait. *Canadian J. Plant Sci* 73, 985-994.

Schutz, W.M., Bernard, R.L., 1967. Genotype x environment interactions in the regional testing of soybean strains. *South African J. of Plant and Soil* 7, 125-130.

Shorter, R., Lawn, R.J., Hammer, G.L., 1991. Improving genotypic adaptation in crops-A role for breeders, physiologists and modellers. *Exploratory Agric.* 27, 155-175.

Shukla, G.K., 1972. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* 29, 237-245.

Simmonds, N.W., 1991. Selection for local adaptation in a *Plant Breeding* programme. *Theor. Appl. Genetics* 82, 363-367.

Sprague, G.F., Federer, W.T., 1951. A comparison of variance components in corn yield trials:II. Error, Year x Variety, Location x Variety and Variety components. *Agron. J.* 43, 535-541.

Steyn, P.J., Visser, A.F., Smith, M.F., Schoeman, J.L., 1993. AMMI analysis of potato cultivar yield trials. *South African J. of Plant and Soil* 10, 28-34.

St-Pierre, C.A., Klinck, H.R., Gauthier, F.M., 1967. Early generation selection under different environments as it influences adaptation of Barley. *Canadian J. Plant Sci.* 47, 507-517.

Swallow, W.H., Wehner, T.C., 1989. Optimum allocation of plots to years, seasons, locations and replications, and its application to once-over-harvest cucumber trials. *Euphytica* 43, 59-68.

van Deventer, C.S., 1986. Genotipe x omgewingsinteraksie by winterkoring in die Oranje-Vrystaat. *South African J. of Plant and Soil* 3, 147-148.

van Eeuwijk, F.A., 1992. Interpreting genotype-by-environment interaction using redundancy analysis. *Theor. Appl. Genetics* 85, 89-100.

Vermeer, H., 1990. Optimising potato breeding. I. The genotypic, environmental and genotype-environment coefficients of variation for tuber yield and other traits in potato(*Solanum tuberosum L.*) under different experimental conditions. *Euphytica* 49, 229-236.

Wallace, D.H., Baudoin, J.P., Beaver, J., Coyne, D.P., Halseth, D.E., Masaya, P.N., Munger, H.M., Myers, J.R., Srlbernagel, M., Yourstone, K.S., Zobel, R.W., 1993. Improving efficiency of breeding for higher crop yield. *Theor. Appl. Genetics* 86, 27-40.

Westcott, B., 1987. A method of assessing the yield stability of crop genotypes. *J. of Agric. Sci.* 108, 267-274.

Westcott,B., 1986. Some methods of analysing genotype x environment interaction. *Heredity* 56, 243-253.

Wright, A.J., 1976. The significance for breeding of linear regression analysis of genotype-environment interactions. *Heredity* 37:83-93.

Wu, R.L., O'Malley, D.M., 1998. Nonlinear genotypic response to macro and micro-environments. *Theor. Appl. Genetics* 96, 669-675.

Yates, F., Cochran, W.G., 1938. The analysis of groups of experiments. *J. Agr. Sci* 28,556-580.

Yau, S.K., 1995. Regression and AMMI analyses of genotype x environment interactions: an empirical comparison. *Agron. J.* 87, 121-126.

Zobel, R.W., Wright, M.J., Gauch Jr, H.G., 1988. Statistical analysis of a yield trial. *Agron. J.* 80, 388-393.