2. REVIEW OF LITERATURE

2.1. Biometrical Methods to Analyze Yield Trials

Numerous methods have been proposed to analyze multi-location yield trials (Finlay and Wilkinson, 1963; Fehr and Russell, 1966; Gauch, 1985; Lin et al., 1986; Westcott, 1986; Huhn, 1987; Nassar and Huhn, 1927; Becker and Leon, 1988). These methods can be broadly classified as ordinary ANOVA, principal component analysis (PCA), linear regression (LR), and additive main effects and multiplicative interaction (AMMI) methods (Gauch, 1988; Gauch & Zobel, 1988; Zobel et al., 1988). Since the last four methods have been chosen for the present investigation, the review will be presented for the literature relevant to these methods only.

2.1.1. Analysis of variance (ANOVA)

In order to present a comprehensible account of the methods used for analyses of yield trials, the following two-way linear model is assumed for convenience:

\[ X_{ij} = \mu + e_j + g_i + (g \times e)_{ij} + e_{ij} \]

where, \( X_{ij} \) is the observed phenotypic mean value of genotype \( i \) \((i = 1, \ldots, G)\) in environment \( j \) \((j = 1, \ldots, E)\), and \( \mu \), \( e_j \), \( g_i \), \((g \times e)_{ij} \) and \( e_{ij} \) represent the overall population mean, the effect of the \( j \)th environment, the effect of the \( i \)th genotype, the effect of the interaction between \( i \)th genotype and \( j \)th environment, and the mean random error of the \( i \)th genotype in the \( j \)th environment, respectively. Symbols \( \bar{X}_i \), \( \bar{X}_j \) and \( \bar{X}_.. \) are used for marginal means of genotype \( i \), environment \( j \) and overall mean, respectively.

If the usual biometrical two-way model, used for ANOVA, is applied for analyzing yield data with genotypes tested in several environments, the estimates of genotypic and environmental effects can be obtained. These effects can explain only part of the variation, if GE interactions exist, or may explain almost whole of the variation if GE interactions are negligible or absent. These unexplained
deviations (GE interactions) in the two-way model represent the uncertainty of an estimation or of a prediction of the phenotypic value of genotype \( i \) in environment \( j \). Therefore, the “GE interaction effects describe the uncertainty or instability in the model”, or vice versa. The genotypes possessing small contributions to GE interactions are considered to be more stable than the genotypes with larger contribution.

Though the ANOVA is a very simple procedure for analyzing yield trial data, it has been criticized on many grounds. The ordinary ANOVA is an additive model and therefore describes only the main effects effectively (Snedecor and Cochran, 1980). ANOVA can test the significance of the GE interaction, but this test may prove to be misleading: an ANOVA test of GE interaction may declare it non-significant even if, in fact, the interaction is agronomically significant. This problem arises because the interaction contains a large number of degrees of freedom: given \( G \), genotypes, and \( E \), environments, the interaction contains \((G-1)(E-1)\) degrees of freedom. Therefore, even if, as is typically the case, the interaction contains 20 to 50% of the total SS—which generally exceeds the SS for genotype main effects—the interaction mean squares (MS) may nearly equal the error MS and can be declared insignificant by an F-test (Gauch, 1988 and Zobel et al., 1988). In any case ANOVA provides no insight into the particular patterns of genotypes or environments that give rise to interaction. In this connection, Lin and Butler (1990) correctly stated as follows: “its (ANOVA) major function is just to test which factor (genotype or environment) or factor combination (e.g., genotype \( \times \) environment) exerts a significant effect, but it does not tell us which factor level or particular factor combination is responsible for exhibiting significance, nor does it show us how their responses differ.”

2.1.2. Linear regression (LR) analysis
Applying the usual biometrical model, it is assumed that the effects are independent of each other. This assumption is fulfilled only when, while taking all
the genotypes together, no covariance exists between the environment effects and GE interactions. However, considering each genotype separately, this covariance may be different from zero (Comstock and Moll, 1963). The regression coefficient is a straightforward description of this covariance.

Stringfield and Slater (1934) were probably the first to calculate linear regression coefficients to characterize the specific response of genotypes to varying climatic factors. Later, this approach was modified and elaborated by a number of workers (Yates and Cochran, 1938; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966; Jinks and Perkins, 1968 a, b; Freeman and Perkins, 1971) to analyze GE interactions. With this approach, the effect of GE interaction may be expressed as follows:

$$(\text{ge})_{ij} = \beta_i e_j + d_i$$

where, $$(\text{ge})_{ij}$$ is the interaction effect of genotype $$i$$ in environment $$j$$; $$e_j$$ is the effect of environment $$j$$, and $$d_i$$ is deviation of $$i$$th genotype in environment $$j$$, and $$\beta_i$$ is the regression of genotype $$i$$ on environmental mean (or index). Two slightly different regression techniques are proposed to explain part of GE interaction: either GE interaction effects may be regressed on environmental effects ($$\tilde{\beta}_i$$, Perkins and Jinks, 1968), or $$X_{ij}$$ values may be regressed on means of environments ($$b_i$$, Finlay and Wilkinson, 1963); both statistics are comparable ($$\beta_i = b_i - 1$$). Finlay and Wilkinson (1963) considered that simply comparing regression slopes ($$b_i$$ values) was not enough—the overall yield level of a genotype also had to be taken into account. According to them, genotypes with slope of 1.0 and a high mean yield are regarded as being well adapted to all environments. As mean yield decreases, genotypes with high or low slopes are regarded as being specifically adapted to favourable or unfavourable environments, respectively.

In addition to coefficient of regression, Eberhart and Russell (1966) proposed that the deviation mean squares ($$S^2_d_i$$) describe the contribution of genotype $$i$$ to GE interactions.
Coefficient of regression ($b_i$) = \[1 + \frac{\sum (x_{ij} - x_{i.} - x_{.j} + x_{..}) (x_{.j} - x_{..})}{\sum (x_{.j} - x_{..})^2}\]

Deviation mean squares ($S^2d_i$) = \[\frac{1}{E - 2} \sum (x_{ij} - x_{i.} - x_{.j} + x_{..})^2 - (b_i - 1)^2 \sum (x_{.j} - x_{..})^2\]

These two statistics ($b_i$, $S^2d_i$) of the regression approach are used in different ways to assess the reaction of genotypes to varying environmental conditions. While $S^2d_i$ is strongly related to the remaining “unpredictable part” of variability of any genotype and therefore is considered as a “stability parameter”, the coefficient of regression $b_i$ characterizes the specific response of genotypes to environments, and may be regarded as “response parameter” (Breese, 1969). Genotypes which do not react to varying environmental factors show “zero” $b_i$ values and would be stable according to the static concept. On the other hand genotypes possessing an average response to changing environmental conditions show $b_i$ values of “one”. Although, for ranking purposes, the choice of desired $b_i$ values depends on the specific goal, independently of the objective, $b_i$ value should be close to unity and deviation mean squares of stable genotypes should be close to zero.

Avian and Shephard (1976) showed that when the genotypic response across environments is predominantly linear, the regression coefficient ($b_i$) for each genotype provides a measure of its adaptation and stability, and when the non-linear component is large, the mean square deviation ($S^2d_i$) is required to specify genotypic stability across environments.

Shukla (1972a) proposed a linear combination of deviation mean squares as a stability statistic. This statistic, denoted as $S^2_h$, is an unbiased estimate of the variance of ($d_{ij} + \bar{e}_j$) and for ranking purposes is equivalent to the deviation mean squares. Pintus (1973) proposed to use the coefficient of determination ($r^2_i$) instead of deviation mean squares to estimate stability of genotypes, because $r^2_i$ is strongly related to $S^2d_i$ (Becker, 1981b).
The application of \( r^2_i \) and \( h_i \) has the advantage that both statistics are independent of the units of measurement.

Bucio-Alanis \textit{et al}. (1966), Byth \textit{et al}. (1976) and DeLacy (1981) suggested that linear regression approach in analyzing multilocation trials can be very informative if GE interactions have high linear association with environmental index—assuming little or no role of non-linear part of GE interaction. However, in case of low linear association, the technique at best may be uninformative and at most misinformative regarding genotypic performance over environments (Byth \textit{et al}.., 1976; DeLacy, 1981) Brennan and Byth (1979) reported that linear model explained less than 40% of the total GE interaction in wheat and thus a general indication of cultivar response to different environments may be obtained in such cases.

The linear regression approach is though widely used in the analysis of multilocation trials (Yates and Cochran, 1938; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966) and is useful in predicting the stability of populations (Breese, 1969; Samuel \textit{et al}., 1970), it has been criticised on various grounds (Westcott, 1987; Becker and Leon, 1988) including the following: a definite model for expected response function is unnecessary and may be misleading (Mungomery \textit{et al}.., 1974), the proportion of genotype-environment interaction sum of squares due to linear regression may be small (Baker, 1969; Byth \textit{et al}., 1976), regression fits may be unduly influenced by performance in relatively few environments (Westcott, 1986); and in practice regression parameters may fail to identify stable/unstable genotypes (Easten and Clements, 1975). Following are the other criticisms of this approach: the sum of squares due to deviations from regression, which was regarded by Eberhart and Russell (1966) as an important component of stability assessment, is not independent of regression slope (Hardwick and Wood, 1972); regression coefficients are biased, because the assumption of the regression
analysis that the independent variables (environmental means) are measured without error is not fulfilled (Sprent, 1969): testing significance in this approach depends upon even more difficult assumptions than those for the usual analysis of variance (Freeman, 1973). In addition, the linear regression model, in general, confounds the interaction with the main effects (Wright, 1971) reducing its power for general significance testing (Zobel et al., 1988).

Crossa (1988) compared the regression method with the ‘spatial’ method (Westcott, 1987) in respect to their consistency in assessing the stability of varieties when certain locations were omitted, and when a subset of varieties were analyzed. He found that stability parameters determined by regression analysis of original set of data varied for some varieties when: (i) extreme sites were excluded from analysis (ii) a subset of entries, out of the original set, was considered in isolation. Westcott (1986), and Hill and Baylor (1983) also reported that regression method does not provide reliable information when extreme points—genotypes with exceptionally high or low means—are either excluded from or included in the analysis. Another problem with the regression of yield on environments is that stability parameters depend on the particular set of genotypes included (Knight, 1970; Lin et al., 1986; Mead et al., 1986). Thus, stability parameters determined by the regression analysis for a given entry will vary according to the mean performance of the genotypes with which the entry is compared. Hence, the same entry in the same set of environments may show varying stability performance when analyzed in subsets.

Easten and Clements (1973) carried out an experiment on wheat genotypes in which differences in the environments were, they claimed, due entirely to measurable amount of nitrogen fertilizer. Some of the genotypes gave atypical (nonlinear) responses to fertilizer amount. The linear regression parameters of Finlay and Wilkinson failed to identify the “aberrant genotypes”. Although the parameters of Plaisted and Peterson (1959), Wricke (1962, 1964) and Eberhart and
Russell (1966) gave poor stability measures to these genotypes, given the
“conditional nature” of these measures (Knight, 1970; Witcombe and Whittington,
1971), they concluded that “caution should be taken in describing as unstable those
genotypes with high values of these parameters”. Here, “conditional nature”
implies that a subset of the entries or genotypes (taken out of a set originally
analyzed) when analyzed separately does not necessarily exhibit same results as the
original set did. Thus a variety could have marked deviations from linear
regression, not because it was inherently irregular, but because it showed a
different response pattern from the majority of the group with which it was being
compared. It is clear then, that the stability measures mentioned above are far
from satisfactory and, consequently, are not ever recommended.

Besides ANOVA and linear regression approaches, multivariate methods
have also been proposed for analyzing yield trials. The wide range of multivariate
methods include multivariate analysis of variance—MANOVA (Calinski et al.,
1987a, b), cluster analysis, principal component analysis, additive main effects and
multiplicative interaction (AMMI) analysis, geometrical methods, stochastic
dominance and methods using external information on environments or genotypes
(for review, see Gauch, 1985; Lin et al., 1986; Westcott, 1986, 1987; Huhn, 1987a;
Becker and Leon, 1988 and Zobel et al., 1988). Considering each ‘environment as a
variable’ and ‘genotypes as objects or replications’, the multivariate techniques
might be used to ‘describe the performance of genotypes’. For the present
investigation, multivariate methods for analyzing yield trials will be described only
in the light of principal component analysis, and additive main-effects and
multiplicative interaction (AMMI) analysis.

2.1.3. Principal component analysis (PCA)
Karl Pearson developed principal component analysis in 1901. He visualized a
matrix of $r$ rows (suppose genotypes here) and $c$ columns (suppose environments
here) as $r$ points in $c$ dimensional space (or the reverse). Like many multivariate
methods, it was not widely used until the advent of electronic computers, but it is now well entrenched in virtually every statistical package.

The central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the original data set. This reduction is achieved by transforming the original set of variables to a new set of variables, the principal components, which are uncorrelated and are ordered so that the first “few” retain most of the variation present in all of the original variables. The original high dimensional data set may be nearly incomprehensible to investigators, but a principal components graph, with only two or three dimensions ($PCA_n, n = 1 - 3$) may reveal important patterns quite clearly (Gauch, 1993).

Principal component analysis can also be given an algebraic interpretation. Each row and column is given a score. The model’s estimate for a given row-and-column (genotype-and-environment) contribution is the product of their scores. This process is repeated for each axis in a model, and the products are summed”. Perkins (1972) applied principal component analysis to the data on two sets of eight inbred lines of *Nicotiana rustica*, one grown in 10 years and the other in 9 years, and revealed that the properties of first two principal components agree with those predicted by a model which was based upon the regression analysis. The first principal component was directly related to the general “response” of lines to environmental differences (variance within genotypes) as measured by average of all the lines in each season. The second principal component was directly related to the average difference between two subgroups of four lines in each set of eight lines.

Mandel (1971) considered the principal components approach further using a multiplicative model for examining interactions in two-way table. This method indicated the number of dimensions necessary to contain the genotypic variation
and gave estimates of the corresponding coefficients, without, however, prior
knowledge of which factors these dimensions represented. When the deviations
from regression on the environmental means are substantial but no environmental
variables have been measured, Mandel’s method may prove particularly valuable

Freeman and Dowker (1973) used principal component analysis to examine
variation between and within genotypes and environments. They revealed that
when the first component of variation within genotypes is that between
environments, the second and subsequent components represent the
genotype-environment interaction. If this interaction has only one component, the
 technique of joint regression analysis, though suffering from some difficulties, can
give a useful guide to the biological interpretation of the results. However, when
more than one statistically important components are needed to describe the
interaction, the value of joint regression analysis is limited.

Principal component analysis is a descriptive and not an explanatory
technique, the chief difficulty with this approach is in the interpretation of the
resulting principal components, which may not bear any obvious relationship to
environmental conditions (Silvey, 1982; Westcott, 1986). The difficulty can be
overcome if the PCA be applied only to the interaction matrix rather than to the
original data which is intermixed with both main effects and interaction. However,
the model is multiplicative and therefore be applied only to the interaction, that is,
to the residual from the additive ANOVA model, which, in turn, is obtained by
removing the genotype and environment main effects from the original data
(Gauch, 1988). Afterwards, several studies have been conducted, in which PCA has
been applied only to the interaction matrix and the analysis was found very
meaningful in interpreting the results (Zobel et al., 1988; Gauch, 1988; Gauch and
Zobel, 1988; Crossa et al., 1991; Nachit et al., 1992; Shaffii et al., 1993; Gauch and
Zobel, 1994).
Zobel et al. (1988) indicated that principal component analysis, a multiplicative model, has a problem of not describing additive main effects. Consequently, the interaction, which is, by definition the residual from the additive ANOVA model, is not even considered, much less effectively analyzed by this model. Therefore, they suggested that PCA should be used only when the interactions are present in the data. The PCA does not identify interaction as a source but it confounds information on genotype means, environment means and interaction in a rather undesirable manner, when the data is intermixed with both main effects and interaction. When principal component analysis was applied to residual from the ANOVA model, they observed the resulting principal components to be meaningful, as the genotype IPCA1 scores for New York soybean yield trial reflected the maturity groups and environment IPCA1 scores reflected growing degree days. In addition, a number of workers including Crossa et al. (1991), Nachit et al. (1992) and Annichiarico (1992) have also reported a clear cut association between the interaction principal components and the features of genotypes and environments.

2.1.4. Additive main-effects and multiplicative interaction (AMMI) analysis

Partitioning and interpretation of GE interaction is generally based on the linear regression techniques (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966) or multivariate technique (Kempton, 1984; Gauch, 1988; Zobel et al., 1988; Crossa et al., 1990; Nachit et al., 1992). As already discussed, the linear regression techniques, however, have shown several deficiencies. On the other hand, multivariate techniques, such as the additive main-effects and multiplicative interaction (AMMI) procedure has been appreciated as a powerful tool in analyzing multilocation trials (Gauch, 1988; Gauch and Zobel, 1988; Zobel et al., 1988; Crossa et al., 1990). AMMI analysis combines the usual analysis of variance (ANOVA) for the genotype and environment main effects with the principal
components analysis (PCA) for GE interaction (Gollob, 1968; Gabriel, 1971, 1978; Bradu and Gabriel, 1978; Kempton, 1984; Freeman, 1985; Gauch, 1985; Zobel et al., 1988; Nachit et al., 1992). Thus, the additive effects are first extracted using usual analysis of variance (ANOVA) and then principal components analysis (PCA) is used to investigate GE interaction.

The additive part of AMMI (i.e., ANOVA) was invented by Fisher in 1918 and the multiplicative part (i.e., PCA) by Karl Pearson in 1901. In 1952, analysis of variance and principal component analysis were combined as AMMI by two groups: Evan Williams of CSIRO, Australia and Eugen Pike and Silverberg of Raytheon Manufacturing Company, USA (Gauch, 1993). Ever since its development, the AMMI has gone under a host of names, including FANOVA, MI, doubly-centered PCA, and biplot analysis (although the last name is a misnomer since “biplot” refers to a “type of graph”, not a particular model)–Gauch and Zobel, 1994. Seminal work on AMMI was done by Gollob (1968), Mandel (1971, 72), Corsten and Van Eijnsbergen (1972), Johnson and Graybill (1972), Bradu and Gabriel (1978), Gabriel (1978), Cornelius (1980) and others. But applications in agricultural research (especially in the yield trial research) began only a decade ago with Kempton (1984), and then were popularized by Zobel et al. (1988) and Gauch (1988), Crossa et al. (1990, 91) and Nachit et al. (1992). The motivation in these former two papers was to model or understand interactions, and Gauch (1985, 1988) introduced another motivation to gain accuracy and efficiency.

Zobel et al. (1988) reported the conclusions of Gollob (1968), Mandel (1971) and Gabriel (1971) that ANOVA, PCA and various regression analyses are in reality the sub-cases of AMMI model. These sub-cases test specific hypotheses about underlying relationship, i.e. only additive effects (ANOVA), only multiplicative effects (PCA), or multiplicative relationship between the genotype yields and the environmental means or vice versa (LR). AMMI is ordinarily the model of first choice when main effects and interaction are both significant.
(Mandel, 1971), which is the most common feature of the yield trials. If, for instance, only the main effects (genotype effects and environment effects) are present in the data, then the AMMI model can be reduced to an ANOVA model; whereas, if only nonadditive component (interaction) is present, then a PCA model is indicated: AMMI results can be readily used to diagnose these and other sub-cases (Bradu and Gabriel, 1978).

AMMI provides a graph (biplot) that shows the additive effects \((G, E)\) on one axis and the first interaction principal component scores (IPCA1 scores) on the other axis, using one type of points for genotypes and another for environments. This so-called "biplot" graph is remarkably informative showing both additive and interaction effects for both genotypes and environments. The parameters of AMMI model often capture 90 percent of the entire treatment variation, providing an accurate and parsimonious presentation of data (Zobel et al., 1988; Crossa et al., 1991; Gauch, 1993).

The 1990 conference on genotype-by-environment interaction, producing 26 papers in the volume edited by Kang (1990), did much to summarize and assess the AMMI models as of four years ago: several papers mentioned the AMMI model and two discussed it in detail. Zobel (1990) found AMMI powerful for analyzing numerous shoots and root traits of soybean. He reported that interactions tend to be larger for traits especially root traits, for which breeders have not imposed strong selection and thereby reduced genetic variability. Gauch (1990b) found AMMI useful for understanding complex interactions, gaining accuracy, improving selections, and increasing experimental efficiency. The results presented in these and other papers published by 1990, now seem widely known and generally accepted (Gauch and Zobel, 1994).

2.1.4.1. Understanding genotypes and environments: In numerous studies, both basic and applied, AMMI has helped to understand genotypes, environments, and their interactions. Crops analyzed include wheat (Kempton, 1984: Crossa et al.,...
1991; Denis and Gower, 1992; Nachit et al., 1992a, 1992b; and Annicchiariico and Perenzin, 1994), corn (Hirotsu, 1983; Cruz, 1992), rice (Okuno et al., 1971; Sutjipto, 1993), barley (Riggs, 1986; Romagosa et al., 1993), grass (Denis and Vincourt, 1982), cabbage (Davik, 1989), drybean (Wallace and Masaya, 1988; Wallace et al., 1991, 1993, 1994a, 1994b) and several other crops. Plant diseases analyzed include fungi on wheat (Kempton, 1984; Van Eeuwijk, 1992b) and corn (Wilson et al., 1993), nematodes on potato (Phillips and McNichol, 1986; Phillips et al., 1987, 1989; Dale and Phillips, 1989). AMMI has also been applied to aphid distributions in England (Kempton, 1984), ant communities in citrus groves (Samways, 1983), and meteorological data (Tamanco and Gabriel, 1984; Gabriel and Mather, 1986; Glassby, 1992).

Most of these papers applied additional statistical analyses besides AMMI, and offered comparisons. Also, many recent papers addressed important matters not covered before 1990. For a great diversity of crops and research objectives, sufficient experience with AMMI has accumulated to reach a balanced assessment of past applications and a realistic expectation of future applications.

(i) General perspective: The data structure typically emerging from a yield trial is a two-way factorial of genotypes by environments. The AMMI algorithm, as already mentioned, has two parts: ANOVA and PCA.

The first part of AMMI analysis uses ANOVA to partition the total variation into three orthogonal sources: genotypes (G), environments (E), and genotype-environment interactions (GE). Two preliminary questions may be asked about this partition (Gauch and Zobel, 1994).

First, what fractions of the treatment variation are typically found in G, E, and GE? DeLacy et al. (1990) presented results for over 100 trials with many crops in numerous locations. Romagosa and Fox (1995) summarized that paper by saying, “In most yield trials the proportion of sum of squares due to differences among sites ranged from 80 to 90% and variation due to G×E interactions was usually
larger than genotypic variation.” One may, therefore, propose a “70-20-10 Rule”
Saying that a median yield trial has about 70% of the variation in E, 20% in GE,
and 10% in G, with of course, wide differences in individual cases. (Gauch &
Zobel, 1994). Note that interaction is frequently larger than the genotype main
effect. Furthermore, the first PCA is usually larger than G. As the genotypes
become more diverse (in contrast to extremely similar elite lines) and the
environments become more diverse, GE tends to increase and easily reaches 40%
to 60% (Gauch & Zobel, 1994).

Second, for which research questions is it appropriate and helpful to impose
this partition? For purposes of selection, only the rankings of genotypes matter,
which are determined by G and GE, with E essentially irrelevant.

Now putting these two facts together, the awkward situation emerges that
most of the variation in yield (usually due to E) is wholly irrelevant to selection
because of the differential environmental influence on the phenotype of a
genotype. Consequently, analyses that provide a parsimonious summary of the
treatment variation in the data (like PCA and most classification procedures)
spend most of their effort on capturing irrelevant E. They impose no fundamental
distinction between variation due to G, E and GE, so parameters of the model
capture overall variation that combines and intermixes these three sources. Hence,
the results are diffuse relative to selection purposes. By contrast, AMMI provides
clear cut distinction for and partitions all these effects, i.e., G, E and GE
effectively, and can produce graphs that focus on data structure relevant to
selection.

The second part of AMMI analysis uses PCA to partition the GE interaction
into several orthogonal IPCA axes. Again two preliminary questions can be asked
about this partition (Gauch & Zobel, 1994).

First, are too many IPCA axes required to allow simple presentation of
results? Romagosa and Fox (1993) expressed concern that “when the best AMMI model includes more than one IPCA axis, assessment and presentation of genetic stability are not as simple as the AMMI1 case”. Two points may be offered in reply. (i) For AMMI2 also, a very satisfactory and simple graph can plot IPCA1 against IPCA2 (as contrasted with the usual main effects and IPCA1 shown for AMMI1; Kempton, 1984; Gauch, 1992b: 85-96). Furthermore, polygons can be superimposed on the AMMI2 biplot to show which genotype is the winner in every region of the biplot (Gauch, 1992b: 213-220). (ii) Experience has shown that only very infrequently there are sufficient grounds for including more than two axes. When AMMI3 and higher models are presented for agricultural data, in most cases, a little scrutiny reveals that the third and higher IPCA axes are dominated by noise and have no predictive value and no biological predictability.

Second, are significant IPCA axes interpretable in terms of known properties of the genotypes and environments? Researchers are concerned that “the chief difficulty with this approach is in the interpretation of the resulting principal components (Westcott, 1986; Becker and Leon, 1988). In this connection, Gauch and Zobel (1994) offered three points in reply: (1) From a statistical perspective, predictively significant model parameters indicate that physical or biological causes are at work, rather than random noise. IPCA axes may be regarded as first terms in the general Taylor expansion of multivariate non-linear relationships, so models like AMMI can be expected to give “a simple linear approximation to relationships in the data” (Aastveit and Martens, 1986). Or, in other words if G captures 7% of the treatment SS and IPCA1 captures 21%, then it would be peculiar to interpret the genotype main effect but to be mystified by the IPCA1 effect, that is, three times as large. To say the least, if the interpretation of a large IPCA is not already evident, then it becomes a mighty good topic for future research. (2) From an agricultural perspective, extensive experience with AMMI is now available, so speculation is hardly necessary any more. The track record for
easy interpretation of large IPCA is quite good. (3) Although interpretation of a
given experiment’s large IPCA axes is usually easy, a couple of related matters
merit close scrutiny: repeatability and heritability.

(ii) Environment interpretation and mega-environments: AMMI analysis of
yield trial data supplies to it no environmental data but just yield data.
Nevertheless, significant AMMI parameters (for both main and interaction effects)
often reflect identifiable causal factors (Gauch, 1992b: 231-236). By various
informal or formal means, the pattern in AMMI parameters or biplot can usually
be interpreted clearly in terms of evident environmental or genetic causal factors.

One strategy for interpreting AMMI results involves analyzing yield and
environmental data separately: AMMI analysis of yields is followed by “correlation
of AMMI parameters with environmental data”. Annicchiarico and Perenzin
(1994) applied AMMI to a bread wheat trial in Italy. Besides yield data, four
environmental factors—rainfall, average daily maximum temperature, average
daily minimum temperature, and number of frost days within a month before mean
heading date—were measured. IPCA1 was strongly correlated with number of frost
days, whereas IPCA2 was related to terminal drought stress. The AMMI2 biplot
revealed four clusters of environments, which related clearly to four geographical
regions in Italy. They concluded: “Besides representing location differences more
faithfully (than FW regression analysis), the use of AMMI analysis can also
contribute to identification of major environmental and genotypic factors related
to GL (genotype-location) interaction occurrence, thereby supporting breeding
programme decisions regarding adaptation targets, selection environments, test
sites, choice of parents, and adaptive traits to breed for”. Whereas AMMI2
captured 38.0% of the interaction (with higher axes nonsignificant noise),
FW-regression captured only 10.4%. Similarly, Annicchiarico (1992) analyzed
alfalfa trials in Italy. IPCA1 reflected day content and summer rainfall. The
AMMI1 biplot revealed three subregions, so cultivar recommendations were given
for each separate subregion. Nachit et al. (1992c) correlated AMMI main effects and IPCA1 scores with 11 environmental factors for a Mediterranean durum wheat trial. The environmental main effect reflected altitude, nitrogen fertilization, and a weather index, while IPCA1 reflected altitude and irrigation. They concluded that "these results show the usefulness of the AMMI model as a statistical tool for interpreting associations between environmental variables and components of GE interaction". Gauch and Zobel (1994) concluded on the basis of these as well as other papers that environmental interpretation of significant AMMI parameters is usually relatively straightforward.

Another strategy, requiring analyses other than AMMI, is joint analysis of yield and environmental data together, such as by redundancy analysis. For an experiment concerning nitrate concentration in lettuce, Van Eeuwijk (1992a) applied both redundancy analysis and AMMI. Remarkably, he observed that "AMMI and redundancy analysis gave comparable results, though the first extracts environmental scores as linear combinations of residuals from additivity (for the yield data), whereas the second forms environmental scores from linear combinations of measured environmental variables". That AMMI, which was not supplied environmental data should find corresponding patterns in the yield data is striking confirmation that significant AMMI parameters reflect underlying environmental and genetic causal factors. Similarly, Van Eeuwijk and Elgersma (1993) compared AMMI and various other analyses. They concluded that the different methods used in this paper to investigate relations between interaction and environmental factors all identified the same variables as important (also see Van Eeuwijk 1992b).

It becomes necessary to subdivide a growing region into several target or mega-environments when interaction is sizeable. Otherwise, differences in yield rankings from environment to environment will imply suboptimal yields in many environments—were the entire region planted to a single variety. However, it is
challenging to provide a clear conception of mega-environments and rational procedure for defining mega-environments for a given crop and region. Breeders tend to develop a negative attitude towards interaction, seeing only an impediment to high heritability and selection gains. But it should be kept in mind, on balance, that positive interactions associated with reasonably predictable features of environments offer an opportunity for higher yields, although at the cost of subdividing a growing region into several mega-environments. Interaction is not merely a problem; it is also an opportunity. Specific adaptations can make difference between a good variety and a superb variety.

A special kind of AMMI biplot graph can help researchers to comprehend and define mega-environments (Gauch 1992b: 213-230). In order to indicate a location’s probable position in an AMMI biplot in future years, circle the points for past results for that location (or give each location a distinctive symbol, or whatever). Various locations will occupy different prediction regions with a small region indicating high predictability and a large region, low predictability. Then superimpose the regions in which various genotypes are winners, which turn out to be horizontal bands in an AMMI1 biplot or polygons in an AMMI2 biplot. Locations situated within a single genotype’s winning turf, or which rarely cross over into another genotype’s turf, have an obvious variety recommendation. But locations that straddle genotype boundaries are inherently less predictable and more problematic. In such cases, income stability may be promoted best by planting two or three different varieties to hedge one’s bets. Those genotype turfs that contain numerous locations (or that represent significant crop regions) are candidates for mega-environments—so, may be corroborated with some environmental features, are defined the mega-environments. Note that, were new varieties developed and introduced, the map of variety winners can change. For example, a remarkably superior new variety might have a winning turf that encompasses two or three predecessors’ turfs, thereby reducing the number of
mega-environments required. That is, useful mega-environments depend not only on the environments, but also on the current genotype roster. Although this outcome may seem paradoxical at first, in fact, it is just right. The name “mega-environment” may suggest that this concept may involves only environments, but it also involves genotypes because interaction is what makes mega-environments necessary, and GE interaction is inherently a concept that involves both genotypes and environments. This specialized biplot showing locations as prediction regions and showing turfs of winning genotypes, is the most powerful tool yet developed for visualizing mega-environments. Such biplot can address effectively the numerous agricultural questions, such as whether certain locations (or mega-environments) are inherently less predictable than others as regards genotype rankings, whether breeding effort should increase or decrease for the various mega-environments, and which crosses might produce new genotypes performing best in a specified region of the biplot. This specialized biplot shows all predictively accurate number of the AMMI family. For rare cases, with higher-order AMMI models, these general concepts can be extended to more dimensions, but a straightforward two-dimensional graph no longer suffices.

AMMI has also been used to guide and modify the selection of test sites in order to increase research efficiency. M’Benga (1989) observed that fewer sites could be suffice for the corn trials in Gambia. Similarly, Saindun and Schaalje (1993) found that fewer sites would be adequate for testing dry bean cultivars in western Canada. Davik (1989) identified redundant sites in the Norwegian cabbage trials. Shafii et al. (1992) applied AMMI to rapeseed, the resulting understanding of GE interactions revealed specific areas that could consistently produce high-quality canola or industrial rapeseed. More efficient sampling of customary growing regions can free up needed resources to better sample potential new growing regions.

(iii) Plant morphology and physiology: Results from AMMI analysis can
illuminate plant physiological processes that can cause genotypes to interact with environments. They can also reveal the relative importance of various environmental factors or stresses (Gauch and Zobel, 1994). Most agricultural papers using AMMI provide a biological interpretation of the AMMI genotype parameters. Zobel et al. (1988) found genotype IPCA1 scores for a New York soybean trial to reflect maturity groups and environment IPCA1 scores, the growing degree days. Cossa et al. (1991) found three major groups of genotypes in an international bread wheat yield trial that reflected pedigree information and identified potential mega-environments. Nachit et al. (1992b) associated AMMI genotype parameters with morpho-physiological traits in durum wheat. Height and tillering explained 59% of the genotype main effect, and height and postmeridiem leaf rolling explained 51% of the genotype IPCA1 scores. The analysis helped to identify morphological and physiological traits related to stress tolerance. In an AMMI analysis of alfalfa trials in Italy, Annicchiarico (1992) found a clear association between the origin of parental materials and the adaptation pattern of progeny. Annicchiarico and Perenzin (1994) reported that genotype IPCA1 scores for a bread wheat trial reflected lateness of heading and frost tolerance (correspondingly the environment IPCA1 scores reflected level of late frosts), and IPCA2 reflected lodging and terminal drought stress. Van Eeuwijk and Elgersma (1993) compared FW regression and AMMI for a ryegrass yield trial. The genotype IPCA1 scores contrasted early and late cultivars, and IPCA2 scores reflected responsiveness to favourable environments. However, FW regression analysis picked up only this second, smaller interaction pattern.

Zobel (1992b) applied AMMI to numerous shoot and root traits of commercial soybean cultivars. A clear pattern emerged from the analysis that root traits typically had considerably larger GE interaction than shoot traits. For instance, the percentage of the treatment SS in interaction was 66.1% for total root number and 64.3% for taproot diameter, but only 24.4% for seed yield and 16.1%
for plant height. “The probable explanation is that the plant breeders have eliminated much of the genetic diversity causing interaction for those traits that have experienced intensive selection”. Such results have important implications for breeding programmes. He states that researchers dealing with traits (or crops) without a prior history of intensive selection are likely to need effective statistical procedures for handling large interactions. Zobel (1992b and 1994) found AMMI helpful in comprehending large interactions for root traits and in relating root systems to stress resistance.

Wallace et al. (1993) recommended a combination of yield system analysis (yield system analysis—YSA—examines rate of biomass accumulation, harvest index, and time to harvest maturity) and AMMI analysis to increase efficiency of breeding for higher crop yield. Annicchiarico and Perenzin (1994) reported that earliness of heading was the main wheat trait related to cultivar adaptation, in conformity with the general conclusions of Wallace et al. (1993).

(iv) Breeding, selection, and heritability: Wallace et al. (1993) concluded: ‘Additive main-effects and multiplicative interaction’ analysis can separate and quantify the genotype x environment interaction (GE) effect on yield and on each physiological component that is caused by each genotype and by the different environment of each yield trial.

Following the AMMI analysis of a bread wheat trial in Italy, Annicchiarico and Perenzin (1994) defined four mega-environments. Heritability was considerably higher in these mega-environments than in the region as a whole.

Zavala-Garcia et al. (1992) expected stability to improve performance for grain sorghum grown in highly unpredictable environments. Accordingly, they examined selection gains from indirect selection based on three stability indices: the FW-regression coefficient, deviations from regression, and genotype IPCA scores from AMMI analysis. Regarding AMMI, for two sorghum populations, they
selected about fifty families having above average genotype main effects and good stability as indicated by small (near zero) IPCA1 scores. They found indirect selection for stability alone to be ineffective, but a combined index using a stability index and genotype means did increase selection efficiency dramatically.

Several other studies are interesting, although they do not concern AMMI. Helms (1993) examined selection for yield and stability (employing several indices) in oats and found that selection solely for yield gave low stability and selection solely for stability gave low yield. Berke et al. (1992) identified quantitative trait loci (QTLs) in wheat for stability of grain yield as assessed by FW regression. More work needs to be done to assess the heritability of AMMI interaction parameters and to associate them with specific QTLs in various crops (Gauch and Zobel, 1994). However, in comparison to other interaction stability indices including FW regression, AMMI interaction parameters may be expected to achieve exceptionally high heritability simply because they are efficient in capturing interaction SS and therefore achieve a high signal-to-noise ratio.

2.2. Number and Choice of Environments

A major problem in selecting for yield stability is the poor repeatability of the statistical parameters estimated. Usually, genotypes rank differently in different years, and results calculated from different sets of locations and different years are only slightly correlated (for literature: Weber and Wricke, 1987; Leon and Becker, 1988). If stability is estimated from several locations within one year only, genotype-year interactions are not taken into consideration and hence a significant difference between two genotypes based on one year results will not necessarily occur again in the following years. In trials with large genotype-year interactions, Leon and Becker (1988) observed a very low repeatability of single year estimates of stability, even if the number of locations in a trial were 15 to 20.

In experiments performed in more than one year it would be desirable to distinguish between stability considering years and stability considering locations in
order to have a more complete characterization of a genotype (Barrah et al., 1981). But most experiments cover only two or three years and thus it is usually assumed that years and locations can replace one another. In a number of experiments the reaction to years and to locations have been observed to be similar (Pederson and Self, 1975; Barrah et al., 1981; Hoppe, 1982; Aastveit and Aastveit, 1984), but this correspondence does not always exist (Wermke, 1966).

In multi-year trials it is possible either to calculate a regression on the locations averaged over the various years (e.g., Yates and Cochran, 1938), or to consider each year-location combination as a “macro-environment” (Eberhart and Russell, 1966). Regressions on locations and regressions on macro-environments usually lead to similar rankings of genotypes: an interpretation of $b_l$ is often easier if calculated from a regression on locations but a regression on macro-environments is more powerful in detecting differences in $S^2_d_l$ and hence will be preferable in most cases (Becker, 1984).

For selection purposes not only the statistical accuracy of the estimated parameters is of importance but also the genetic variability available. This is taken into consideration when applying the concept of heritability. The heritability of stability parameters is rather low; in performance trial with wheat and barley at 12 locations in Germany, F.R., heritability of $b_l$ and $S^2_d_l$ was between 0.11 to 0.27 if based on one-year results and between 0.27 to 0.52 if based on three-year results (Becker, 1987).

Becker and Leon (1988) concluded that the useful stability measures are usually impossible to calculate from a few environments only. Locations, years and cultural practices will sometimes result in similar reactions of a genotype and thus can replace each other, but this depends upon the material and the geographic region and should not generally be taken for granted.