DRAFT

Training material version 1.3

Advanced Statistical Analysis of Multi-location Variety Trials

M. Singh

Computer and Biometric Services Unit

International Center for Agricultural Research in the Dry Areas
P.O. Box 5466, Aleppo, Syria

September 2000

Statistical Analysis of Multi- location Variety Trials

This Module

Objectives and Output

To develop the participant 's skill in designing multi-locational plant variety trials, carrying out the statistical analyses of the data generated from such designs, interpretation and presentation of the results from such analyses. Each participant will be expected to a prepare a draft scientific manuscript using the data from his/her own experiments.

Scope

The course covers:

- Design and analysis of data from Multi-locational Variety Trials and interpretation of GxE interaction, in general. In particular it focuses on - analysis of data from individual environments, test for homogeneity of error variances, combined analysis of data (evaluation of GxE interaction and tests for parallelism of the regression lines), common stability statistics, clustering methods (hierarchical and non-hierarchical cluster analysis of genotypes/environments), principal component analysis (of genotypes/environments) and heritability of the traits in broad sense (from individual environments as well as all the environments combined), additive main-effects and multiplicative interaction model, and inter-site transferability of crop varieties.
- Programs coded in GENSTAT 5 Release 4.2 for analysing data from multi-locational trials conducted in block designs.

In this note we have adapted the materials from several sources and have been cited in the reference/bibliography.

Statistical Analysis of Multi- locational Variety Trials

1. Introduction

Crop improvement process is long and involved several stages of germplasm collections, selection of desired material types, developing crosses between desired parents, preliminary evaluation of (generally large number of) genotypes/lines, selection, evaluations of selected genotypes in replicated trials followed by further (advanced) yield trials at multi-environments/locations representing the target domain for which the varieties are ultimately developed for production. When a number of varieties of a crop is grown over several environments (locations, years), their relative responses on various characters may show variation over the environments. This happens due to differential interplay between genetic and non-genetic (environmental) factors and therefore such a variation is said to occur due to interaction between genotype and environment or genotype x environment interaction (GEI). The major aspect of multi-locational trials is to identify stable and/or adaptable genotypes to the changing environments. Search for such genotypes requires a careful examination and exploitation of the GEI.

Keeping above aspect in view, this manuscript discusses commonly used experimental designs, data analysis from individual environments, combined analysis of data in Section 2, stability analyses Section 3, partitioning of GEI in Section 4, stochastic dominance in Section 5, a brief introduction to additive main-effects and multi-plicative interaction (Section 6) and inter-site transferability of crop varieties (Section 7). Various analyses covered in these sections have been coded in GENSTAT 5 (Genstat 5 Committee 1993) and a sample printout has been presented in Section 8.

2. Estimation of genotype means and GEI

2.1 Experimental design and data

We consider for that a set of p genotypes have been evaluated in q environments in replicated trials conducted in complete or incomplete block designs. A checklist of concerns in planning of an experiment is given by Jeffers (1978). The number of replications at a site depends on the variability in the experimental material (plots) and on the precision required of the estimates. For multi-locational trials, Kempthorne (1952, p583) provides an expression for an optimum number of replication in terms of error

variance, genotype variance and genotype x environment interaction variance and the cost factors. Although two replicates are absolutely minimum to estimate experimental error variance, in many case this is also the optimum number.

A number of statistical packages, such as GENSTAT 5, ALPHGEN, ALPHA+, GENDEX, etc. can facilitate generating randomized plans for various types of experimental designs. Plot-wise records on the response variable, generally taken as yield, are required for statistical analysis.

2.2. Analysis of variance from individual locations

Analysis of variance would be generated by fitting the model

yield = general mean + genotype effect + replicate (or complete block) effect + error

for randomized complete block design (RCBD), and

yield = general mean + genotype effect + replicate effect
+ effect of incomplete block within replicate + error

for the incomplete block design used.

The above model for data from RCBDs can be expressed using the following notations.

$$y_{ij} = \mu + \tau_i + b_j + e_{ij}$$

where y_{ij} = yield corresponding to

$j-th$ block ($j = 1, 2, ..., r$), $i-th$ variety ($i = 1, 2, ..., v$);

μ = general mean, τ_i = effect of $i-th$ genotype, b_j = effect of $j-th$ block,

e_{ij} = normally and independently distributed random variables with mean zero and

variance σ^2.

If the varieties are selected randomly to represent a population, we then assume that the τ_i s are normally and independently distributed random variables with mean zero and variance σ_g^2. Analysis of variance for the data with expected mean squares is in the following:

Table. Analysis of variance and expected mean squares from a single location data

Source	*df*	*Mean Square*	*Expectation of Mean Square*
(a) Variety effects assumed fixed.			
Blocks	$r-1$	–	
Varieties	$v-1$	V	$\sigma^2 + \frac{r}{v-1}\sum(\tau_i - \bar{\tau})^2$
Error	$(r-1)(v-1)$	E	σ^2
(b) Variety effects assumed random			
Blocks	$r-1$	–	–
Varieties	$v-1$	V	$\sigma^2 + r\sigma_g^2$
Error	$(r-1)(v-1)$	E	σ^2

The estimates of variance components are

$$\hat{\sigma}^2 = E$$

$$\hat{\sigma_g}^2 = (V - E)/r$$

Gain due to selection

The gain in selecting a chosen proportion p of the lines is the difference between mean of the selected lines and the population mean. This difference has the expected value

$$= \frac{K\sigma_g^2}{\sqrt{\sigma^2 + \sigma_g^2}}$$

The constant K, a function of p, is given by

$K = Z/p$; Z = ordinate of the standard normal distribution at the point which covers the area p in the right tail of the distribution and is given by

$$Z = (1/(2\pi)^{1/2})exp(-x^2/2)$$

where $p = \int_x^{\infty}(1/(2\pi)^{1/2})exp(-u^2/2)du$.

For example $p = 0.2$ (i.e. 20%) gives $Z = 10.28$ and $K = 1.40$.

If the selection is based on means of r replications then the expected difference is

$$= \frac{K\sigma_g^2}{\sqrt{\sigma_g^2 + \sigma^2/r}}$$

Models can similarly be written for other type of experimental designs. These models can be fitted using statistical packages (e.g. GENSTAT, SAS). The analysis of variance provides an estimate of experimental errors variance and a test of significance if the varietal differences in the yield response are real rather than arising from experimental error or chance. Means of varieties can be estimated (with adjustment for incomplete blocks where used) for individual environments. The experimental error variances may be examined for their homogeneity over the environments using Bartlett's chi-square test. If the error variances are found homogeneous then a pooled error variance can be obtained.

2.3 Combined analysis of variance over all locations

The combined analysis of variance to study GEI should distinguish the two cases- i) error variances homogeneous and ii) error variances heterogeneous.

Homogeneous error variances

One may estimate the GEI by fitting the model for RCBDs
Yield = General mean +Environment effect + Replication effects (within environments) + genotype effect + genotype x environment interaction effect + error
This will produce a common (pooled) error for testing significance of GEI.

Expressing the above model in notations

$$y_{ijk} = \mu + \mathrm{E}_j + \beta_{kj} + \tau_i + (GE)_{ij} + e_{ijk} \qquad (8)$$

where

variety, $i = 1, 2, ..., v$
environment $j = 1, 2, ... L$
blocks $k = 1, 2, ..., r$
E_j = effect of $j-th$ environment; β_{kj} = effect of $k-th$ block in $j-th$ environment;
τ_i = effect of $i-th$ variety;
$(GE)_{ij}$ = interaction term for $i-th$ genotype and $j-th$ environment;
e_{ijk} = plot error assumed independent and normally distributed with mean zero and constant variance σ^2.

Let us assume that the environment and variety effects are fixed and the genotype x environment interaction are independent and normally distributed with means zero and variance σ^2_{ge}. We have the following ANOVA structure.

Table. Analysis of variance of data from several experiments conducted in RCBDs

Source	*df*	*Mean Square*	*Expectation of Mean Square*
Environment	$L-1$	–	–
Blocks within Envs.	$L(r-1)$	–	–
Variety	$v-1$	V	$\sigma^2 + r\sigma^2_{ge} + \frac{rL}{v-1}\sum(\tau_i - \bar{\tau})^2$
Variety x Envs.	$(v-1)(L-1)$	I	$\sigma^2 + r\sigma^2_{ge}$
Error	$L(r-1)(v-1)$	E	σ^2
Total	$Lrv-1$		

(Assuming environment effects fixed, genotype effects and the interaction effects random)

Envs. =Environments

Gain due to selection

The gain in selecting a chosen proportion p of the lines (based on means over the replications) has the expected value

$$= \frac{K\sigma_g^2}{\sqrt{\sigma_g^2 + \sigma_{ge}^2/L + \sigma^2/(Lr)}}$$

Example 1. Combined ANOVA over RCBDs

Data: *Chickpea yields*

Identifier	Type	Length	Values	Missing
Location	Factor	460	Present	0
Rep	Factor	460	Present	0
Geno	Factor	460	Present	0
Yield	Variate	460	Present	0

******* Analysis of variance *******

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Location.Rep stratum					
Location	4	1.127E+08	2.817E+07	22.10	<.001
Residual	15	1.911E+07	1.274E+06	8.30	
Location.Rep.Geno stratum					
Geno	22	7.644E+06	3.474E+05	2.26	0.001
Location.Geno	88	3.968E+07	4.509E+05	2.94	<.001
Residual	330	5.064E+07	1.535E+05		
Total	459	2.297E+08			

***** Tables of means *****

Variate: Yield

Grand mean 1016.

Location	1.00	2.00	3.00	4.00	5.00		
	1012.	322.	880.	1867.	998.		
Geno	1.00	2.00	3.00	4.00	5.00	6.00	7.00
	1124.	976.	942.	854.	1057.	1254.	979.
Geno	8.00	9.00	10.00	11.00	12.00	13.00	14.00
	1128.	951.	959.	1196.	1121.	733.	1040.
Geno	15.00	16.00	17.00	18.00	19.00	20.00	21.00
	1073.	1081.	856.	1030.	1028.	1186.	1069.
Geno	22.00	23.00					
	761.	965.					

Location	Geno	1.00	2.00	3.00	4.00	5.00	6.00
1.00		828.	1054.	1016.	1023.	880.	1111.
2.00		172.	284.	401.	132.	381.	312.
3.00		933.	792.	750.	625.	818.	917.
4.00		2579.	1954.	1852.	1238.	2207.	2931.
5.00		1107.	798.	691.	1250.	1000.	1000.

Location	Geno	7.00	8.00	9.00	10.00	11.00	12.00
1.00		773.	1078.	918.	1240.	1153.	1000.
2.00		328.	555.	281.	276.	318.	427.
3.00		932.	875.	766.	826.	792.	719.
4.00		1875.	2598.	2149.	1704.	2942.	2556.
5.00		988.	536.	643.	750.	774.	905.

Location	Geno	13.00	14.00	15.00	16.00	17.00	18.00
1.00		979.	905.	1005.	943.	882.	1036.
2.00		271.	293.	326.	481.	286.	349.
3.00		818.	693.	906.	875.	891.	1110.
4.00		906.	2285.	2117.	1783.	1446.	1729.
5.00		691.	1024.	1012.	1322.	774.	929.

Location	Geno	19.00	20.00	21.00	22.00	23.00
1.00		1183.	1262.	1059.	926.	1014.
2.00		280.	333.	354.	228.	339.
3.00		1099.	1219.	1026.	959.	896.
4.00		803.	1903.	1442.	929.	1016.
5.00		1774.	1214.	1465.	762.	1560.

*** Standard errors of means ***

Table	Location	Geno	Location Geno
rep.	92	20	4
e.s.e.	117.7	87.6	224.8
d.f.	15	330	151.47
Except when comparing means with the same level(s) of Location			195.9

```
 d.f.                                                330

*** Standard errors of differences of means ***

Table              Location        Geno    Location
                                                Geno
rep.                     92          20           4
s.e.d.                166.4       123.9       318.0
d.f.                     15         330      151.47
Except when comparing means with the same level(s) of
 Location                                       277.0
 d.f.                                             330

*** Least significant differences of means ***

Table              Location        Geno    Location
                                                Geno
rep.                     92          20           4
l.s.d.                354.7       243.7       628.2
d.f.                     15         330      151.47
Except when comparing means with the same level(s) of
 Location                                       544.9
 d.f.                                             330
```

Example 2. Combined analysis of data from triple lattices

Data: *Barley yield:* $v = 64$ genotypes, 10 environments.

```
***** REML Variance Components Analysis *****

Response Variate : Yield

Fixed model      : Constant+Loc+Geno+Loc.Geno
Random model     : Loc.Rep+Loc.Rep.Blk

Number of units  : 1920
No absorbing factor


*** Estimated Variance Components ***

Random term                Component        S.e.

Loc.Rep                      34618.       13020.
Loc.Rep.Blk                  41528.        5621.
*units*                      83456.        3642.

*** Approximate stratum variances ***

                                                Effective d.f.
Loc.Rep                                2631205.         20.00
Loc.Rep.Blk                             304939.        210.00
*units*                                  83456.       1050.00

* Matrix of coefficients of components for each stratum *

      Loc.Rep        64.00         8.00         1.00
```

Loc.Rep.Blk	0.00	5.33	1.00
units	0.00	0.00	1.00

***** Wald tests for fixed effects *****

Fixed term	Wald statistic	d.f.
Loc	1521.7	9
Geno	517.4	63
Loc.Geno	1314.0	567

***** Table of predicted means for Constant *****

*** Table of predicted means for Geno ***

Geno	1	2	3	4	5
	2132	2332	2251	2315	2028
Geno	6	7	8	9	10
	1955	1993	2228	2209	1936
Geno	11	12	13	14	15
	2099	2314	2261	2288	2093
Geno	16	17	18	19	20
	2079	1969	2102	2124	2131
Geno	21	22	23	24	25
	2108	2113	1663	1916	1784
Geno	26	27	28	29	30
	1980	2166	1738	2027	1951
Geno	31	32	33	34	35
	2061	2333	1933	2296	1861
Geno	36	37	38	39	40
	2167	1946	2273	2247	2314
Geno	41	42	43	44	45
	2243	2189	2068	2164	2216
Geno	46	47	48	49	50
	2146	2159	2302	2327	2154
Geno	51	52	53	54	55
	2283	2296	2260	1847	2257
Geno	56	57	58	59	60
	2304	2335	2245	2300	2263
Geno	61	62	63	64	
	2288	2164	2194	2176	

Standard error of differences:	Average	78.98
	Maximum	79.51
	Minimum	78.23
Average variance of differences:		6237.

***** Table of predicted means for Loc.Geno *****

Geno	1	2	3	4	5
Loc					
1	421	360	379	357	353
2	819	715	716	759	808
3	1980	2442	2372	1927	2305
4	1180	1379	1090	996	1225
5	3327	3716	3386	3858	2737
6	4422	4764	4100	4935	3178
7	1080	1142	1423	1246	1302
8	4918	5038	5160	4872	5052
9	1570	1485	1712	2191	1296
10	1601	2278	2171	2008	2027

..... .

Geno	61	62	63	64
Loc				
1	430	533	416	475
2	805	989	823	864
3	2325	2501	2374	2237
4	1246	1102	1196	778
5	3347	3046	3236	3774
6	4167	3646	4181	4216

```
           7         1324          825         1264         1305
           8         5102         5413         5013         4410
           9         1701         1437         1412         1599
          10         2436         2149         2024         2100
Standard error of differences:     Average          293.2
                                   Maximum          298.0
                                   Minimum          246.4

Average variance of differences:                   86200.

Standard error of differences for same level of factor:

                    Loc        Geno
Average           249.7       298.0
Maximum           251.4       298.0
Minimum           246.4       298.0
Average variance of differences:
                 62372.      88806.
```

2.4 Combined analysis of data from experiments in RCBDs conducted over several locations and years

We shall consider the analysis of a trial on v varieties evaluated in r randomized blocks at each of L locations in each of the same Y years. We shall use is the following model.

$$y_{ijkl} = \mu + L_j + Y_k + (LY)_{jk} + \tau_i + (L\tau)_{ij} + (Y\tau)_{ik} + (LY\tau)_{ijk} + \beta_{jkl} + e_{ijkl}$$

where the various terms in right hand side represent general mean, location effect, year effect, location x year interaction, variety effects, location x variety interaction, year x variety interaction, location x year x variety interaction, blocks within location and year, and plot error respectively and associated with the suffixes representing the following:

$i = 1, 2, ..., v$: variety

$j = 1, 2, ..., L$: locations

$k = 1, 2, ..., T$: years

$l = 1, 2, ..., r$: blocks.

Further we assume that variety and replication effects are fiexd while location effect and all other factor effects are random. The interactions with variety i.e. location x variety interaction, year x variety interaction, location x year x variety interaction, and the plot errors are assumed independently and normally distributed with means zero and variances σ^2_{lv}, σ^2_{yg}, σ^2_{lyg} and σ^2 respectively. We have the following ANOVA structure.

Table Analysis of variance of experiments combined over several locations and years

Source	*df*	*Mean Square*	*Expectation of Mean Square*
Location	L − 1	Lo	–
Years	Y − 1	*Ye*	–
Location × years	(L − 1)(Y − 1)	*LoYe*	–
Varieties	$v-1$	V	$\sigma^2 + r\sigma^2_{lyg} + rL\sigma^2_{yg} + rY\sigma^2_{lg} + \frac{rLY}{v-1}\sum(\tau_i - \bar{\tau})^2$
Places × varieties	$(L-1)(v-1)$	*LoV*	$\sigma^2 + r\sigma^2_{lyg} + rY\sigma^2_{lg}$
Years × varieties	$(Y-1)(v-1)$	*YeV*	$\sigma^2 + r\sigma^2_{lyg} + rL\sigma^2_{yg}$
Places × years × varieties	$(L-1)(Y-1)(v-1)$	*LoYeV*	$\sigma^2 + r\sigma^2_{lyg}$
Replications	$LY(r-1)$	–	
Error	$LY(v-1)(r-1)$	E	σ^2
Total	$LYvr-1$		

Gain due to selection

The gain in selecting a chosen proportion p of the lines (based on means over the replications) has the expected value

$$= \frac{K\sigma_g^2}{\sqrt{\sigma_g^2 + \sigma_{lg}^2/L + \sigma_{yg}^2/Y + \sigma_{lyg}^2/(LY) + \sigma^2/(LYr)}}$$

Heterogeneous error variances

We can fit the following model on the genotype x environment data on means (or adjusted means for incomplete blocks) using a weighted analysis of variance with weights being inversely proportional to the variance of the means. The weight corresponding to a mean (for a combination of genotype and environment) may be estimated by 1/(standard error of the mean)^2.

Mean=Environment effect + Genotype effect + residual

The residual sum of squares produced by the weighted least-squares would be the weighted GEI sum of squares and would be approximately distributed as chi-square with GEI degrees of freedom.

Once there is a significant GEI, we may carry out further analyses to identify the causes of interaction inters of the responsiveness of the genotypes to the environments.

3 Exploitation of G x E Data: Stability Analysis

Genotypes performance changes due to environmental pressures or stresses (due to the population heterogeneity or population buffering and changes in the genetic make up taking place over generations) and differences in their ability to adapt to the stress factors (short-term acclimatization). A number of statistical models to study genotypic adaptation based on phenotypic performance have been discussed in literature. Byth and Mungomery (1981) discussed the following three concepts *Stability, adaptability, and predictability.*

"

Phenotypic *stability* refers to the ability of a genotype to maintain a near constant phenotype for the character of interest over variable environments. Such a genotype would be regarded as having *wide adaptation.* But certain genotypes may also show predictably superior performance in particular types of environments indicating that broad adaptation inevitably involves sacrifice of performance in specific environments. Thus the strategies of plant improvements for broad adaptation (minimizing G x E interaction) and specific adaptation (emphasizing favorable interaction) are in direct conflict.

Predictability refers to the extent to which response is systematic.

Responsiveness is the ability of a genotype to respond in a particular manner to a general change in the environmental potential.

Sensitivity (also *stability)* refers to the extent of unpredictable variation in response. Some researchers relate stability to variability of performance over time (temporal variation) at a location while adaptability to variability in performance across locations (spatial variation).

"

We shall in the present chapter discuss various concepts of stability using statistical measures in common practice. The various concepts and measures of stability, originating due to different outlooks of experimenter to their specific problems, have added to the difficulty of choosing a stability parameter (s) for a given situation. We include the two approaches discussed by Lin et al (1986).

3.1 Parametric Approach

Stability statistics are derived (computed) for each genotype from two-way tables of genotype and environment data. These statistics are based on either of the following three types of stability concepts. A genotype is considered to be stable if

- its among environments variance is small (Type I stability),
- its response environments is parallel to mean response of all genotypes in trial (Type II stability),
- the residual mean square from regression model on environment index is small(Type III stability).

In order to list various statistics, we shall use the following notations. Let y_{ij} denote the mean value of i-th genotype in the j-th environment (i=1,2, ...p, j=1,2...q). Let

$$\bar{y}_{i.} = \sum_{j} y_{ij}/q; \qquad \bar{y}_{.j} = \sum_{i} y_{ij}/p, \quad \bar{y}_{..} = \sum_{i}\sum_{j} y_{ij}/(pq)$$

represent respectively, means of i-th genotype, j-th environment and overall mean. The nine statistics and one more in current use are briefly described as follows:

1. The variance of a genotype across environments

$$S_i^2=\sum_{j=1}^{q}(y_{ij} - \bar{y}_{i.})^2/(q-1),$$

2. The coefficient of variation

$CV_i=S_i/\bar{y}_{i.}$.

Francis and Kannenberg (1978) used the conventional CV% of each genotype as a stability measure.

3. Plaisted and Peterson's (1959) mean variance component for pair- wise G x E interaction ($\bar{\theta}_i$)

$$\tilde{\theta}_i = (p\sum_{j=1}^{q}(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2+\sum\sum(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2)/(2(p-1)(q-1)).$$

The mean of the estimated variance components of the G x E interaction for all pairs of genotypes that include genotype i is the stability measure of genotype i.

4. Plaisted's (1960) variance component for G x E interaction ($\theta_{(i)}$)

$$\theta_{(i)}=(-p\sum_{j=1}^{q}(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2/(p-1)+\sum\sum(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2)/((p-2)(q-1)).$$

One genotype i is deleted from the entire set of data and the G x E interaction variance from this subset is the stability index for genotype i.

5. Wricke's (1962) ecovalence (w_i^2)

$$w_i^2=\sum_{j=1}^{q}(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2 .$$

This G x E interaction effects for genotype i , squared and summed across all environments, is the stability measure for genotype i.

6. Shukla's (1972a) stability variance (σ_i^2)

$$\sigma_i^2= (p\sum(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2-\sum\sum(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})^2/(p-1))/((p-2)(q-1)).$$

Based on residuals in a two-way classification, the variance of a genotype across environments is the stability measure.

7. Finlay and Wilkinson's (1963) regression coefficients(b_i)

$$b_i =\sum_{j}(y_{ij}-\bar{y}_{i\cdot})(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})/\sum(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2.$$

The observed values are regressed on environmental indices environments and the overall means. The regression coefficient of each genotype is taken as its stability parameter.

8. Perkins and Jinks' (1968) regression coefficient (β_i)

$$\beta_i = \sum(y_{ij}-\bar{y}_{i\cdot}-\bar{y}_{\cdot j}+\bar{y}_{\cdot\cdot})(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})/\sum_{j}(\bar{y}_{\cdot j}-\bar{y}_{\cdot\cdot})^2.$$

Similar to (7) except that the observed values are adjusted for environment effects for computing regression coefficients ($\beta_i = b_i - 1$).

9. Eberhart and Russell's (1966) deviation parameter (δ_i^2)

$$\delta_i^2 = (\sum_j (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 - \beta_i^2 \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2)/(q-2)$$

This is the residual mean square (MS) of deviation from regression defined in (7) or (8) is the measure of stability.

10. Variance of genotypes across environments on the ratios of yields to environment means

Yau (1972) gave an other statistics, denoted here by ξ_i as the variance across environments of the ratios of yields to the mean under respective environment

$$\xi_i = \sum_{j=1} (r_{ij} - \bar{r}_{i.})^2/(q-1), \text{ where } \bar{r}_{i.} = \sum_{j=1} r_{ij}/q, \text{ and } r_{ij} = y_{ij}/\bar{y}_{.j}$$

This statistic can be see to measure Type-II stability.

Grouping of the indices and their similarity

The first nine statistics are based either on the deviation from average genotype effect (DG) = $y_{ij} - \bar{y}_i$ or on the G x E interaction term $I_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$ (in form of their sums of squares SS , regression coefficient or deviation from regression) and were classified into four groups (A, B, C, D):

Group A: DG, SS:	(S_i^2 , CV_i) – Type I stability
Group B: GE, SS:	(θ_i, $\theta_{(i)}$, W_i^2, σ_i^2) – Type II stability
Group C: DG or GE regression coefficient :	(b_i, β_i) – Type II stability
Group D: DG or GE regression deviation.	(δ_i^2) - Type III stability

Lin et al (1986) noted:

(i) Since Var (log(y)) ~ Var(y)/(mean(y))^2 = (CV(y))^2. Thus the two statistics in group A are equivalent, except for data transformation.

(ii) The four statistics in group B are equivalent for the purpose of ranking genotypes. σ_i^2, also is an unbiased estimate of variance of genotype i. An approximate test for homogeneity of σ_i^2 has been given by Shukla (1972b).
(iii) Since $\beta_i = b_i - 1$, the two statistics in group C are equivalent. Similarly, the statistics of group D are equivalent.
(iv) When variability in response can be satisfactorily expressed by a regression model, the regression coefficient (of group C) can serve as stability parameters and could be preferred to variability measures (group B) since, they (of Group C measures) provide information on shape of response along with its variation.

Stability indices and stability measures
(i) The statistics in Group A measure Type I stability; those of Group B and Yau's index measure Type II while those of Group D measure Type III. The statistics of Group C are of Type I or Type II stability measures depending on the nature of the stable genotype. If stable genotype are defined by having $b_i=1$ ($\beta_i=0$) Type II is implied; but if they are defined by $b_i=0$ ($\beta_i=-1$), then Type I is implied.
(ii) Type I stability: Type I indicates homeostasis, a biological concept (Becker 1981). It differs from agronomic concept of stability given by Type II. Although Type I is theoretically sound, but breeder do not use it frequently, for a breeder would like to select cultivars with high yields besides having Type I stability. Type I stability is associated with relatively poor yield in environments which are high yielding for other cultivars. Also, b_i and yield are positively correlated (Finlay and Wilkinson, 1963). Although wide (broad) adaptation may be desirable but difficult to achieve in practice. A still more convenient way would be to breed cultivars with (specific) adaptation to different environments to maximize the production. Since Type I stability does not depend on the presence of other genotypes, it has broad inference base. However, it does not provide information on the response structure.
(iii) Type II stability: The inferences from Type II stability measures are relative to the genotypes included in the test. For example, a genotype A may be assessed stable and B unstable if A resembles majority of genotypes in the set more closely than does B. In an another set of genotypes, if B resembles majority of genotypes more closely than does A, then B is stable and A unstable. This measure is useful for comparing a specific set of genotypes and thus, does not have a broad inference base for general assessment.
(iv) Type III stability: Eberhart and Russell (1966) suggested another measure of stability based on mean square of deviation (σ_i^2). Thus, there are two measures of stabilty, (b_i, σ_i^2) for a single character. Use of σ_i^2 was advocated by Breese (1969) as he considered

'stability' should refer the unpredictable variation (irregularities) in response to environment. The variability of response to environment can be divided into predictable variation (given by regression mean squares) and unpredictable variation (measured by deviation MS, σ_i^2). This argument is sound but the measure of stability by deviation MS is inappropriate as it represent the goodness of fit of the model we choose. To support the argument in practice, one must measure independent variables explaining environment and a prediction model be made with them. The environmental index (based on means of all genotypes) can not provide an independent measure of environment potential. Thus low value of percent variance accounted for or high σ_i^2 or heterogeneous MS simply indicate that regression model is not adequate for stability and some other methods should be investigated. Type III is useful only when the prediction model is considered and is based on independently measured environmental variables.

3.2 Non-parametric Approach

In sequel to our previous discussion, we now consider the non-parametric approach to study genotype x environment interaction. This approach is used to search pattern in the genotypes and or in the environments.

The statistics under parametric approach express multivariate information (responses over multi-environments are considered multi-variate) in terms of a univariate, and measure only individual aspects (Types I, II, or III) of stability. It is possible to arrive at a contradiction, i.e. a genotype may be found stable for one type of stability measure but could be found unstable for the other measure(s). These do not provide any interrelationship among the genotype exploiting the response patterns from these (common) environments. Classifying genotypes into quantitatively homogeneous stability subsets, based on similarity of their responses to the environments, is another line of thought to evaluate interrelationships among the genotypes, and such an approach is considered to be non-parametric. The classification method has an advantage in the sense that although the genotypes are grouped on the basis of a specific data set, the relative relationship among genotypes can be independent of it or any specific data set. For instance, two genotypes say A and B, with dissimilar response patterns (unrelated) can always be grouped into two different stability sets, irrespective of the presence of genotypes resembling A and or B. We now consider methods for classification.

3.2.1 Cluster Analysis

Several commonly used methods for clustering genotypes (or environments) based on similarity of response characteristics are available in references cited in the end of this material. Every clustering technique has two considerations. (i) a definition of the similarity matrix, (ii) a strategy for grouping. Two cases arise. In one case, similarity is based on genetic effect and G x E interaction (means of p x q table) while in the other case similarity is based on G x E interaction only. Strategies for grouping could be, for instance incremental sums of squares (ISS) fusion strategy and group average (GA) fusion strategy (see, Cormak, 1971). This is the case of hierarchical clustering.

Limitations of clustering methods

The particular choices of similarity matrices and clustering strategies give rise to different cluster groups and this may lead to problems of preferring one method of cluster to other. Another criticism of clustering method is that it can also force unwanted structure on a data set suggesting misleading results.

2.2.2 Non-hierarchical clustering

In non-hierarchical clustering, the purpose is to group the units (genotypes or locations) in a number of disjoint classes chosen in advance using the information on a number of variables on them. The units with a class are expected to be homogeneous on the basis of some criterion. In the hierarchical clustering one can cut the dendrogram at a level of similarity to provide a selected number of groups but the statistical properties of such a grouping is not yet clear. In non-hierarchical clustering the groups of units are obtained by optimizing the selected criterion. Some of these are

i. maximization of between -group sum of squares
ii. maximal predictive classification,
iii. minimizing the determinant of the pooled within-class dispersion matrix,
iv. maximizing the total Mahalanobis squared distance between the groups.

2.2.3 Ordination techniques

Ordination techniques are used to simplify multivariate data for a set of individuals by summarizing relationships among individuals or among attributes describing them. This is done by producing a simple visual representation of the individuals as points which can be plotted to portray their relationships acceptably free of distortion. The ordination techniques try to reduce the dimensionality of the multivariate systems efficiently to

preserve the relationships among individuals as far as possible, but to provide a simplified view of those relationships in fewer dimensions than specified by original variables. There are two methods of ordination.

3.2.4 Principal Component Analysis (PCA).

PCA considers finding a new set of coordinate axes which accounts more effectively for the variation among individuals than do those based on original variables. PCA represent a transformation of data from one set of coordinate to another. This may not necessarily lead to reduction of dimensionality. However, when only (first) few principal components account for most of the variation, then it becomes effectively useful. Algebraically, the principal axes are determined by the latent vectors from the matrix of corrected sums of squares and products among variables. Elements of each vector specify the linear combinations of original variables necessary to give the corresponding PC and the associated latent root give the variation attributable to the component. PCA can also be applied on environment in same way as it could be done to genotypes. Mandel (1969) considered it for G x E interaction effect.

3.2.5 Principal Coordinate Analysis of genotypes. (PCO)

PCO analysis requires finding a set of rectangular coordinate axes which accounts as efficiently as possible for variation among individuals and may lead subsequently to a reduction in dimensionality for simplification. These objectives are similar to PCA but PCO is based on a much more general approach. It does not automatically assume that original variables define a multidimensional Euclidian space, in which relationship between pairs of individuals are indicated by Euclidian distance. Many similarity measures (e.g. correlation coefficients) or dissimilarity measures (distance) could be used. PCO involves two steps for computation.

1) presentation of the set of individuals as points in a coordinate space derived from the original matrix of measures. Gower (1966) showed that the interpoint Euclidian distances in this space are a simple function of the original measures of relationship between individuals. The significance of this method is that it refers individuals to Euclidian coordinate axes even when an initial coordinate framework is unavailable, and it represents original measures of relationships as Euclidian distances even if they are non-Euclidian.

2) carrying out a PCA on the data derived in step (1).

The two steps of PCO combined in one are given by Gower (1967). The only requirement to guarantee a distortion-free representation by PCO is that the original matrix of measures must be symmetric (so that no negative latent roots are obtained). In general, principal axes will not be a linear combination of original variables as in usual PCA. However, it is possible to investigate the relationship of original variables to each principal axis by correlating the set of principal coordinate scores for each axis with each of original variables. A correlation of large magnitude for a particular variable implies that it is strongly reflected in the axis concerned. Gower (1966) also showed that PCA is a special case of PCO when measures used in PCO are squared Euclidian distances.

4. Partitioning of GxE interaction

We may present

a) the results of cluster analysis employed for zoning the environments and grouping the genotypes.

b) partitioning of the GxE interaction using these groupings.

The care must be taken in justifying the groups resulting from a methods in terms of the number of groups, and the nature of locations and genotypes within. It is recommended that the groupings must be looked into the light of some other (independent) variables reflecting the physical properties of the environments and phenological and morphological traits of the genotypes.

A complete hierarchy should be presented with help of dendrogram when agglomerative methods of forming groups are used. Let n^e and n^g be the number of environment groups and number of genotypes groups respectively. Also let n_i^g be the number of genotypes in the i-th genotype group (i=1...n^g) and n_j^e be the number of environments in the j-th environment group (j=1...n^e). Note that $\sum_{j=1}^{n^e} n_j^e = q$ and $\sum_{i=1}^{n^g} n_i^g = p$. Further, with the reduced G x E data matrix one may present:

(i) Partitioning of the variation related to grouping model.

Analysis of variance skeleton.

Source	d.f.	SS	MS
Environments (E)	q-1		
Among E groups	n^e-1		
Within E groups	$\sum(n_j^e-1)$		
Genotypes (G)	p-1		
Among G groups	n^g-1		
Within G groups	$\sum(n_i^g-1)$		
G x E	(q-1)(p-1)		
Among G groups x among E groups	-		
Among G groups x within E groups	-		
Within G groups x among E groups	-		
Within G groups x within E groups	-		
Residual	-		

(ii) Group performance plots and

(iii) Patterns of (G x E) interactions on grouped sets.

5. Stochastic Dominance of Varieties

This procedure emphasizes the riskiness of (new) genotype or variety. New crop varieties (or new technologies, in general) may often be regarded by farmers more risky than traditional ones. Risk may, therefore tend to act as an impediment to their adoption. Improved varieties that would be preferred by "risk-averse" farmers can be identified by stochastic dominance procedure under certain assumptions. Anderson (1974) used this procedure for analyzing data from the Sixth International Spring Wheat Yield Nurseries administered by CIMMYT. He made following three assumptions.

(i) it makes sense to talk about (or large regional) probability distribution of wheat yields,

(ii) the selection of sites, cooperators, fields and growing and disease conditions is representative of the relevant world (or regional) domain of production, and

(iii) yield per se provides a reasonable surrogate for the argument of the average farmer's utility function.

Menz (1980) used cluster analysis of Byth et al (1976) to analyse CIMMYT International Spring Wheat Yield Nurseries over five years and also used stochastic dominance. He found considerable degree of agreement in the results based on the two methods.

6. Additive Main Effects and Multiplicative Interaction Model

The AMMI model stands for additive main effects and multiplicative interaction model. The data on GxE are fitted using

i. main effects of genotypes and environments,

ii. the interaction GxE is fitted as sum of multiplicative PCA scores for genotype and environments.

Main advantage of this method is that it facilitates examination of the pattern of GxE interaction as expressed by a general number of principal components. Further details are available in a recent series of articles including Gauch (1988), and Gauch and (1988).

7. Inter-site Transferability of Crop Varieties

Development of varieties and their evaluation often takes place on a limited range of environments (e.g. experimental stations) but they are actually targeted for production in much larger set of environments (e.g. farmers' field). Therefore, transferability of variety response to a new location is an important aspect of variety recommendation. Singh et al (1996) provided a statistical measure of the transferability of a variety using multi-

locational data. The approach is as follows. For a given variety say i, its response to the environment can be modeled as a linear regression on environmental index (often considered to be sound biological measure and is taken as mean of all the genotypes at that location). To evaluate transferability of the genotype response to an environment say j-th, fit the linear regression of yield on environmental index using data on (response, index) pairs for all locations except the j-th location and compute the difference in yield response obsevered and predicted response at the j-th location using the above linear regression. Such a difference has been called inter-site residual (Wood and Cady, 1981) and predicted residual (Cook and Wiesberg 1982). Such differences can be obtained by leaving one location at a time. Their (weighted) sum of squares gives inter-site residual sum of squares. For assessing the inter-site residuals, we may consider plot- residuals as within-site residuals. A measure (P) of transferability for the genotype (i under consideration) then is the ratio of inter-site transfer residual sum of square to within-site residual sum of squares weighted with replications. Statistical distributions of linear functions of P has been worked out when error variances over locations are homogeneous/ heterogeneous. Six trials with number of locations varying from 16-53 and variety varying from 21-23 have been presented for barley and wheats in Singh et al (1996).

8. An Illustration

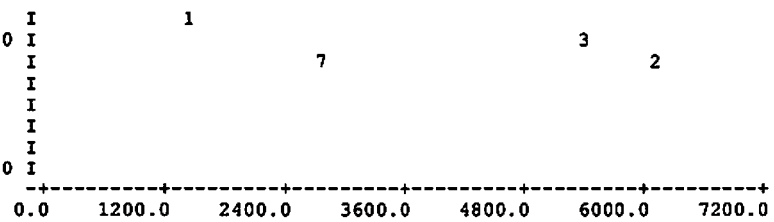
We list in the following printout from a GENSTAT 5 program written for analyzing data from multi-locational variety trials conducted in randomized complete block designs. The program codes are available for designs in complete blocks as well as in incomplete blocks on diskette.

```
  3  "
 -4
 -5
 -6   ************************************************************************
 -7         GENSTAT program for analyzing multi-locational variety trials conducted
 -8          in complete blocks. Data from all locations are in a single file.
 -9
-10             This includes
-11               i.    analysis of data from individual environments
-12               ii.   tests for homogeneity of error variances,
-13               iii.  combined analysis of data for GxE interaction
-14                      under homogeneous/heterogeneous errors
-15               iv.   tests for parallelism of regression lines
-16               v.    common stability statistics
-17               vi.    hierarchical cluster analysis of genotypes
-18               vii.   hierarchical cluster analysis of environments
-19               viii.  clustering of genotypes and environments into groups which
maximizes
-20                       GxE interaction between the groups of genotype and
-21                       groups of environments (Corsten and Denis 1990)
```

```
-22              ix.    non-hierarchical cluster analysis of genotypes
-23              x.     non-hierarchical cluster analysis of environments
-24              xi.    principal component analysis of genotypes
-25              xii.   principal component analysis of environments
-26              xiii.  heritability of the traits
-27
-28            Software: GENSTAT 5 Rel 4.1
-29
-30  ************************************************************************
-31   "
 32
 33
 34
 35
 36  Open ch=2; fi=in ; "Give the name of the text data file"   Name = 'mlvt1_2.txt'
 37
 38
 39
 40
 41  Scal Alpha; 0.05
 42  Scal NRoots;3     " for PCA "
 43  Scal  GxEI% ; 60   " % of GxEI explained by between G-group and E-group"
 44  Scal NLoc, NRepMax, NGeno
 45  Skip[ch=2]2 : Read [ch=2] NLoc,NRepMax,NGeno

  Identifier   Minimum      Mean   Maximum    Values   Missing
       NLoc     10.00     10.00     10.00         1         0
    NRepMax     3.000     3.000     3.000         1         0
      NGeno     15.00     15.00     15.00         1         0

 46  Fact[leve=NLoc] Loc : Fact[leve=NRepMax]Rep : Fact[leve=NGeno]Geno
 47  Skip[ch=2]1 : Read[ch=2] Loc, Rep, Geno, Yield : Clos 2

  Identifier   Minimum      Mean   Maximum    Values   Missing
       Yield     150.0      3692      9000       450         1

  Identifier    Values   Missing    Levels
         Loc       450         0        10
         Rep       450         0         3
        Geno       450         0        15

 48
 49  Scal NRep[1...NLoc]  " replications under individual environments "
 50
 51  For i=1...NLoc ; dr=NRep[1...NLoc]
 52  Rest Rep ; Loc==i : Calc dr=Max(Rep): Rest Rep : Endf
 53
 54
 55  Scal NRepAvrg :  Calc NRepAvrg=VMean(!p(NRep[1...NLoc]))
 56  Scal NObs, NGxNL : Calc NObs=NGeno*VSum(!p(NRep[1...NLoc]))
 57    Calc NGxNL=NGeno*NLoc : Prin NGxNL,NObs  ; deci=0

     NGxNL        NObs
       150         450

 58
 59  Units[NObs]
 60
 61  " =======   Below is only for statistical programmers use ========  "
 62  " =======   Below is only for statistical programmers use ========  "
 63  " =======   Below is only for statistical programmers use ========  "
 64
 65
 66  " 1.  Individual locations analysis "
 67  " 1.  Individual locations analysis "
 68
 69  Scal sigma2, ss,df
 70  Vari[Nval=NGeno] Mean[1...NLoc], GenoMean
 71  Vari[Nval=NLoc]ErrMS,ErrDF, Weight
 72
 73  Vari[Nval=NGeno] Mean[1...NLoc], GenoMean
 74  Vari[Nval=NLoc]CV%,SEM, LocMean
 75  Fact[Leve=NLoc; Valu=1...NLoc] LocNum
 76
 77  Block Rep/Geno : Treat Geno
```

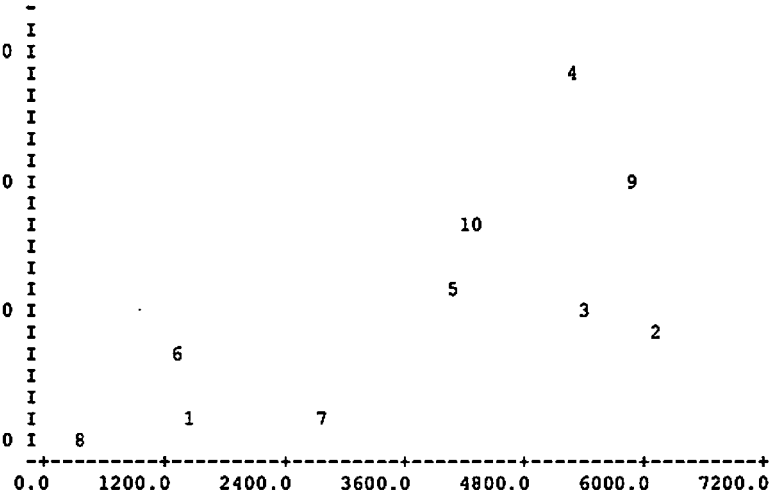
```
 78  For I=1...NLoc     ; MN=Mean[1...NLoc]
 79  Print '******* Location number  is  =   ', I,  '*********'
 80  Rest Yield; Cond=Loc.EQ.I
 81  Anova[prin=a; fpro=y] Yield
 82  Akeep Rep.Geno; ss=ss; df=df
 83  Akeep Geno; Means=TDum
 84  Calc sigma2=ss/df
 85  Calc ErrMS$[I]=sigma2 : Calc ErrDF$[I]=df : Calc Weight$[I]=NRep[I]/sigma2
 86  Equa TDum;MN
 87  Dele[Rede=Y] TDum
 88  Calc LocMean$[I]=Mean(Yield)
 89  Calc CV%$[I]=100*Sqrt(sigma2)/Mean(Yield)
 90  Calc SEM$[I]=sqrt(sigma2/NRep[I])
 91
 92  Rest Yield
 93
 94  Endf

 ******* Location number  is  =          1.000   *********

94.....................................................................

***** Analysis of variance *****

Variate: Yield

Source of variation     d.f.       s.s.       m.s.    v.r.  F pr.

Rep stratum                2     27563.     13781.    0.23

Rep.Geno stratum
Geno                      14    666583.     47613.    0.78  0.682
Residual                  28   1711141.     61112.

Total                     44   2405287.


 ******* Location number  is  =          2.000   *********


94.....................................................................

***** Analysis of variance *****

Variate: Yield

Source of variation     d.f.       s.s.       m.s.    v.r.  F pr.

Rep stratum                2    165032.     82516.    0.18

Rep.Geno stratum
Geno                      14   5153189.    368085.    0.79  0.672
Residual                  28  13060300.    466439.

Total                     44  18378521.


....  other locations' ANOVA dropped  .......

 ******* Location number  is  =          10.00   *********


94.....................................................................

***** Analysis of variance *****

Variate: Yield
```

```
Source of variation      d.f.       s.s.       m.s.    v.r.  F pr.

Rep stratum                 2     933760.    466880.    0.47

Rep.Geno stratum
Geno                       14    9569067.    683505.    0.69  0.769
Residual                   28   27909973.    996785.

Total                      44   38412800.

  95
  96  Prin LocNum,LocMean,CV%,SEM,ErrMS,ErrDF; fiel=9

   LocNum  LocMean      CV%       SEM     ErrMS     ErrDF
        1     1409    17.54     142.7     61112     28.00
        2     6069    11.25     394.3    466439     28.00
        3     5417    14.43     451.4    611252     28.00
        4     5324    24.61     756.6   1717324     28.00
        5     4086    20.23     477.3    683349     28.00
        6     1264    47.34     345.4    357816     28.00
        7     2816    13.59     221.0    146458     27.00
        8      415    20.48      49.1      7220     28.00
        9     5872    18.84     638.7   1223871     28.00
       10     4227    23.62     576.4    996785     28.00

  97  Hist[ngroup=5] CV% :& ErrMS : & LocMean

Histogram of CV%


            -  16  3 ***
        16 -  24  5 *****
        24 -  32  1 *
        32 -  40  0
        40 -      1 *

Scale:  1 asterisk represents 1 unit.


Histogram of ErrMS


         -   400000  4 ****
   400000 -   800000  3 ***
   800000 -  1200000  1 *
  1200000 -  1600000  1 *
  1600000 -          1 *

Scale:  1 asterisk represents 1 unit.

Histogram of LocMean

            -  1500  3 ***
        1500 -  3000  1 *
        3000 -  4500  2 **
        4500 -  6000  3 ***
        6000 -        1 *

Scale:  1 asterisk represents 1 unit.

  98  Graph[nrows=20; ncolumn=60] CV%; LocMean; symb=LocNum

            -
            I           6
       45.0 I
            I
            I
            I
            I
            I
            I
       30.0 I
            I
            I                                          4
            I                                  10
            I   8                          5               9
```



CV% v. LocMean using factor LocNum

```
 99  Graph[nrows=20; ncolumn=60] ErrMS; LocMean; symb=LocNum
```



ErrMS v. LocMean using factor LocNum

```
100
101  " 2. Bartlette Test for homogeneity of error variances "
102  " 2. Bartlette Test for homogeneity of error variances "
103  Scal Prob, PoolMS,PoolDF
104  Calc PoolDF=Sum(ErrDF): &  PoolMS=Sum(ErrMS*ErrDF)/PoolDF
105  Calc Prob=( PoolDF*Log(PoolMS)-Sum(ErrDF*Log(ErrMS)) )/   \
106           ( 1+ (Sum(1/ErrDF)-1/PoolDF)/3/(NLoc-1)  )
107  Calc Prob=Cuchi(Prob;NLoc-1)
108  Prin PoolMS,PoolDF,Prob

     PoolMS       PoolDF        Prob
     628886        279.0           0

109  If Prob.lt.Alpha
110  Print    ' Location error variances heterogeneous  at ' , Alpha,' probability'

                                              Alpha
 Location error variances heterogeneous  at      0.05000  probability
 111  Else
112  Print ' Location error variances homogeneous at ' , Alpha, ' probability'
113  Endif
114
115  Vari[Nval=NLoc] LocMean
116
117  " 3. Combined analysis over locations "
118  " 3. Combined analysis over locations "
119
120  Bloc Loc.Rep/Geno : Trea Loc*Geno
121  Anova[prin=a,m; pse=m; pfact=1;fpro=y] Yield

121.....................................................................

***** Analysis of variance *****

Variate: Yield
```

Source of variation	d.f.(m.v.)	s.s.	m.s.	v.r.	F pr.
Loc.Rep stratum					
Loc	9	1.759E+09	1.955E+08	148.92	<.001
Residual	20	2.625E+07	1.313E+06	2.09	
Loc.Rep.Geno stratum					
Geno	14	2.097E+07	1.498E+06	2.38	0.004
Loc.Geno	126	1.193E+08	9.470E+05	1.51	0.003
Residual	279(1)	1.755E+08	6.289E+05		
Total	448(1)	2.101E+09			

***** Tables of means *****

Variate: Yield

Grand mean 3691.

Loc	1	2	3	4	5	6	7
	1409.	6069.	5417.	5324.	4086.	1264.	2821.

Loc	8	9	10
	415.	5872.	4227.

Geno	1	2	3	4	5	6	7
	3167.	3751.	3744.	3737.	4019.	3599.	3784.

Geno	8	9	10	11	12	13	14
	3584.	3713.	3635.	3694.	4122.	3810.	3542.

Geno	15
	3458.

*** Standard errors of means ***

Table	Loc	Geno
rep.	45	30
d.f.	20	279
e.s.e.	170.8	144.8

(Not adjusted for missing values)

```
122
123  Fact[Leve=NGeno; Valu=1...NGeno] GenoNum
124  Fact[Leve=NLoc; Valu=1...NLoc] LocNum
125  Calc GenoMean=VMean(!P(Mean[1...NLoc]))
126  For i=1...NLoc : Calc LocMean$[i]=Mean(Mean[i]) : Endf
127
128  Prin GenoNum,Mean[1...NLoc],GenoMean; field=7
```

GenoNum	Mean[1]	Mean[2]	Mean[3]	Mean[4]	Mean[5]	Mean[6]	Mean[7]	Mean[8]	Mean[9]	Mean[10]
1	1167	6178	4000	5167	4026	660	2107	441.7	3567	4360
2	1333	6044	4933	6635	3173	1457	2308	391.7	6300	4933
3	1341	6556	6100	5967	3423	1050	2611	316.7	5813	4267
4	1411	5700	5292	6245	4314	699	2785	283.3	6958	3680
5	1464	5511	5867	5617	5551	1437	2833	333.3	7558	4013
6	1511	5595	6050	5489	3699	1168	3063	275.0	6003	3133
7	1411	5933	5350	5737	4987	1451	2740	358.3	5428	4440
8	1495	5867	4700	5445	3070	987	3181	375.0	6837	3880
9	1516	6389	5133	5700	3897	1899	2596	491.7	5387	4120
10	1576	5611	5167	5056	3423	1565	3061	408.3	6200	4280
11	1511	6189	5400	3833	5135	1254	3048	575.0	5637	4360
12	1161	6455	6600	5958	5019	1391	3501	483.3	6280	4373
13	1357	6389	6117	4278	4750	1469	2845	641.7	6360	3893
14	1533	6456	4800	4481	3494	1264	2816	366.7	5200	5013
15	1354	6167	5750	4259	3333	1203	2817	483.3	4558	4653

GenoMean
3167
3751

```
   3744
   3737
   4019
   3599
   3784
   3584
   3713
   3635
   3694
   4122
   3810
   3542
   3458

 129  Prin[orie=a] LocMean ;fiel=7

      LocMean    1409    6069    5417    5324    4086    1264    2821     415    5872

      LocMean    4227

 130
 131  " Transpose data matrix for the sake of convenience"
 132  Matr [Rows=NGeno; Colu=NLoc] GE : & [Rows=NLoc; Colu=NGeno] EG
 133  Equa !P(Mean[1...NLoc]) ; EG
 134  Calc GE=Tran(EG)
 135
 136  Vari[Nval=NLoc] GMean[1...NGeno]
 137  Equa GE;  !P(GMean[1...NGeno])
 138
 139       Dele GE, EG
 140
 141  Vari[ Nval=NGxNL] GEData
 142  Equa  !P(Mean[1...NLoc]); GEData
 143
 144  " ******** Analysis of variance ***************"
 145
 146  Fact[ Leve=NGeno; Nval=NGxNL] Genol : Fact[ Leve=NLoc; Nval=NGxNL] Locl
 147  Gene Locl, Genol
 148  Bloc
 149  Treat Locl*Genol
 150  Anov[prin=a;fpro=y] GEData


150.....................................................................

***** Analysis of variance *****

Variate: GEData

Source of variation     d.f.       s.s.       m.s.    v.r.  F pr.
Locl                       9 5.8640E+08 6.5156E+07
Genol                     14 6.9898E+06 4.9927E+05
Locl.Genol               126 3.9774E+07 3.1566E+05
Total                    149 6.3317E+08

 151
 152
 153  Vari[Nval=NGxNL]AllWet: Equa !p(#NGeno(#Weight)); AllWet
 154  Anov[weight=AllWet; prin=a;fpro=y] GEData


154.....................................................................

***** Analysis of variance *****

Variate: GEData
Weight variate: AllWet

Source of variation     d.f.       s.s.       m.s.    v.r.  F pr.
Locl                       9   9442.689   1049.188
Genol                     14     61.356      4.383
Locl.Genol               126    201.326      1.598
Total                    149   9705.370
```

```
 155  Dele[Rede=Y]Loc1,Geno1 : dele[rede=y] AllWet
 156
 157  " Note--- Mean[1...NLoc] of length NGeno and GMean[1...NGeno] of length NLoc "
 158
 159  " 4. Partition GxE Int in heterogeneity of linear regressions"
 160  " 4. Partition GxE Int in heterogeneity of linear regressions"
 161
 162  " 4.1 Test for heterogeneity of linear regressions: unweighted analysis "
 163  Fact[ Leve=NGeno; Nval=NGxNL] Geno1 : Fact[ Leve=NLoc; Nval=NGxNL] Loc1
 164  Gene Loc1, Geno1
 165
 166  Bloc Loc.Rep/Geno  : Trea Geno*Pol(Loc;1;LocMean) : Anov[prin=a;fpro=y]Yield

166.....................................................................

***** Analysis of variance *****

Variate: Yield

Source of variation     d.f.(m.v.)      s.s.        m.s.     v.r.  F pr.

Loc.Rep stratum
Loc                          9       1.759E+09  1.955E+08  148.92  <.001
  Lin                        1       1.759E+09  1.759E+09 1340.28  <.001
  Deviations                 8       0.535E-21  0.669E-22    0.00  1.000
Residual                    20       2.625E+07  1.313E+06    2.09

Loc.Rep.Geno stratum
Geno                        14       2.097E+07  1.498E+06    2.38  0.004
Loc.Geno                   126       1.193E+08  9.470E+05    1.51  0.003
  Lin.Geno                  14       1.147E+07  8.192E+05    1.30  0.205
  Deviations               112       1.079E+08  9.630E+05    1.53  0.003
Residual                   279(1)    1.755E+08  6.289E+05

Total                      448(1)    2.101E+09

 167
 168
 169  Vari[Nval=NGxNL]AllLoc  : Equa !P(#NGeno(#LocMean)); AllLoc
 170
 171  Model GEData : Fit[Prin=m,s,a;fpro=yes]AllLoc+Geno1+AllLoc.Geno1


171.....................................................................


***** Regression Analysis *****

 Response variate: GEData
     Fitted terms: Constant + AllLoc + Geno1 + AllLoc.Geno1


*** Summary of analysis ***

             d.f.         s.s.         m.s.      v.r.  F pr.
Regression     29    5.972E+08    20593622.     68.74  <.001
Residual      120    3.595E+07      299591.
Total         149    6.332E+08     4249436.

Percentage variance accounted for 92.9
Standard error of observations is estimated to be 547.
* MESSAGE: The following units have large standardized residuals:
         Unit     Response    Residual
           56         3833.       -2.73
          121         3567.       -3.13
* MESSAGE: The error variance does not appear to be constant:
           large responses are more variable than small responses


*** Accumulated analysis of variance ***

Change                          d.f.         s.s.         m.s.      v.r.  F pr.
+ AllLoc                           1    5.864E+08    5.864E+08   1957.35  <.001
+ Geno1                           14    6.990E+06    4.993E+05      1.67  0.072
+ AllLoc.Geno1                    14    3.823E+06    2.731E+05      0.91  0.549
```

```
Residual                          120     3.595E+07     2.996E+05

Total                             149     6.332E+08     4.249E+06

 172  "Fit[Prin=*;Cons=o;fpro=yes; tpro=y]Genol/AllLoc"
 173
 174  "4.2 Test for heterogeneity of linear regressions: weighted analysis "
 175  Vari[Nval=NGxNL]AllWet: Equa !p(#NGeno(#Weight)); AllWet
 176  Model[weight=AllWet;disp=1]GEData
 177  Fit[Prin=m,s,a;fpro=yes]AllLoc+Genol+AllLoc.Genol

177.....................................................................


***** Regression Analysis *****

 Response variate: GEData
   Weight variate: AllWet
     Fitted terms: Constant + AllLoc + Genol + AllLoc.Genol


*** Summary of analysis ***

             d.f.         s.s.         m.s.      v.r. chi pr
Regression     29       9539.8      328.959     328.96  <.001
Residual      120        165.6        1.380
Total         149       9705.4       65.137
* MESSAGE: ratios are based on dispersion parameter with value 1

Percentage variance accounted for 97.9
Standard error of observations is fixed at 1.00
* MESSAGE: The following units have large standardized residuals:
        Unit     Response    Residual
          12      1161.00       -3.03
          16      6178.00        3.03
          20      5511.00       -2.82
          65      5551.00        2.72
* MESSAGE: The following units have high leverage:
        Unit     Response    Leverage
         106       441.67        0.92
         107       391.67        0.92
         108       316.67        0.92
         109       283.33        0.92
         110       333.33        0.92
         111       275.00        0.92
         112       358.33        0.92
         113       375.00        0.92
         114       491.67        0.92
         115       408.33        0.92
         116       575.00        0.92
         117       483.33        0.92
         118       641.67        0.92
         119       366.67        0.92
         120       483.33        0.92


*** Accumulated analysis of variance ***

Change                           d.f.         s.s.         m.s.      v.r. chi pr
+ AllLoc                            1     9442.689     9442.689   9442.69  <.001
+ Genol                            14       61.356        4.383      4.38  <.001
+ AllLoc.Genol                     14       35.762        2.554      2.55  0.001
Residual                          120      165.563        1.380

Total                             149     9705.370       65.137
* MESSAGE: ratios are based on dispersion parameter with value 1

 178  "Fit[Prin=*;Cons=o;fpro=yes; tpro=y]Genol/AllLoc"
 179
 180  dele AllLoc, AllWet
 181
 182  " 5. Compute stability indices "
 183  " 5. Compute stability indices "
```

```
 184
 185  Vari[nvalu=NGeno] GenoCV
 186  Calc GenoCV=100.*Sqrt(Vvar(!P(Mean[1...NLoc])))/GenoMean
 187
 188  Vari[ Nval=NGeno] Slope,SeSlop,DeviMS, Wricke,Pla_Pet,Plaisted, Shukla, YauH
 189  Vari[Nval=NGeno] DevRegDF, Probb1,ProbDev
 190  Vari[ Nval=NGeno] SlopeW,SeSlopW, Probb1W,DeviSSW,ProbDevW, DevRgDFW
 191  Vari[Nval=NGeno]RSq%,RSqW%  " Goodness-of-fit % R-squares adjusted for df"
 192  For I=1...NGeno ; Y=GMean[1...NGeno]
 193  Scal YMeanSq, YMeanSqW : Calc YMeanSq=Var(Y)
 194  Calc YMeanSqW=Sum(Weight*(Y-Sum(Weight*Y)/Sum(Weight))**2 )/(Nval(Y)-1)
 195
 196  " Yau and Hamblin (1995)'s stability index"
 197  Calc YauH$[I]=Var(Y/LocMean)
 198
 199  " Unweighted regression analysis"
 200  Model Y ; Fitt=F
 201  Fit[prin=*] LocMean
 202  RKeep ; Est=Est; Se=Se ; Devi=SS ; DF=df
 203  Calc Slope$[I] =Est $[2] : & SeSlop$[I]=Se$[2] : &  DeviMS $[I] =SS/df
 204  Calc DevRegDF$[I]=df  : & RSq%$[I]=100*(1-SS/df/YMeanSq)
 205  "
-206  Graph[nrows=20; ncolumn=60] Y,F; LocMean; symb='o','.'; Meth=p,c
-207  "
 208
 209  " Weighted regression analysis"
 210  Model[Weight=Weight] Y ; Fitt=F
 211  Fit[prin=*] LocMean
 212  RKeep ; Est=Est; Se=Se ; Devi=SS ; DF=df
 213  Calc SlopeW$[I] =Est $[2] : & SeSlopW$[I]=Se$[2] : &  DeviSSW$[I] =SS
 214  Calc DevRgDFW$[I]=df : & RSqW%$[I]=100*(1-SS/df/YMeanSqW)
 215
 216  Endf
 217
 218  Calc Probb1=Abs(Slope-1)/SeSlop : & Probb1=Cut(Probb1;DevRegDF)
 219  Calc ProbDev=DeviMS/(PoolMS/NRepAvrg)
 220  Calc ProbDev=CuF(ProbDev;DevRegDF;PoolDF)
 221  " above is based on an average number of replications. Use the weighted
-222    analysis results
-223  "
 224
 225  Calc Probb1W=Abs(SlopeW-1)/SeSlopW : & Probb1W=Cut(Probb1W;DevRgDFW)
 226  Calc ProbDevW=Cuchi(DeviSSW;DevRgDFW)
 227
 228  "  Get GxE interactions for stability indices"
 229
 230  Vari[ Nval=NGxNL] GEInt
 231  Bloc  Loc1.Geno1
 232  Trea Loc1+Geno1
 233  Anov[prin=*] GEData; Res=GEInt
 234  AKee  Loc1.Geno1 ;  ss =GxEISS
 235
 236  Prin GxEISS

      GxEISS
    39773742

 237
 238  Calc GEInt=GEInt*GEInt
 239  Tabu [Class=Geno1] GEInt; Tota=TDum
 240
 241  Equa TDum; Wricke
 242       Dele[Rede=Y] TDum
 243
 244  Scal SsGE
 245  Calc SsGE=Sum(Wricke)
 246
 247  Calc Pla_Pet=(NGeno*Wricke+SsGE)/(2*(NGeno-1)*(NLoc-1))
 248  Calc Plaisted=(-NGeno*Wricke/(NGeno-1)+SsGE)/((NGeno-2)*(NLoc-1))
 249  Calc Shukla = (NGeno*Wricke - SsGE/(NGeno-1))/(NGeno-2)/(NLoc-1)
 250
 251  "     Correlation between indices "
 252  Corr[Print=c] GenoMean, Slope, DeviMS,GenoCV,Wricke,Pla_Pet,Plaisted,Shukla, \
 253                YauH, SlopeW,DeviSSW
*** Correlation matrix ***
```

```
GenoMean     1.000
   Slope     0.763    1.000
  DeviMS    -0.367   -0.175    1.000
  GenoCV    -0.142    0.508    0.442    1.000
  Wricke    -0.352   -0.133    0.987    0.484    1.000
 Pla_Pet    -0.352   -0.133    0.987    0.484    1.000    1.000
Plaisted     0.352    0.133   -0.987   -0.484   -1.000   -1.000    1.000
  Shukla    -0.352   -0.133    0.987    0.484    1.000    1.000   -1.000
   YauH     -0.117    0.004    0.501    0.285    0.497    0.497   -0.497
  SlopeW     0.876    0.794   -0.472    0.005   -0.443   -0.443    0.443
 DeviSSW    -0.151    0.048    0.788    0.478    0.810    0.810   -0.810
          GenoMean    Slope   DeviMS   GenoCV   Wricke  Pla_Pet Plaisted

  Shukla     1.000
   YauH      0.497    1.000
  SlopeW    -0.443   -0.319    1.000
 DeviSSW     0.810    0.306   -0.223    1.000
           Shukla     YauH   SlopeW  DeviSSW
```

```
254  Print GenoNum, GenoMean, Slope,SeSlop,Probbl, DeviMS,ProbDev,RSq%, GenoCV,YauH;\
255       field=8
```

GenoNum	GenoMean	Slope	SeSlop	Probbl	DeviMS	ProbDev	RSq%	GenoCV	YauH
1	3167	0.875	0.12606	0.1757	621262	0.0034	83.99	62.19	0.03663
2	3751	1.062	0.10712	0.2893	448611	0.0323	91.53	61.36	0.02380
3	3744	1.107	0.06156	0.0597	148129	0.6857	97.29	62.39	0.01585
4	3737	1.113	0.08666	0.1149	293599	0.1959	94.79	63.55	0.03950
5	4019	1.097	0.11631	0.2150	528832	0.0116	90.71	59.38	0.02798
6	3599	1.001	0.08054	0.4966	253568	0.2931	94.46	59.44	0.02250
7	3784	0.988	0.06195	0.4260	150046	0.6776	96.57	55.27	0.01113
8	3584	1.007	0.09379	0.4698	343905	0.1132	92.70	60.58	0.01921
9	3713	0.945	0.05525	0.1744	119352	0.8028	97.01	53.77	0.03144
10	3635	0.929	0.05189	0.1041	105270	0.8543	97.26	53.93	0.01285
11	3694	0.920	0.10187	0.2266	405679	0.0549	89.94	54.37	0.03123
12	4122	1.122	0.05947	0.0370	138238	0.7270	97.53	57.37	0.01526
13	3810	1.007	0.08812	0.4703	303553	0.1764	93.50	56.74	0.03940
14	3542	0.925	0.08468	0.2013	280336	0.2248	92.93	56.23	0.01360
15	3458	0.902	0.09509	0.1671	353476	0.1015	90.82	56.75	0.01741

```
256  Print GenoNum, GenoMean, SlopeW,SeSlopW,ProbblW, DeviSSW,ProbDevW, RSqW%; fiel=8
```

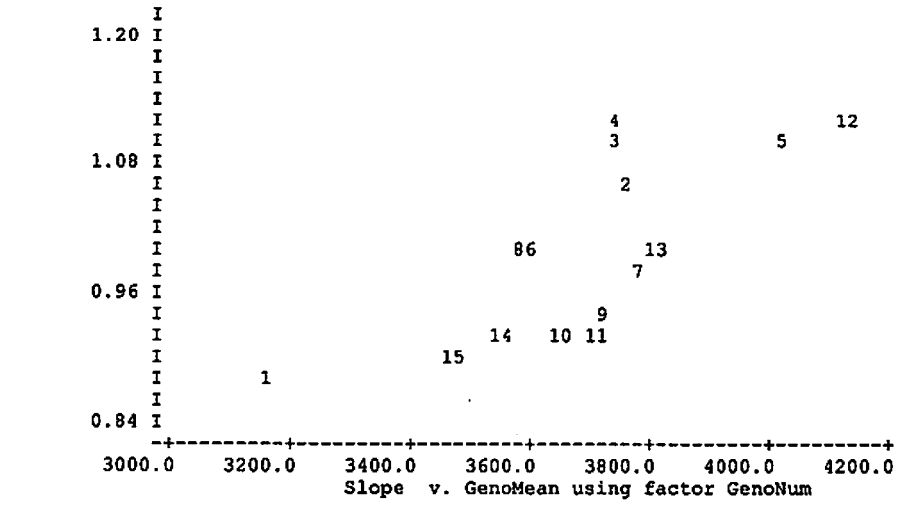
GenoNum	GenoMean	SlopeW	SeSlopW	ProbblW	DeviSSW	ProbDevW	RSqW%
1	3167	0.841	0.06573	0.0209	21.76	0.0054	94.76
2	3751	0.968	0.05322	0.2809	14.26	0.0751	97.34
3	3744	1.052	0.03527	0.0890	6.27	0.6175	99.00
4	3737	1.036	0.04158	0.2080	8.71	0.3677	98.57
5	4019	1.079	0.05987	0.1105	18.05	0.0208	97.30
6	3599	1.034	0.04771	0.2508	11.46	0.1768	98.11
7	3784	1.022	0.03232	0.2558	5.26	0.7294	99.11
8	3584	0.992	0.05240	0.4402	13.83	0.0863	97.54
9	3713	0.969	0.03524	0.2012	6.26	0.6187	98.82
10	3635	0.982	0.03735	0.3218	7.03	0.5338	98.71
11	3694	0.984	0.04245	0.3550	9.08	0.3359	98.35
12	4122	1.117	0.05394	0.0308	14.65	0.0663	97.94
13	3810	0.986	0.04448	0.3816	9.97	0.2674	98.20
14	3542	0.993	0.04376	0.4367	9.65	0.2908	98.28
15	3458	0.946	0.04308	0.1227	9.34	0.3141	98.16

```
257  Prin GenoNum, GenoMean,Wricke,Pla_Pet,Plaisted,Shukla; fiel=8
```

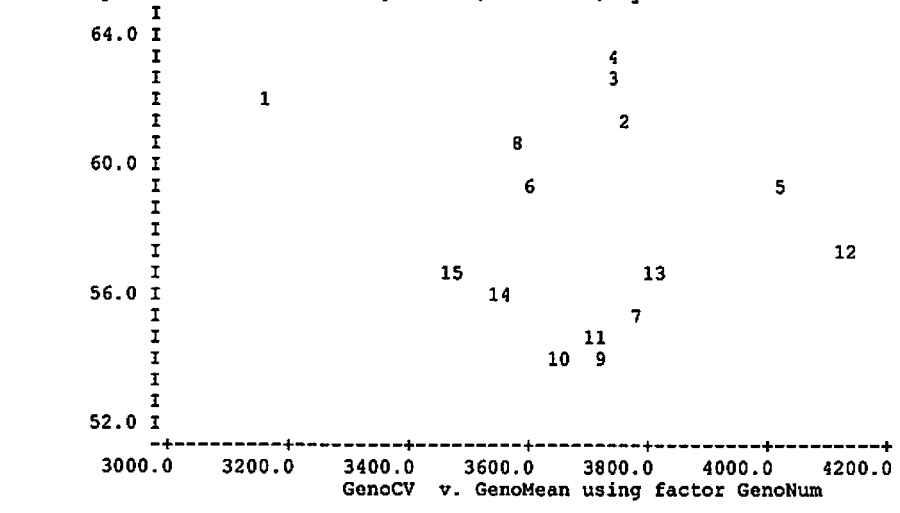
GenoNum	GenoMean	Wricke	Pla_Pet	Plaisted	Shukla
1	3167	5578639	489894	288860	690928
2	3751	3739272	380408	305704	455112
3	3744	1635091	255159	324973	185345
4	3737	2844787	327165	313895	340434
5	4019	4596131	431412	297857	564966
6	3599	2028564	278580	321370	235790
7	3784	1205943	229615	328903	130326
8	3584	2753349	321722	314733	328712
9	3713	1073016	221702	330120	113284
10	3635	1039503	219707	330427	108988
11	3694	3497666	366027	307917	424137
12	4122	1688955	258365	324480	192251
13	3810	2430222	302488	317692	287285
14	3542	2461612	304357	317404	291309

15 3458 3200992 348368 310633 386102
258

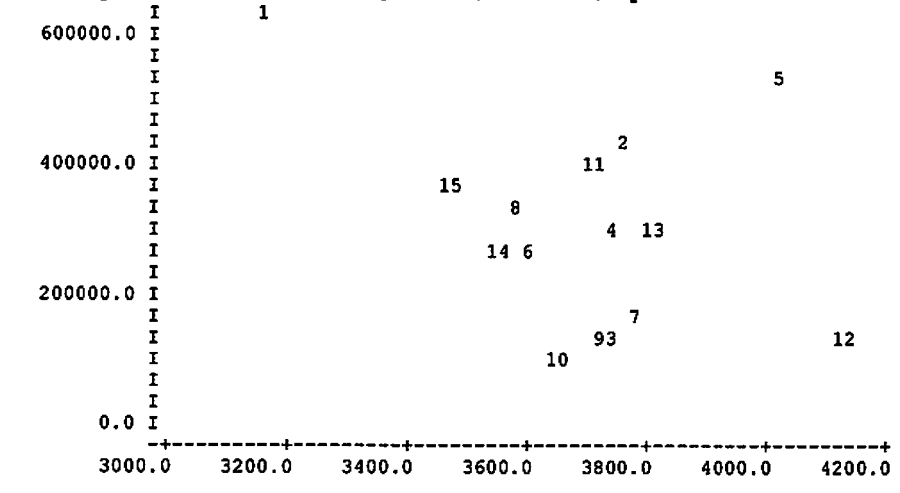
```
259  Graph[nrows=20; ncolumn=60] Slope; GenoMean; Symb=GenoNum
```




```
260  Graph[nrows=20; ncolumn=60] GenoCV; GenoMean; Symb=GenoNum
```

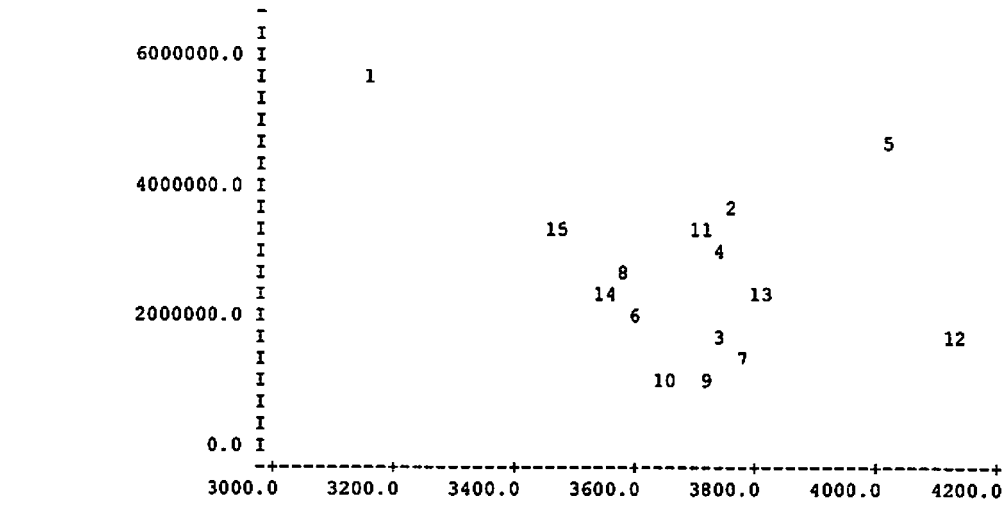



```
261  Graph[nrows=20; ncolumn=60] DeviMS; GenoMean; Symb=GenoNum
```



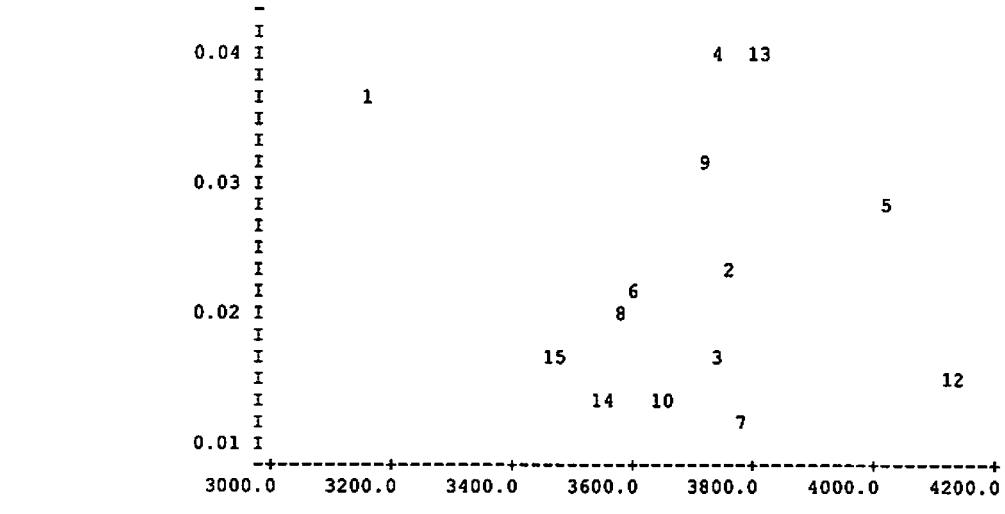
DeviMS v. GenoMean using factor GenoNum

```
262  Graph[nrows=20; ncolumn=60] Wricke; GenoMean; Symb=GenoNum
```



Wricke v. GenoMean using factor GenoNum

```
263  Graph[nrows=20; ncolumn=60] YauH; GenoMean; Symb=GenoNum
```

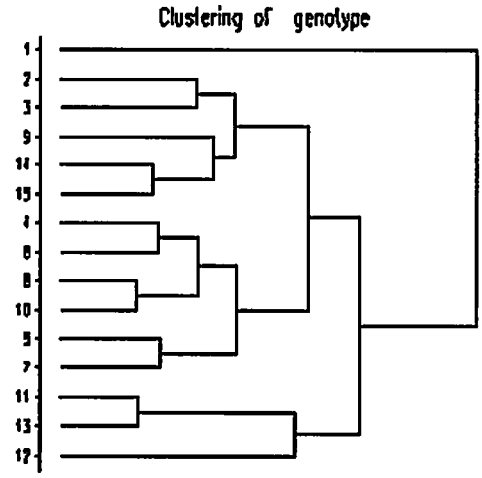


YauH v. GenoMean using factor GenoNum

```
 Points coinciding with 9
  11

264
265
266  Vari[nval=NGeno]RGenoMn,RSlope,RDeviMS,RGenoCV,RWricke, \
267           RPla_Pet,RPlaist, RShukla, RYauH
268  For D= GenoMean, Slope, DeviMS,GenoCV,Wricke,Pla_Pet,Plaisted, \
269              Shukla,YauH ; \
270      DD= RGenoMn, RSlope, RDeviMS,RGenoCV,RWricke, \
271                           RPla_Pet,RPlaist,RShukla,RYauH
272  Vari[valu=1...NGeno]Order, DD
273  Sort[dire=d] D, Order
```

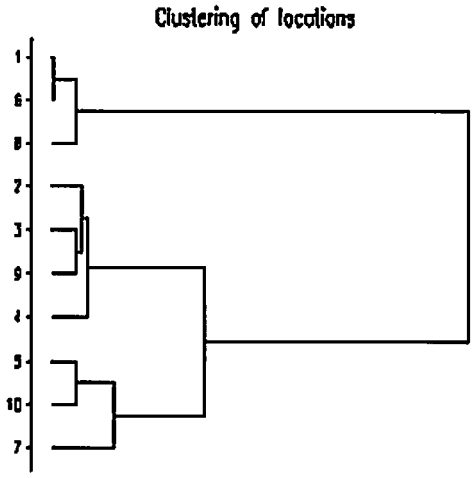
```
274  Sort Order, D, DD
275  endf
276
277
278
279  "       Correlations between ranks "
280  Corr[Prin=c] \
281   RGenoMn,RSlope,RDeviMS,RGenoCV,RWricke,RPla_Pet,RPlaist,RShukla,RYauH

*** Correlation matrix ***
      RGenoMn     1.000
       RSlope     0.739     1.000
      RDeviMS    -0.189    -0.207     1.000
      RGenoCV    -0.014     0.486     0.404     1.000
      RWricke    -0.200    -0.136     0.968     0.475     1.000
     RPla_Pet    -0.200    -0.136     0.968     0.475     1.000     1.000
      RPlaist     0.200     0.136    -0.968    -0.475    -1.000    -1.000     1.000
      RShukla    -0.200    -0.136     0.968     0.475     1.000     1.000    -1.000
        RYauH     0.018     0.046     0.493     0.304     0.507     0.507    -0.507

                RGenoMn    RSlope   RDeviMS   RGenoCV   RWricke RPla_Pet   RPlaist

      RShukla     1.000
        RYauH     0.507     1.000

                RShukla     RYauH

282
283  dele RSlope,RDeviMS, RWricke,RPla_Pet, RPlaisted, RShukla, RYauH
284  dele DevRegDF, Probb1,ProbDev
285  dele SlopeW,SeSlopW, Probb1W,DeviSSW,ProbDevW, DevRgDFW
286  dele RSq%,RSqW%
287
288  " 6. Hierarchical Clustering of Genotypes "
289  " 6. Hierarchical Clustering of Genotypes "
290    Symm[Rows=NGeno] Simi
291     Fsim[Simi=Simi] Mean[1...NLoc]; Test=Eucl
292      Hclus[prin=a,d; method=average] Simi ; Amalg=MatSimi ; Permu=PermSimi

**** Average linkage cluster analysis ****

** Merging clusters **

     8    10  95.3
    11    13  95.2
    14    15  94.3
     4     6  93.9
     5     7  93.8
     2     3  91.5
     4     8  91.5
     9    14  90.5
     2     9  89.2
     4     5  89.2
    11    12  85.6
     2     4  84.7
     2    11  81.5
     1     2  74.3

**** Hierarchical clusters ****

  ** Level     95.0

   8   10

  11   13
** Ungrouped
   1    2    3    9   14   15    4    6    5    7
  12


  ** Level     90.0

   2    3

   9   14   15
```

```
   4    6    8   10

   5    7

  11   13
** Ungrouped
   1   12


  ** Level      85.0

   2    3    9   14   15

   4    6    8   10    5    7

  11   13   12
** Ungrouped
   1


  ** Level      80.0

   2    3    9   14   15    4    6    8   10    5
   7   11   13   12
** Ungrouped
   1


  ** Level      75.0

   2    3    9   14   15    4    6    8   10    5
   7   11   13   12
** Ungrouped
   1


  ** Level      70.0

   1    2    3    9   14   15    4    6    8   10
   5    7   11   13   12

**** Dendrogram ****
** Levels   100.0  90.0  80.0  70.0

           1  ..................
           2  .....             )
           3  .....)..          )
           9  .....   )         )
          14  .....)  )         )
          15  .....)..)..       )
           4  .....       )     )
           6  .....)      )     )
           8  ..   )      )     )
          10  ..)..)..    )     )
           5  .....   )   )     )
           7  .....)..)..)      )
          11  ..          )     )
          13  ..).....    )     )
          12  ........)..).....).....

293       DDENDROGRAM [order=given; style=Average]   MatSimi ; permu=PermSimi ; \
294         Title=' Clustering of  genotype'
295  Dele[Rede=Y] Simi , MatSimi, PermSimi
296
```

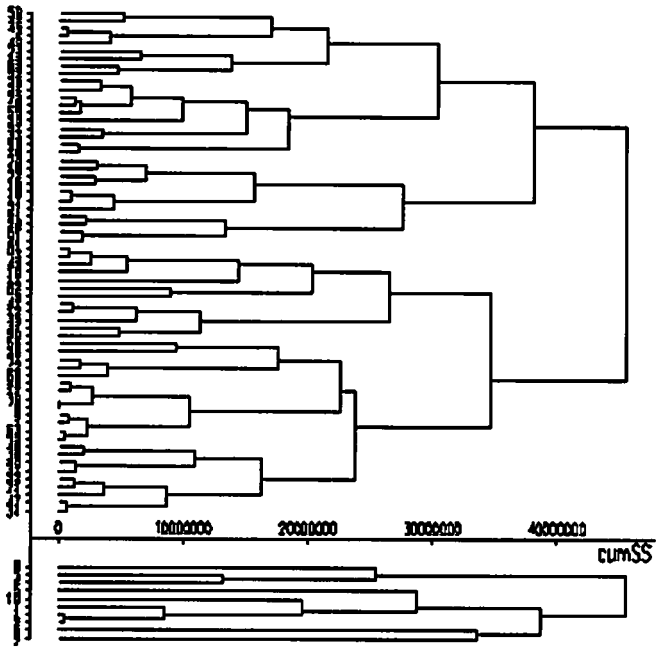



```
297   " 7. Hierarchical Clustering of Environments "
298   " 7. Hierarchical Clustering of Environments "
299
300   Symm[Rows=NLoc] Simi
301   Fsim[Simi=Simi] GMean[1...NGeno]; Test=Eucl
302   Hclus[prin=a,d; method=average] Simi ; Amalg=MatSimi ; Permu=PermSimi

**** Average linkage cluster analysis ****

** Merging clusters **

     1      6  99.7
     3      9  97.5
     1      8  97.5
     5     10  97.5
     2      3  96.9
     2      4  96.3
     5      7  93.7
     2      5  84.5
     1      2  57.9

**** Hierarchical clusters ****

   ** Level      95.0

    1    6    8

    2    3    9    4

    5   10
** Ungrouped
    7


   ** Level      90.0

    1    6    8

    2    3    9    4
```

```
  5   10    7

 ** Level     85.0

  1    6    8

  2    3    9    4

  5   10    7

 ** Level     80.0

  1    6    8

  2    3    9    4    5   10    7

 ** Level     75.0

  1    6    8

  2    3    9    4    5   10    7

 ** Level     70.0

  1    6    8

  2    3    9    4    5   10    7

 ** Level     65.0

  1    6    8

  2    3    9    4    5   10    7

 ** Level     60.0

  1    6    8

  2    3    9    4    5   10    7

 ** Level     55.0

  1    6    8    2    3    9    4    5   10    7

**** Dendrogram ****
** Levels   100.0  90.0  80.0  70.0  60.0

           1  ..
           6  ..)
           8  ..)....................
           2  ..                     )
           3  ..)                    )
           9  ..)                    )
           4  ..)........            )
           5  ..         )           )
          10  ..)..      )           )
           7  .....).....)...........).........
```




```
 303
 304  DDENDROGRAM [order=given; style=Average]   MatSimi ; permu=PermSimi ; \
 305   Title=' Clustering of locations'
 306
 307  Dele[Rede=Y] Simi , MatSimi, PermSimi
 308
 309
 310  "
-311   8. Clusters genotypes and environments into groups which maximizes
-312      GxE interaction between the groups of genotype and groups of environments
(Corsten and Denis 1990)
-313   8. Clusters genotypes and environments into groups which maximizes
-314      GxE interaction between the groups of genotype and groups of environments
(Corsten and Denis 1990)
-315  "
 316
 317  "
-318  Scal VarMean  : Calc VarMean=PoolMS/NRepAvrg
-319  Scal  SSThres  : Calc SSThres=(1-GxEI%/100)*GxEISS
-320
-321  Tabu[Class=Genol, Locl] GEData; Means=TabGE
-322  CINTERACTION [prin=sort,aov,summ,vari,dend ; \
-323      Vari=VarMean; DF=PoolDF; SSTHRES=SSThres] Table=TabGE
-324  "
```




```
325
326
327  " 10. Non-hierarchical clusters of environments "
328        " Changes are required if more than 3 groups are required"
329
330  If NLoc.lt.NGeno
331  Print  '        Number of environments less than number of genotypes '  , \
332           '          resulting in a singular variance-covariance matrix  '
```

```
   Number of environments less than number of genotypes


    resulting in a singular variance-covariance matrix
```

```
333  endi
334
335
336  If NLoc.ge.NGeno
337  Fact[leve=3; nval=NLoc] Grp[3] : & [leve=2; nvalu=NLoc] Grp[2]
338  Pointer[Valu=GMean[1...NGeno]] EG_data
339  Cluster[prin=c,o; data=EG_data; cri=maha]Ngroups=3,2 ; groups=Grp[3,2]
340
341  For i=1...3
342  Rest LocNum,GMean[1...NGeno] ; Grp[3].eq.i
343  Prin Grp[3],LocNum,GMean[1...NGeno] ;fiel=6
344  Rest LocNum,GMean[1...NGeno]
345  Endf
346
347  For i=1...2
348  Rest LocNum,GMean[1...NGeno] ; Grp[2].eq.i
349  Prin Grp[2],LocNum,GMean[1...NGeno] ;fiel=6
350  Rest LocNum,GMean[1...NGeno]
351  Endf
352
353  Endi
354
355
356
357
```

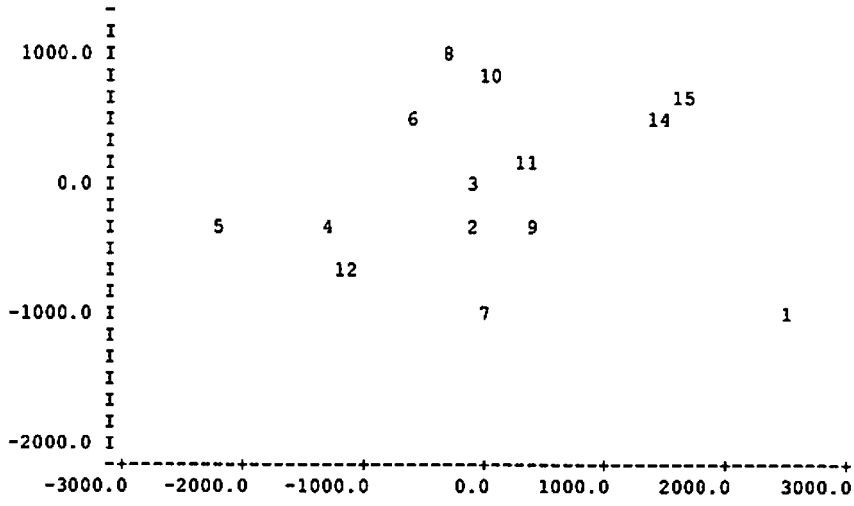
```
 358  " 9. Non-hierarchical clusters of genotypes
-359   9. Non-hierarchical clusters of genotypes
-360  "
 361     " Changes are required if more than 3 groups are required   "
 362
 363  If NGeno.lt.NLoc
 364    Print '         Number of genotypes less than number of environments' , \
 365        '           resulting in a singular variance-covariance matrix'
 366  Endif
 367
 368   If  NGeno.ge.NLoc
 369    Fact[leve=3; nval=NGeno] Grp[3] : & [leve=2; nvalu=NGeno] Grp[2]
 370   Pointer[Valu=Mean[1...NLoc]] GE_data
 371   Cluster[prin=c,o; data=GE_data; cri=maha]Ngroups=3,2 ; groups=Grp[3,2]

371.....................................................................

*****  Non-hierarchical Clustering  *****

**** Mahalanobis distance criterion  ***


***  Optimum classification  ***

***  Number of classes = 3

***  Class contributions not printed  ***

***  Criterion value = 8934.36114

***  Classification of units  ***

     3     1     2     1     2     2     2     2     2     2     3     2     3
     3     3



***  Optimum classification  ***

***  Number of classes = 2

***  Class contributions not printed  ***

***  Criterion value = 2444.85203

***  Classification of units  ***

     2     2     1     1     2     1     2     1     2     2     2     1     2
     2     2
  372
 373  For i=1...3
 374   Rest GenoNum,Mean[1...NLoc] ; Grp[3].eq.i
 375   Prin Grp[3],GenoNum,Mean[1...NLoc] ;fiel=6
 376   Rest GenoNum,Mean[1...NLoc]
 377   Endf

 Grp[3] GenoNum Mean[1] Mean[2] Mean[3] Mean[4] Mean[5] Mean[6] Mean[7] Mean[8]
Mean[9] Mean[10]
     1      2   1333   6044   4933   6635   3173   1457   2308 391.7   6300   4933
     1      4   1411   5700   5292   6245   4314    699   2785 283.3   6958   3680


 Grp[3] GenoNum Mean[1] Mean[2] Mean[3] Mean[4] Mean[5] Mean[6] Mean[7] Mean[8]
Mean[9] Mean[10]
     2      3   1341   6556   6100   5967   3423   1050   2611 316.7   5813   4267
     2      5   1464   5511   5867   5617   5551   1437   2833 333.3   7558   4013
     2      6   1511   5595   6050   5489   3699   1168   3063 275.0   6003   3133
     2      7   1411   5933   5350   5737   4987   1451   2740 358.3   5428   4440
     2      8   1495   5867   4700   5445   3070    987   3181 375.0   6837   3880
     2      9   1516   6389   5133   5700   3897   1899   2596 491.7   5387   4120
     2     10   1576   5611   5167   5056   3423   1565   3061 408.3   6200   4280
     2     12   1161   6455   6600   5958   5019   1391   3501 483.3   6280   4373
```

```
 Grp[3] GenoNum Mean[1] Mean[2] Mean[3] Mean[4] Mean[5] Mean[6] Mean[7] Mean[8]
Mean[9] Mean[10]
     3     1  1167  6178  4000  5167  4026   660  2107 441.7  3567  4360
     3    11  1511  6189  5400  3833  5135  1254  3048 575.0  5637  4360
     3    13  1357  6389  6117  4278  4750  1469  2845 641.7  6360  3893
     3    14  1533  6456  4800  4481  3494  1264  2816 366.7  5200  5013
     3    15  1354  6167  5750  4259  3333  1203  2817 483.3  4558  4653

 378
 379    For  i=1...2
 380   Rest GenoNum,Mean[1...NLoc] ; Grp[2].eq.i
 381   Prin Grp[2],GenoNum,Mean[1...NLoc] ;fiel=6
 382   Rest GenoNum,Mean[1...NLoc]
 383    Endf

 Grp[2] GenoNum Mean[1] Mean[2] Mean[3] Mean[4] Mean[5] Mean[6] Mean[7] Mean[8]
Mean[9] Mean[10]
     1     3  1341  6556  6100  5967  3423  1050  2611 316.7  5813  4267
     1     4  1411  5700  5292  6245  4314   699  2785 283.3  6958  3680
     1     6  1511  5595  6050  5489  3699  1168  3063 275.0  6003  3133
     1     8  1495  5867  4700  5445  3070   987  3181 375.0  6837  3880
     1    12  1161  6455  6600  5958  5019  1391  3501 483.3  6280  4373


 Grp[2] GenoNum Mean[1] Mean[2] Mean[3] Mean[4] Mean[5] Mean[6] Mean[7] Mean[8]
Mean[9] Mean[10]
     2     1  1167  6178  4000  5167  4026   660  2107 441.7  3567  4360
     2     2  1333  6044  4933  6635  3173  1457  2308 391.7  6300  4933
     2     5  1464  5511  5867  5617  5551  1437  2833 333.3  7558  4013
     2     7  1411  5933  5350  5737  4987  1451  2740 358.3  5428  4440
     2     9  1516  6389  5133  5700  3897  1899  2596 491.7  5387  4120
     2    10  1576  5611  5167  5056  3423  1565  3061 408.3  6200  4280
     2    11  1511  6189  5400  3833  5135  1254  3048 575.0  5637  4360
     2    13  1357  6389  6117  4278  4750  1469  2845 641.7  6360  3893
     2    14  1533  6456  4800  4481  3494  1264  2816 366.7  5200  5013
     2    15  1354  6167  5750  4259  3333  1203  2817 483.3  4558  4653

 384
 385    Endif
 386
 387
 388
 389  " 11. Principal component analysis   for genotypes "
 390  " 11. Principal component analysis   for genotypes "
 391
 392  Pointer[values=Mean[1...NLoc]]Data_Loc
 393  Matr[rows=NGeno; Colu=NRoots]PCScore
 394  Vari[nval=NGeno] PCS[1...NRoots]
 395
 396  PCP[Print=1,r,t; nroots=NRoots] Data_Loc; Scores=PCScore

396.....................................................................


*****  Principal components analysis  *****

***  Latent Roots  ***
            1           2           3
     19652260    11339087     5721755
***  Percentage variation  ***
            1           2           3
         42.02       24.25       12.24

***  Trace  ***
    46763498

***  Latent Vectors (Loadings)  ***
                  1           2           3
     Mean[1]    -0.01595    -0.00287     0.11513
     Mean[2]     0.12483     0.08662    -0.06460
     Mean[3]    -0.34907     0.30239     0.10301
     Mean[4]    -0.29299    -0.66406    -0.60624
     Mean[5]    -0.31794     0.63791    -0.63810
```

```
     Mean[6]     -0.05731      0.06721      0.02572
     Mean[7]     -0.14832      0.13131      0.24152
     Mean[8]      0.02140      0.07354      0.01506
     Mean[9]     -0.78072     -0.16108      0.34760
    Mean[10]      0.20030     -0.00525     -0.13124

***  Significance tests for equality of final K roots  ***

  No. (K)      Chi
   Roots    squared      df

      2       11.51       2
      3       32.02       5
      4       37.26       9
      5       42.67      14
      6       59.62      20
      7       73.35      27
      8       84.70      35
      9      105.01      44
     10      129.82      54

397
398  Calc PCS[1...NRoots] =PCScore$[*;1...NRoots]
399  Graph[nrows=20; ncolumn=60] GenoMean;PCS[1] ; Symbol=GenoNum
           I
     4500.0 I
           I
           I
           I
           I
           I                    12
     4000.0 I         5
           I
           I                        13
           I                  4           27   9
           I                               10 11
           I                        6  8               14
     3500.0 I
           I                                              15
           I
           I
           I                                                    1
           I
     3000.0 I
           -+---------+---------+---------+---------+---------+---------+
       -3000.0   -2000.0   -1000.0        0.0    1000.0    2000.0    3000.0
                          GenoMean  v. PCS[1] using factor GenoNum

            Points coinciding with 2
              3
400  Graph[nrows=20; ncolumn=60] PCS[2];PCS[1] ; Symbol=GenoNum
           -
           I                                    11
     1500.0 I
           I                        13
           I
           I
           I         5         12                           15
           I                                14
        0.0 I                                                    1
           I                        6     10  9
           I
           I                  4           3
           I                            8
           I
    -1500.0 I
           I                              2
           I
           I
           I
           I
    -3000.0 I
           -+---------+---------+---------+---------+---------+---------+
       -3000.0   -2000.0   -1000.0        0.0    1000.0    2000.0    3000.0
```

PCS[2] v. PCS[1] using factor GenoNum

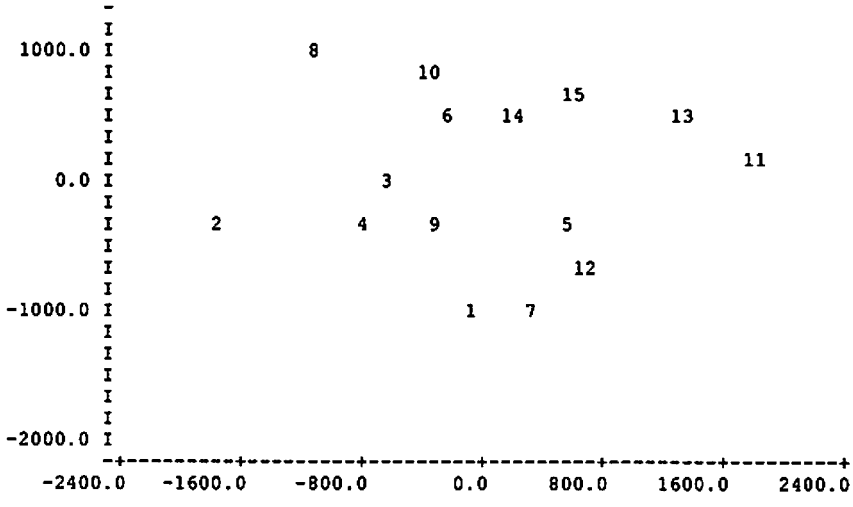
```
401  Graph[nrows=20; ncolumn=60] PCS[3];PCS[1] ; Symbol=GenoNum
```



PCS[3] v. PCS[1] using factor GenoNum

Points coinciding with 6
13

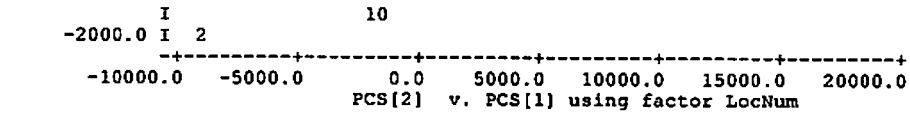
```
402  Graph[nrows=20; ncolumn=60] PCS[3];PCS[2] ; Symbol=GenoNum
```



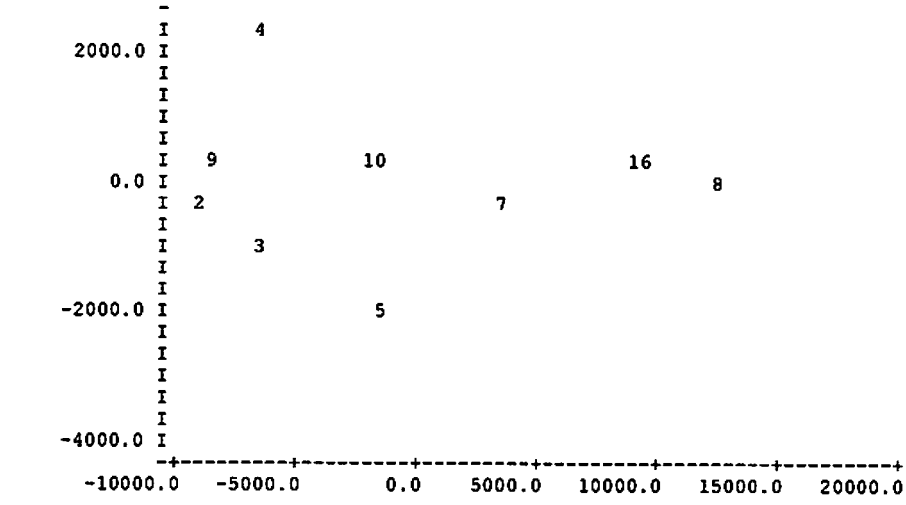
PCS[3] v. PCS[2] using factor GenoNum

```
403  Dele Data_Loc
404
405  " 12. Principal component analysis   for environments "
406  " 12. Principal component analysis   for environments "
407  Pointer[values=GMean[1...NGeno]]Data_Gen
408  Matr[rows=NLoc; Colu=NRoots]PCScore
409  Vari[nval=NLoc] PCS[1...NRoots]
410
411  PCP[Print=1,r,t; nroots=NRoots] Data_Gen; Scores=PCScore
```

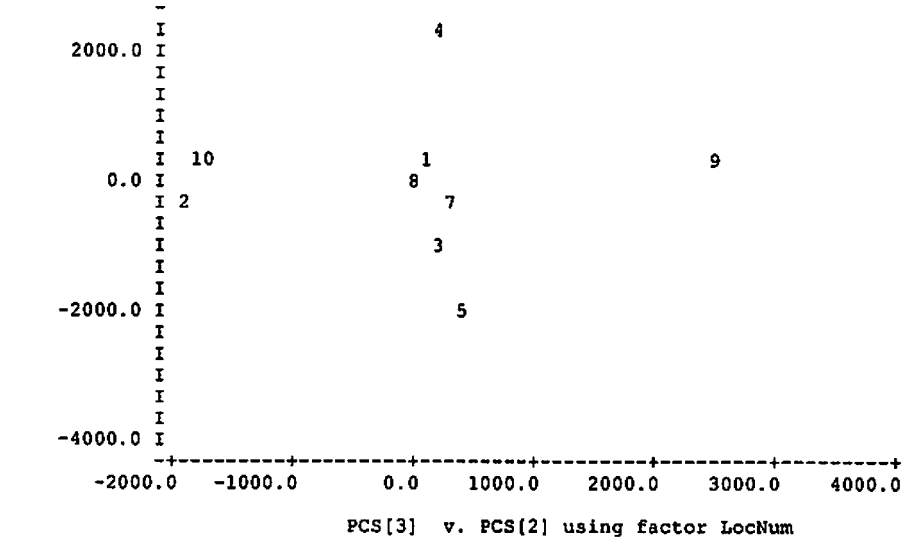
```
411.....................................................................

*****  Principal components analysis  *****

***  Latent Roots  ***

            1           2           3
    590271617    13246608    10841600

***  Percentage variation  ***

            1           2           3
        94.27        2.12        1.73

***  Trace  ***

   626176149

***  Latent Vectors (Loadings)  ***

                        1           2           3
    GMean[1]    -0.22466    -0.52192     0.03860
    GMean[2]    -0.27343    -0.05691     0.53187
    GMean[3]    -0.28493    -0.10886     0.17753
    GMean[4]    -0.28683     0.32571     0.16908
    GMean[5]    -0.28277     0.46215    -0.20718
    GMean[6]    -0.25788     0.23251     0.01987
    GMean[7]    -0.25422    -0.07736    -0.08669
    GMean[8]    -0.25951     0.20937     0.29682
    GMean[9]    -0.24309    -0.13974     0.13184
   GMean[10]    -0.23917     0.08706     0.11551
   GMean[11]    -0.23638    -0.07437    -0.51634
   GMean[12]    -0.28886     0.05121    -0.17524
   GMean[13]    -0.25907     0.09627    -0.40939
   GMean[14]    -0.23759    -0.33745    -0.00923
   GMean[15]    -0.23171    -0.36133    -0.14441

***  Significance tests for equality of final K roots  ***
The last 6 latent roots are (effectively) zero
and have been excluded from the tests.
   No. (K)      Chi
    Roots    squared      df
      2         0.92       2
      3         2.40       5
      4         4.48       9
      5        11.95      14
      6        17.98      20
      7        27.14      27
      8        33.37      35
      9       135.61      44
 412
 413  Calc PCS[1...NRoots] =PCScore$[*;1...NRoots]
 414  Graph[nrows=20; ncolumn=60) PCS[2];PCS[1] ; Symbol=LocNum
              -
              I
      4000.0  I
              I
              I
              I
              I
              I   9
      2000.0  I
              I
              I
              I
              I
              I       3         5         7
         0.0  I       4                             16     8
              I
              I
              I
              I
```




```
415  Graph[nrows=20; ncolumn=60] PCS[3];PCS[1] ; Symbol=LocNum
```




```
416  Graph[nrows=20; ncolumn=60] PCS[3];PCS[2] ; Symbol=LocNum
```




```
 Points coinciding with 1
  6

417
418  " 13. Estimation of variance components and heritabilities"
419  " 13. Estimation of variance components and heritabilities"
420
421
422  " 13.1 From individual environments"
```

```
423
424  Vari[Nvalu=NLoc]RCBSGg2,RCBSGe2,RCBHerit,RCBBias, RCBSeh2
425
426  SCAL SGg2,SGe2,h2
427  Scal Vgg,Vge,Vee,Bias, Seh2
428  symm[2] Vcov_r
429
430  For i=1...NLoc
431
432  Rest Yield; Loc.eq.i
433
434  VCOMP[fixed=Rep] RANDOM=Rep+Geno ; cons=pos
435  REML[print=*] Yield
436
437  VKEEP[SIGMA2=SGe2;vcov=Vcov_r] Geno; COMP=SGg2
438  EQUA Vcov_r ; !p(Vgg,Vge,Vee)
439  CALC h2=SGg2/(SGg2+SGe2)
440  CALC One_h22=(1-h2)**2
441  CALC Bias=One_h22*((1-h2)*Vgg-h2*Vge)/(h2*SGe2*SGe2)
442  CALC Seh2=(1-h2)*SQRT(One_h22*Vgg-2*h2*(1-h2)*Vge+Vee*h2**2)/SGe2
443  Calc  (RCBSGg2,RCBSGe2,RCBHerit,RCBBias, RCBSeh2)$[i]= SGg2,SGe2,h2,Bias,Seh2
444  rest Yield
445  Endf

**** G5W0001 **** Warning (Code CA 7). Statement 10 in For Loop
Command: CALC Seh2=(1-h2)*SQRT(One_h22*Vgg-2*h2*(1-h2)*Vge+Vee*h2**2)/SGe2
Invalid value for argument of function
The first argument of the SQRT function in unit 1 has the value 0.0000

446  Corr[prin=c]LocMean,RCBHerit,RCBSGg2,RCBSGe2

*** Correlation matrix ***
      LocMean    1.000
     RCBHerit   -0.168    1.000
      RCBSGg2    0.483    0.483    1.000
      RCBSGe2    0.704   -0.210    0.418    1.000

               LocMean RCBHerit  RCBSGg2  RCBSGe2

447  Print LocNum,LocMean,RCBHerit,RCBBias, RCBSeh2,RCBSGg2,RCBSGe2 ; fiel=10

   LocNum   LocMean  RCBHerit   RCBBias   RCBSeh2   RCBSGg2   RCBSGe2
        1      1409    0.0000    0.0000    0.0000         0     56612
        2      6069    0.0000    0.0000         *         0    433654
        3      5417    0.2913    0.1184    0.1731    251218    611261
        4      5324    0.0424    0.6085    0.1603     75969   1717368
        5      4086    0.3887    0.0924    0.1677    434517    683351
        6      1264    0.0000    0.0000    0.0000         0    343770
        7      2821    0.3243    0.1100    0.1735     69855    145541
        8       415    0.5432    0.0645    0.1471      8587      7220
        9      5872    0.3166    0.1105    0.1722    567058   1223885
       10      4227    0.0000    0.0000    0.0000         0    892358

448
449  " 13.2 Overall environments"
450
451  SCALAR SGg2,SGe2,SGi2,SGb2,Vgg,Vee,Vii,Vgi,Vge,Vie, h2, Bias,Seh2
452       Symm[3] Vcov_r
453
454  VCOMP[abso=Loc; Fixed=Loc/Rep] RANDOM=Loc+Loc.Rep+Geno+Geno.Loc ; cons=pos
455            REML[print=*] Yield
456
457  VKEEP[SIGMA2=SGe2;vcov=Vcov_r] Geno+Geno.Loc; COMP=SGg2,SGi2
458     EQUA Vcov_r  ;  !p(Vgg,Vgi,Vii,Vge,Vie,Vee)
459  CALC h2=SGg2/(SGg2+SGi2+SGe2)
460     CALC Bias =h2*(Vgg-h2*(Vgg+Vgi+Vge))/SGg2/SGg2
461       CALC Seh2=Vgg+Vii+Vee+2*(Vgi+Vge+Vie)
462         CALC Seh2=Vgg+h2*h2*Seh2-2*h2*(Vgg+Vgi+Vge)
463          CALC Seh2=h2*SQRT(Seh2)/SGg2
464  Print[iprin=*] 'heritability = ',  h2, ' Bias = ', Bias, '  SError= ', Seh2

heritability =      0.02438     Bias =      0.02661    SError=      0.02535

465
466  Clos
```

REFERENCES

Anderson, J.R. 1974. Risk-efficiency in the interpretation of agricultural production research.Review of Marketing and Agricultural Economics, 42, 131-184.

Becker, H.C. 1981. Correlations among some statistical measures of phenotypic stability Euphytica 30:835-840.

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

Byth, D.E. and Mungomery, V.E. (Editor) (1981) Interpretation of Plant Response and Adaptation to Agricultural environments. Refresher Training Workshop held at University of Queensland, St Lucia, Brisbane, Feb 2-6, 1981; Australia.

Byth, D.E., Eisemann, R.L. and DeLacy, I.H. (1976). Two-way pattern analysis of a large data set to evaluate genotypic adaptation. Heredity 37(2):215-230.

Cook, R.D. and Wiesberg, S. (1982). Residuals and Influence in Regression. Chapman and Hall, New York.

Cormack, R.M 1971. A review of classification (with discussion) Journal of the Royal Statistical Society, A134, 321-367.

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

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

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

Gauch, HG (1988). Model selection and validation for yield trials with interaction. Biometrics 88: 705-715.

Gauch, HG and RW Zobel . (1988). Predictive and postdictive success of statistical analyses of yield trials. Theoretical and Applied Genetics 76: 1-10.

Genstat 5 Committee (1993). Genstat^{TM} 5 Release 3 Reference Manual. Clarendron Press: Oxford.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 53: 325-328.

Gower, J.C. 1967. The Statistician 17:13-28

Kempthorne O (1952). Design and Analysis of Experiments. Wiley, New York.

Lin, C.S., Binns, M. R. and Lefkovitch, L.P. (1986). Stability Analysis: where do we stand? Crop Science, 26:894-900.

Mandel, J. (1969). Journal of Research of the National Bureau of standards, Section B, Mathematical Sciences 73B, 309-28.

Menz, K. M. 1980. A comparative analysis of wheat adaptation across international environments using stochastic dominance and pattern analysis. Field Crops Research, 3:33-41.

Plaisted, R.L. 1960. A shorter method for evaluating the abilty of selections to yield consistently over locations. Am. Potato J. 37:166-172.

Plaisted, R.L. and L.C. Peterson. 1959. A technique for evaluating the ability of selections to yield consistently in different locations or seasons. Am. Potato J. 36:381-385.

Perkins, J.M. and J.L. Jinks. 1968. Environmental and genotype-environmental

components of variability. III. Multiple lines and crosses. Heredity 23:339-356.
Shukla, G.K. 1972a. Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29:237-245.
Shukla, G.K. 1972b. An invariant test for the homogeneity of variances in a two-way classification. Biometrics 28:1063-1072.
Singh, M., Yau, S.K., Hamblin, J. and Porceddu, E. (1996). Inter-site transferability of crop varieties: another approach for analyzing multi-locational variety trials. Euphytica 89:305-311.
Wood, C.L. and Cady, F. B. (1981). Inter-site transfer of estimated response surfaces. Biometrics 37:1-10.
Wricke, G. 1962. Urbereine Methode zur Erfassung der okolgischen streubreite in Feldversuchen z. Pflanzenzuecht. 47:92-96.
Yates, F. and W.G. Cochran (1938). The analysis of groups of expriments. Journal of Agricultural Science, 28:556-580.
Yau, SK and J. Hamblin (1994). Relative yield as a measure of entry performance in variable environments. Crop Science 34:813-817.

BIBLIOGRAPHY

Abou-El-Fittouh, HA. , J.O. Rawlings, and P.A. Millet, 1969. Classification of environments to control genotype by environment interactions with an application to cotton. Crop Sci. 9:135-140.
Grubs, F.E. 1948. On estimating precision of measuring instruments and product variability J. A m. Stat. Assoc. 43:243-264.
Guitard, A.A. 1960. The use of diallel correlations for determining the relative locational performance of varieties of barley. Can. J. Plant Sci. 40:645-651
Habgood, R.M. 1977. Estimation of genetic diversity of self-fertilizing cereal cultivars based on genotype-environment interactions. Euphytica 26:485-489.
Hanson, W.D. 1970. Genotypic stability. Theor. Appl. Genet. 40:226-231.
Johnson, G.R. 1977. Analysis of genotypic similarity in terms of mean yield and stability of environmental response in a set of maize hybrids. Crop Sci. 17:837-842.
Knight, R. 1970. The measurement and interpretation of genotype environment interactions. Euphytica 19:225-235.
Lefkovitch, L.P. 1980. Conditional clustering. Biometrics 36:43-58.
Lin, C.S. 1982. Grouping geneotypes by a cluster method directly related to genotype-environment interaction mean square. Theor. Appl. Genet. 62:277-280.
Lin, C.S. and B. Thompson. 1975. An empirical method of grouping genotypes based on
a linear function of the genotype-environment interaction. Heredity 34:255-263.
Tai, G.C.C. 1971. Genotypic stability analysis and its application to potato regional. Crop Sci. 11:184-190.
Williams, E.J. 1952. The interpretation of interactions in factorial Experiments. Biometrika, 39:65-81.
Westcott, B. (1986). Some methods of analysing genotype-environment interaction, Heredity 56:243- 253.
Zobel, R.W., AJ Wright and HG Gauch .(1988). Agronomy Journal. 80: 388-393