CLASSIFYING WINTER WHEAT ENVIRONMENTS INTO ADAPTIVE ZONES AS A BASIS FOR RECOMMENDING A REDUCTION IN THE NUMBER OF INTERNATIONAL WINTER WHEAT PERFORMANCE NURSERY TEST SITES

by

Davies Maleka Lungu

A THESIS

Presented to the Faculty of
The Graduate College in the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Master of Science

Department of Agronomy

Under the Supervision of Professor John W. Schmidt

Lincoln, Nebraska

July, 1978

ACKNOWLEDGEMENTS

The author expresses sincere gratitude and appreciation to his major advisor, Dr. J. W. Schmidt, for suggesting this topic for his thesis, for advice and guidance during the course of this study and for substantial counsel and advice throughout the period of graduate study.

Appreciation is expressed to Dr. K. D. Wilhelmi for some valuable suggestions during the initial phase of this study. Deep gratitude is expressed to Dr. V. A. Johnson for serving on the oral examining committee and to Dr. C. O. Gardner for the discussions that brought about a fuller understanding of some statistical concepts to the author and also for serving on the oral examining committee. Sincere gratitude is expressed to Mr. G. H. Semuguruka for his untiring help with computer programming.

Special thanks are expressed to the University of Zambia for providing the opportunity and the financial support for this post-graduate study and to the Department of Agronomy, University of Nebraska for making the facilities available in this study.

Sincere gratitude is expressed to his dear wife, Beatrice, for her invaluable help and encouragement during the graduate studies and for typing the initial copy of this thesis. Gratitude is also expressed to his loving daughter, Zaliwe, who inspires his working.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
LITERATURE REVIEW	5
Genotype x Environment Interaction	5
Method of Studying GE Interaction	8
Ways to Reduce GE Interaction	14
MATERIALS AND METHODS	20
Materials	20
Methods	20
Statistical Analysis of Data	29
RESULTS	44
Estimates of Components of Variance from Analyses of Variance	44
Results of Classification	49
Estimates of Stability Parameters	60
DISCUSSION	67
SUMMARY AND CONCLUSIONS	83
LITERATURE CITED	86
APPENDIX	91

LIST OF TABLES

Table		Page
1.	Varieties entered in the fourth and fifth International Winter Wheat Performance Nursery. 1972-1973.	21
2.	Locations for the fourth and fifth IWWPN used in the study.	22
3.	Table showing breakdown of locations into regions.	28
4.	Form of the individual year analysis of variance for Southern Europe, United States, and Far East.	30
5.	Form of the individual year analysis of variance for all IWWPN locations combined and separately for regions; Northern Europe, and Near East.	31
6.	Form of combined year analysis of variance for United States, Southern Europe and the Far East.	33
7.	Form of combined year analyses of variance for all the test locations combined and, separately, for regions, Northern Europe and Near East.	34
8.	Estimation of variance components for individual year analysis of variance.	35
9.	Estimation of variance components for combined year analysis of variance.	36
10.	Methods of estimating standard errors of variance components for individual year analysis of variance.	38
11.	Method of estimating standard errors of variance components for combined year analysis of variance.	39
12.	Form of analysis of variance for genotype stability.	43

Table Page

13. Estimates of pertinent variance components from the analysis of variance across all locations included in the study. 45

14. Estimates of pertinent variance components from the individual regional analyses of variance. 47

15. Analysis of variance of grain yield measured on twenty-five varieties of wheat grown in 66 environments (1972-1973) used for stability parameter estimation. 61

16. Stability parameters for yield means of twenty-five wheat cultivars tested in sixty-six environments in fourth and fifth IWWPN, 1972-1973. 63

17. The pertinent variance components obtained from analyzing two-year data for separate regions and for clusters formed from individual regions. 77

A1. Analysis of variance for individual regions and all the regions combined. 92

LIST OF FIGURES

Figure		Page
1.	Geographical distribution of IWWPN test locations.	24
2.	Diagram of relationships among locations in the United States region of the current zoning system.	50
3.	Diagram of relationships (dendrogram) among locations in Northern Europe.	52
4.	Diagram of relationships (dendrogram) among locations in Southern Europe.	54
5.	Diagram of relationships among 18 locations from combining two regions--N. Europe and S. Europe.	55
6.	Diagram of relationships (dendrogram) among locations in two regions--Near East and Far East.	57
7.	Diagram of relationships (dendrogram) among locations in all five regions.	58
8.	Regression of yield on environmental indices for three cultivars, Maris Nimrod, Dacia and Kirac 66.	66
9.	Map showing distribution of locations in the United States	71
10.	Map showing locations in Northern and Southern Europe boundary between Northern and Southern Europe.	72
11.	Map showing locations in Near and Far East. . .	75
12.	Yield fluctuations across a set of locations shown to be similar by cluster analysis. . . .	80
13.	Yield fluctuations of four varieties across a randomly selected set of locations.	81

Figure Page

A1. Two-dimensional diagram showing the distance between two hypothetical species with reference to two characters, X and Y. 95

A2. Two-dimensional projection of the distance between two species with reference to three characters, X, Y, and Z. 96

INTRODUCTION

Testing of new lines and varieties in areas of intended production is a standard procedure in every breeding program. The term "Regional Variety Trial" is very familiar to every plant breeder. Regional tests furnish information on the performance of the genotypes in different environments and on the interaction of genotypes with environment. Genotype x environment interaction is the failure of the differences between genotypes to be the same from one environment to another. From such trials the plant breeder can establish areas of adaptation for particular genotypes and recommend for production those cultivars that prove adapted to an environment; that is, those that perform very well in a particular environment.

International testing is a relatively recent phenomenon. Most international agriculture research centers conduct international testing nurseries of some kind. International testing, the testing of lines and varieties in an international array of environments, is designed to:

(a) Identify high yielding and disease and insect resistant varieties with broad adaptability

(b) Serve as a channel for the exchange of germplasm for other national programs

(c) Disseminate germplasm.

The International Winter Wheat Performance Nursery (IWWPN) is one such program. It was begun in 1969 by the University of Nebraska-Lincoln under contract with the Agency For International Development, U. S. Department of State. Varieties included for testing in this nursery are submitted by participating countries. The IWWPN has helped identify superior winter wheat genotypes that are being used in many wheat breeding programs for improvement of varieties, productivity and nutritional quality.[1]

The magnitude of genotype x environment interaction gives an indication of the heterogeneity of the environments under which the genotypes are tested. Testing in similar environments would give a small genotype x environment (GE) interaction component of variance. Any method that would permit classification of locations with similar patterns of varietal performance averaged over yearly fluctuations into groups would reduce GE interaction within that group of locations. Horner and Frey (1957) described procedures to restrict the diversity of test environments in order to reduce the magnitude of the GE interaction. They, also, determined whether certain test sites contributed more than others to the GE mean square. Abou-El Fittouh (1967) used

---

[1]J. E. Stroike and V. A. Johnson, International Nurseries and Cooperation. Paper presented at the International Winter Wheat Conference, Zagreb, Yugoslavia, June 1975.

cluster analysis to subdivide the cotton growing areas of the United States into subregions or adaptive zones within which GE interaction was minimal.

In this study the fourth and fifth International Winter Wheat Performance Nursery data were used with the objective of classifying the IWWPN locations with similar patterns of varietal performance into adaptive zones by applying the method of cluster analysis. Cluster analysis can be defined as a method for subdividing a population into a series of homogeneous groups based on some measure of similarity. Such a classification offers a possibility of reducing the number of test sites for the IWWPN. The nursery seed could be sent for testing to key areas that are representative of other locations within the zone. If there is good cooperation on the scientific level among countries within a zone of adaptation, the best performing varieties would still circulate among scientists for inclusion in their national programs.

Twenty-five cultivars included in the fourth and fifth IWWPN, also, were studied for stability. The model proposed by Eberhart and Russell (1966) provided a way of estimating stability parameters for each cultivar. These parameters were used to study the pattern of cultivar performance across locations within regions formed by the method of cluster analysis, thereby providing a measure of

effectiveness of cluster analysis in grouping similar environments together.

LITERATURE REVIEW

Genotype x Environment Interaction

Genotype x environment interaction can be defined as the failure of the differences between genotypes to be the same from one environment to another. The existence of interactions between genotypes and environmental factors has long been recognized, long before the analysis of variance technique came into use. Fisher and Mackenzie (1923), in considering the manurial responses of different potato varieties concluded that the yields of different varieties under different manurial treatments were better fitted by a product formula than by a sum formula.

GE interactions act to confound comparisons among lines in most field situations. Their effects in breeding and genetic studies have caused the plant breeder problems in evaluating differences among improved varieties and hybrids. Haldane (1946) recognized GE as one of the central and complex problems of genetics. He stated that although several genotypes may maintain their ranks when exposed to a series of environments they may respond to this change in the same direction or in opposite directions. The same types of responses may be found when the genotypes change their ranks from one environment to another. He developed a formula to compute the number of possible types of

interactions between m number of genotypes and n number of environments. He used interaction not in the statistical sense but rather to refer to any change in varietal performance resulting from environmental differences. He recommended testing genotypes over a wide range of environmental conditions which the selected material is expected to encounter. Those genotypes which show relatively low performance should not be destroyed because they may be needed in one or more of the following cases:

1. Emergence of new diseases
2. Changes in agricultural practices
3. Introducing the crop to new areas of production.

From genetic studies Mather and Morley Jones (1958) concluded that the growing mass of evidence for interaction between genotypes and environments lends urgency to the need for its close consideration. They explained how GE interactions influence variances and covariances used for measuring variation in biometrical genetical models. When interactions exist the measures of genetic effects apply only to the range of environments studied and conversely.

The significance of genotype x environment interactions and the role they play in most problems of quantitative genetics and plant breeding has also been realized by other workers. Comstock and Moll (1963) discussed GE interactions, how serious they are, and whether or not their

explanations should be sought. In their lengthy paper they aimed at: (a) clarifying the ways in which GE interactions are involved in problems faced by quantitative geneticists and plant breeders, (b) demonstrating how logical considerations lead to useful operational decisions which in turn lead to constructive simplification in breeding programs. They expressed GE interaction as:

$$(xy)_{ij} = P_{ij} - y_i - x_j - \mu$$

where $(xy)_{ij}$ = the effect of interaction between the i^{th} genotype and the j^{th} environment.

P_{ij} = the phenotypic value of the i^{th} genotype,

y_i = the effect of the i^{th} genotype,

x_j = the effect of the j^{th} environment,

μ = the general mean of the population.

The environment is made up of many factors whose differences and interplay result in the complex variation of the environment itself. These variable factors may be of many different kinds and we cannot expect all the genes affecting the manifestation of a character to respond equally to all their changes. Some genes will react chiefly to changes in certain of the factors and other genes will react more to changes in other factors, though of course, expression of any gene may be affected in its own particular way by changes in more than one factor. Allard and Bradshaw (1964) categorized environmental factors

as follows: (1) Predictable. These are those environmental factors which include all permanent characters of the environment such as general features of the climate and soil types as well as those characteristics of the environment which fluctuate in a systematic manner such as day length. Also included in this category are those aspects of the environment that are determined by man and can therefore be fixed more or less at will, such as planting date, sowing density, methods of harvest and other agronomic practices. (2) Unpredictable environmental factors. These include fluctuations in weather such as rainfall and temperature as well as disease outbreaks and insect infestation.

Predictable variation can be easily identified and taken care of in the breeding program; it's the unpredictable variations which cause the greater problems.

Method of Studying GE Interaction

The study of GE interaction has been approached in two different ways. One is the estimation of components of variance and the other is the estimation of stability parameters. The components of variance technique was first proposed by Sprague and Federer (1951). These workers showed how variance components could be used to separate out the effects of genotypes, environments and their expectations on the random model. Thus, in terms of a mathematical model, the yield, y_{ijk}, of the k^{th} replicate

of the i^{th} genotype in the j^{th} environment is regarded as made up of a general mean μ, a genotypic effect d_i, an environmental effect ε_j, an interaction effect g_{ij} and a random error e_{ijk}, i.e.

$$y_{ijk} = \mu + d_i + \varepsilon_j + g_{ij} + e_{ijk}$$

where

$$\sum_i d_i = \sum_j \varepsilon_j = \sum_{ij} g_{ij} = o$$

The component of variance technique as proposed by Sprague and Federer (1951) has been used to study genotype x environment interactions by researchers in oats (Horner and Frey 1957), in potatoes (Plaisted and Peterson 1959), in cotton (Miller et al. 1959, 1962), in barley (Rasmusson and Lambert 1961), in barley, oats, hard red spring and durum wheats (Baker, 1968), in wheat, barley and oats (Liang et al., 1966), in corn (Rojas and Sprague, 1952, Akposoe, 1967), in fall rye (Kaltsikes 1970), and tobacco (Gupton et al., 1974). Baihaki, Stucker and Lambert (1976) did a study to determine the relationship of GE interaction to yield level (high, medium or low) in a preliminary yield test of soybean lines. GE interactions were estimated within each group; that is, the three yield groups: high, medium and low. They found that about 50% of the total GE for yield was contributed by the low yielding group, 25% by medium yielding group and 25% by the high yielding group.

Other workers have used regression methods in studying GE interactions. Finlay and Wilkinson (1963) used a regression technique for an analysis of adaptation of 277 barley varieties in trials across Australian environments. For each variety a linear regression of yield on the mean yield of all varieties in each environment was computed to measure variety adaptation. Variety yields were measured on a logarithmic scale and this induced a high degree of linearity. The stablity parameter used was the regression coefficient. A regression coefficient of one indicated average stability. Regression coefficients greater or less than one indicated the varieties had less than or greater than average stability. A regression coefficient of zero would indicate absolute stability, in which case the variety would yield the same no matter what the environmental conditions. It was concluded that the regression method was particularly effective in emphasizing the actual trend of varietal yield responses to a range of natural environments.

It is interesting to note that a similar technique yielding similar results was reported by Yates and Cochran (1938) but was apparently not further developed or used until Finlay and Wilkinson (1963) rediscovered it.

Eberhart and Russel (1966) further developed the regression method to include the estimation of a second stability parameter--the variance due to deviations from regression. When the linear regressions fail to remove a

large proportion of the GE interaction, the deviations from regression parameter can be used to evaluate yield stability. They added together the sum of squares for environments and GE interactions and repartitioned them into a linear component between environments with one degree of freedom, a linear component of the GE interaction with (t-1) degrees of freedom and deviations from regression. The deviations being found separately for each of the t genotypes with (e-2) degrees of freedom. They proposed the following linear model for the study of GE interactions and stability:

$$Y_{ij} = \mu_i + \beta_i I_j + \delta_{ij}$$

where

Y_{ij} = the variety mean of the i^{th} variety at the j^{th} environment,

μ_i = the i^{th} variety mean over all environments,

β_i = the regression coefficient that measures the response of the i^{th} variety to varying environments,

δ_i = the deviation from regression of the i^{th} variety at the j^{th} environment,

I = the environmental index.

This model defines stability parameters that may be used to describe the performance of a variety over a series of environments. A stable variety was defined as one with

regression coefficient of one (b=1.0) and with deviations from regression as small as possible (s^2_d=0). What this means is that a desirable variety would have a nonsignificant deviation mean square as tested by the pooled error mean square, its regression coefficient would be close to unity and its varietal response would exceed the average yield of all genotypes. In other types of acceptable responses, the regression coefficients may be greater than or less than unity (b>1.0 or b<1.0) but the mean yield must be well above average in such instances. The authors note that an independent index obtained from environmental factors (rainfall, temperature, etc.) would be desirable; however present knowledge does not permit the computation of such an index.

Breese (1969) applied the regression technique to yield data of five genotypes of cocksfoot, Dactylis glomerata grown in two environments over two years. The grass genotypes were subjected to several different harvest schedules during the study. The regression technique transformed a complex tangle of GE interactions into an orderly series of linear and hence predictable responses. Breese summarized several experiments that were analyzed using the regression approach and concludes that linear relationships have held up for many different crop varieties grown under a wide range of environments. The use of the

organism itself to determine the quality of the environment was justified by the fact that the phenotype is a product of both the genotype and environment, and it is just as appropriate to numerically grade an environment according to the mean expression of a range of genotypes as it is to quantify a genotype by its average expression over a range of environments. Breese also concluded that a stability concept must include a measure of the deviations from regression as proposed by Eberhart and Russell (1966).

Multivariate analysis techniques have, also, been applied in the analysis of GE interactions and genotypic adaptation. Principal component analysis was used by Suzuki (1968) in studying strain adaptibility and Perkins (1972) compared this technique with linear regression for Nicotiana rustica data. Freeman and Dowker (1973) analyzed carrot variety performance with a two way principal component analysis. The objective of principal component method is to produce of the original variable a linearly transformed set of hypothetical variables called principal components. The new variables are mutually independent and can therefore be considered separately. They have a decreasing order of variability so that only the first few components may be needed to condense the information contained in the original variables. Details of this method were described by Seal (1964).

Chuang-Sheng and Thompson (1975) concluded that the relatively new principal component approach in the study of GE interaction is a powerful tool for analysis of interactions, especially in situations where models are not easily defined. However, one disadvantage of this approach is that the components may be mathematical artifacts without any obvious direct relationship to environmental conditions. If this is true then unless these artifacts can be replaced by highly correlated environmental factors, the method is useful only for identifying genotypes with similar interactions but cannot provide information about relationships of genotypes to environments.

Ways to Reduce GE Interaction

By testing genotypes in different locations the plant breeder may select stable varieties, that is, those that show low interaction with environment but give consistently high yields. Allard and Bradshaw (1964) explained that a variety can achieve stability in two ways: (1) It can be made up of a family of genotypes each adapted to somewhat different range of environments. They called this stability population buffering. (2) If the individuals themselves are well buffered such that each member of the population is adapted to a range of environments this would be individual buffering.

Jensen (1952) proposed the use of a multiline variety of oats rather than a pure line in order to have higher and more stable production, wider adaptation to environment and greater resistance to disease. Allard (1961), working on lima beans, recommended mixture or rational blends of pure lines chosen for stability in yield, uniformity of appearance and quality. He suggested that the yields of such a well planned blend could be even greater than that of the best adapted variety. Work at the International Maize and Wheat Improvement Center (CIMMYT)[2] has shown multilines to be more stable in their yield and disease reaction and also to perform better in an international array of environments than standard varieties.

Zoning or grouping sites can be used to reduce GE interaction within the zone. Horner and Frey (1957) examined environmental aspects of GE interactions and adaptation. In a five year experiment, yield data from oat varieties grown at nine locations in Iowa were examined from the standpoint of subdividing the state into subareas which minimized variety x location interaction within the sub-areas. Optimum patterns of divisions into two, three, four and five sub-regions resulted, respectively in a 11, 21, 30 and 40% decrease in the average variety x location interaction

---

[2]CIMMYT review 1976.

component of variance within sections compared to the same statistic for the state as a whole. Liang et al. (1966) estimated variety x environment interaction components of variance in yield tests conducted for three years involving 10 wheat varieties at 13 locations, 4 barley varieties at 10 locations and 5 oat varieties at 5 locations. They emphasized that the significant variety x location interaction for wheat and barley indicated that Kansas should be divided into sub-areas for varietal testing. By grouping appropriate locations together, the magnitude of the variety x location component of variance was reduced.

Pattern analysis is a general term encompassing the use of both cluster analysis and ordination in examining data structure. Cluster analysis, which is a method for identifying natural arrangement of objects into homogeneous groups, was used by Chuang-Sheng Lin and Thompson (1975) to delineate groups of genotypic performance types, based on differences among linear regression coefficients. Abou -El-Fittouh and Rawling (1969) used cluster analysis to classify locations for cotton variety tests in United States. As a result of this classification of locations into fairly homogeneous regions, they could reduce the within-region genotype x location interactions in regional cotton variety tests. They then suggested some modification in the recognized zones of adaptation for cotton.

Classificatory methods found their initial application in the field of ecology and numerical taxonomy (Goodall 1953; Sokal and Sneath 1963). Web et al. (1967a, 1967b) considered pattern analysis to be useful in elucidating the structure of tropical rain-forests, while Clifford and Goodall (1967) and Clifford et al. (1969) utilized numerical taxonomic procedures in the classification of Poacea. Mungomery et al. (1974) applied pattern analysis procedure to analyze adaptation of 58 soybean cultivars across environments in south-eastern Queensland (Australia). They managed to classify the 58 cultivars into 10 groups on the basis of yield performance and also into groups depending on protein content. In their classification scheme, the mean performance of each line in each of the p environments was measured. It was considered that the p environments defined a p-dimensional space. Each cultivar tested was regarded as a point in this space, the coordinates of which were its performance in each of the p environments. The mean performances of each line in the p environments were thus regarded as p attributes and were used to classify the cultivars into groups. Each group contained cultivars which were relatively close to each other but more distant from members of other groups in the p dimensional space. The coefficients of resemblance used were inter-individual unstandardized squared Euclidean distances.

Hussaini, Goodman and Timothy (1977) used principal component analysis to group a world collection of 640 entries of cultivated finger millet (Eleusine coracana L.) from Uganda, Ethiopia, Pakistan, Sikkim and 16 states of India. They identified 12 broad groups into which these finger millet cultivars fell. Indian material showed clinal variation with Southern and Eastern Indian samples forming the extremes. The pattern of variation in African material was distinctly different from that found in India. Ugandan material exhibited distinct separation and isolation from the rest of the material studied.

Byth, Eisemann and De Lacy (1976) applied pattern analysis and numerical classification as proposed by Mungomery et al. (1974) on 49 cultivars of spring wheat and 63 locations to define separately groups of cultivars and groups of environments based on similarity in yield performance. Cultivar classification identified groups of cultivars which differed meaningfully in environmental response, and the hierarchy reflected the degree of similarity of cultivar group response. Classification of the 63 environments produced groups of environments which differed in their average level of performance or in the pattern of response they elicited from the 49 cultivars, or both. Within groups environments were relatively consistent in average yield and cultivar response.

Pattern analysis was used by Shorter et al. (1977) to characterise environmental contributions to differences among soybean cultivars in performance and response across environments.

MATERIALS AND METHODS

Materials

Twenty-four of the thirty hexaploid winter wheats, Triticum aestivum L. em Thell., included in the fourth and fifth IWWPN were used in this study. Cultivars tested in the two sets of nurseries were provided by some of the participating countries. The 24 varieties and their origins are given in Table 1.

Data from 33 locations for the fourth and fifth IWWPN were used to classify the locations into adaptive zones. The locations are listed in Table 2 and the geographical distribution of the locations are shown in Fig. 1.

Methods

The 24 cultivars were grown in a randomized complete block design with four replications in each of 30 of the 33 locations in each of the two years. Only three replications were used for location 31 in 1972 and locations 12 and 30 in 1973. Row length was variable since each cooperator was encouraged to adjust row length and spacing to achieve a seeding rate most compatible with local practices. Yield was obtained from center rows and it was expressed in quintals per hectare. Observations taken were: days to ripening, protein content, lodging, yield, flowering date, winter survival and test weight. However, only yield has been used in this study.

Table 1. Varieties entered in the fourth and fifth International Winter Wheat Performance Nursery. 1972-1973.

Cultivar†	Origin
Hokuei	Japan
Bezostaia 1	USSR
Probstdorfer Extrem	Austria
Blueboy	USA
TX 62A4793-7	USA
Atlas 66	USA
NB 68513 (C.I. 15074)	USA
Jyva	Finland
Centurk (NB 66425)	USA
Vakka	Finland
Starke	Sweden
Sava (NS 611)	Yugoslavia
Strampelli	Italy
Lerma Rojo 64	Mexico
Backa	Yugoslavia
Clarion	Netherlands
Victor I	Italy
Marimp 3	Italy
Dacia	Romania
Golden Valley Zg 5994/66	Yugoslavia
Maris Nimrod	England
Zenith	Switzerland
Roussalka	Bulgaria
Caribo	W. Germany
Diplomat	W. Germany
Kirac 66	Turkey
Lilifen	Chile
Lancota (NB 701132)	USA
Carifen 12	Chile
Moldova	Romania

†Varieties Strampelli, Victor 1, Marimp 3, Lilifen, Golden Valley, and Lerma Rojo 64 were not included in this study.

Table 2. Locations for the fourth and fifth IWWPN used in the study.

Location No	Location	Country	Latitude	Longitude	Altitude
4	Wageningen	Netherlands	N51° 28'	E05° 38'	7 m
6	Milano	Italy	N45° 30'	E09° 30'	68 m
7	Novi Sad	Yugoslavia	N45° 05'	E19° 08'	84 m
8	Kabul	Afghanistan	N34° 33'	E69° 12'	1803 m
9	Suwon	Korea	N37° 16'	E126° 59'	37 m
10	Karaj	Iran	N35° 48'	E50° 58'	1300 m
12	Eskisehir	Turkey	N36° 45'	E30° 95'	789 m
13	Ankara	Turkey	N39° 40'	E32° 40'	850 m
14	Fundulea	Romania	N44° 03'	E24° 10'	66 m
15	Sulaimaniya	Iraq	N35° 05'	E46° 05'	700 m
16	Svalof	Sweden	N55° 35'	E13° 06'	50 m
17	North Carolina (Raleigh)	U.S.A.	N35° 42'	W80° 37'	825 m
18	Oklahoma (Stillwater)	U.S.A.	N36° 06'	W97° 04'	270 m
21	Rieti	Italy	N42° 24'	E12° 52'	402 m
24	New York (Ithaca)	U.S.A.	N42° 05'	W76° 05'	366 m
25	Washington (Pullman)	U.S.A.	N46° 42'	W117° 08'	777 m
27	Zagreb	Yugoslavia	N49° 49'	E15° 59'	177 m
28	Monsheim	West Germany	N49° 35'	E08° 20'	160 m
29	Weihenstephan	West Germany	N48° 24'	E11° 44'	467 m
30	Vienna	Austria	N48° 12'	E16° 45'	147 m
31	Cambridge	England	N52° 10'	E00° 08'	20 m
32	Zurich	Switzerland	N47° 39'	E08° 32'	445 m
33	Martonvasar	Hungary	N47° 21'	E18° 49'	150 m
37	Tolbukhin	Bulgaria	N43° 40'	E28° 10'	236 m
38	Colorado (Fort Collins)	U.S.A.	N40° 35'	W105° 10'	1475 m

Table 2. (Concluded)

Location No	Location	Country	Latitude	Longitude	Altitude
43	Szeged	Hungary	N46° 10'	E20° 00'	80 m
44	Malé Ripnany	Czechoslovakia	N40° 29'	E17° 39'	172 m
45	Sedlec	Czechoslovakia	N50° 14'	E14° 30'	300 m
47	Toluca	Mexico	N19° 16'	W99° 51'	2640 m
48	Krasnodar	USSR	N45° 00'	E38° 55'	38 m
49	Morioka Iwate	Japan	N39° 45'	E141° 08'	167 m
53	Oregon (Corvallis)	U.S.A.	N44° 32'	W123° 15'	70 m
54	Hamadan	Iran	N35° 12'	E48° 43'	1644 m

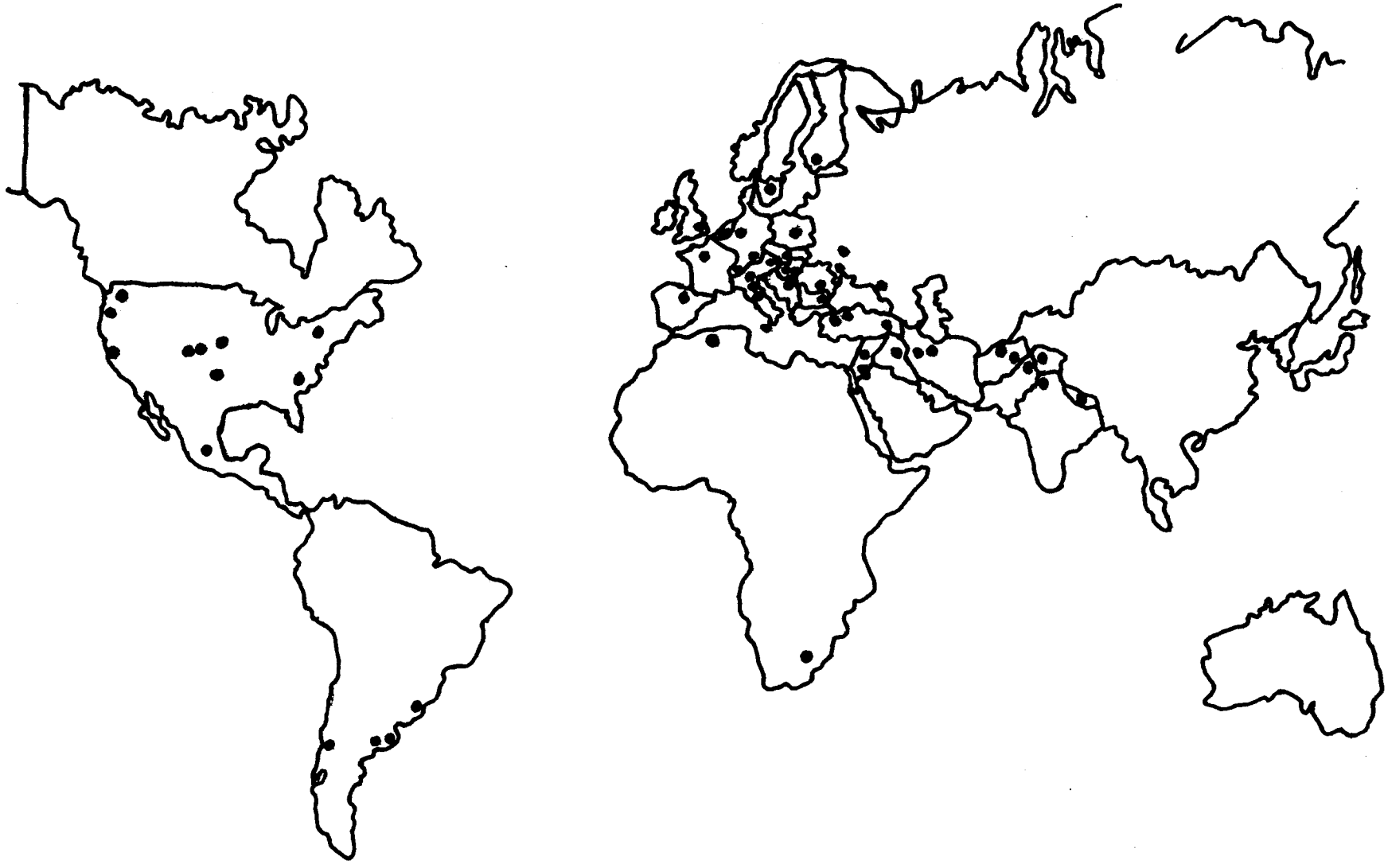


Figure 1. Geographical distribution of IWWPN test locations.

Cluster analysis

Cluster analysis, which can be defined as a technique for dividing a population into a series of homogeneous groups based on some measure of similarity, was used in this study to group similar environments (locations). In order to perform a clustering process a numerical measure of similarity among locations had to be calculated. This involved computing genotype x environment interaction effects for individual genotypes and using them to compute interindividual distance coefficients. It was this matrix of distance coefficients which was used in actual cluster analysis.

A computer was employed in computing GE effects for individual cultivars using the following formulas:

$$(g\ell)_{ij} = P_{ij} - G_i - L_j + \mu$$

Where

$(g\ell)_{ij}$ = the interaction effect of the i^{th} genotype in the j^{th} environment,

P_{ij} = mean yield of the i^{th} genotype in the j^{th} environment,

$G_i = \frac{1}{\ell_j} \Sigma P_{ij}$; the value of i^{th} genotype,

$L_j = \frac{1}{n_i} \Sigma P_{ij}$; the value of the j^{th} location,

$\mu = \frac{1}{mnij} \Sigma\Sigma P_{ij}$; the grand mean of the population of genotypes over all locations

ℓ = number of locations,

g = number of genotypes used in the study.

GE interaction effects were used to compute inter-individual distance coefficients between all pairwise comparisons of locations. Distance as a measure of similarity has been used extensively in taxonomy by Sokal and Michener (1960). Abou-El-Fittouh (1969) was first to apply it to problems of cultivar adaptation. In taxonomy where it has been widely used, the distance coefficient determines similarity between two taxa as a function of their distance in an n-dimensional space whose coordinates are the characters. (For details on coefficients, see appendix.) In this study similarities between pairs of locations were determined by calculating the distance between the two locations in a g-dimensional space. It was considered that the g (g=24) cultivars defined a g-dimensional space. Distance coefficients were all caclulated by computer using the formula:

$$\delta_{1,2} = \sum_{i}^{g} [(g\ell)_{i1} - (g\ell)_{i2}]^2 \}^{\frac{1}{2}}$$

where

$\delta_{1,2}$ = *distance coefficient between location* 1 and 2,

g = total number of cultivars used in the study,

$(g\ell)_{i1}$ = genotype x environment interaction effect of the i^{th} cultivar in first location,

$(g\ell)_{i2}$ = genotype x environment interaction effect of the i^{th} cultivar in second location.

A computer program BMDPIM-cluster analysis on variables was employed in the classification of locations into homogeneous groups. This program uses average linkage method based on Sneath and Sokal (1973). The program employs a hierarchical, agglomerative, clustering strategy (Williams 1971) by which the most similar individuals (locations) and groups are successively fused until the complete population has been incorporated into a single group. Groups are presented in the form of a tree diagram called a dendrogram.

Cluster analyses were performed for: (1) each current region. The currently recognized regions are presented in Table 3. This analysis was aimed at studying the level of similarity among locations in the region, that is, to find out whether these locations are similar enough to have been grouped into a region. (2) Adjacent regions combined. This analysis was for the purpose of studying inter-regional similarity of locations. (3) For all locations combined. This classification was for the purpose of identifying pockets of locations with similar patterns of varietal performance and also for identifying locations which are similar and yet quite distant apart geographically.

Table 3. Table showing breakdown of locations into regions.

Region	Location †
Northern Europe	4
	16
	28
	29
	30
	31
	32
	44
	45
Southern Europe	6
	7
	14
	21
	27
	33
	37
	43
	48
United States	17
	18
	24
	25
	38
	47
	53
Near East	8
	10
	12
	13
	15
	54
Far East	49
	9

†Locations included in the regions are only those which were included in the study, that is, for which the fourth and fifth IWWPN data were available.

Statistical Analysis of Data

An IBM 360 computer and the statistical analysis system (SAS) programs as designed and implemented by Barr and Goodnight were employed for data analysis. Experimental data were analyzed according to the method for randomized complete block design with experiments repeated in different locations and over years. Statistical procedures and interpretation were based on standard methods outlined by Cochran and Cox (1957) and Steel and Torrie (1960). Where data were unbalanced because the experiments being analysed had unequal numbers of replications, the analysis was computed from unweighted treatment means. This was done to facilitate computation of variance components. Where data were balanced, individual plot yields were used in analyses and variance components were estimated in the usual manner.

The linear additive model for individual year analyses of variance may be written as follows:

$$P_{ijk} = \mu + g_i + \ell_j + (g\ell)_{ij} + r_{jk} + \varepsilon_{ijk}$$

where

P_{ijk} = The observed value of the plot containing the i^{th} genotype in the k^{th} replication in the j^{th} location,

μ = the grand mean,

g_i = the effect of the i^{th} genotype,

ℓ_j = the effect of the j^{th} location,

$(g\ell)_{ij}$ = the interaction effect of the i^{th} genotype with the j^{th} location,

r_{jk} = the effect of the k^{th} replication nested in the j^{th} location,

ϵ_{ijk} = the random error associated with the observation on the plot containing the i^{th} genotype in the k^{th} replication of the j^{th} location.

Form of the individual year analyses of variance are given in tables 4 and 5.

Table 4. Form of the individual year analysis of variance for Southern Europe, United States, and Far East. Data from these regions were balanced.

Source of variation	d.f.	Mean square	
		Observed	Expected
Locations (L)	$(\ell-1)$		
Replication/L	$\ell(r-1)$		
Genotypes (G)	$(g-1)$	MS_g	$\sigma^2_\epsilon + r\sigma^2_{g\ell} + r\ell\sigma^2_g$
G x L	$(g-1)(\ell-1)$	$MS_{g\ell}$	$\sigma^2_\epsilon + r\sigma^2_{g\ell}$
Error	$\ell(g-1)(r-1)$	MS_ϵ	σ^2_ϵ

Where g = the number of genotypes included in the analysis,

ℓ = the number of environments (test locations),

r = the number of replications within each location,

σ^2_ϵ = the error variance due to environmental differences among plots within replications at

individual locations.

$\sigma_{g\ell}$ = the variance due to interaction of genotypes with environments,

σ_g^2 = the variance due to genetic differences among genotypes.

Table 5. Form of the individual year analysis of variance for all IWWPN locations combined and separatly for regions; Northern Europe, and Near East. Data were unbalanced because of unequal number of replications.

Source of variation	d.f.	Mean square	
		Observed	Expected
Locations (L)	$(\ell-1)$		
Genotypes (G)	$(g-1)$	MS_g	$\frac{\sigma_\varepsilon^2}{n} + \sigma_{g\ell}^2 + \ell\sigma_g^2$
G x L	$(g-1)(\ell-1)$	$MS_{g\ell}$	$\frac{\sigma_\varepsilon^2}{n} + \sigma_{g\ell}^2$
Error	$\sum^{\ell}(g-1)(r-1)$	MS_ε	$\frac{\sigma_\varepsilon^2}{n}$

Where n is the harmonic mean of the number of replications and is given by the formula: $n = \frac{P}{\sum \frac{1}{r_i}}$

P = number of experiments,

r_i = number of replications in the i^{th} experiment,

g = number of genotypes included in the analysis,

ℓ = number of test locations,

σ_ε^2 = error variance,

$\sigma^2_{g\ell}$ = the variance due to interaction of genotypes with locations,

σ^2_g = the variance due to genetic differences among genotypes.

For combined over years analyses of variance the linear additive model may be written as follows:

$$P_{ijk\ell} = \mu + g_i + \ell_j + (g\ell)_{ij} + y_k + (gy)_{ik} + (\ell y)_{jk} + (g\ell y)_{ijk} + r_{jk\ell} + \varepsilon_{ijk\ell}$$

where

$P_{ijk\ell}$ = the measurement of yield of the i^{th} genotype grown in the ℓ^{th} replication at the j^{th} location in the k^{th} year,

μ = the grand mean of all genotypes,

g_i = effect of the i^{th} genotype,

ℓ_j = effect of the j^{th} location,

$(g\ell)_{ij}$ = interaction effect of the i^{th} genotype with the j^{th} location,

y_k = effect of the k^{th} year,

$(gy)_{ik}$ = interaction effect of the i^{th} genotype and the k^{th} year,

$(\ell y)_{jk}$ = interaction effect of the j^{th} location and the k^{th} year,

$(g\ell y)_{ijk}$ = the second order interaction effect of the i^{th} genotype, the j^{th} location and the k^{th} year,

$r_{jk\ell}$ = the effect of the ℓ^{th} replication in the k^{th} year at the j^{th} location,

$\varepsilon_{ijk\ell}$ = the random environmental effect associated with the plot containing the i^{th} genotype in the ℓ^{th} replication in the k^{th} year at the j^{th} location.

Form of the combined year analysis of variance is given in Tables 6 and 7.

Table 6. Form of combined year analysis of variance for United States, Southern Europe and the Far East. Data were balanced.

Source of variation	d.f.	Mean square	
		Observed	Expected
Locations (L)	$(\ell-1)$		
Years (Y)	$(y-1)$		
L x Y	$(\ell-1)(y-1)$		
Reps in L and Y	$\ell y(r-1)$		
Genotypes	$(g-1)$	MS_g	$\sigma^2_\varepsilon + r\sigma^2_{g\ell y} + ry\sigma^2_{g\ell} + r\ell\sigma^2_{gy} + r\ell y\sigma^2_g$
G x L	$(g-1)(\ell-1)$	$MS_{g\ell}$	$\sigma^2_\varepsilon + r\sigma^2_{g\ell y} + ry\sigma^2_{g\ell}$
G x Y	$(g-1)(y-1)$	MS_{gy}	$\sigma^2_\varepsilon + r\sigma^2_{g\ell y} + r\ell\sigma^2_{gy}$
G x L x Y	$(g-1)(\ell-1)(y-1)$	$MS_{g\ell y}$	$\sigma^2_\varepsilon + r\sigma^2_{g\ell y}$
Error	$\ell y(r-1)(g-1)$	MS_ε	σ^2_ε

Where σ^2_ε = the variance due to environmental differences among plots within replications.

$\sigma^2_{g\ell y}$ = the variance due to interaction of genotypes, locations and years.

σ^2_{gy} = the variance due to interaction of genotypes with years,

$\sigma^2_{g\ell}$ = the variance due to interaction of genotypes with test locations,

σ^2_g = the variance due to genetic differences among genotypes.

Table 7. Form of combined year analyses of variance for all the test locations combined and, separately, for regions, Northern Europe and Near East. Unbalanced data.

Source of variation	d.f.	Mean square	
		Observed	Expected
Locations (L)	$(\ell-1)$		
Years (Y)	$(y-1)$		
L x Y	$(\ell-1)(y-1)$		
Genotypes (G)	$(g-1)$	MS_g	$\frac{\sigma^2_\varepsilon}{n} + \sigma^2_{g\ell y} + y\sigma^2_{g\ell} + \ell\sigma^2_{gy} + \ell y\sigma^2_g$
G x L	$(g-1)(\ell-1)$	$MS_{g\ell}$	$\frac{\sigma^2_\varepsilon}{n} + \sigma^2_{g\ell y} + y\sigma^2_{g\ell}$
G x Y	$(g-1)(y-1)$	MS_{gy}	$\frac{\sigma^2_\varepsilon}{n} + \sigma^2_{g\ell y} + \ell\sigma^2_{gy}$
G x L x Y	$(g-1)(\ell-1)(y-1)$	$MS_{g\ell y}$	$\frac{\sigma^2_\varepsilon}{n} + \sigma^2_{g\ell y}$
Error	$\sum\sum^{\ell Y}(r-1)(g-1)$	MS_ε	$\frac{\sigma^2_\varepsilon}{n}$

Where n = harmonic mean number of replications,

g = number of genotypes,

y = number of years of testing,

ℓ = number of test locations,

r = number of replications per location,

σ^2_ε = error variance

$\sigma^2_{g\ell y}$ = the variance due to interaction of genotypes with locations and years,

σ^2_{gy} = variance due to interaction of genotypes with years,

σ^2_g = the variance due to the genetic differences among genotypes.

Tables 8 and 9 present a summary of the methods for estimating variance components.

Table 8. Estimation of variance components for individual year analysis of variance.

Variance component	Method of estimation		F.†
	Balanced data	Unbalanced data	
$\hat{\sigma}^2_g$	$\frac{MS_g - MS_{g\ell}}{r\ell}$	$\frac{MS_g - MS_{g\ell}}{\ell}$	$\frac{MS_g}{MS_{g\ell}}$
$\hat{\sigma}^2_{g\ell}$	$\frac{MS_{g\ell} - MS_\varepsilon}{r}$	$MS_{g\ell} - MS_{\varepsilon/n}$	$\frac{MS_{g\ell}}{MS_\varepsilon}$
$\hat{\sigma}^2_\varepsilon$	MS_ε	MS_ε	

†The F-test of significance to test the hypothesis that each component is equal to zero. The F-test that $\hat{\sigma}^2_{g\ell} = o$ for unbalanced data is obtained by the formula: $Ms_{g\ell} / (MS_{\varepsilon/n})$.

Table 9. Estimation of variance components for combined year analysis of variance.

Variance component	Method of estimation: Balanced data	Method of estimation: Unbalanced data	F.
$\hat{\sigma}^2_g$	$\frac{MS_g - MS_{g\ell} - MS_{gy} + MS_{g\ell y}}{r\ell y}$	$\frac{MS_g - MS_{g\ell} - MS_{gy} + MS_{g\ell y}}{\ell y}$	Complex F †
$\hat{\sigma}^2_{g\ell}$	$\frac{MS_{g\ell} - MS_{g\ell y}}{ry}$	$\frac{MS_{g\ell} - MS_{g\ell y}}{y}$	$\frac{MS_{g\ell}}{MS_{g\ell y}}$
$\hat{\sigma}^2_{gy}$	$\frac{MS_{gy} - MS_{g\ell y}}{r\ell}$	$\frac{MS_{gy} - MS_{g\ell y}}{\ell}$	$\frac{MS_{gy}}{MS_{g\ell y}}$
$\hat{\sigma}^2_{g\ell y}$	$\frac{MS_{g\ell y} - MS_\varepsilon}{r}$	$MS_{g\ell y} - MS_{\varepsilon/n}$	$\frac{MS_{g\ell y}}{MS_\varepsilon}$
$\hat{\sigma}^2_\varepsilon$	MS_ε	MS_ε	

† A complex F-test was done for σ^2_g. The F-test that $\hat{\sigma}^2_{g\ell} = o$ for unbalanced data is obtained by the formula: $MS_{g\ell}/(MS_{\varepsilon/n})$.

The expectations of the mean squares indicate the correct test term to use for each component when testing the components of variance for significance of difference from zero. In the case of σ_g^2 there exists no single mean square that can be used when a combined over years analysis of variance is performed. A complex F-test involving linear functions of mean squares was used.

$$Ho : \sigma_g^2 = 0 \qquad F = \frac{MS_g}{MS_{g\ell} + MS_{gy} - MS_{gy\ell}}$$

The degrees of freedom for the numerator is in the analysis of variance table. The degrees of freedom for the denominator was estimated by the formula:

$$d.f. = \frac{L^2}{\sum_i \frac{M_i^2}{f_i}}$$

Where L = Linear function of mean squares used,

M_i = the mean squares in the linear function,

f_i = the degrees of freedom for the i^{th} mean square.

In order to indicate the degree of precision with which components of variance are estimated, approximate standard errors for the variance components were calculated as shown in Tables 10 and 11.

Analysis of variety stability was based on the model proposed by Eberhart and Russell (1966). By this model, the regression of each variety in an experiment on an

Table 10. Methods of estimating standard errors of variance components for individual year analysis of variance.

Variance component	SE †	Method of estimation: Balanced data	Method of estimation: Unbalanced data
$\hat{\sigma}^2_{\varepsilon}$	SE $\hat{\sigma}^2_{\varepsilon}$	$= \sqrt{\frac{2MS^2_{\varepsilon}}{f_{\varepsilon} + 2}}$	$\sqrt{\left[\frac{2MS^2_{\varepsilon}}{f_{\varepsilon} + 2}\right]}$
$\hat{\sigma}^2_{g\ell}$	SE $\hat{\sigma}^2_{g\ell}$	$= \sqrt{\left(\frac{1}{r}\right)^2 \left[\frac{2MS^2_{g\ell}}{f_{g\ell} + 2} + \frac{2MS^2_{\varepsilon}}{f_{\varepsilon} + 2}\right]}$	$\sqrt{\left[\frac{2MS^2_{g\ell}}{f_{g\ell} + 2} + \frac{2MS^2_{\varepsilon}}{f_{\varepsilon} + 2}\right]}$
$\hat{\sigma}^2_{g}$	SE $\hat{\sigma}^2_{g}$	$= \sqrt{\left(\frac{1}{r\ell}\right)^2 \left[\frac{2MS^2_{g}}{f_{g} + 2} + \frac{2MS^2_{g\ell}}{f_{g\ell} + 2}\right]}$	$\sqrt{\left(\frac{1}{\ell}\right)^2 \left[\frac{2MS^2_{g}}{f_{g} + 2} + \frac{2MS^2_{g\ell}}{f_{g\ell} + 2}\right]}$

† Standard error of variance components.

Table 11. Method of estimating standard errors of variance components for combined year analysis of variance

Variance component	SE	Method of estimation: Balanced data	Method of estimation: Unbalanced data
$\hat{\sigma}^2_{\varepsilon}$	SE $\hat{\sigma}^2_{\varepsilon}$	$=\sqrt{\frac{2MS^2_{\varepsilon}}{f_{\varepsilon}+2}}$	$\sqrt{\left[\frac{2MS^2_{\varepsilon}}{f_{\varepsilon}+2}\right]}$
$\hat{\sigma}^2_{g\ell y}$	SE $\hat{\sigma}^2_{g\ell y}$	$=\sqrt{\left(\frac{1}{r}\right)^2\left[\frac{2MS^2_{g\ell y}}{f_{g\ell y}+2}+\frac{2MS^2_{\varepsilon}}{f_{\varepsilon}+2}\right]}$	$\sqrt{\left[\frac{2MS^2_{g\ell y}}{f_{g\ell y}+2}+\frac{2MS^2_{\varepsilon}}{f_{\varepsilon}+2}\right]}$
$\hat{\sigma}^2_{gy}$	SE $\hat{\sigma}^2_{gy}$	$=\sqrt{\left(\frac{1}{r\ell}\right)^2\left[\frac{2MS^2_{gy}}{f_{gy}+2}+\frac{2MS^2_{g\ell y}}{f_{g\ell y}+2}\right]}$	$\sqrt{\left(\frac{1}{\ell}\right)^2\left[\frac{2MS^2_{gy}}{f_{gy}+2}+\frac{2MS^2_{g\ell y}}{f_{g\ell y}+2}\right]}$
$\hat{\sigma}^2_{g\ell}$	SE $\hat{\sigma}^2_{g\ell}$	$\sqrt{\left(\frac{1}{ry}\right)^2\left[\frac{2MS^2_{g\ell}}{f_{g\ell}+2}+\frac{2MS^2_{g\ell y}}{f_{g\ell y}+2}\right]}$	$\sqrt{\left(\frac{1}{y}\right)^2\left[\frac{2MS^2_{g\ell}}{f_{g\ell}+2}+\frac{2MS^2_{g\ell y}}{f_{g\ell y}+2}\right]}$

Table 11. (Concluded)

Variance component	SE	Method of estimation
		Balanced Data
$\hat{\sigma}^2_g$	$SE\,\hat{\sigma}^2_g$	$= \sqrt{\left[\frac{1}{r\ell y}\right]^2 \left[\frac{2MS_g^2}{f_g + 2} + \frac{2MS_{gy}^2}{f_{gy} + 2} + \frac{2MS_{g\ell}^2}{f_{g\ell} + 2} + \frac{2MS_{g\ell y}^2}{f_{g\ell y} + 2}\right]}$
		Unbalanced Data
$\hat{\sigma}^2_g$	$SE\,\hat{\sigma}^2_g$	$= \sqrt{\left[\frac{1}{\ell y}\right]^2 \left[\frac{2MS_g^2}{f_g + 2} \quad \frac{2MS_{gy}^2}{f_{gy} + 2} \quad \frac{2MS_{g\ell}^2}{f_{g\ell} + 2} \quad \frac{2MS_{g\ell y}^2}{f_{g\ell y} + 2}\right]}$

f is degrees of freedom for particular mean square.

environmental index and a function of the squared deviations from this regression provide estimates of two stability parameters for each variety. The parameters which define the stability of each genotype are the linear regression coefficient and the variance due to deviations from regression.

The model of Eberhart and Russell (1966) provides a means of partitioning the GE interaction of each variety into two parts: (1) the variation due to the response of the variety to varying environmental indices (sums of squares due to regression) and (2) the unexplainable deviations from the regression on the environmental index.

Varieties in the fourth and fifth IWWPN were analyzed for stability in 66 environments (varieties were grown in 33 locations for two years) using the following model:

$$Y_{ij} = \mu i + \beta_i I_j + \delta_{ij},$$

Where

Y_{ij} = the variety mean of the i^{th} variety at the j^{th} environment ($i = 1,2,\ldots\ldots, g$; $j = 1,2,\ldots\ldots, e$)

μ_i = the mean of the i^{th} variety over all environments

β_i = the regression coefficient that measures the response of the i^{th} variety to varying environments

I_j = the environmental index, obtained as the mean

of all genotypes at the jth environment minus the grand mean of environments,

δ_{ij} = the deviation from regression of the i^{th} genotype grown in the j^{th} environment.

The form of analysis of variance for stability is shown in Table 12.

The parameters provided by the stability analysis that are used to evaluate the genotypes are: b_i, the estimate of the linear regression coefficient for the regression of the performance of genotype i on the environmental index, $S^2 d_i = \left[\Sigma \hat{\sigma}^2_{ij}/(e-2) \right] - MS_\varepsilon/n$, which is the mean square deviation from regression for the i^{th} variety minus the pooled error. This measures the variance due to genotype x environment interaction with the random error subtracted out.

The hypothesis that there are no genetic differences among varieties for their regression on the environmental index, Ho: $\beta_1 = \beta_2 = ... \beta_g$ can be tested approximately by the F test, $F = MS_2/MS_d$. The significance of the differences among variety means, Ho: $\mu_1 = \mu_2 ..... \mu_g$ can be tested by the F test, $F = MS_1/MS_d$.

Table 12. Form of analysis of variance for genotype stability.

Source of variation	d.f.	s.s	M.S.
Total	eg-1	$\Sigma\Sigma_{ij} i_j^2$ - C.F.	
Genotypes (g)	g-1	$\frac{1}{e}\Sigma Y_{i.}^2$ - C.F.	MS_1
Environments (E) V x E	e-1 (g-1)(e-1) }g(e-1)	$\Sigma\Sigma_{ij} Y_{ij}^2$ - $\Sigma Y_i^2/e$	
E (linear)	1	$\frac{1}{g}(\Sigma_j Y_{.j} I_j)^2/\Sigma_j I_j^2$	
g x E (linear)	g-1	$\Sigma_i[(\Sigma_j Y_{ij} I_j)^2/\Sigma_j I_j^2$ - E (linear) S.S	MS_2
Pooled deviations	g(e-2)	$\Sigma\Sigma_{ij} \delta_{ij}^2$	MS_d
genotype 1	e-2	$\Sigma_j Y_{1j}^2$ - $\frac{(Y_{1.})^2}{e}$ - $(\Sigma_j Y_{1j} I_j)^2/\Sigma_j I_j^2$	MS_{d_1}
⋮	⋮	⋮	⋮
genotype g	e-2	$\Sigma_j Y_{gj}^2$ - $\frac{Y_{g.}^2}{e}$ - $(\Sigma_j Y_{gj} I_j)^2/\Sigma_j I_j$ = $\Sigma\delta_{gj}^2$	MS_{d_g}
Pooled error	Σ_e(r-1)(g-1)		MS_ε

RESULTS

The results of the analyses of variance, estimates of variance components and their standard errors and coefficients of variation involving twenty-four wheat cultivars included in the fourth and fifth IWWPN are presented in this section. In the next section results of environment classification by cluster analysis are presented and discussed. Stability parameters for the 25 cultivars are presented in the last section.

Estimates of Components of Variance from Analyses of Variance

The within year and combined year analyses of variance for all locations combined are presented in Table 13. The analyses show that both genotype and genotype x location interaction components were highly significantly greater than zero in 1972 and 1973. The genotype x location interaction variance component ($\hat{\sigma}^2_{g\ell}$) was biggest in magnitude in both years. The significant genotypic component of variance suggests that the varieties included in the fourth and fifth IWWPN contributed a sufficiently large variation to the total variation in the experiment which could be attributed to differences in the genetic constitution of the varieties. Genotype x location interaction components were greatest in magnitude in both

Table 13. Estimates of pertinent variance components from the analyses of variance of yield across all locations included in the study.

Years	Component estimated	Magnitude	S.E.	C.V.
1972	$\hat{\sigma}^2_g$	22.48**	6.70	
	$\hat{\sigma}^2_{g\ell}$	56.48**	3.29	
	$\hat{\sigma}^2_\epsilon$	26.55	0.19	13.62
1973	$\hat{\sigma}^2_g$	23.84**	7.17	
	$\hat{\sigma}^2_{g\ell}$	42.96**	2.58	
	$\hat{\sigma}^2_\epsilon$	25.63	0.19	12.72
1972-1974	$\hat{\sigma}^2_g$	22.28**	6.77	
	$\hat{\sigma}^2_{g\ell}$	23.32**	2.24	
	$\hat{\sigma}^2_{gy}$	0.87**	0.53	
	$\hat{\sigma}^2_{g\ell y}$	26.40**	1.72	
	$\hat{\sigma}^2_\epsilon$	26.09	0.14	13.16

** Means significance at the .01 level of probability.

years and were highly significant, which means that the cultivars failed to perform the same relative to each other in different locations. The analysis thus indicates that different genetic backgrounds exist among the cultivars used in the study and that they differ in their adaptation to different environments.

A combined analysis of variance for all the locations for the years 1972 and 1973 shows similar results, $\hat{\sigma}^2_g$, $\hat{\sigma}^2_{g\ell}$, $\hat{\sigma}^2_{gy}$ and $\hat{\sigma}^2_{g\ell y}$ were all highly significant. The genotype x location x year interaction component was the largest in magnitude, indicating that the cultivars not only performed differently relative to each other in different locations but also performed differently relative to each other in the two years. Genotype x year interaction component of variance was smallest in magnitude. The size of the standard errors relative to the component indicates that reasonably adequate precision was attained in estimating the components.

Results of analyses of variance for the currently defined regions are presented in Table 14. In 1972, genotypic variance component and genotype x location interaction component were both highly significant. Genotypic components varied nearly two-fold among the regions. In 1973, genotypic components were highly significant in all but one region--Far East in which the component was non-significant. The non-significant genotypic component in the Far East could have been due to chance. The standard error for this component in the Far East was larger than the component itself so that the degree of precision with which the component was estimated was not very good. Genotype x location interaction components were all highly significant.

Table 14. Estimates of pertinent variance components from the individual regional analyses of variance.

Years	Component estimated	Region									
		N. Europe		S. Europe		Near East		Far East		U.S.A.	
		Comp.	SE	Comp.	SE	Comp.	SE	Comp.	SE	Comp.	SE
1972	σ^2_g	47.86**	15.20	50.04**	15.58	23.49**	8.97	14.67*	9.05	24.89**	9.25
	$\sigma^2_{g\ell}$	47.63**	5.47	41.08**	4.69	40.24**	6.31	17.99**	8.05	43.63**	6.44
	σ^2_ϵ	19.49	0.29	16.54	0.94	31.47	2.18	40.63	4.86	39.94	2.56
C.V.		9.62		11.28		18.29		21.94		16.33	
1973	$\hat{\sigma}^2_g$	26.96**	8.61	47.81**	14.61	10.81**	4.49	18.06	27.07	26.49**	10.23
	$\hat{\sigma}^2_{g\ell}$	25.90**	3.21	25.94**	3.59	24.36**	3.88	58.59**	32.84	59.61**	7.92
	$\hat{\sigma}^2_\epsilon$	19.61	0.29	33.54	1.90	19.79	0.37	29.96	3.58	26.41	1.69
C.V.		9.51		13.19		14.89		12.11		15.68	
1972-1973	$\hat{\sigma}^2_g$	34.87**	11.06	49.46**	14.69	15.40**	6.07	14.40	14.27	19.30**	8.21
	$\hat{\sigma}^2_{g\ell}$	9.27**	3.14	8.86**	2.98	20.30**	4.05	28.94*	15.54	20.87**	5.36
	$\hat{\sigma}^2_{gy}$	2.54*	1.78	-0.53	0.89	1.75*	1.43	1.96	9.03	6.39**	3.44
	$\hat{\sigma}^2_{g\ell y}$	27.49**	3.38	24.66**	3.22	11.91**	2.44	34.34**	12.23	30.64**	4.67
	$\hat{\sigma}^2_\epsilon$	19.55	0.21	25.04	1.00	25.80	0.33	35.30	2.99	33.17	1.51
C.V.		9.56		12.52		16.77		16.00		16.12	

* Means significance at .05 level of probability; ** means significance at .01 level of probability.

in 1973 also. The magnitude of genotype x location component was greatest in the U.S.A.

In a two year (1972-1973) analysis, genotypic variance components were highly significant in all regions except the Far East where it was non-significant. However, the magnitude of this component in the Far East does not mean much because of poor precision of estimation as shown by the large standard error associated with it. Genotype x location interaction components varied nearly three-fold among regions. They were highly significant in all but the Far East where it was only significant at the .05 level of probability. There was one negative estimate of the genotype x year interaction component of variance. This component was largest in magnitude and highly significant in the U.S. indicating that the cultivars showed the most varied performance relative to each other over the two years in the United States. Genotype x year interaction components were significant in Northern Europe and Near East but non-significant in Southern Europe and Far East.

Genotype x location x year interaction components were highly significant in all regions. They varied nearly three-fold among the regions.

Genotype x location interaction components were highly significant in all regions. This means that locations within regions represented sufficiently different

environmental conditions to which the cultivars reacted differently relative to each other as expressed by their different yield performances in different locations. The results of cluster analysis, one of the methods by which the magnitude of the genotype x location interaction variance component can be reduced are presented in the next section.

Results of Classification

Two-year yield data of 24 cultivars included in the fourth and fifth IWWPN were used to compute genotype x location interaction effects for each of the cultivars. Distance coefficients as measures of resemblances between pairs of locations were then computed using the interaction effects. On the basis of distance coefficients, groups of locations with similar patterns of interaction were identified by cluster analysis. Results of cluster analysis are presented in form of a diagram of relationships among locations known as a dendrogram.

Figure 2 shows the relationships among locations in the United States which included location 47-Toluca, Mexico. The dendrogram shows the presence of two major groupings of locations. Locations 17, 18 and 38 form a cluster of locations possessing a similar pattern of interaction. In this cluster, locations 17 and 18 appear to have a closer affinity to each other. Location 17 is Raleigh, North

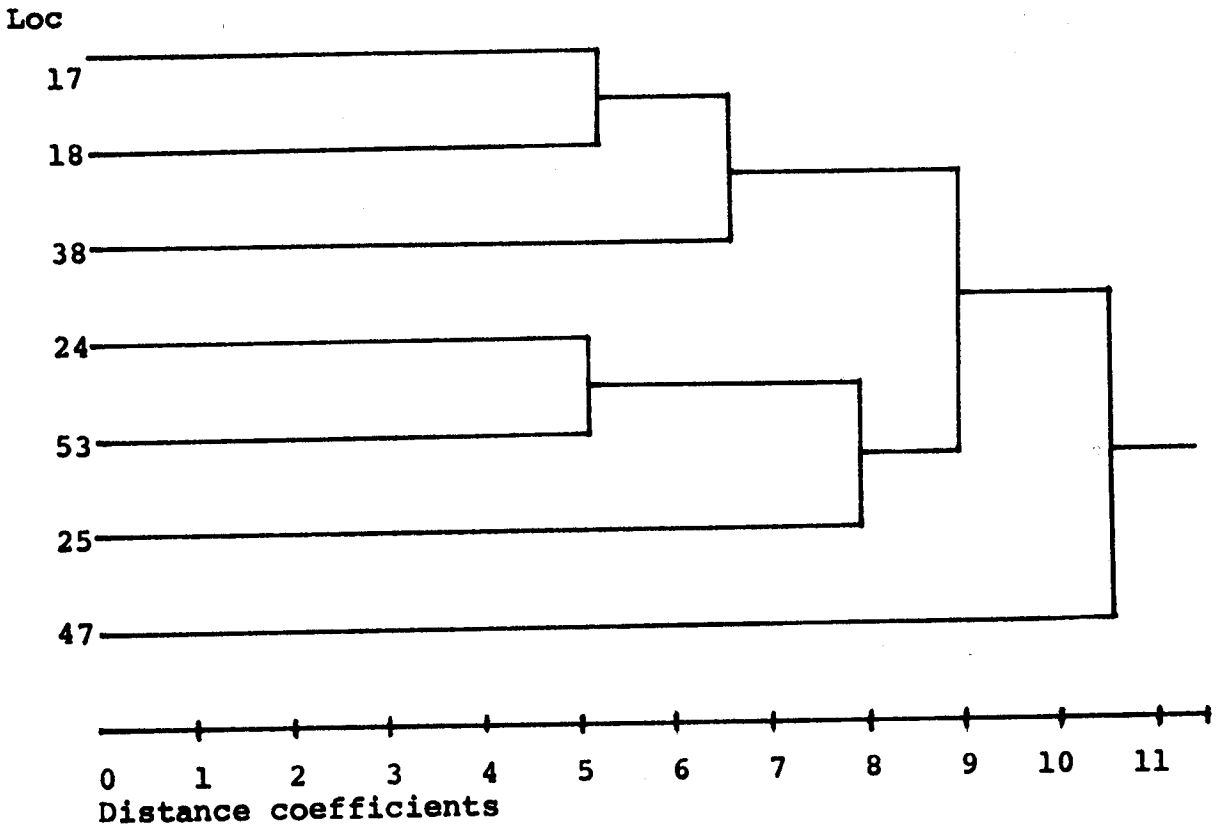


Figure 2. Diagram of relationships among locations in the United States region of the current zoning system.

Carolina and location 18 is Stillwater, Oklahoma. These locations are near the same latitude. Location 17 is on latitude N35° 42' while location 18 is on latitude N36° 06'. The other major cluster indicated in the dendrogram consists of locations 24, 53 and 25. Locations 24 and 53 show a considerable similarity to each other while these two have a weak resemblance to location 25. In fact, location 25 which is Pullman, Washington does not show an acceptable degree of similarity to 24 and 53 that it would be excluded as different from locations 24 and 53. Location 24 is Ithaca, New York and 53 is Corvallis, Oregon. These locations are also near the same latitude. Location 24 is on latitude N42° 05' and Corvallis is on latitude N44° 32'. This could explain their close similarity suggested in the dendrogram. Location 47 which is Toluca, Mexico shows very little similarity to any of the locations in the region. Location 47 is closer to the equator than the other locations and also it is at the highest altitude (2640 m) compared to other locations.

Figure 3 presents results of similarity analysis (cluster analysis) for Northern Europe. Two major clusters are indicated in the diagram of relationships. One cluster comprises locations 4, 16, 29, 31, 32 and 45. The other cluster is made up of locations 28, 44 and 30. The dendrogram shows the former cluster to possess some locations that show a high degree of affinity to each other. Location 4

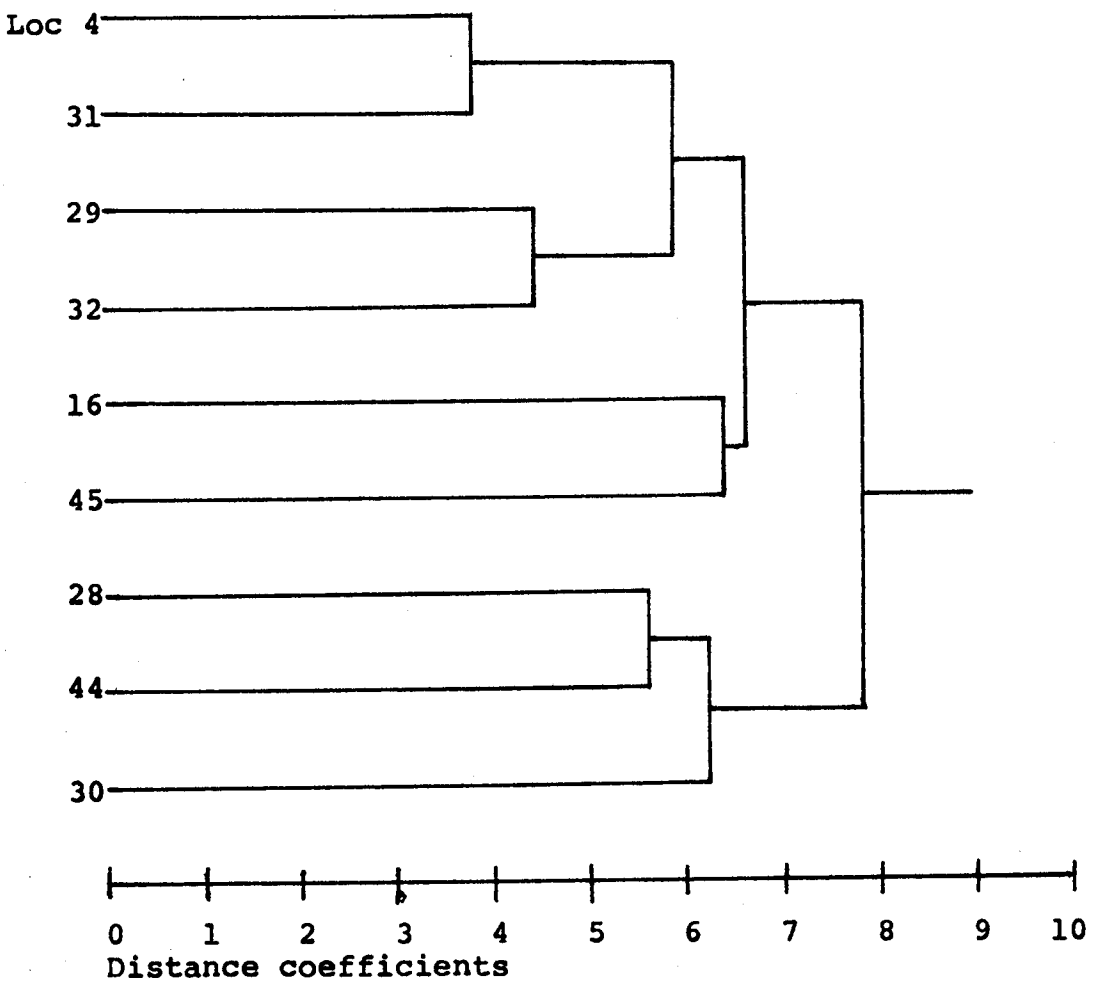


Figure 3. Diagram of relationships (dendrogram) among locations in Northern Europe.

(Wageningen, Netherlands) shows a high degree of resemblance to location 31 (Cambridge, England) in their pattern of interaction. Location 29 (Weihenstephan, West Germany) appears to have a close similarity to location 32 (Zurich, Switzerland).

Cluster analysis results for Southern Europe are presented in Figure 4. Figure 4 indicates the presence of three clusters with locations that are similar in their patterns of interaction. Locations 6 and 43 make up a cluster of their own; locations 7, 21, 27, 37 and 48 belong to one cluster. Locations 14 and 33 also make up a group.

Inter-regional cluster analysis was performed on locations in adjacent regions--Northern Europe and Southern Europe. The purpose for this was to study similarities among locations across the two regions. Results of this analysis are presented in Figure 5. The diagram suggests that locations in these two regions fall into four clusters. Locations 4, 16, 29, 31, 32 and 45 form one cluster. Intra-regional analysis for Northern Europe showed these same locations to form a group. Another group indicated in the diagram consists of locations 6, 28, 30, 43 and 44. It will be remembered that location 28 formed a cluster with locations 44 and 30 in intra-regional cluster analysis for Northern Europe. It still forms a cluster with locations 44 and 30 but the dendrogram suggests that location 28 has

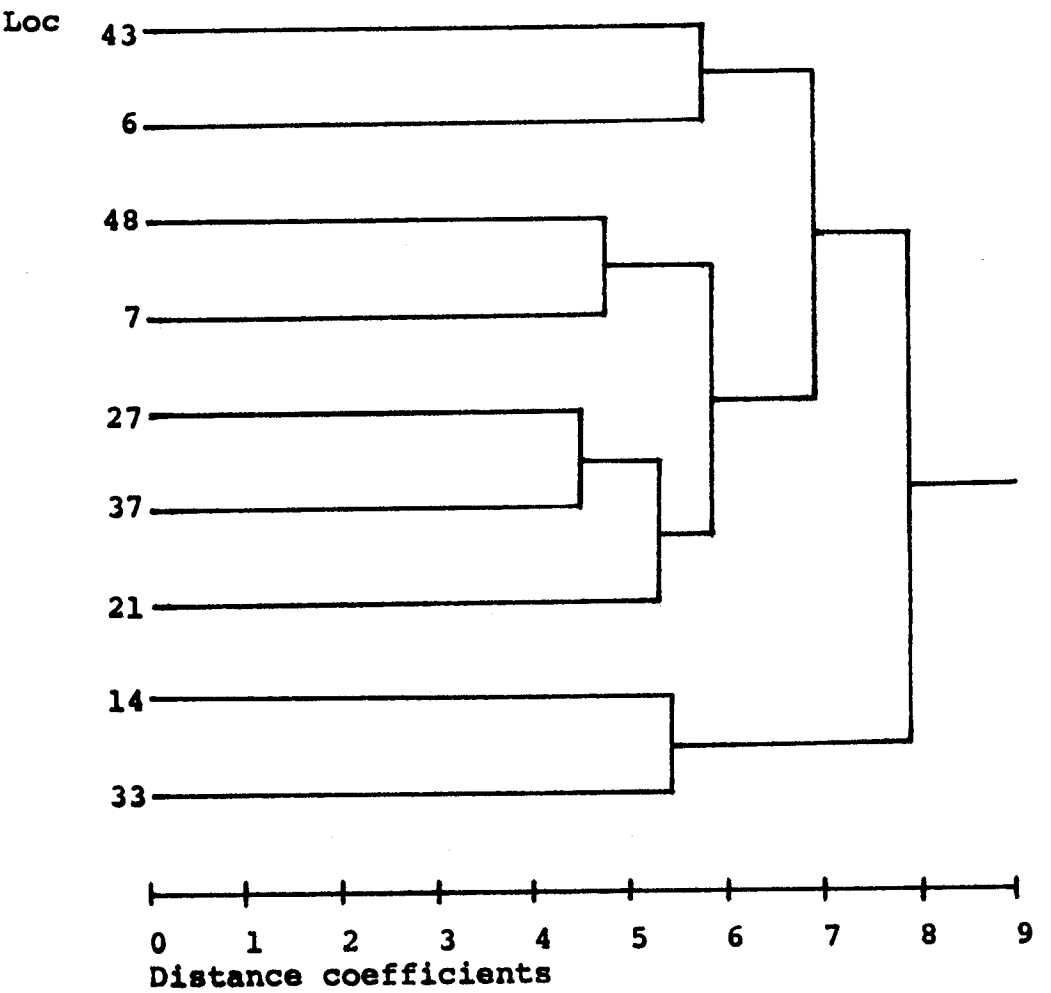


Figure 4. Diagram of relationships (dendrogram) among locations in Southern Europe.

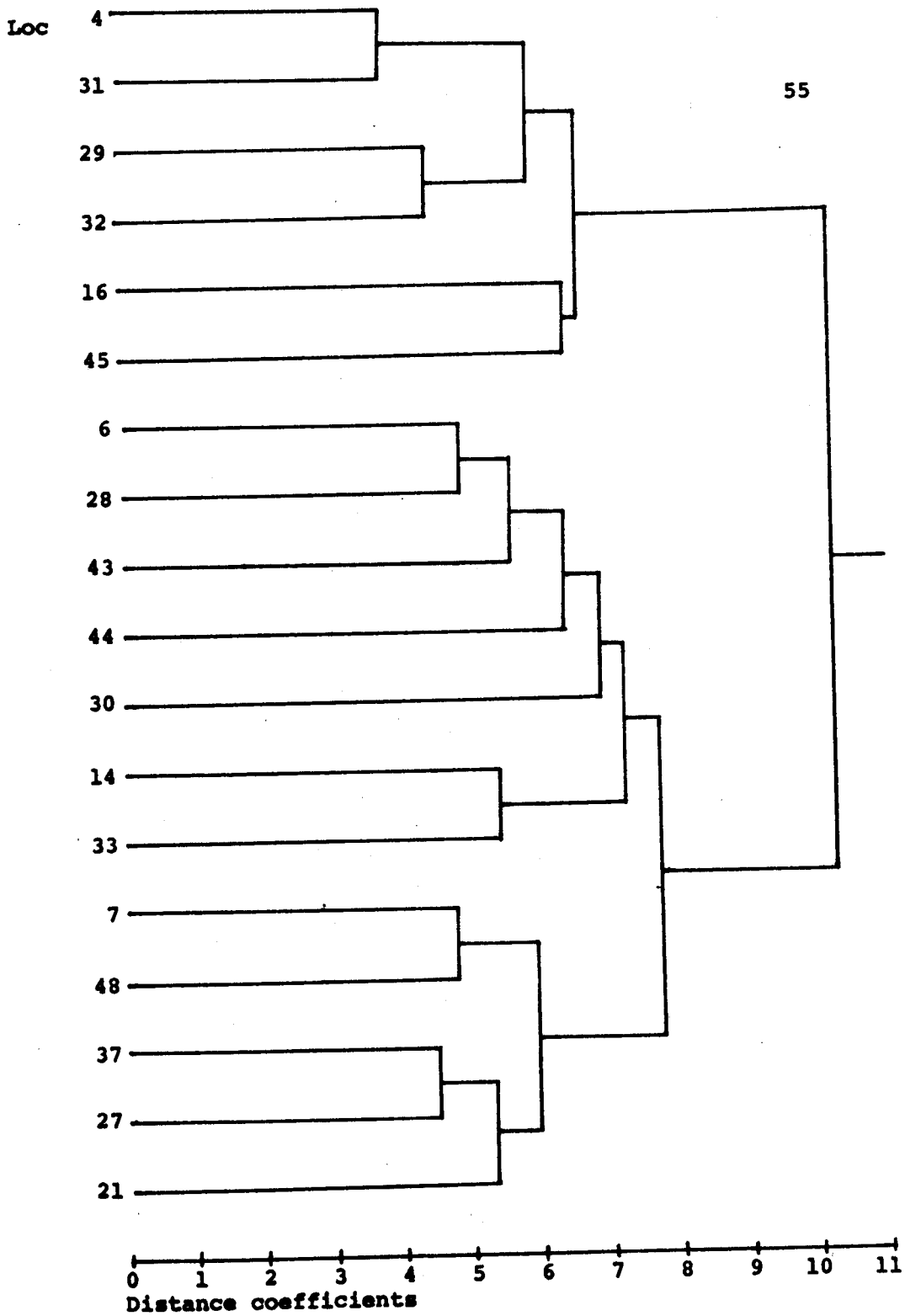


Figure 5. Diagram of relationships among 18 locations from combining two regions--N. Europe and S. Europe.

a closer affinity to location 6 which belongs to Southern Europe than with either 44 or 30 all of which belong to Northern Europe. Locations 14 and 33 still maintain a close affinity as in intra-regional analysis. The Southern European locations, locations 7, 21, 27, 37, and 48 still remain as a cluster.

The Far East was made up of only two locations. Because of this, the Far East was combined with the Near East. Figure 6 presents results of this inter-regional cluster analysis. The dendrogram shows a grouping of the locations in the two regions into three clusters. Location 8 and 9 form a group; locations 10, 13 and 15 form another group and lastly locations 12, 54 and 49 form a cluster which is too variable that location 49 should not be considered a member of the group. Thus the groups with an acceptable degree of similarity are 8 and 9; 10, 13 and 15; and 12 and 54. The diagram suggests a strong similarity between locations 12 and 54. The diagram also shows that there is very little similarity between location 9 and 49, the only locations representing the Far East.

Cluster analysis for all locations in the fourth and fifth IWWPN combined was performed for the purpose of studying the manner of similarity among all the locations. The results are presented in Figure 7. Several groupings are revealed in this analysis. Locations 4, 16, 29, 31, 32 and 45 form a cluster of locations that have a similar

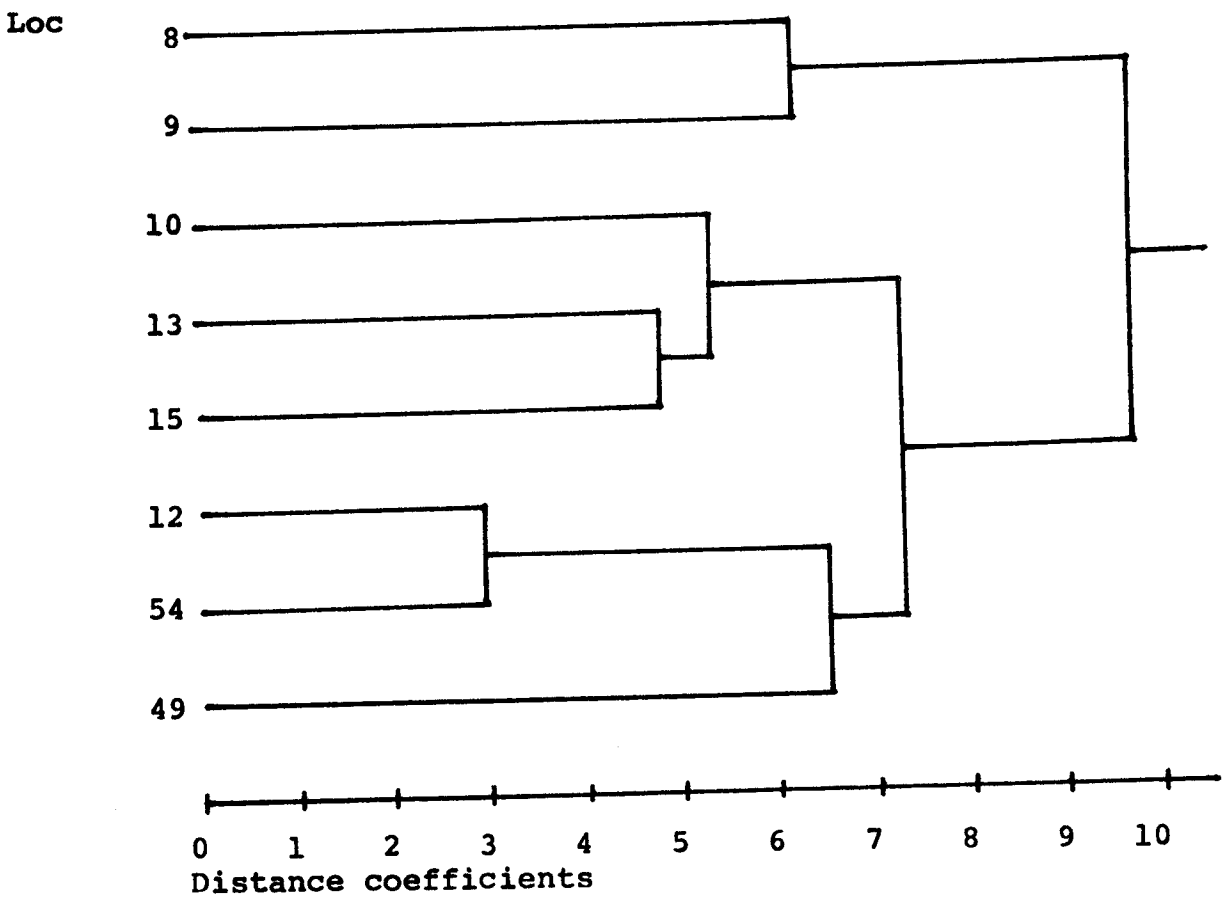


Figure 6. Diagram of relationships (dendrogram) among locations in two regions--Near East and Far East.

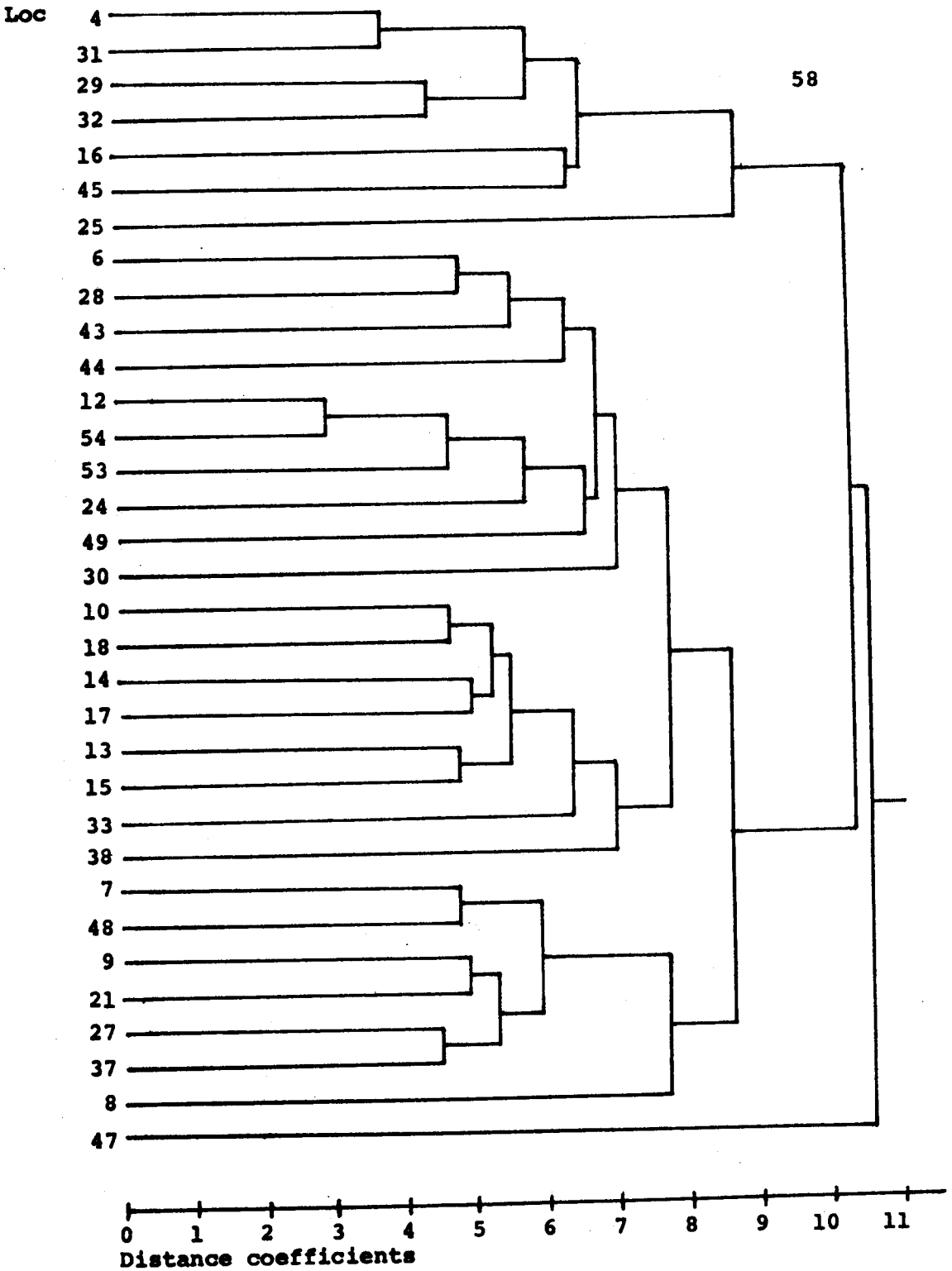


Figure 7. Diagram of relationships (dendrogram) among locations in all five regions.

pattern of interaction. These six locations belong to Northern Europe and interestingly they have maintained their similarity which they exhibited in intra-regional cluster analysis for Northern Europe. The next group comprises locations 6, 28, 43 and 44. Locations 6 and 43 belong to Southern Europe and location 28 and 44 belong to Northern Europe. The next group comprises locations 12, 24, 49, 53, and 54. Locations 12 and 54 are shown to have a very close affinity to each other. Location 12 is Eskisehir, Turkey and location 54 is Hamadan, Iran. Both locations belong to the Near Eastern region. Location 53 shows a close affinity to locations 12 and 54. Location 53 is Corvallis, Oregon in the United States. Such a close similarity of location 53 to locations 12 and 54 suggests that places quite removed from each other in terms of geographical distance can possess similar environmental conditions to which cultivars react similarly. The last member of this group is location 49 which is Morioka, Japan. Locations 10, 18, 14, 17, 13 and 15 form a cluster. This cluster shows a close relationship among locations belonging to different regions. Location 10 is Karaj, Iran and is closely similar to location 18 which is Stillwater, Oklahoma in the United States. Location 14 is Fundulea, Romania and is closely related to location 17 (North Carolina, U.S.). Locations 13 and 15 belong to the Near Eastern region.

Another group comprises locations 7, 9, 21, 27, 37, 48. All but location 9 belong to Southern Europe. This group indicates that Suwon (location 9) is similar to some locations in Southern Europe.

Estimates of Stability Parameters

Twenty-five cultivars were used in studying stability parameters. Through stability analysis, stable varieties, that is, those varieties that show a consistency of performance in different environments were identified. The stability parameters we are concerned with in the study are the regression coefficient (b_i) and deviation mean square (S^2d_i) of each variety. The analysis of variance is shown in Table 15. The results indicate that differences exist among the linear regression coefficients of the varieties meaning that the varieties have different responses to improving environmental indices.

The five-year mean yield of each cultivar and estimates of the stability parameters are presented in

Table 15. Analysis of variance of grain yield measured on 25 varieties of wheat grown in 66 environments (1972-1973) used for stability parameter estimation.

Source of variation	df	SS	MS
Genotypes (G)	24		
Environments (E) G x E	1625		
E (linear)	1	312620.75	
G x E linear	24	4504.04	187.66**
Pooled deviations	1600	84533.55	52.83**
Pooled error	4680	31767.27	6.78

** indicates significance at the .01 level of probability.

Table 16. As shown in Table 16, yields ranged from a low of 29.51 quintals per hectare for Jyva to a high of 47.66 quintals per hectare for Sava. Starke with a yield of 29.77 had second lowest yield. Fifteen of the 25 varieties had mean yields greater than the grand mean. Regression coefficients of yield means on evironmental indices ranged from 0.71 to a high of 1.20. The smallest regression coefficient of 0.71 was given by Kirac 66. Maris Nimrod and Clarion possessed the largest regression coefficients. Deviation mean squares (S d_i) were quite variable. The smallest deviation from regression mean square (20.88) was expressed by Hokuei. Starke showed the greatest interaction with environments as suggested by its large deviation from regression mean square of 91.94.

Kirac can be described as a variety that does not respond much to favorable environments. It is expected to yield better than average in poor environments but lower than average in more favorable environments.

Clarion and Maris Nimrod had the highest regression coefficients (1.20). They both yielded better than average. Maris Nimrod responded very well to more favorable environments but the large deviation mean square of 67.70 suggests that it is unstable. Clarion responded well to an increasing environmental index. It also can be considered to be unstable in this set of varieties due to its relatively large deviation mean square.

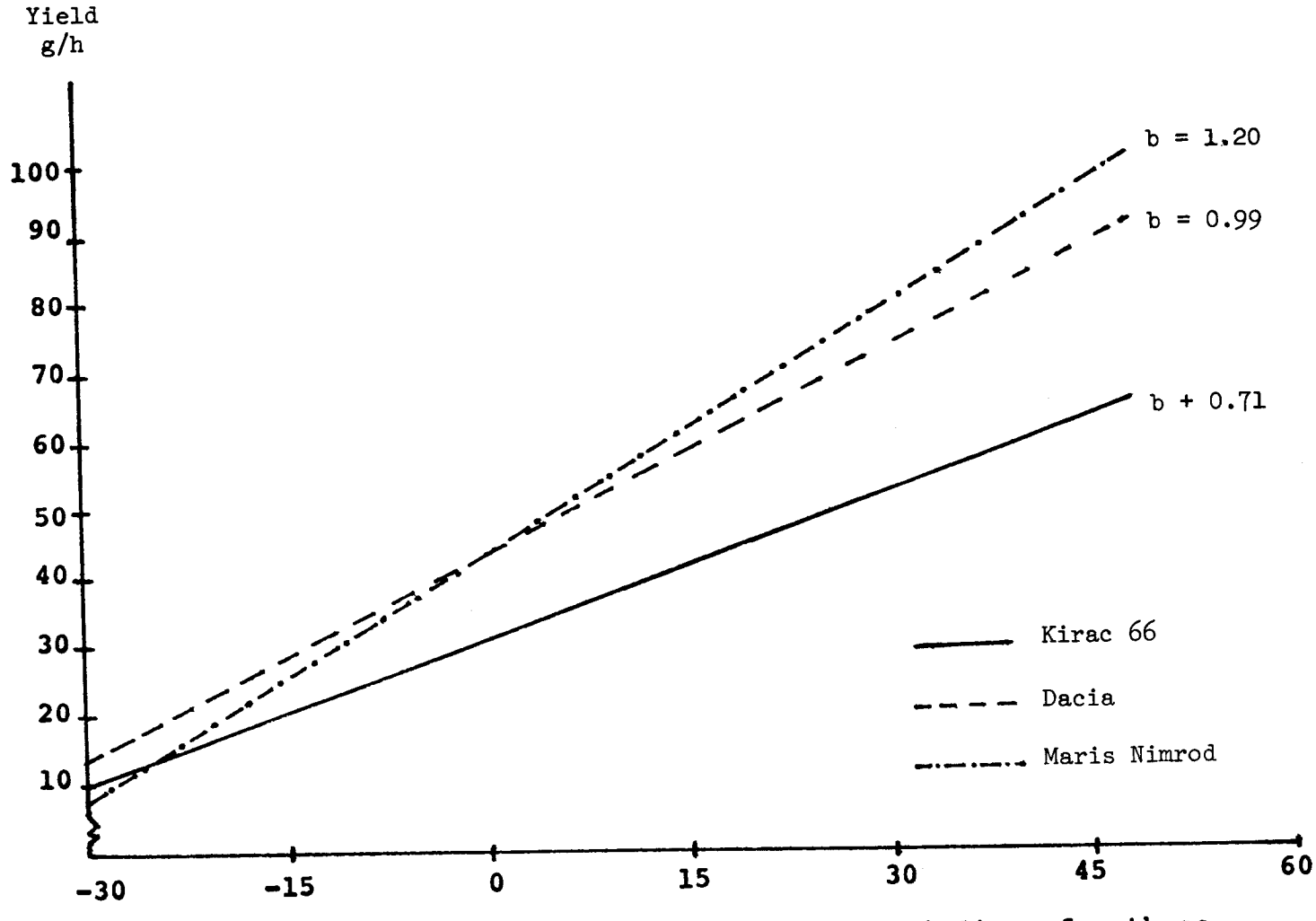
Table 16. Stability parameters for yield means of twenty-five wheat cultivars tested in sixty-six environments in fourth and fifth IWWPN, 1972-1973.

Entry	Mean quint/Hect.	Regression Coef.	Regression Dev. MS. (s^2d_i)
Caribo	40.82	1.11	45.97
Diplomat	35.34	0.92	43.45
Clarion	39.12	1.20	50.87
Starke	29.77	1.07	91.94
Atlas 66	36.87	0.93	25.13
Blueboy	42.68	0.98	24.03
Centurk	42.96	1.04	46.95
C.I. 15074	34.31	0.83	47.01
Tamwheat 102 (Tx62A4793-F)	38.73	0.83	57.59
NE 701132	39.21	0.98	49.27
Bezostava 1	43.82	0.95	29.92
Probstadorfer Extrem	43.33	1.12	26.00
Jyva	29.51	1.04	57.40
Vakka	31.86	0.91	47.88
Hokuei	37.28	0.94	20.88
Backa	41.44	0.93	57.63
Sava	47.66	1.18	48.72
Golden Valley (zg 5994/66)	44.66	1.03	63.14
Maris Nimrod	43.93	1.20	67.70
Dacia	43.69	0.99	24.37
Moldova	39.94	0.92	28.97
Zenith	39.86	1.17	29.14
Roussalka	42.95	0.96	73.32
Kirac 66	31.65	0.71	42.22
Carifen 12	36.09	1.03	48.58
Overall mean:	39.05		

Although Starke had a regression coefficient of 1.07 which is slightly above 1.00, it yielded lower than average and had the biggest deviation mean square, suggesting great fluctuations in yield in different environments. Hokuei had a regression coefficient close to 1.00 (0.95) and possessed the lowest deviation mean square. This suggests that in this set of varieties studied, Hokuei could be considered the most stable variety. It, however, yielded lower than average. Sava gave the highest yield. Its regression coefficient was 1.18 which indicates that this variety responds well to improving environmental indices; however, a deviation mean square of 48.72 indicates that the variety interacts considerably with environments. Probstdorfer Extrem, Dacia, Blueboy and Atlas 66 had small deviation mean squares. Of these four varieties only Atlas 66 yielded lower than average. Atlas 66 could be described as a variety, that does not respond much in favorable environments ($b<1.00$). Its relatively small deviation mean square suggests that it does not interact greatly with environments. It thus can be considered a stable variety. Probstdorfer Extrem had a regression coefficient greater than one ($b>1.00$). This variety yielded higher than average and showed relatively little interaction with environments as suggested by its small deviation mean square. It too can be described as a stable variety.

Dacia and Blueboy both had regression coefficients that were close to 1.00. They too showed stability. C.I. 15074 and Tamwheat 102 (Tx62A4793-7) had a regression coefficient of 0.83 and they both had lower than average yield.

Three varieties, Maris Nimrod, Dacia and Kirac 66 were compared in Figure 8. Maris Nimrod gave lowest yield in poor environments as indicated by Figure 8. It, however, gave highest yields in more favorable environments. Dacia yielded better than both Maris Nimrod and Kirac 66 at poor environments; however, it was second to Maris Nimrod in yielding in more favorable environments. Kirac showed the lowest response to improving environmental conditions.



	S^2d_i	Mean
Kirac 66	42.22	31.65
Dacia	24.37	43.69
Maris Nimrod	67.70	43.93

Figure 8. Regression of yield on environmental indices for three cultivars, Maris Nimrod, Dacia and Kirac 66.

DISCUSSION

The occurrence of genotype x environment interaction presents the breeder with two problems. In the first place he has to decide whether, in spite of the interaction, to try to produce a single variety with good general adaptation to the whole range of environmental and agronomic conditions of importance, or to breed varieties adapted to specific subsets of these environments (such as distinct regions or specified managements). Secondly after the breeder has chosen the range of environmental conditions within which to operate his selection program he has to decide how best to evaluate his material with respect to its adaptability to the environments. The linear regression model proposed by Eberhart and Russell (1966) can be utilized by the breeder to study genotype stability and hence adaptability.

If selection for general adaptation is considered an inefficient procedure, the breeder should consider the possibility of arranging environments into more homogeneous groups. This can be achieved on the basis of physical relationships of environments such as locations within a region, differential management, etc. or by observing the performance of genotypes and empirically combining in one group those environments exhibiting a similar pattern of genotypic performance. The method of cluster analysis is emerging as a potent tool for measuring similarities among

environments and classifying them into homogeneous groups. Abou-El-Fittouh et al. (1969) used cluster analysis to classify locations used for cotton variety tests in the United States cotton belt. Byth et al. (1976) used pattern analysis (cluster analysis) to classify sixty-three international environments into homogeneous groups. He also classified forty-nine wheat cultivars grown in each of the sixty-three environments into groups possessing a similar manner of performance. Johnson (1977) applied cluster analysis to arrange 49 maize hybrids into groups that were differentiated in terms of means and stability.

Classifying environments into homogeneous groups is effective in reducing the variance due to genotype x environment interaction but not genotype x year interaction component (Wright 1976). Genotype x year interactions are caused by unpredictable environmental conditions such as insect and disease outbreaks, fluctuations in amount and distribution of precipitation and in temperature patterns.

Estimates of the pertinent variance components obtained in intra-regional analysis of variance show the genotype x year interaction component of variance to be in general smaller than other components in each region. This suggests that in general environmental conditions did not change much in the two years. The three factor interaction component--genotype x location x year was large in magnitude

in all locations except those in the Near East.

The magnitude of the genotype x location component of variance in all the currently defined regions indicates the presence of enough differences among locations within regions to warrant classification of the locations into adaptive groups. Intra-regional cluster analyses were performed to study similarities among locations within each region.

Results of classification for the United States indicated that there exist diverse environmental conditions within this region such that the seven sites used in the study do not adequately sample all the conditions in this big region. More test locations are required to properly identify clusters of similar locations. Location 47 is very different from the rest of the locations in this region, it should belong to a different region; since it is not a winter wheat area after all. However, Figure 2 suggested that location 17, 18 and 38 had a similar pattern of interaction and also locations 24 and 53 showed a high degree of similarity. Genotype x location interaction component was reduced 100% within the cluster consisting of locations 17, 18 and 38 compared to the U. S. region as a whole. The genotype x location interaction component estimate was negative (-2.44) which is considered to be zero. In the group made up of two locations, 24 and 53, genotype x location

interaction variance component was 10.42, which compared to that of the region (20.97) was about 50% lower. However, it was highly significant indicating that this group did not significantly reduce the genotype x location interaction. Figure 9 shows the locations in this region.

In Northern Europe, classification reveals a close resemblance among locations. Figure 3 indicated that locations 4, 16, 29, 37, 32 and 45 had a similar pattern of interaction and locations 28, 30 and 44 formed a cluster of locations with a similar pattern of interaction. Although location 45 is between locations 28 and 44 (see Figure 10) the analysis suggests that it is not adequately similar to either location 28 or 44 to be grouped with them. Genotype x location interaction component for the group consisting of locations 4, 16, 29, 31 and 45 was non-significantly different from zero and almost 50% lower in magnitude compared to that of the whole region. The genotype x location interaction component in this cluster was 3.59 while that for the whole region was 9.27. The group consisting of locations 28, 30 and 44 gave an even smaller and non-significant genotype x location interaction component of variance. The component had a magnitude of 1.37.

Results of classification for Southern Europe suggested that there are three groups into which

Figure 9. Map showing distribution of locations in the United States.

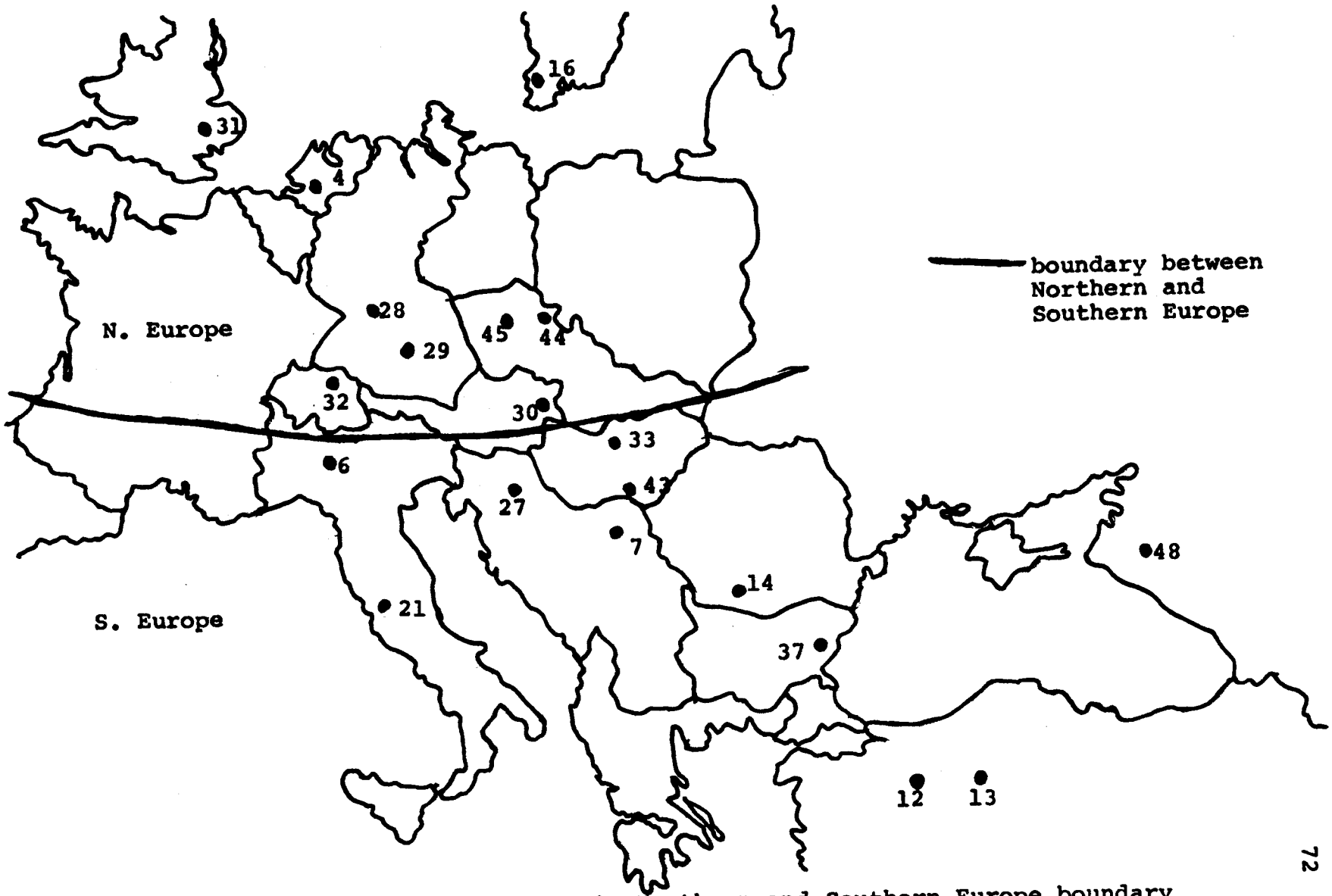


Figure 10. Map showing locations in Northern and Southern Europe boundary between Northern and Southern Europe.

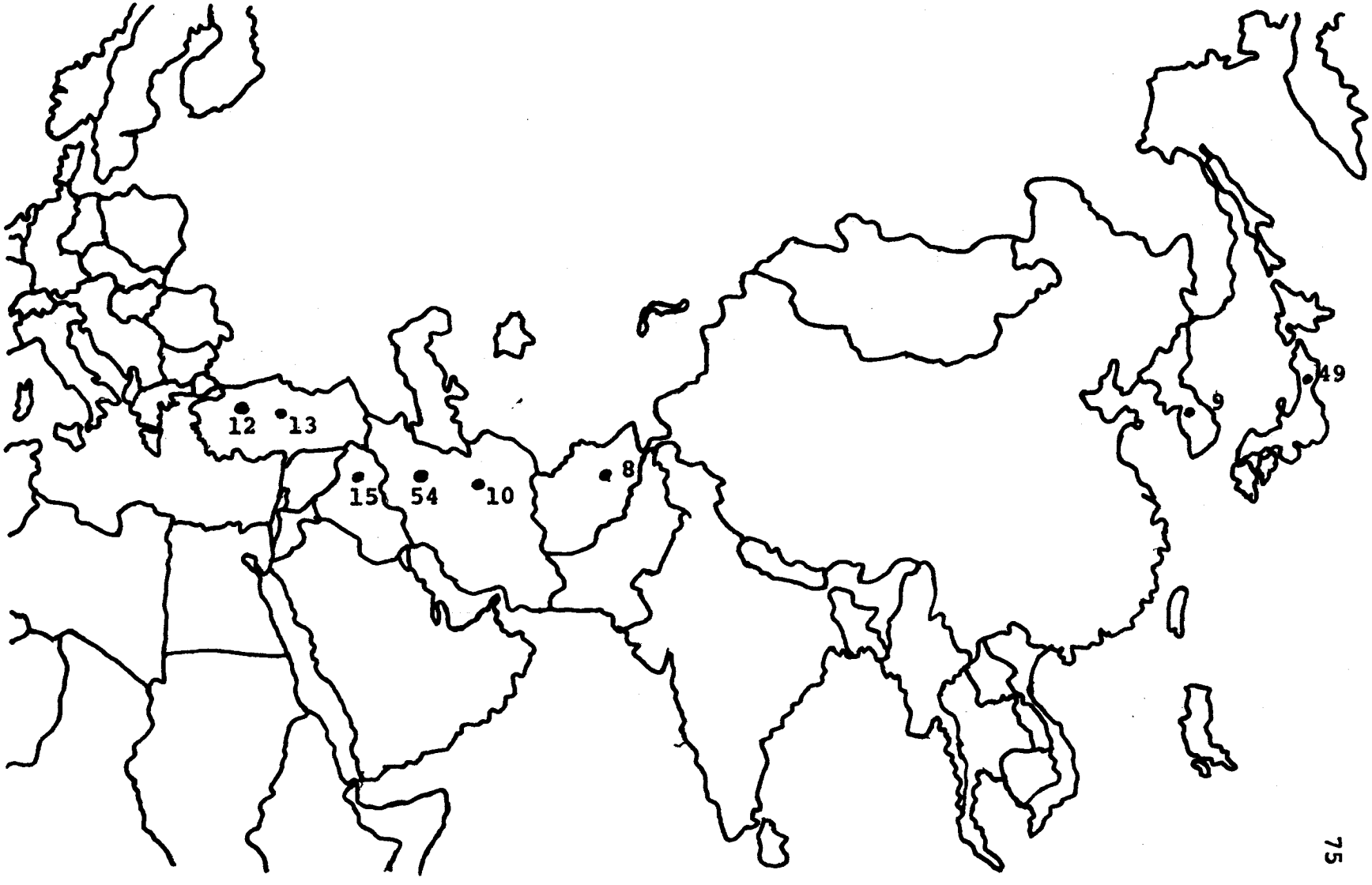
locations fall. The groups indicated in Figures 4 and 5 are locations 6 and 43; locations 7, 21, 27, 37 and 48; and locations 14 and 33. Estimates of genotype x location interaction components of variance were non-significant in all three subgroups. Genotype x location interaction component for the whole region was 8.86. The magnitude of this interaction in the cluster comprising locations 6 and 43 was 0.17, a 98% reduction in the genotype x location interaction compared to that of the region as a whole. Locations 7, 21, 27, 37 and 48 form a homogeneous group that gave a non-significant genotype x location interaction variance component. The magnitude of the component (4.07) represents a 54% reduction compared to that of the region as a whole. A group made up of location 14 and 33 had a similar pattern of interaction. Genotype x location interaction component expressed by this pair was 2.36. It was non-significantly different from zero and represents a 70% reduction in the interaction component compared to the same statistic for the region as a whole.

Cluster analysis for Northern Europe and Southern Europe combined showed a clear separation between the two regions for most locations. Figure 5 showed locations 4, 16, 29, 31, 32 and 45 which belong to Northern Europe to have maintained a close resemblance to each other just as in within regional analysis for Northern Europe. Locations

7, 21, 27, 37 and 48; and locations 14 and 33 all of which belong to Southern Europe, maintained their grouping, as they did in intraregional cluster analysis for Southern Europe. Had geographical proximity among locations been the classification criterion location 33 would have been grouped into the same group as location 43 (see Figure 10), but this analysis suggests that location 33 is more similar to location 14 than to 43. The cluster in Figure 6 revealed across-region similarity among locations. In the cluster comprising locations 6, 28, 30, 43 and 44 location 6 showed a closer affinity to location 28. Location 6 is in Southern Europe and location 28 is in Northern Europe. Location 43 belongs to Southern Europe while locations 44 and 30 belong to Northern Europe.

Three major clusters are suggested in the results of classification analysis for Near and Far East combined. Figure 6 indicates locations 8 and 9; locations 10, 13, 15 and locations 12, 49 and 54 form clusters of locations with similar patterns of interaction. There appears to be very little similarity between location 9 and 49, the only locations representing the Far East. These two locations would be expected to have a closer similarity to each other than the results suggest because they are close to each other geographically (see Figure 11) and also they are both surrounded by oceans. The dendrogram suggests that there is a

Figure 11. Map showing locations in Near and Far East.



close similarity between locations 12 and 54. Location 12 is geographically so close to location 13 that a closer similarity between these two would be expected. However, cluster analysis suggested a close similarity between locations 12 and 54. These two locations gave a non-significant genotype x location interaction. The genotype x location interaction component of variance was 89% smaller than the same statistic for Near East as a whole. Genotype x location interaction component for the Near East region was 20.30 while that for the two locations was 2.21. Locations 10, 13 and 15 form a cluster with a much smaller genotype x location interaction component of variance compared to that of Near East as a whole. The magnitude of the interaction component was 5.54 which was a 70% reduction compared to that of the region as a whole, which was 20.30. Pertinent variance components from groupings formed by cluster analysis are presented in Table 17.

An overall cluster analysis for all the locations combined suggests interesting similarities among the international environments. Locations 4, 16, 29, 31, 32 and 45 of Northern Europe still maintain their similarities. Except for lcoation 9 in the cluster comprising locations 7, 9, 21, 27, 37 and 48 all of these locations belong to Southern Europe. The Southern European locations in this

Table 17. The pertinent variance components obtained from analyzing two-year data for separate regions and for clusters formed within individual regions.

	U.S.A.						North Europe					
	Whole region		Loc 17, 18, 38		Loc 24, 53		Whole region		Loc 4, 16, 29, 31, 32, 45		Loc 28, 30, 44	
Var comp	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE
$\hat{\sigma}^2_{g\ell y}$	30.64**	4.62	38.69**	9.06	5.93**	3.07	27.49**	3.38	26.22**	4.19	29.62**	6.75
$\hat{\sigma}^2_{gy}$	6.39**	3.45	22.32**	10.92	0.17	2.18	2.54*	1.78	-0.71	1.48	9.61*	6.25
$\hat{\sigma}^2_{g\ell}$	20.97**	5.36	-2.44	6.06	10.42**	4.72	9.27**	3.14	3.59	3.31	1.37	4.97

	South Europe								Near East					
	Whole region		Loc 6, 43		Loc 7, 21, 27, 37, 48		Loc 14, 33		Whole region		Loc 10, 13, 15		Loc 12, 54	
Var comp	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE	Var comp	SE
$\hat{\sigma}^2_{g\ell y}$	24.65**	3.22	21.55**	9.50	17.76**	3.39	21.85**	7.14	11.91**	2.44	8.64	3.43	1.77	1.67
$\hat{\sigma}^2_{gy}$	-0.53	0.89	1.86	7.06	-2.07	1.41	14.39*	9.73	1.75*	1.42	14.56**	5.85	2.21*	1.66
$\hat{\sigma}^2_{g\ell}$	8.85**	2.98	0.17	6.71	4.07	2.81	2.36	5.53	20.31**	4.05	5.54	3.30	1.67	1.53

cluster maintained their similarity as they did in both intra-regional and across regional (Northern Europe and Southern Europe combined) cluster analysis. The classification results over these two regions indicates the presence of pockets of locations having strong similarities. Reducing the number of IWWPN test locations could be best done in Northern Europe and Southern Europe. These two regions appear (from cluster analysis) to have the largest clusters of locations possessing similar patterns of interaction. Thus only a few test sites are required to represent most locations in Northern and Southern Europe.

The overall cluster analysis also revealed the presence of similarities among locations which are widely separated geographically (see Figure 7). Two members of the cluster made up of locations 6, 28, 43 and 44 belong to Southern Europe (locations 6 and 43) and the other three belong to Northern Europe. Locations 12, 24, 49, 53 and 54 form a cluster with locations that even though far removed geographically, possess quite a close affinity to each other. Location 12 is Eskisehir, Turkey; location 24 is Ithaca, New York in the United States; location 49 is Morioka Iwate, Japan; location 53 is Corvallis, Oregon in the United States and location 54 is Hamadan, Iran. Another cluster indicated in Figure 7 contains locations 10, 13, 14, 15, 17, 18 and 33. Locations 10, 13 and 15 are

of the Near Eastern region. Location(s) 17 and 18 belong to the United States. Location 33 belongs to Southern Europe.

Similarities among international environments as suggested by the results of this study reinforce the importance of international testing of cultivars and cooperation among crop scientists. A wheat breeder in the United States may be inadvertently breeding for some areas in Turkey or Iran or some other area on both sides of the globe.

From stability analysis the yields of four varieties (two stable and two unstable) were plotted on graphs to study their mode of performance in locations considered similar by cluster analysis and in randomly selected locations. The varieties Maris Nimrod and Starke (considered unstable by stability analysis) and varieties Hokuei and Blueboy (considered stable by stability analysis) were used in this study. The results are presented in Figure 12 and Figure 13. Figure 12 shows the fluctuations in yield of the four varieties across five locations of one cluster (see Figure 5). The fluctuations in yield of the four varieties are similar. Thus both stable and unstable varieties behave similarly in similar environments. A marked difference between stable and unstable varieties appears when their yield performances are studied across randomly selected locations. The yield curves for Maris Nimrod and

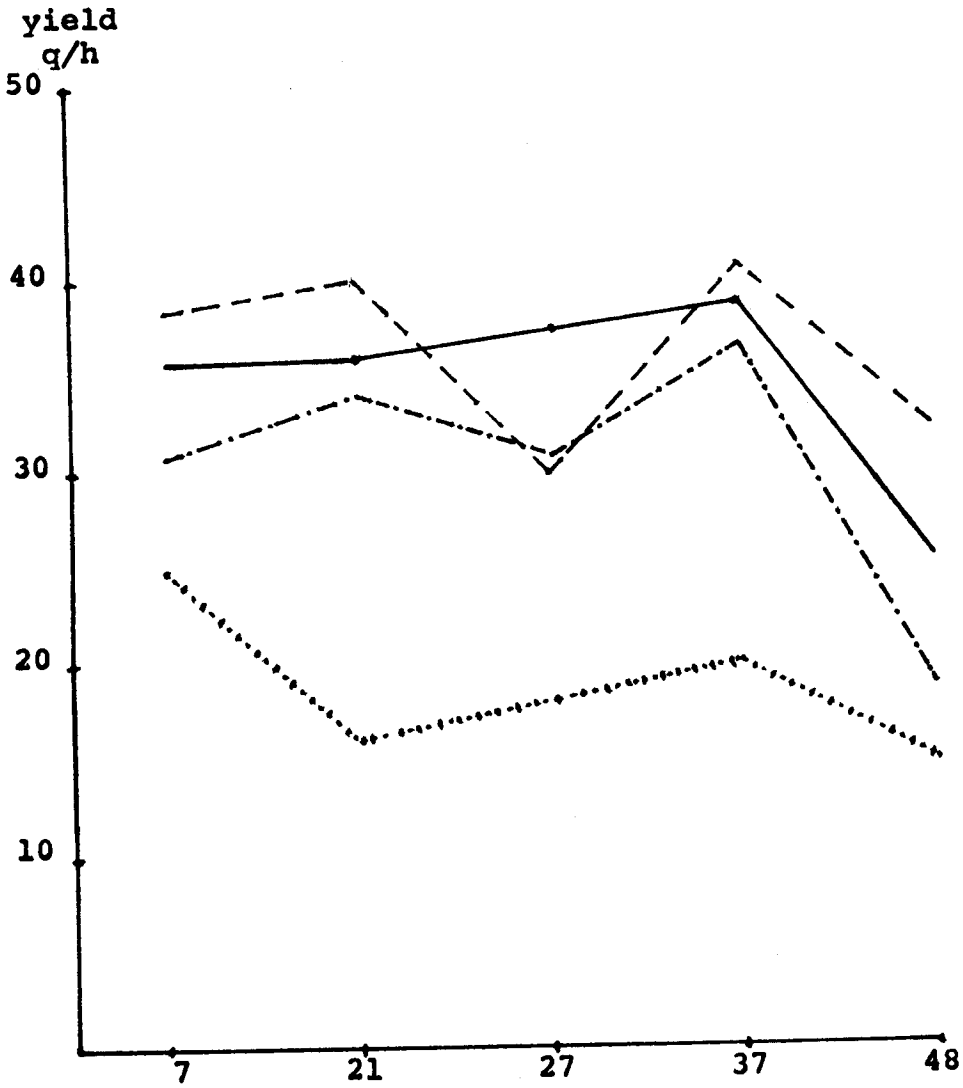


Figure 12. Yield fluctuations across a set of locations shown to be similar by cluster analysis.

-·-·-·- Hokuei

——— Maris Nimrod

- - - - Blueboy

····· Starke

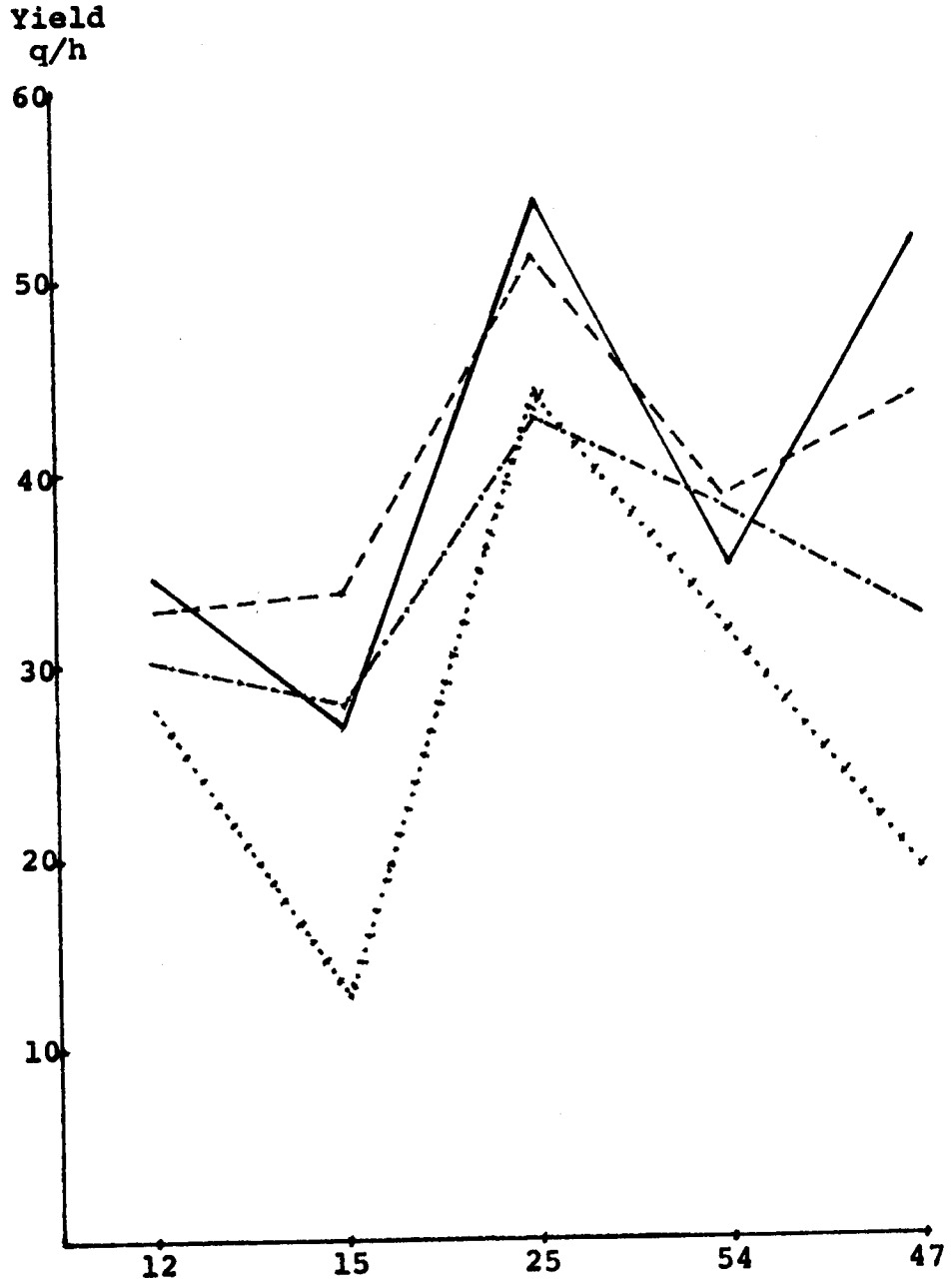


Figure 13. Yield fluctuations of four varieties across a randomly selected set of locations.

Starke are more erratic than those for the stable varieties Hokuei and Blueboy. The amplitudes of the graphs are different; the unstable varieties having the greatest amplitude. The difference between Figures 12 and 13 indicate that cluster analysis was effective in grouping similar locations together such that the amplitudes of yield fluctuations of unstable varieties are minimized enough for the varieties to appear to be stable in a cluster of these locations.

SUMMARY AND CONCLUSIONS

The fourth and fifth IWWPN data were used to classify 33 international sites into clusters of locations possessing similar patterns of interaction, and also to study the possibility of reducing the number of IWWPN test sites. Analysis of variance for each of the currently defined regions gave highly significant genotypic components of variance and also genotype x location interaction components in all the regions. This indicates the presence of environmental diversity among locations within each region which could be reduced by classifying the locations into homogeneous groups by cluster analysis.

The method of cluster analysis was employed to classify locations within each region, across-regions, and all the regions combined into adaptive zones (clusters of locations possessing a close affinity to each other).

The results of classification revealed that the United States region has quite diverse environmental conditions. Only locations 17, 18 and 38 and locations 24 and 53 showed affinity to each other. Thus, it would not be advisable to reduce the number of test sites in the United States. Results of classification for Northern Europe and Southern Europe revealed the presence of large pockets of closely similar locations. Thus if the number of IWWPN test locations are to be reduced, Northern Europe and Southern Europe would be

the best regions to reduce number of test sites. According to this study, only a few test locations are needed in Northern and Southern Europe to represent the rest of the locations in the two regions. This study also shows a demarcation between Northern Europe and Southern Europe so that the regions are currently well defined. Locations in the Near East fall into homogeneous groups. There is a possibility of reducing the number of locations but the test locations are so few that it probably would not be best to reduce them any further. The results also showed that the two locations in the Far East possess very little similarity to each other. These need to be retained. The study also suggested the presence of a close affinity among locations that are quite far removed from each other geographically.

Cluster analysis in this study showed it to be very effective in reducing the genotype x location interaction component as shown in Table 17 but did not seem to be effective in reducing the three factor interaction of genotype x year x location. Analysis of independent data collected at later dates would give a clearer test of effectiveness of the grouping.

A study of yield fluctuations across grouped locations as shown in Figures 12 and 13 suggests that cluster analysis is effective in grouping homogeneous groups together.

The clusters formed in this study might not be repeated with a different set of varieties. There is then a need to do further studies with different sets of varieties to study the consistency in the grouping patterns.

LITERATURE CITED

Abou-El-Fittouh, H. A. 1967. Genotype by environment interactions in cotton--their nature, related environmental factors and implication in zoning. Ph.D. Thesis. North Carolina State University.

_____, J. O. Rawlings and P. A. Miller. 1969. Classification of environments to control genotype by environment interactions with application to cotton. Crop Sci. 9:135-140.

Akposoe, M. K. 1967. Genotype x environment interaction in the development of maize hybrid varieties. M.S. Thesis. Univ. of Minnesota.

Allard, R. W. 1961. Relationship between genetic diversity and consisting of performance in different environments. Crop Sci.

_____ and A. D. Bradshaw. 1964. Implications of genotype environmental interactions in applied plant breeding. Crop Sci. 4:503-507.

Baihaki, A., R. E. Stucker, and J. W. Lambert. 1976. Association of genotype x environment interactions with performance level of soybean lines in preliminary yield tests. Crop Sci. 16:718-721.

Baker, J. L. and L. M. Verhalen. 1973. The inheritance of several agronomic and fiber properties among selected lines of upland cotton, Gossypium hirsutum L. Crop Sci. 13:144-450.

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

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

Chuang-Sheng Lin, and B. Thompson. 1975. An empirical method of grouping genotypes based on a linear function of the genotype-environment interaction. Heredity, 34:255-263.

Clifford, H. T., and D. W. Goodal. 1967. A numerical contribution to the classification of the poaceae. Austral. J. Bot. 15:499-519.

_____, W. T. Williams, and G. N. Lance. 1969. A further numerical contribution to the classification of the Poaceae. Austral. J. Bot. 17: 119-31.

Cochran, W. G. and G. M. Cox. 1957. Experimental Designs. John Wiley and Sons, Inc.

Comstock, R. E. and R. H. Moll. 1963. Genotype x environment interactions. Statistical genetics and plant breeding. NAC-NRC. pub. 982:164-196.

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

Finlay, K. W. and G. W. Wilkinson. 1963. The analysis of adaptation in a plant-breeding programme. Austral. J. of Agr. Res. 14:742-754.

Fisher, R. A., and Wa. S. Mackinzie. 1923. Studies in crop variation. II. The manurial response of different potato varieties. J. Agric. Sci., Camb., 13:311-320.

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

Goodall, D. W. 1953. Objective methods for the classification of vegetation. 1. The use of positive interspecific correlation. Austral. J. Bot. 1:39-63.

Gupton, C. L., P. D. Legg, L. A. Link, and H. F. Ross. 1974. Genotype x environment interactions in burley tobacco variety tests. Crop Sci. 14:811-814.

Haldane, J. B. S. 1946. The interaction of nature and nurture. Ann, Eugen. 13:197-205.

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

Hussaini, S. H., M. M. Goodman, and D. H. Timothy. 1977. Multivariate Analysis and the Geographical Distribution of the World Collection of Finger Millet. Crop Sci. 17:257-263.

Jensen, N. F. 1952. Intra-varietal diversification in oat breeding. Agron. J. 44:30-34.

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.

Kaltsikes, P. J. 1970. Genotype-environment interaction variances in yield trials of fall rye. Can. J. Pl. Sci. 50:77-80.

Liang, G. H., E. G. Heyne and T. L. Walter. 1966. Estimates of variety x environment interactions in yield tests of three small grains and their significance on the breeding programs. Crop Sci. 6:135-139.

Mather, K. and R. Morley Jones. 1958. Interaction of genotype and environment in the continuous variation. 1. Description. Biometrics. 14:343-359.

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

_____, H. F. Robinson and O. A. Pope. 1962. Cotton variety testing: additional information on variety x environmental interactions. Crop Sci. 2:349-352.

Mungomery, V. E., R. Shorter, and D. E. Byth. 1974. Genotype x environment interactions and environmental adaptation. I. Pattern Analysis--Application to Soybean Populations. Austral. J. Agric. Res., 1974 25:59-7

Perkins, J. M. 1972. The principal component analysis of genotype-environmental interactions and physical measures of the environment. Heredity; 29:51-70.

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

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

Rojas, B. A. and G. F. Sprague. 1952. A comparison of variance components in corn yield trials. III. General and specific combining ability and their interactions with locations and years. Agron. J. 44:462-466.

Seal, H. L. 1964. Multivariate statistical analysis for biologists. John Wiley and Sons, Inc., New York.

Shorter, R., D. E. Byth and V. E. Mungomery. 1977. Genotype x Environment Interactions and Environmental Adaptation. II. Assessment of Environmental Contributions. Aust. J. Agric. Res., 28:223-235.

Sokal, R. R. and C. D. Michener. 1958. A statistical method for evaluating systematic relationships. Kans. Univ. Sci. Bull. 38:1409-1438.

_____ and P. H. A. Sneath. 1963. Principles of Numerical Taxonomy. W. H. Freeman and Company, San Francisco.

_____ and P. H. A. Sneath. 1973. Numerical Taxonomy. The principles and practice of numerical classification.

Steel, R. G. D. and J. H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York.

Suzuki, S. 1968. Studies on the evaluation of strain adaptability. Proc. 12th Int. Congr. Genet., Tokyo. 1:270.

Webb, L. J., J. G. Tracey, W. T. Williams and G. N. Lance. 1967a. Studies in the numerical analysis of complex rain-forest communities. I. A comparison of methods applicable to site/species data. J. Ecol. 55:171-191.

_____. 1967b. Studies in the numerical analysis of complex rain-forest communities. II. The problem of species sampling. J. Ecol. 55:525-538.

Williams, W. T. 1971. Principles of clustering. A. Rev. Ecol. System. 2:303-326.

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

Yates, F. and W. G. Cochran. 1938. The analysis of groups of experiments. J. Agric. Sci., Camb., 28:556-580.

APPENDIX

Table A1. Analysis of variance for individual regions and all the regions combined.

Source	U.S.A. df	MS.	Var. comp.	SE	South Europe df	MS.	Var. comp.	SE	North Europe df	MS.	Var. comp.	SE
Genotypes	23	1583.53**	19.30	8.21	23	3736.31**	49.46	14.69	23	701.60**	34.87	11.06
G x Y	23	334.74**	6.39	3.44	23	104.44	-0.53	0.89	23	55.38**	2.54	1.78
G x L	138	323.54**	20.87	5.36	184	194.52**	8.86	2.99	184	51.10**	9.27	3.14
G x L x Y	138	155.75	30.64	4.67	184	123.66**	24.66	3.21	184	32.56**	27.49	3.38
Pooled error[1]	966	33.17	33.17	1.51	1242	25.04	25.04	1.00	1196	5.07	19.55	0.21
mean (q/h)		35.72				39.97				46.37		
C.V. (%)		16.12				12.52				9.56		

Source	Near East df	MS.	Var. comp.	SE	Far East df	MS.	Var. comp.	SE	All Regions Combined df	MS.	Var. comp.	SE
Genotypes	23	254.47**	15.40	6.07	23	650.41	14.40	14.27	23	1579.35**	22.28	6.77
G x Y	23	29.08**	1.75	1.43	23	188.37	1.96	9.03	23	61.91**	0.87	0.53
G x L	115	59.16**	20.30	4.05	23	404.23*	28.94	15.54	736	79.67**	23.32	2.24
G x L x Y	115	18.54**	11.91	2.44	23	172.68**	34.34	12.23	736	33.02**	26.40	1.72
Pooled error	805	6.63	25.80	0.33	276	35.29	35.30	2.99	4485	6.62	26.09	0.14
mean (q/h)		30.20				37.11				38.86		
C.V. (%)		16.77				16.00				13.16		

[1]The mean square for the pooled error term was obtained by dividing the original pooled error by the harmonic mean for N. Europe, and Near East and also for the overall analysis. The original pooled error was used for the rest.

Table A1. (continued)

* indicates significance at the .05 level of probability.

** indicates significance at the .01 level of probability.

The Computation of Distance

An example from the field of numerical taxonomy will be used to illustrate the method of distance as a measure of similarity. Assuming that two species, 1 and 2 are described on the basis of two characters X and Y (characters X and Y are measured if continuous and numerically coded in other cases) then a measure of similarity between species 1 and 2 based on the two characters would be the distance in figure A1. The two species may have any value for X and Y within the range of variation of these characters. The data for species 1 and species 2 can be plotted as shown in the figure where each of the two species assumes a fixed position in the two dimensional space defined by the two axes X and Y. A measure of the similarity between species 1 and 2 is the distance between the species and this is easily computed using Pythagoras' theorem.

$$\delta_{1,2} = \sqrt{(X_1-X_2)^2 + (Y_1-Y_2)^2}$$

When similarity between two species is to be computed based on three characters, the exact positions of the species are in a three-dimensional space as shown in figure A2. The computation of the distance between the two species which are now suspended at fixed points in the three-dimensional space is by an extension of the distance formula. The distance is now computed thus:

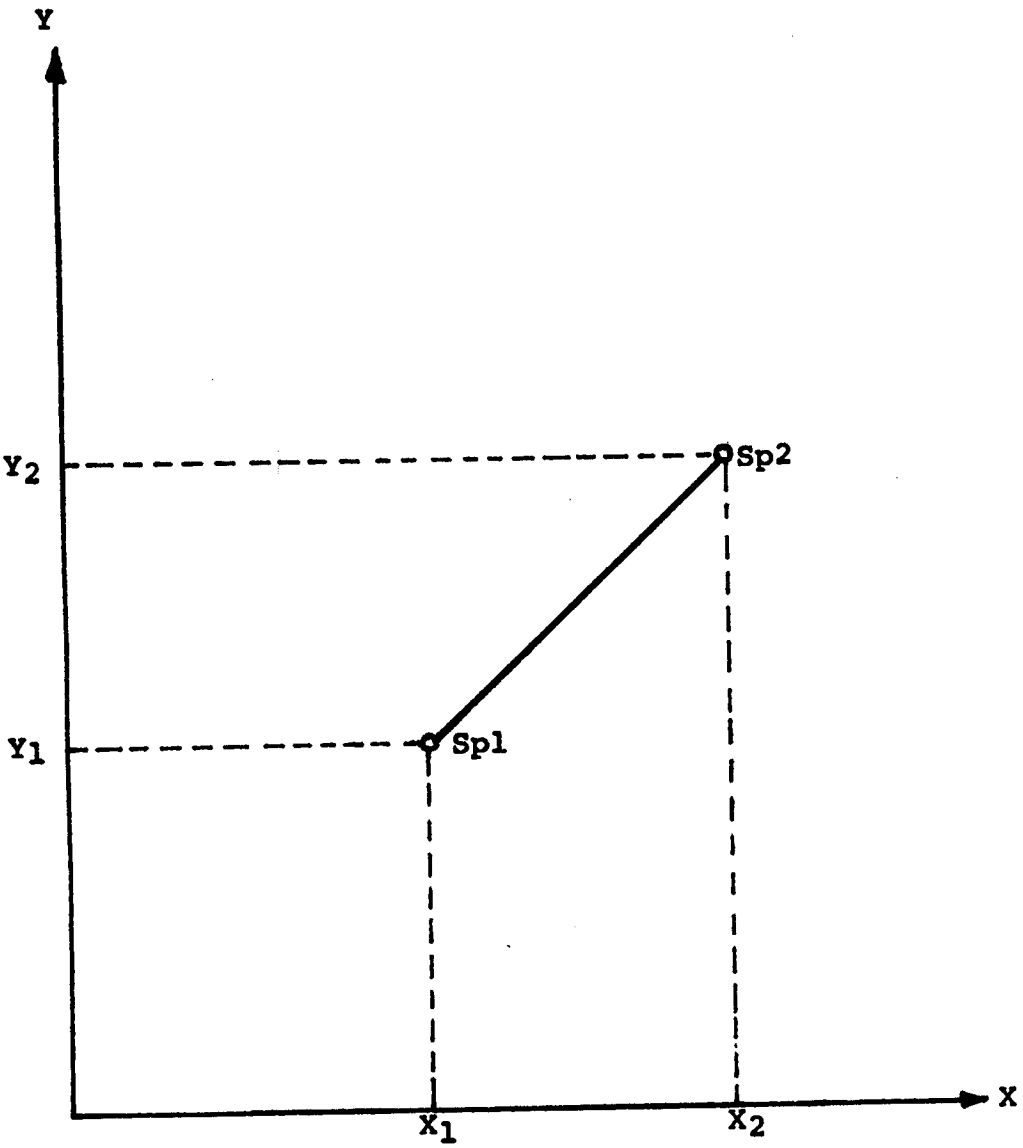


Figure A1. Two-dimensional diagram showing the distance between two hypothetical species with reference to two characters, X and Y.

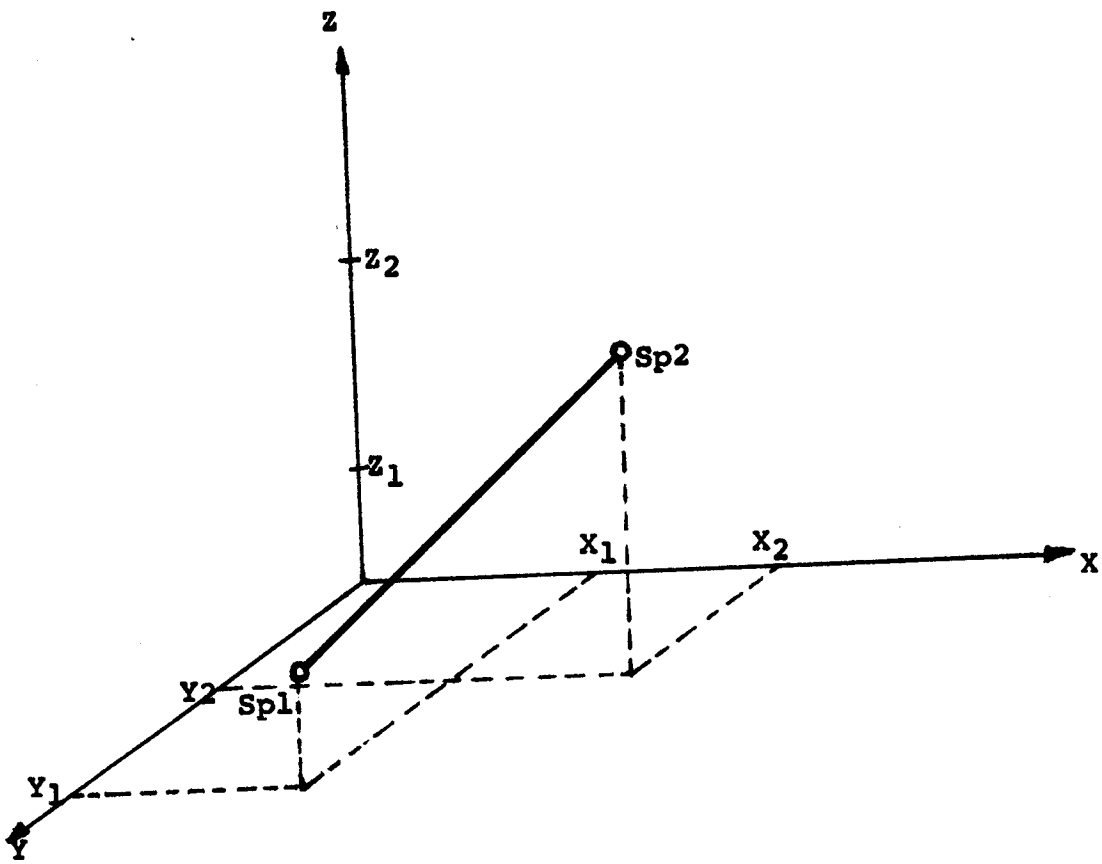


Figure A2. Two-dimensional projection of the distance between two species with reference to three characters, X, Y, and Z.

$$\delta_{1,2} = \sqrt{(X_1-X_2)^2 + (Y_1-Y_2)^2 + (Z_1-Z_2)^2}$$

The distance formula is equally valid for any number of characters, or in an n-dimensional space (hyper space). The distance between two species in such a hyper space of n characters can be computed very simply by summing the squares of the distances along each of the n axes (characters). The general formula for the distance of two species and n characters can be written thus:

$$1,2 = \sqrt{\sum_{i=1}^{n} (X_{i1}-X_{i2})^2}$$

In the study distances between pairs of locations using varieties' genotype x location interaction effects as characters were computed to measure similarity between locations.