*Euphytica* **58**: 5–12, 1991.

# Classification of chickpea growing environments to control genotype by environment interaction

R.S. Malhotra & K.B. Singh
*International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syria*

Received 1 February 1991; accepted 17 October 1991

*Key words: Cicer arietinum*, cluster analysis, environments, G × E interaction, seed yield, time to flowering

## Summary

Cluster analysis was used as a tool to classify chickpea growing environments. Data on time to flowering (days) and seed yield (kg ha$^{-1}$) for two chickpea international yield trials developed by ICARDA and ICRISAT, and conducted by cooperating scientists during 1985–86 and 1986–87 were used for this study. The GENSTAT hierarchical, agglomerative clustering programme was employed with correlation coefficient as the distance measure and single linkage as the clustering strategy. Results revealed that by characterization of locations, the genotype × location interaction within a cluster/zone was minimized. From the classification, it appears that selection for performance at Tel Hadya – the main research station at ICARDA in Syria – should be relevant to much of Syria, the drier areas of Algeria and parts of the Iberian Peninsula. In absence of sufficient data and high degree of season-to-season variability in weather patterns it was not possible to indicate other key sites which could provide an opportunity for selection of materials for specific adaptation in a group of environments or a zone.

## Introduction

Chickpea (*Cicer arietinum* L.) is grown in diverse conditions of moisture supply, temperature, soil type, crop management, and biotic stress in West Asia, North Africa, Mediterranean Europe, and Latin America. Within these environments, it is cultivated primarily in rainfed areas receiving 350 mm to 600 mm of average annual precipitation. ICARDA (International Center for Agricultural Research in the Dry Areas), together with ICRISAT (International Crops Research Institute for the Semi-Arid Tropics), are involved in kabuli chickpea improvement. One way to improve the efficacy of a breeding programme to help national agricultural research systems in such diverse environmental conditions would be to identify relatively homogeneous regions and develop cultivars to meet the specific requirements of each region. To date, no objective sub-division of chickpea growing areas seems to have been carried out, although Malhotra et al. (1985) identified major divisions on the basis of feed-back of results from various cooperators from different parts of the world.

Mulitze et al. (1987) reviewed the various approaches to stratify environments into sub-regions. Cluster analysis based on differential grain yield responses of a set of genotypes has been the most widely used technique. Abou-El-Fittouh et al. (1969) were among the first scientists who applied

Joint contribution from ICARDA, P.O. Box 5466, Aleppo, Syria and ICRISAT (International Crops Research Institute for the Semi-Arid Tropics), Patancheru P.O., A.P. 502 324, India. ICRISAT Journal Article No. 1238.

cluster analysis to classify the cotton variety trial sites in the U.S.A. and advocated the use of the technique for stratification of environments into sub-regions. Later this technique was used by others (Byth et al., 1976; Ghaderi et al., 1980; Imrie et al., 1981; Fox et al., 1985; Imrie & Shanmugasundaram, 1987). Since such information is not available for chickpea, an attempt was made to classify the chickpea growing environments of West Asia, North Africa, Mediterranean Europe, and Latin America into regions based on differential yield responses and if possible to select a group of representative locations as key selection sites within or across countries and concentrate selection work on each group of environments independently.

## Materials and methods

Data from the Chickpea International Yield Trial – Winter-Mediterranean Region (CIYT-W-MR), which was assembled at ICARDA by the ICARDA/ICRISAT chickpea breeding programmes and distributed from Aleppo, Syria, to various cooperators during 1985–86 and 1986–87, were used. The CIYT-W-MR was grown by national programmes in a randomized complete block design with four replicates and mostly with a plot size of 4.8 m$^2$ (4 rows each, 4-m long, spaced 30 cm apart). There were 24 entries, including a local check determined by each cooperator. The entries had performed well in at least two years of yield testing at Aleppo, Syria, and also in unreplicated screening nurseries in some countries. The data on seed yield (kg ha$^{-1}$) and time to flowering (days) from 37 sites in 1985–86 and 39 sites in 1986–87 and for time to flowering from 29 sites in 1985–86 and 32 sites in 1986–87 reported by cooperators from West Asia, North Africa, Mediterranean Europe, and Latin America were used for the analyses (Table 1). Eleven of the test entries and 22 of the growing locations were common in the two seasons. A detailed description of the trial sites and entries as well as data reported by cooperators are given in the International Nursery Reports (Anon. 1988, 1989). Throughout this text, the code numbers assigned to each location in Table 1 have been used instead of location names.

The GENSTAT hierarchical, agglomerative clustering programme (Genstat 5 Committee 1988) was employed. The data analyses have shown that clusters formed on the basis of correlation account more for G × E interaction compared to normally-used Euclidean distance. Furthermore the single linkage is most frequently used technique for clustering and GENSTAT works for even negative values of correlation setting them to zero as similarity values, while average linkage as a technique does not work. Thus correlation coefficient as the distance measure and single linkage as the clustering strategy were used. Similarity values were estimated from the correlations such that correlation (r) with value 1.00 is equated with 100% similarity and $r \leq 0.0$ is equated with 0% similarity. The data on seed yield and time to flowering of each entry in each site, excluding the national or local check, were analyzed. To have a manageable number of clusters, the dendrogram was truncated or cut at an arbitrarily chosen level of similarity keeping to some extent agrometerological similarity of environments in mind.

## Results and discussion

The combined ANOVA for seed yield for 23 genotypes at 37 locations and 39 locations, during the 1985–86 and 1986–87 seasons, respectively, is given in Table 2. The mean squares due to genotypes (G), locations (L), and genotype by location (G × L) interaction were highly significant in both years. Further, the combined ANOVA for 11 common genotypes across 22 common and 30 non-common locations in two years (Table 3) also confirmed the presence of highly significant differences due to locations, genotype × location (G × L) and genotype × location × year (G × L × Y) interactions. The outcome in further analysis minimized the effect of genotype × location interaction and maximized the recovery of information on the performance of genotypes within environments.

*Table 1.* Locations, latitude, elevation, annual rainfall+, seed yield and time to flowering+ of chickpea, 1985–86 and 1986–87

| Code No. | Country | Location | Latitude (N) | Elevation (m) | Annual Rainfall (mm) | Yield (kg ha$^{-1}$) | Time to Flowering (days) |
|---|---|---|---|---|---|---|---|
| 1. | Algeria | Guelma | 36.29 | 300 | 301 | 1530 | 84 |
| 2. | Algeria | Khroub | 36.25 | 640 | 475 | 1367 | 137 |
| 3. | Algeria | Quadah | NA | NA | NA | 869 | 83 |
| 4. | Algeria | Setif | 36.09 | 1023 | 301 | 845 | 129 |
| 5. | Algeria | Sidi Bel Abbes | 35.11 | 488 | 300 | 2050 | 111 |
| 6. | Bulgaria | Toshevo | 43.40 | 236 | 561 | 1072 | 194 |
| 7. | Columbia | Surbata | 05.49 | 2540 | 527 | 649 | 87 |
| 8. | Cyprus | Laxia | 35.06 | 150 | 254 | 358 | 112 |
| 9. | France | Montboucher | 44.34 | 136 | 340 | 870 | 167 |
| 10. | France | Montpellier | 43.37 | 49 | 487 | 1839 | 177 |
| 11. | Greece | Larissa | 39.07 | 70 | 320 | 2718 | 151 |
| 12. | Iraq | Sulaimaniyah | 36.05 | 700 | NA | 1118 | 179 |
| 13. | Italy | Capalbio | NA | NA | NA | 2166 | 190 |
| 14. | Italy | Catania | 37.28 | 700 | 341 | 2333 | 149 |
| 15. | Italy | Metaponto | 40.24 | 18 | 317 | 3853 | 124 |
| 16. | Italy | Tarquinea | 42.15 | 50 | 157 | 4664 | 131 |
| 17. | Jordan | Irbid | 32.33 | 620 | 510 | 244 | 124 |
| 18. | Jordan | Marow | 32.33 | 580 | 414 | 1303 | 126 |
| 19. | Lebanon | Beqa's | 33.55 | 995 | 657 | 1546 | 150 |
| 20. | Lebanon | Terbol | 33.49 | 890 | 529 | 2319 | 144 |
| 21. | Libya | El-Safsaf | 32.49 | 580 | NA | 1090 | NA |
| 22. | Morocco | Marchouch | 33.33 | 450 | 416 | 2055 | 139 |
| 23. | Morocco | Zememra | NA | 450 | 196* | 889 | 105 |
| 24. | Pakistan | Islamabad | 33.29 | 683 | 373 | 1315 | 172 |
| 25. | Portugal | Elvas | 38.53 | 208 | 484 | 1653 | 145 |
| 26. | Portugal | Oeiras | 38.41 | 50 | 484 | 1521 | 97 |
| 27. | Spain | Badajoz | 38.49 | 219 | 178 | 2020 | 124 |
| 28. | Spain | Cordoba | 37.51 | 110 | 481 | 2065 | NA |
| 29. | Spain | Granada | 37.20 | 950 | NA | 617 | 154 |
| 30. | Spain | Madrid | 40.30 | 599 | 340 | 2172 | 155 |
| 31. | Spain | Sevilla | 37.30 | 20 | 410 | 1163 | NA |
| 32. | Syria | Al-Ghab | 35.30 | 170 | 872 | 3017 | 126 |
| 33. | Syria | Deir-Ez-Zor | 32.50 | NA | 43* | 1658 | 127 |
| 34. | Syria | Gelline | 32.80 | 421 | NA | 1647 | 124 |
| 35. | Syria | Hama | 35.08 | 316 | 324 | 2577 | 113 |
| 36. | Syria | Heimo | 37.03 | 426 | 341 | 2252 | 145 |
| 37. | Syria | Homs | 34.45 | 485 | 333 | 2495 | 130 |
| 38. | Syria | Idleb | 36.56 | 446 | 614 | 2595 | 126 |
| 39. | Syria | Izra'a | 32.51 | 575 | 405 | 1983 | 132 |
| 40. | Syria | Jableh | 35.40 | 7 | 970 | 5958 | NA |
| 41. | Syria | Jindiress | 36.24 | 210 | 505 | 3288 | 134 |
| 42. | Syria | Tel Hadya | 36.01 | 284 | 337 | 1939 | 112 |
| 43. | Tunisia | Beja | 36.52 | NA | NA | 2656 | NA |
| 44. | Tunisia | El-Kef | 36.10 | NA | NA | 1617 | NA |
| 45. | Tunisia | Mateur | 37.03 | NA | 377 | 208 | NA |
| 46. | Tunisia | Menzel Temime | 36.45 | NA | 386 | 1953 | NA |
| 47. | Tunisia | Oued Meliz | 37.55 | NA | NA | 2628 | NA |
| 48. | Tunisia | Ras Rajel | 37.21 | NA | NA | 202 | NA |
| 49. | Turkey | Adana | 37.00 | 35 | 537 | 1965 | 124 |
| 50. | Turkey | Balikhesin | 40.19 | NA | NA | 1033 | NA |
| 51. | Turkey | Diyarbakir | 37.55 | 660 | 475 | 2110 | 132 |
| 52. | Turkey | Izmir | 38.05 | 100 | 452 | 1594 | 131 |

* Additional irrigation provided, but exact amount not reported here. + when two year data were available for a location, average over two years was presented here.

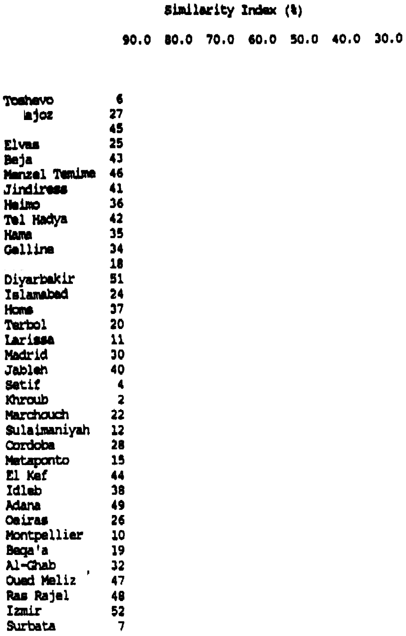

*Fig. 1.* Amalgamation of locations based on seed yield, 1985–86.

## *Cluster analysis*

The dendrogram (Fig. 1), which illustrates the outcome of cluster analysis based on yield correlations as similarity indices for the 1985–86 trials, shows that two locations, Jindiress and Heimo in Syria, were the first to form a cluster at the 85% similarity value; at the 75% similarity value, two more locations, Menzel Temime (Tunisia) and Tel Hadya (Syria) were added into this cluster. There were large differences between mean seed yields at these four locations (Table 1) indicating that location mean yield did not play an important role in grouping of environments into a cluster.

Four clusters were formed at the 55% similarity value: Cluster I with 13 locations, Elvas in Portugal (25), Beja (43) and Menzel Temime (46) in Tunisia, Jindiress (41), Heimo (36), Tel Hadya (42), Hama (35), and Gelline (34) in Syria, Marow in Jordan (18), Diyarbakir in Turkey (51), Islamabad in Pakistan (24), Homs in Syria (37) and Terbol in Lebanon (20); Cluster II with three locations, Madrid in Spain (30), Jableh (40) in Syria and Setif (4)

*Table 2.* Analysis of variance for seed yield and estimated components of variance for 23 chickpea genotypes at 37 locations in 1985–86 and 39 locations in 1986–87

| Source of variation | 1985–86 | | | 1986–87 | | |
|---|---|---|---|---|---|---|
| | df | SS ($\times 10^4$) | MS ($\times 10^4$) | df | SS ($\times 10^4$) | MS ($\times 10^4$) |
| Genotypes (G) | 22 | 1039.6 | 47.25** | 22 | 794.4 | 36.11** |
| Locations (L) | 36 | 119090.0 | 3308.00** | 38 | 65417.0 | 1721.50** |
| G × L | 791 | 6048.8 | 7.65** | 836 | 5084.4 | 6.08** |
| G × L between clusters | 131 | 1298.0 (21.4)++ | | 132 | 1389.0 (27.3) | |
| G × L within cluster 1 | 264 | 1106.3 (18.2) | | 308 | 1044.8 (20.5) | |
| G × L within cluster 2 | 44 | 520.4 ( 8.6) | | 66 | 557.0 (10.9) | |
| G × L within cluster 3 | 22 | 60.0 ( 1.0) | | 44 | 321.1 ( 6.3) | |
| G × L within cluster 4 | 22 | 24.3 ( 0.4) | | 22 | 75.2 ( 1.5) | |
| G × L within cluster 5 | 44 | 187.6 ( 3.1) | | 22 | 76.5 ( 1.5) | |
| G × L within cluster 6 | 264 | 2852.5 (47.1) | | 66 | 21.3 ( 4.2) | |
| G × L within cluster 7 | - | – | | 176 | 140.7 (27.7) | |
| Pooled Error | 2346 | 7502.20 | 3.20 | 2499 | 8194.28 | 3.28 |
| | $S^2g = 1.0704 \times 10^4$ | | | $S^2g = 0.8515 \times 10^4$ | | |
| | $S^2g \times l = 4.4492 \times 10^4$ | | | $S^2g \times l = 2.8028 \times 10^4$ | | |
| | $S^2e = 12.7915 \times 10^4$ | | | $S^2e = 13.1161 \times 10^4$ | | |
| | $S^2g : S^2g \times l = 1 : 4.2$ | | | $S^2g : S^2g \times l = 1 : 3.45$ | | |

++ Values in parentheses are the contribution to SS as per cent of SS due to G × L. ** Significance a he 0.01 level of probability. $S^2$ Variance or sigma square.

in Algeria; Cluster III with two locations, Khroub in Algeria (2) and Marchouch in Morocco (22); and Cluster IV with two locations, Cordoba in Spain (28) and Metaponto in Italy (15). The remaining ungrouped locations were assigned to two clusters – those amalgamating at 40% similarity index to Cluster V, Mateur in Tunisia (45), Larissa in Greece (11), Sulaimaniyah in Iraq (12), El-kef in Tunisia (44), Idleb in Syria (38), Adana in Turkey (49), Oeiras in Portugal (26), and the remainder to Cluster VI which included Badajoz in Spain (27), Montpellier in France (10), Beqa'a in Lebanon (19), Al-Ghab in Syria (32), Toshevo in Bulgaria (6), Oued Meliz (47) and Ras Rajel (48) in Tunisia, Izmir in Turkey (52), and Surbata in Columbia (7). In the grouping on yield, three locations, namely Surbata in Columbia (7), Ras Rajel in Tunisia (48), and Izmir in Turkey (52), were quite distinct and were amalgamated at low similarity index values.

The dendrogram based on time to flowering (Fig. 2) revealed that a large number of locations amalgamated in a similar way to that for seed yield. A few locations such as Beqa'a in Lebanon (19), Badajoz in Spain (27) and Al-Ghab in Syria (32) amalgamated at earlier stages on time to flowering (Fig. 2) than on a seed yield basis (Fig. 1). The strongest grouping for both characters, but especially for time to flowering, included mainly locations in West Asia (Jordan, Syria, Lebanon) and southern Europe (the Iberian Peninsula and Greece). As for its seed yield, Surbata was also an outlier for time to flowering.

*Table 3.* Analysis of variance for seed yield for 11 common chickpea genotypes in different locations combined over 1985–86 and 1986–87 seasons

| Source of variation | df | SS (× 10⁴) | M.S. (× 10⁴) |
|---|---|---|---|
| Genotype (G) | 10 | 744.80 | 74.48** |
| Location (L) | 51 | 74060.00 | 1452.16** |
| G × L | 510 | 3933.00 | 7.71** |
| G × L between clusters | 40 | 583.10 (14.8) | |
| G × L within cluster 1 | 210 | 1601.00 (40.7) | |
| G × L within cluster 2 | 10 | 38.73 ( 1.0) | |
| G × L within cluster 3 | 10 | 28.58 ( 0.7) | |
| G × L within cluster 4 | 50 | 504.20 (12.8) | |
| G × L within cluster 5 (ungrouped) | 190 | 1176.00 (29.9) | |
| Year (Y) | 1 | 11030.80 | 11030.80** |
| Y × L | 21 | 20710.00 | 986.19** |
| G × Y/Loc | 241 | 1210.00 | 5.02** |
| Pooled error | 2210 | 7228.58 | 3.27 |

** Significance at the 0.01 level of probability.

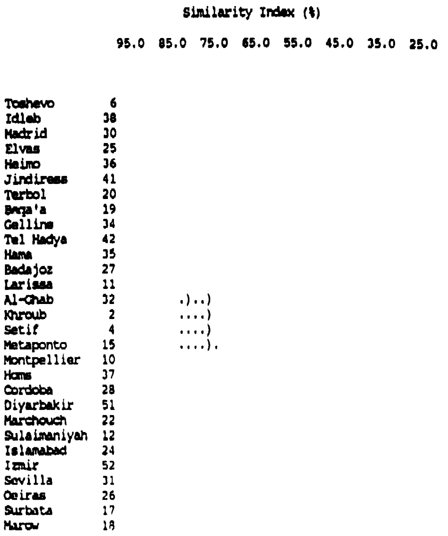

*Fig. 2.* Amalgamation of locations based on time to flowering, 1985–86.

In the seed yield data for 1986–87, clustering of locations was initiated earliest at 70% similarity values (Fig. 3) when two clusters were formed each with two locations, Terbol in Lebanon (20), and Catania in Italy (14), in the first case, and Madrid in Spain (30), and Izra'a in Syria (39), in the second. All four locations represented medium elevation sites. At the 55% similarity value, a large number of locations were grouped giving rise to five clusters: Cluster I including Terbol in Lebanon (20), Catania in Italy (14), Montboucher (9) and Montpellier (10) in France, Gelline (34), Hama (35), Homs (37) and Izra'a (39) in Syria, Madrid (30) and Sevilla (31) in Spain, Laxia in Cyprus (8), Sidi Bel Abbes in Algeria (5), Elvas in Portugal (25), Toshevo in Bulgaria (6); Cluster II including Oued Meliz in Tunisia (47), Heimo in Syria (36), El-Safsaf in Libya (21); Cluster III including Tarquinia in Italy (16), Tel Hadya (42) and Deir-Ez-Zor (33) in Syria, and Beja in Tunisia (43); Cluster IV included

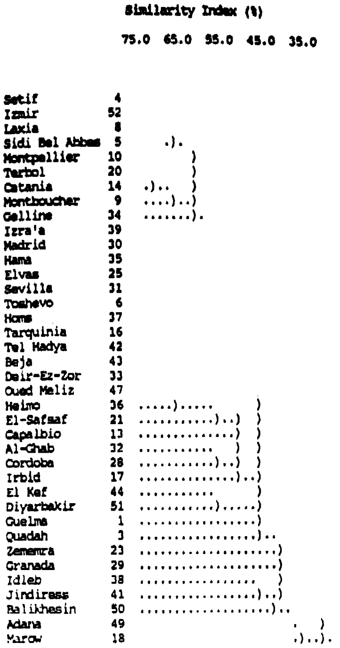

*Fig. 3.* Amalgamation of locations based on seed yield, 1986–87.

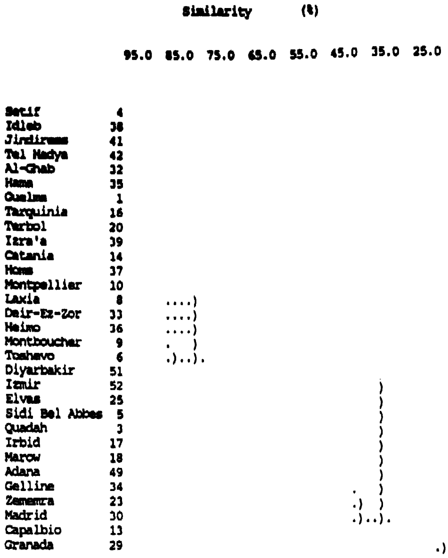

*Fig. 4.* Amalgamation of locations based on time to flowering, 1986–87.

Al-Ghab in Syria (32) and Cordoba in Spain (28); and, Cluster V included El-kef in Tunisia (44) and Diyarbakir in Turkey (51), which later merged at the 50% similarity value. Based on the amalgamation process at similarity values lower than 55%, the ungrouped locations were put into two clusters, Cluster VI included Setif in Algeria (4), Izmir in Turkey (52), Capalbio in Italy (13), Irbid in Jordan (17); and Cluster VII included, Guelma (1) and Quadah (3) in Algeria, Idleb (38) and Jindiress (41) in Syria, Zememra in Morocco (23), Granada in Spain (29), Balikhesin (50) and Adana (49) in Turkey, and Marow in Jordan (18).

The dendrogram based on time to flowering revealed that a large number of locations amalgamated in a way similar to that for seed yield (Fig. 4). However, some of locations like Jindiress (41) and Idleb (38) in Syria, amalgamated much earlier on time to flowering basis (Fig. 4) than seed yield (Fig. 3).

In general, the grouping was weaker in 1986–87 than in 1985–86, especially for time to flowering. Nonetheless, a similar pattern emerged in that the closest associations occurred amongst locations in West Asia and southern Europe. Of the 22 locations that were common to both analyses, half grouped in the same way in both years. Locations in North Africa appear distinct from those in the above regions, but showed no consistent groupings amongst themselves.

The seed yield data for the entries common to both years were pooled and a further analysis was carried out with data from all 52 locations that were used in one or both years. Four clusters formed at the 60% similarity index level (Fig. 5) and confirmed the groupings of the more diverse set of genotypes of the individual years. These clusters, encompassing 32 of the 52 locations, amalgamated at the 55% level of similarity index.

### *Analysis of variance*

When the clusters were incorporated into an ANOVA, 21, 27 and 15 percent of the total SS due to G × L interaction could be separated into a between-cluster G × L interaction in the first and second year's, and the combined years', analysis respectively (Tables 3 and 4). The corresponding figures for the proportion of the total SS remaining in the major cluster were 18, 21 and 41 percent. These values were disappointingly small in the first case and large in the third. They reflect the relatively low level of the similarity index at which clusters formed in the cluster analyses, and illustrate that while the grouping was relatively consistent, a good deal of diversity remains within the clusters.

That notwithstanding, the inclusion of clusters in the analysis of variance allowed the ratio of the 'component of genotypic variance' ($S^2g$) to the 'genotype × location interaction component' ($S^2g \times l$) to be reduced within the clusters to 1 : 1. Furthermore, in the clusters which were formed by grouping the ungrouped locations at much lower similarity indices, this ratio increased showing that the environments within these clusters were quite diverse. This suggested that environmental effects which were associated with locations in determining differential genotypic responses were reduced when relatively homogeneous locations were clustered together. Thus, by clustering, the predictability of performance of different genotypes within clusters should increase and the identification of specific types of genotypes suitable for specific clusters or zones should be possible.

From these results tentative guidelines could be proposed for selection of material with specific adaptation for inclusion in international nurseries. It appears that selection for performance at Tel Hadya, Syria, should be relevant to much of Syria, the drier areas of Algeria and parts of the Iberian Peninsula. This is our first attempt to carry out a spatial analysis of the data from these trials. Further work, and probably further information, is needed to make full use of the information in the data sets. In the course of exploration of the data, significant correlation was observed between latitude (a surrogate for photoperiod) and the time to flowering (0.4794). This indicates that photoperiod at various locations may help in characterization of the environments. Furthermore, the temperature data are unavailable for most sites, and the role of temperature in classificatory analysis couldn't be studied. But we have taken initial steps to remedy that by providing minimum-data set automatic weather stations to selected locations. Because of the limited size of the data set, and in particular its continuity, it was not possible to identify another cluster, group or zone which can help identification of other key sites.

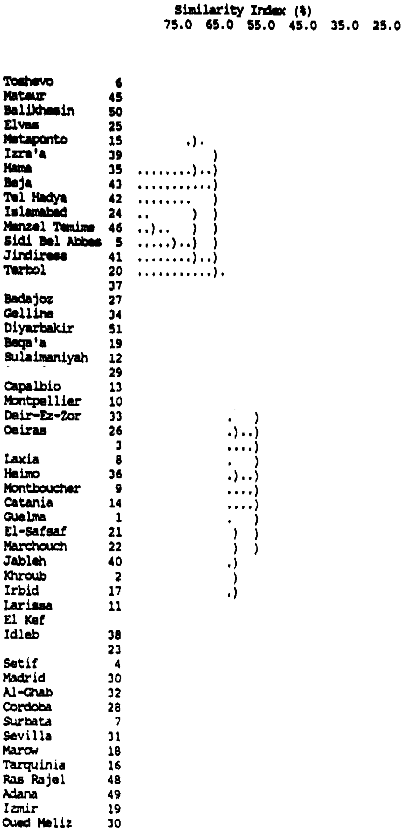

*Fig. 5.* Amalgamation of locations based on seed yield combined over years, 1985–86 and 1986–87.

## References

Abou-El-Fittouh, H.A., J.O. Rawlings & P.A. Miller, 1969. Classification of environments to control genotype by environment interaction with an application to cotton. Crop Science 9: 135–140.

Anonymous, 1988. International Nursery Report No. 10. Food Legume Nurseries 1985–86. ICARDA, Aleppo, Syria, pp 525.

Anonymous, 1989. International Nursery Report No. 11. Food Legume Nurseries 1986–87. ICARDA, Aleppo, Syria, pp 498.

Byth, D.E., R.L. Eisemann & I.H. De Lacy, 1976. Two way pattern analysis of large data set to evaluate genotype adaptation. Heredity 37: 215–230.

Fox, P.N., A.A. Rosielle & W.J.R. Boyd, 1985. The nature of genotype × environment interactions for wheat yield in Western Australia. Field Crops Research 11: 387–398.

Genstat 5 Committee, 1988. Genstat 5 Reference Manual. Oxford, Clarendon Press.

Ghaderi, A., E.H. Everson & C.E. Gress, 1980. Classification of environments and genotypes in wheat. Crop Science 20: 707–710.

Imrie, B.C., D.W. Drake, I.H. De Lacy & D.E. Byth, 1981. Analysis of genotypic and environmental variation in international mungbean trials. Euphytica 30: 301–311.

Imrie, B.C. & S. Shanmugasundaram, 1987. Source of variation in yield in international mungbean trials. Field Crops Research 16: 197–208.

Malhotra, R.S., L. Robertson, K.B. Singh, W. Erskine & M.C. Saxena, 1985. Cooperative international testing program on faba beans, kabuli chickpeas and lentils. Pages 297–313. In: M.C. Saxena & S. Varma (Eds.) Proc. Int. Workshop on Faba Beans, Kabuli Chickpeas and Lentils in the 1980s, ICARDA, Aleppo, Syria.

Mulitze, D., R.S. Malhotra & P. Goldsworthy, 1987. Relating production and environment: The International Nursery Networks. Pages 85–96. In: A.H. Bunting (Ed.), Agricultural Environments-Characterization, Classification and Mapping. C.A.B. International, U.K.