GROUPING CROP SYSTEMS BASED ON GROUNDWATER EXTRACTION

INDEX

Section	Description	Page No.
6.1	Overview	113
6.2	Technique	113
6.2.1	Procedure	113
6.3	Results	115
6.3.1	Specification of the number of clusters	115
6.3.2	Grouping cropping systems in to specified clusters	115
6.4	Discussions and policy implications	118

6.1 Overview

The significant relationship between cropping system and groundwater extraction has been well established in the previous sections. Further, the fact that the mean groundwater extraction discriminates between the cropping systems practiced in the tube well command in the watershed reaffirms the relationship. It, therefore, logically follows to group the cropping systems based on the mean extraction level in order to identify the distinct groups of crops/ cropping system to be able to manipulate the existing crop production scenario and, thus, the groundwater use for irrigation, in the watershed. The manipulation of the crops grown in the watershed can be done based on the water productivity of the crops, as discussed in the subsequent section.

6.2 Technique used

Cluster analysis was performed to identify homogeneous groups based on groundwater extraction. The analysis attempts to identify the sets of crop systems wherein, the within group variation in terms of groundwater extraction is minimum, but the variation between groups is maximum.

6.2.1 Procedure

There are two approaches to cluster analysis. *Hierarchical clustering* allows to select a definition of distance, select a linking method of forming clusters, then determines how many clusters best suit the data. In *k-means clustering*, the number of clusters is specified in advance, then cases are assigned to the K clusters.

'Hierarchical clustering' analysis was first performed to specify the appropriate numbers of clusters before using 'k-means clustering' approach to identify significantly different groups based on the different cropping systems practiced in the watershed. Dendrograms, which show the relative size of the proximity coefficients at which cases are combined, were used for specifying the number of clusters. The bigger the distance coefficient or the smaller the similarity coefficient, the more clustering involves combining unlike entities, which may be undesirable. Dendrograms are usually depicted horizontally with each row representing a case on the Y axis, while the X axis is a rescaled version of the proximity coefficients. Cases with low distance/high similarity are close together.

'k-means clustering' approach was, then, used to finalize the groups and identify the cropping systems that fall in particular groups. In this approach, initial cluster centers are chosen in a first pass of the data, then in each additional iteration group observations are based on nearest Euclidean distance to the mean of the cluster. Cluster centers change at each pass. The process continues until iteration limit is reached. Cluster centers are the average value on all clustering variables of each cluster's members. The initial cluster center gives the average value of each variable for each cluster for the k well-spaced cases when no initial file is supplied. The final cluster center gives the same thing for the last iteration step. The iteration history shows the change in cluster centers when the usual iterative approach is taken. When the change drops below a specified cutoff, the iterative process stops and cases are assigned to clusters according to which cluster center they are nearest.

There are a variety of different measures of inter-observation distances and inter-cluster distances to use as criteria when merging nearest clusters into broader groups or when considering the relation of a point to a cluster. A euclidean distance criterion was considered to measure this difference. Euclidean distance is the most common distance measure. The Euclidean distance is the square root of the sum of the square of the x distance plus the square of the y distance. *UPGMA* (un-weighted pair-group method using averages) method was used to determine how cases or clusters are combined at each step. This method computes the distance between two clusters as the average distance between all inter-cluster pairs and was considered appropriate. From the cluster membership table, which shows the number of the cluster to which the case belongs, the cropping systems belonging to particular cluster were identified.

6.3 Results

6.3.1 Specification of number of cluster

Hierarchical cluster analysis was performed, with 2003-04 data, to specify the number of clusters to be finalized. The dendrodram, which is also called tree diagram, confirmed that five numbers of clusters would be appropriate for the data set (Fig.6.1). Cases showing lowest distance are closest, showings the line linking them to a short distance from the left of the dendogram. This indicated that they are agglomerate into a cluster at a low distance coefficient, indicating alikeness. Five clusters were considered to be appropriate for grouping the data.

6.3.2 Grouping cropping systems in to specified clusters

'k-means' cluster analysis was done to group the cropping systems. The three years differed in terms of rainfall and mean groundwater extraction. The cluster analysis was performed for individual years and also for the pooled data to examine the change in cluster formation with respect to the mean groundwater extraction.

Cropping system groups, 2003-04

The initial cluster centers for the two variables are given in Table 6.1.

With the initial cluster centers, five iterations were performed for convergence (Table 6.2). Convergence was achieved due to small distance change. The minimum distance between initial centers was 419.43.

Based on the Euclidean distance between the mean groundwater extraction, cluster membership data set were obtained. The final cluster centres and the distance between final clusters computed at the end of analysis are given in Tables 6.3 and 6.4, respectively.

Analysis of variance test was performed for the groundwater extraction variable (Table 6.5). Though, the F tests are used only for descriptive purposes as the clusters have been chosen to maximize the differences among cases in different clusters, yet the significant variables indicate that they contribute to the differentiation of clusters.

The final grouping is depicted in Figure 6.2. Depending upon the number of cases falling in a group, the particular cropping system was considered to be a member of that group, where majority of its members are clustered. Cotton based cropping systems formed separate groups based on the individual crops that make the cropping system. Group 2 consisted of cropping systems such as 'pure cotton', 'cotton_castor-fennel', 'cotton- paddy-wheat' and 'cotton-fennel'. Some cases of 'pure cotton' clustered in group 1 also but such numbers were only a few. Hence, this crop was considered to be a member of group 2. Likewise, though double cropping systems where cotton along with castor, paddy in kharif is followed by fennel and wheat in rabi season, appeared in group 2, actually belong to group 5 based on more number of cases clustering in that group. In other words, cotton-fennel, cotton-cumin, cotton-sunflower form one group, that is group 5. The mono crops such as paddy, fennel, summer pearl millet and sunflower formed another group.

Cropping system groups, 2004-05

Similar analysis was also performed for the year 2004-05. Initial cluster centers are depicted in Table 5.6.

Convergence was achieved in three iterations (Table 6.7). The minimum distance between initial centers was 411.67. Final cluster centers and distance between them are given in Tables 6.8 and 6.9, respectively. The analysis of variance test revealed that groundwater extraction significantly (P < 0.05) affected the cluster formation.

The final clusters of cropping systems are as depicted in Figure 6.3. Cotton mono crop, cotton_castor and cotton-cumin formed group 1, cotton-fennel and cotton-wheat formed group 3. Mono crop like cumin, summer pearl millet formed a separate group 5. The other crop systems were distributed in different groups and hence, these could not be made member of any particular group.

Cropping system groups, 2005-06

With five initial clusters (Table 6.11), the analysis again converged in 4 iterations (Table 6.12) with final cluster centers and the distance between them as given in tables 6.13 and 6.14, respectively. The F tests showed (Table 6.15) that the groundwater extraction variable significantly (P < 0.05) explained the grouping among the cropping systems.

The cropping systems finally grouped are depicted in Figure 6.4. The number of crop combinations sown by farmers was more in this year as compared to the crops grown in the previous normal years. A favourable moisture condition in soil and the assured availability of groundwater for meeting deficit irrigation requirements explained this behaviour on the part of farmers. Most of the cotton based cropping systems were clustered into groups 1 and 2. Group 1 had higher number of crop combinations as compared to group 2, which had mostly double cropping. Group 4 consisted of mono crops, whether taken in kharif or rabi.

6.4 Discussion and policy implications

The cluster analysis revealed that cotton based cropping systems clearly dominated in all the years. The results are summarized as under,

1) In a year of normal rainfall (2003-04 and 2004-05), the mono crops such as paddy, fennel, summer pearl millet and sunflower formed one group, while double cropping system such as cotton-fennel, cotton-cumin, cotton-sunflower formed another group. Cotton mono system was altogether a separate group.

2) The exceptionally high rain fall year (2005-06) revealed a different trend. Firstly, the number of crop combinations tried by farmers increased as compared to the crops grown in the previous normal year, though as in the normal rainfall year, cotton based cropping systems formed a distinct group. Though most of the mono crop and double crops formed separate groups, some were distributed in different groups.

3) The analysis of pooled data (2003 – 05) indicated that cotton mono crop, cotton-fennel, cotton-cumin and cotton_castor clearly dominated and formed one group. These crops and crop combinations are distinct in groundwater extraction also. Therefore, manipulation of irrigation water use in these crops, based on water productivity, has implications for water saving in the watershed.

Table 6.1. *Initial Cluster Centers, 2003-04*

Description	Cluster				
	1	2	3	4	5
G.WATER	517.30	1577.32	.00	936.64	2213.89
CROPS	10	6	0	1	6

Table 6.2. *Iteration History, 2003-04*

Iteration	Change in Cluster Centers				
	1	2	3	4	5
1	140.39	48.07	80.76	112.30	121.23
2	35.72	26.49	8.30	0.00	76.34
3	58.46	0.00	42.15	0.00	0.00
4	13.85	0.00	19.06	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00

Table 6.3. *Final Cluster Centers, 2003-04*

Description	Cluster				
	1	2	3	4	5
G.WATER	269.17	1598.79	11.17	1048.88	2016.33
CROPS	10	2	2	5	8

Table 6.4. *Distances between Final Cluster Centers, 2003-04*

Cluster	1	2	3	4	5
1		1329.64	258.11	779.73	1747.16
2			1587.62	549.91	417.59
3				1037.72	2005.17
4					967.45
5					

Table 6.5. *ANOVA F test for the two variables, 2003-04*

Description	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
G.WATER	10497640.02	4	8560.36	74	1226.30	.000
CROPS	220.63	4	28.32	74	7.73	.000

Table 6.6. *Initial Cluster Centers, 2004-05*

Description	Cluster					
	1	2	3	4	5	6
Groundwater extraction	1467.31	981.58	570.45	.00	5290.53	2163.02
Cropping system	2	2	23	0	12	10

Table 6.7. *Iteration History, 2004-05*

Iteration	Change in Cluster Centers					
	1	2	3	4	5	6
1	101.554	38.586	45.257	37.928	.000	143.214
2	9.010	16.648	.000	.000	.000	.000
3	.000	.000	.000	.000	.000	.000

Table 6.8. *Final Cluster Centers, 2004-05*

Description	Cluster					
	1	2	3	4	5	6
Groundwater extraction	1356.79	926.91	525.42	37.69	5290.53	2019.88
Cropping system	5	9	19	4	12	6

Table 6.9. *Distances between Final Cluster Centers, 2004-05*

Cluster	1	2	3	4	5	6
1		429.899	831.481	1319.100	3933.751	663.091
2			401.604	889.231	4363.625	1092.976
3				487.939	4765.117	1494.516
4					5252.850	1982.191
5						3270.660
6						

Table 6.10. *ANOVA F test for the two variables, 2004-05*

Description	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Groundwater extraction	11059143.41	5	10895.15	73	1015.05	.000
Cropping system	180.694	5	51.370	73	3.518	.007

Table 6.11. *Initial Cluster Centers, 2005-06*

Descroption	Cluster				
	1	2	3	4	5
Groundwater extraction	10000.00	.00	3300.00	13440.00	6432.00
Cropping system code	5.00	.00	29.00	11.00	3.00

Table 6.12. *Iteration History, 2005-06*

Iteration	Change in Cluster Center				
	1	2	3	4	5
1	693.35	517.62	74.53	0.00	109.61
2	0.00	47.50	49.24	0.00	0.00
3	0.00	48.38	48.47	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00

Table 6.13. *Final Cluster Centers, 2005-06*

Description	Cluster				
	1	2	3	4	5
Groundwater extraction	10693.33	613.50	3325.79	13440.00	6322.84
Cropping system code	10.33	3.54	9.29	11.00	13.00

Table 6.14. *Distances between Final Cluster Centers, 2005-06*

Cluster	1	2	3	4	5
1		10079.836	7367.548	2746.667	4370.492
2			2712.292	12826.502	5709.350
3				10114.214	2997.059
4					7117.158
5					

Table 6.15. *ANOVA F test for the two variables, 2005-06*

Description	Cluster Mean Square	df	Error Mean Square	df	F	Sig.
Groundwater extraction	159893698.36	4	795635.31	74	200.96	0.000
Cropping system code	275.35	4	63.01	74	4.36	0.003

Rescaled Distance Cluster Combine

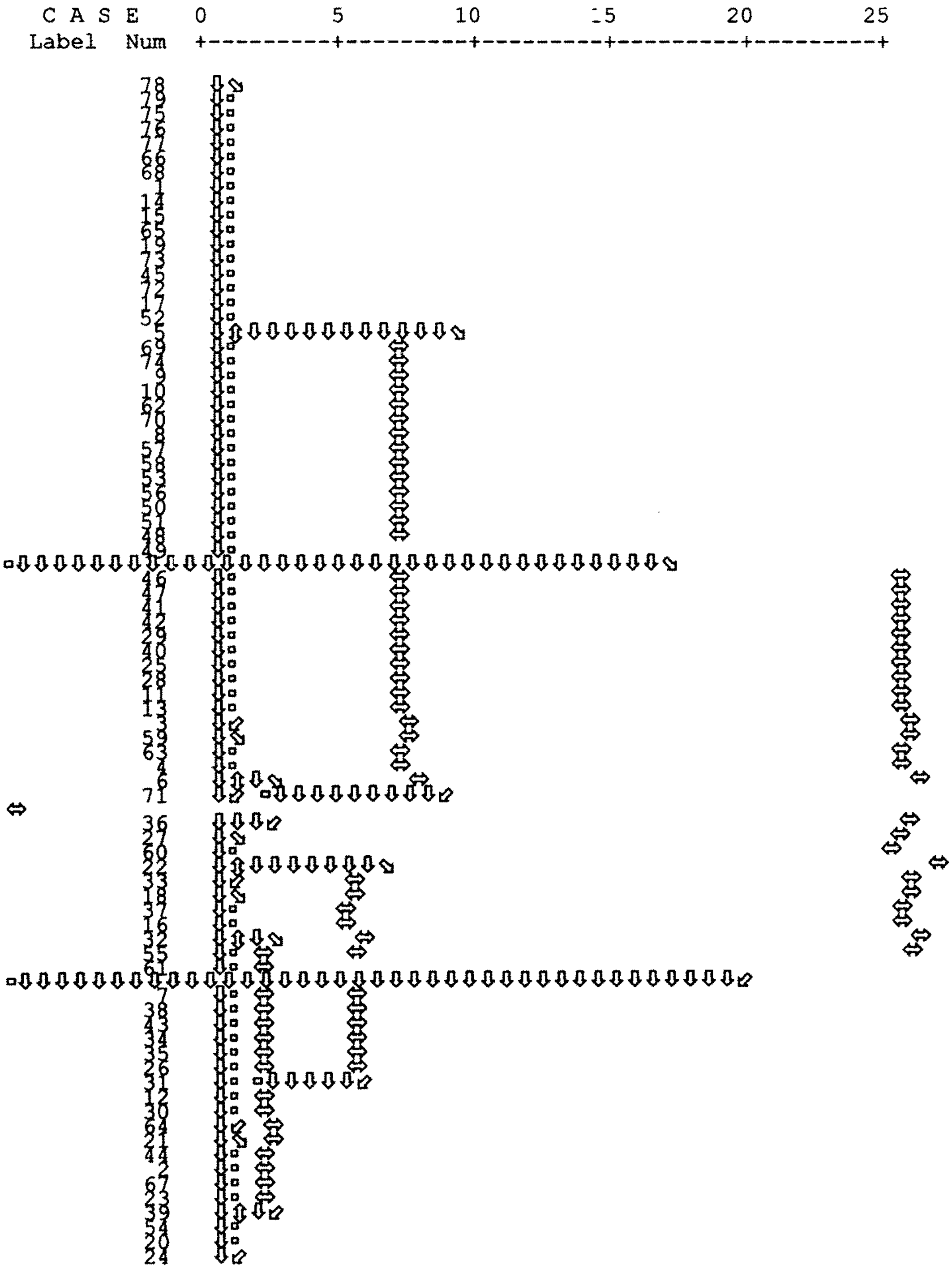

Fig. 6.1.*Dendrogram using average linkage between Groups with scaled distance*

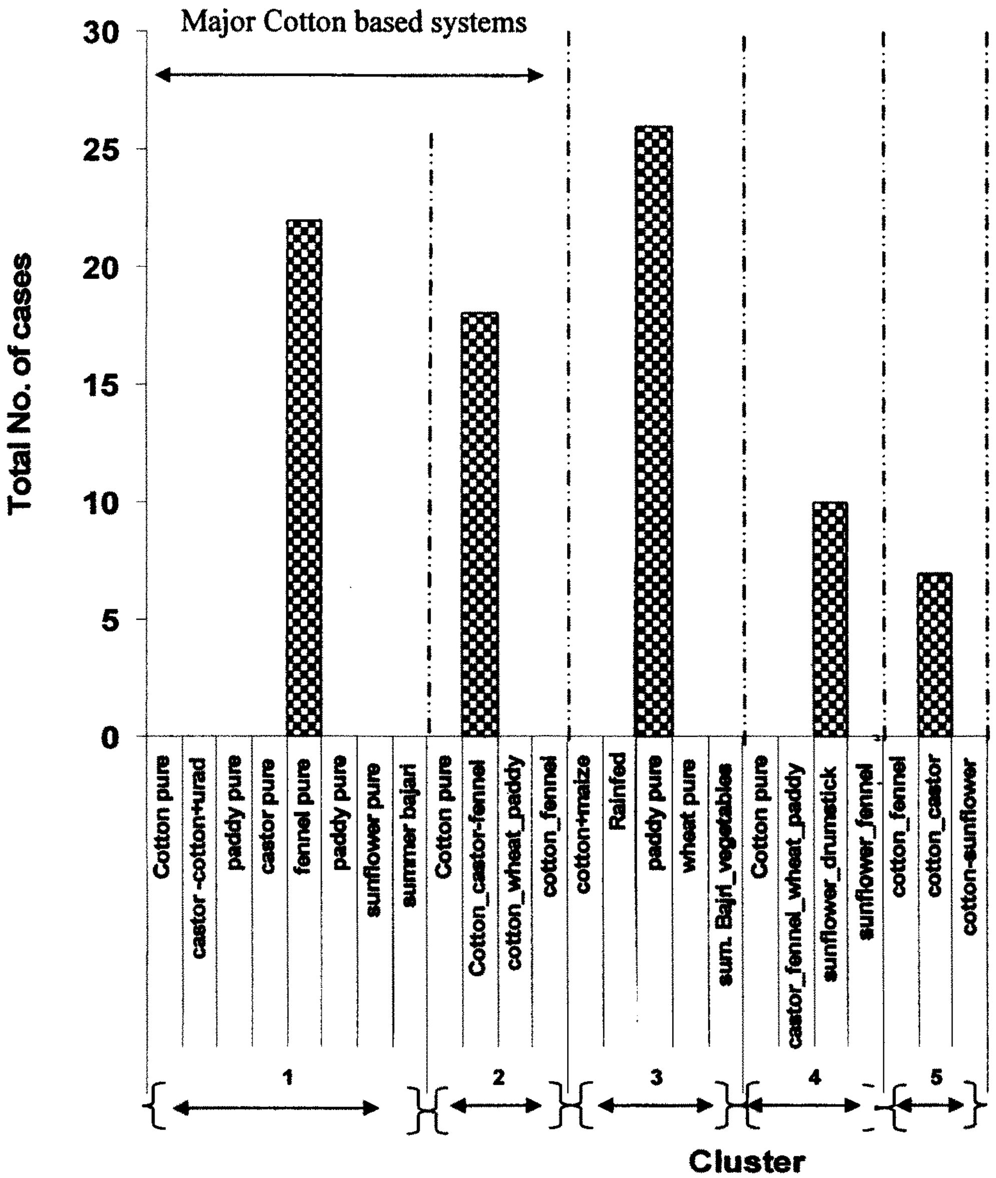

Fig. 6.2. *Clusters of cropping systems practiced in tube well command based on mean groundwater extraction from the well, 2003-04*

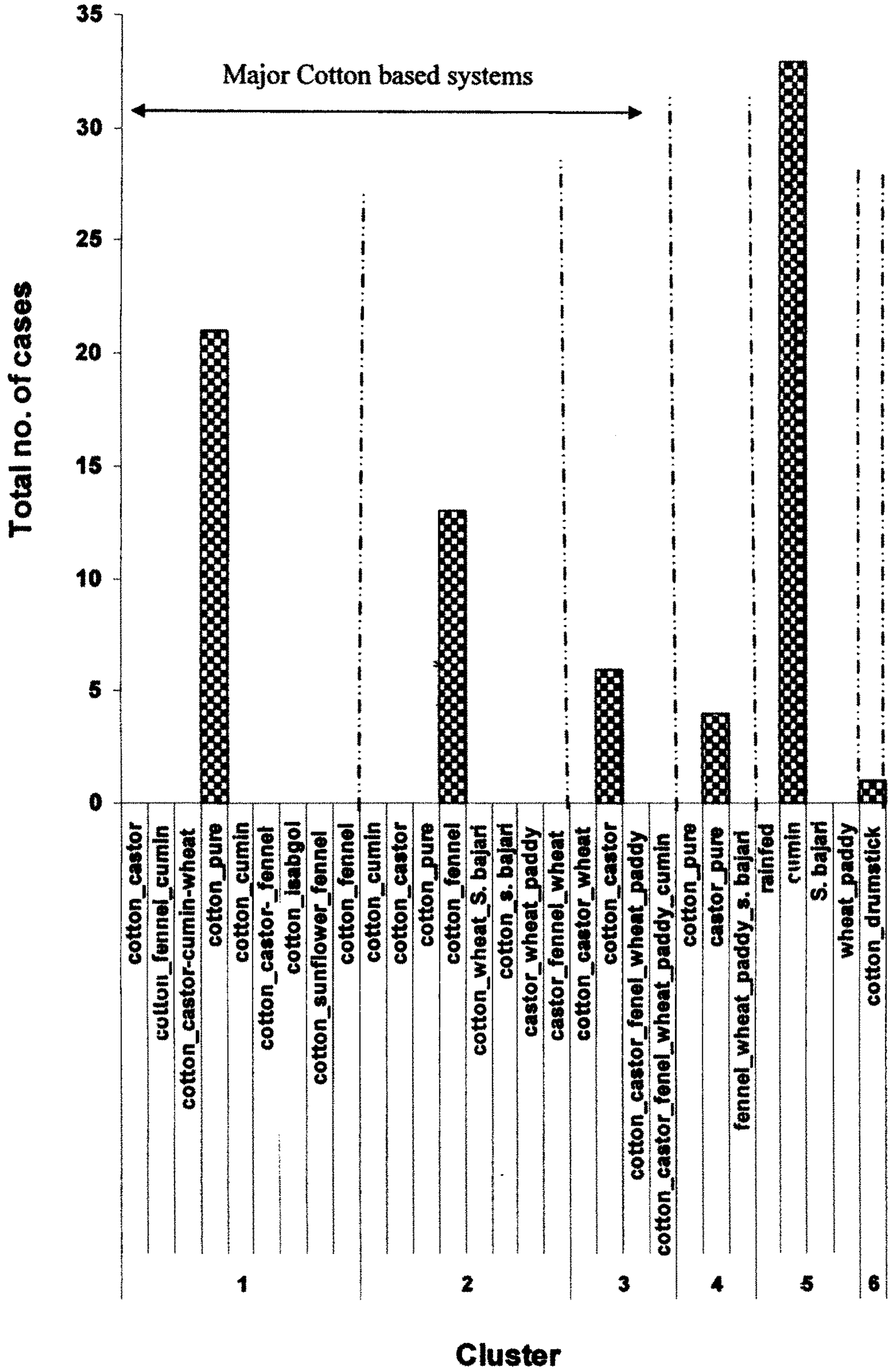

Fig. 6.3. *Clusters of cropping systems practiced in tube well command based on mean groundwater extraction from the well, 2004-05*

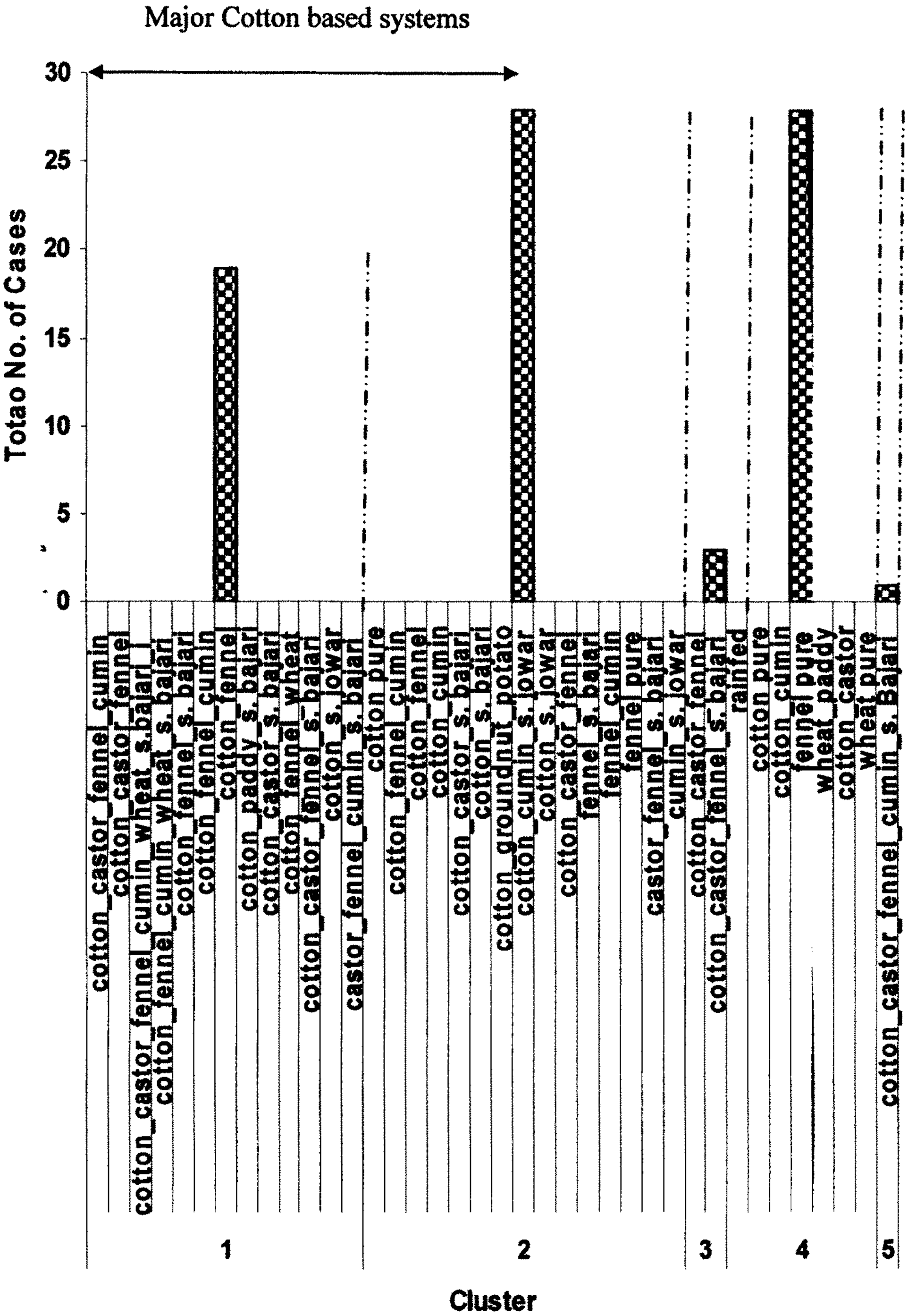

Fig. 6.4. *Clusters of cropping systems practiced in tube well command based on mean groundwater extraction from the well, 2005-06*

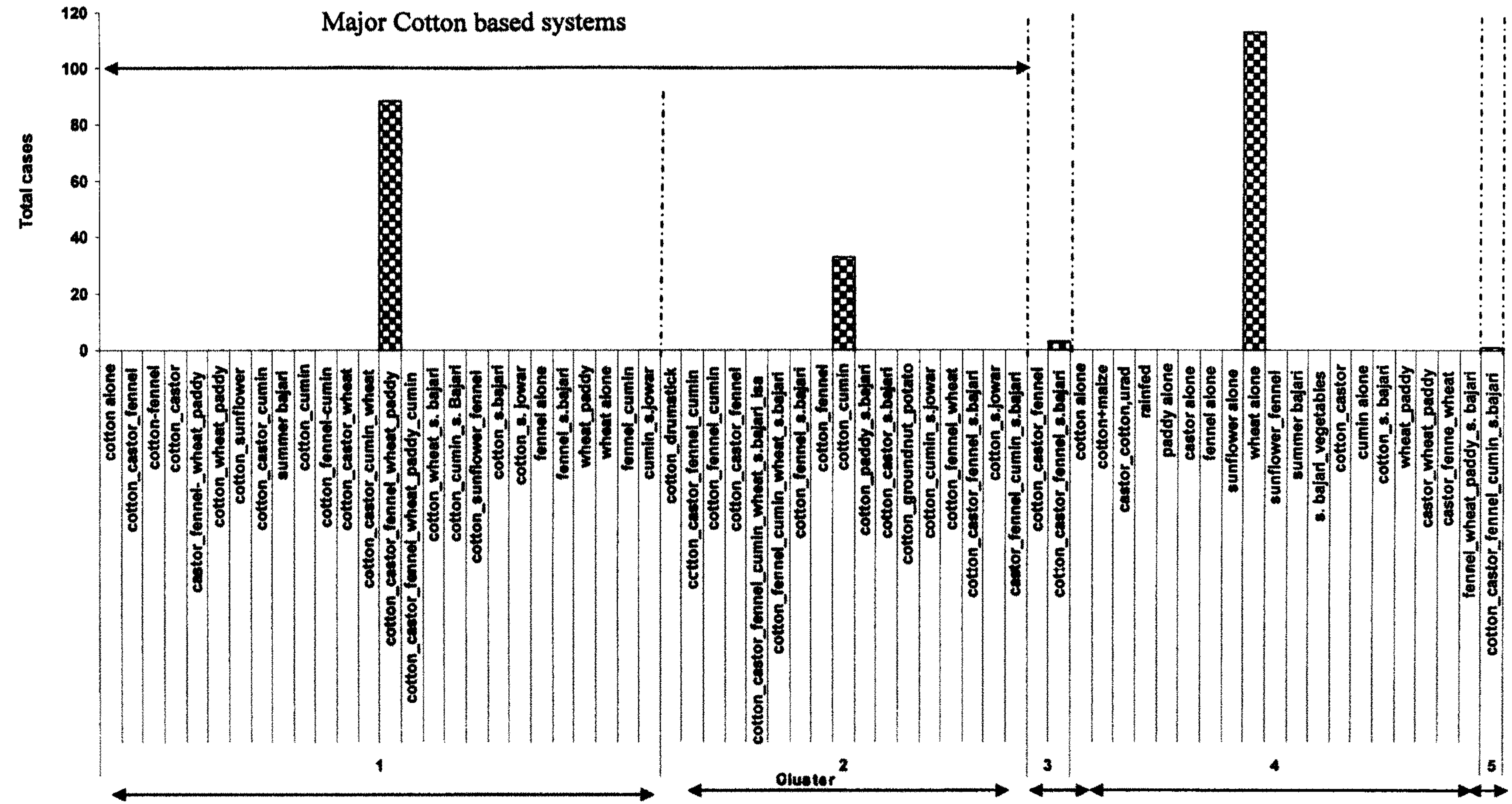

Fig. 6.5. *Clusters of cropping systems practiced in tube well command based on mean groundwater extraction from the wells, 2003-05 pooled data*