International Journal of Tropical Agriculture
Volume 42, Numbers (1) • 2024 : pp. 17-26
ISSN: 0254-8755

Statistical Analysis to form Clusters based on Area, Production and Productivity for Major Crops in Karnataka

G. Padmapriyadarshini and H. S. Surendra
GKVK, UAS, Bangalore. E-mails: Padmapriya964@gmail.com , hssurendra@rediffmail.com

Abstract: To analyze the factors influencing the diversification in agriculture, eighteen related variables were considered. Multiple linear regression analysis result showed factors influencing the diversification were rainfall, area sown, area irrigated, rural banks and literacy rate were found significant for both the periods

Cluster analysis was carried out based on area, production, productivity of selected crops. For first period cropped area of sorghum, cotton, paddy, groundnut, and ragi showed similarity in area. For second period paddy, maize, mango, ragi and groundnut had similarity in area. Sugarcane and sorghum had similarity in production for first period and sugarcane alone for second period. Based on productivity major cluster formed by horticulture crops.

INTRODUCTION

Karnataka state forms the South Western part of the Deccan Peninsula and lies between 11.5°and 18.6° North latitude and 74.0° and 78.4° East longitudes. It is the eighth largest state in the country having an area of 1,91,791 Sq. Kms (6.25% of India's total area of 3,065,027 Sq.Kms,). As per the census of 2011, the State has a total population of 6.12 crores accounting for 5.13 per cent of the country's total population of 121.21 crores. 72 per cent of the total population resides mainly in rural areas, whose main occupation focusing on agriculture and allied activities. Out of the total population, 44.6 per cent concise to working population, of which 69.36 lakhs are cultivators and 62.09 lakhs are agricultural labourers. One important feature, of agricultural labourers is that the percentage of women (58.19%) overrides the percentage of men (41.81%). The literacy rate of the State found to be 67.04 per cent, while in urban areas it is 81.05 per cent, which is comparatively higher than rural areas (59.68%). The State has totally 30 districts, 176 taluks, 745 hoblies, 29,483 Villages (27,575 inhabited and 1908 uninhabited) and 5692 grama panchayaths (KARNATAKA AGICULTURAL POLICY, (2006) Farmer Centric Agriculture Policy Document, 2010-2011).

As per the Agricultural Census of 2005-2006, the State has about 123.85 lakh hectares of cultivable area out of total geographical area of 190.50 lakh hectares, accounting for 65.0 per cent. The total number of operational holdings is 75.89 lakh with 1.63 hectares, as average size operational land holding. Small and marginal farmers account for 72.9 per cent of the total holdings, cultivating only 34.4 per cent of the total cultivable area. The number of holdings increased by 8.58 lakh due to fragmentation of the land in the last five years. As result the average size of holding has decreased from 1.74 hectares to 1.63 hectares. Out of the total cultivable area of 123.85 lakh hectares, as per the statistics of 2005-06, the gross cultivated area was 116.70 lakh hectares and the net cultivated area was 100.31 lakh hectares, indicating a cropping intensity of 116 per cent. Out of the gross cultivated area, the area under irrigation was 30.89 lakh hectares (26.5%). The State receives normal annual rainfall of 1139 mm, mainly through southwest monsoon (June to September – 806 mm) and Northeast

Article History: Received: 12 November 2023; Revised: 02 December 2023; Accepted: 11 January 2024; Published: 25 March 2024

monsoon (October to December – 195 mm). Accordingly, the state has three agricultural seasons – Kharif (April to September), Rabi (October to December) and Summer (January to March). Agricultural crops are cultivated in an area of about 107 lakh hectares annually. Out of which, in Kharif season is around 69 lakh hectares (64%), in Rabi season is around 32 lakh hectares (30%) and the rest 6 lakh hectares (6%) during summer season. Out of gross cultivated area of agricultural crops an area of about 22 lakhs hectares (20.5%) comes under irrigation. Karnataka State with a food grains production of about 138.25 lakh tones contributing nearly 5 per cent to the national food grains production.

REVIEW OF LITERATURE

Brett *et al.* (1987) examined that the Regression and Cluster Analysis of Environmental Responses of Hybrid and Pure line Winter Wheat Cultivars1. And he observed that the commercial development and release of F1 hybrid wheat (Tritium aestivum L.) cultivars should enable wheat producers to achieve higher levels of grain production. This study was conducted to compare yield responses of hybrid and pure line cultivars across a wide range of environments. Regression and cluster analyses were applied to data obtained from four winter wheat performance trials from 1982 to 1985. A different set of either 30 or 40 genotypes with approximately half of which were hybrids was evaluated each year in Oklahoma. Because the environmental responses of all genotypes were highly correlated in each year, environment means were employed as the index in the regression analysis. In 1982 the average regression coefficient (b) for hybrids (1.12) was significantly greater than that for either semi dwarf pure lines (0.99) or tall pure lines (0.79), indicating eater responsiveness of hybrids to improving production conditions. The classification of genotypes into homogeneous-response clusters according to actual responses indicated high similarity between environmental responses of hybrids and semi dwarf pure lines.

Mo Huidong (1987) studied that the cluster analysis for agronomic characters of barley varieties in Jiangsu-zhejiang shanghai area and noticed that the hierarchical and dynamic clustering were conducted for 100 barley varieties in the area of Jiangsu and Zhejiang provinces and Shanghai district. Based on the Euclidean distances of principal components among varieties, these samples were divided into several groups according to the number of characters participated in clustering. The results showed that in clustering for single, two or three characters, each group had its explicit features. It is helpful to use them selectively in plant breeding. As more characters were included, the difference of some characters between groups became indistinct and the outcome tended to be grouped under natural types. On the other, the more the characters participated hand in, the more stable the grouping. Thus it may be necessary to include more characters in the classification of germplasm resource.

Cali ski and Harabasz (1992) studied that the A Dendrite Method for Cluster Analysis and they observed that the method for identifying clusters of points in a multidimensional Euclidean space was described and its application to taxonomy considered. It reconciles, in a sense, two different approaches to the investigation of the spatial relationships between the points, viz., the agglomerative and the divisive methods. A graph, the shortest dendrite of Florek et al. (1951), was constructed on a nearest neighbor basis and then divided into clusters by applying the criterion of minimum within-cluster sum of squares. This procedure ensures an effective reduction of the number of possible splits. The method may be applied to a dichotomous division, but was perfectly suitable also for a global division into any number of clusters. An informal indicator of the "best number" of clusters was suggested. It was a "variance ratio criterion" giving some insight into the structure of the points. The method was illustrated by three examples, one of which was original. The results obtained by the dendrite method are compared with those obtained by using the agglomerative method of Ward (1963) and the divisive method of Edwards and Cavalli-Sforza (1965).

Qian Shiqiang and Guo Jiaming (1994) showed that the application of Cluster analysis for dividing flue-cured tobacco growing regions. The result examined that the by systematical clustering method,18 testing locations in Huang-Huai area were divided into 8 sub regions in the light of the values which came from 7 flue-cured tobacco varieties. The condensation point was established according to the results of the variance analysis in the sub regions. The differences among the testing locations were more significant than the ones among the testing locations in the same sub region. The negative correlation between the averages of values and the averages of CV in the 8 sub regions was estimated to be r =-0.7526. The conclusion could be of some values and interest for dividing growing regions, the layout and introduction of Variety using Cluster analysis.

Ana Iglesias *et al.* (2000) noticed that the Agricultural impacts of climate change in Spain: developing tools for a spatial analysis. The study examined that the CERES-Wheat, a dynamic process crop growth model, was specified and validated for seven sites in the major wheat-growing regions of Spain. Variables explaining a significant proportion of simulated yield variance are crop water and temperature during the growing season. A multiple linear regression model was developed to represent simulated yield response to these variables. Seven agro-climatic regions are defined based on K-mean cluster analysis of temperature and precipitation data from 329 meteorological stations and provincial crop yield data. The yield functions derived from the validated crop model were then used with the gridded agro-climatic data base to conduct a spatial analysis of climate change impacts on national wheat production. Climate change scenarios with and without sulfate aerosols developed from the Hadley Centre (HCGG and HCGS) and Canadian Climate Centre (CCGG and CCGC) are tested.

Keim *et al.* (2000) notified that Evaluation of soybean RFLP marker diversity in adapted germ plasma. And he observed that the Soybean RFLP markers have been primarily developed and genetically mapped using wide crosses between exotic and adapted genotypes. We have screened 38 soybean lines at 128 RFLP marker loci primarily to characterize germ plasma structure but also to evaluate the utility of RFLP markers identified in un adapted populations. Of these DNA probes 70% detected RFLPs in this set of soybean lines with an average polymorphism index of 0.30. This means that only 1 out of 5 marker loci was informative between any particular pair of adapted soybean lines. The variance associated with the estimation of RFLP genetic distance (GDR) was determined, and the value obtained suggested that the use of more than 65–90 marker loci for germ plasma surveys will add little precision. Cluster analysis and principal coordinate analysis of the GDR matrix revealed the relative lack of diversity in adapted germ plasma.

Machado *et al.* (2002) studied that the Spatial and Temporal Variability of Corn Growth and Grain Yield. The result noticed that Effects of factors influencing spatial and temporal variability of crop yields are usually expressed in crop growth patterns. Consequently, monitoring crop growth can form the basis for managing site-specific farming (SSF). This experiment was conducted to determine whether crop growth patterns forecast grain yields. Effects of irrigation rates, elevation, soil texture, soil NO3-N, arthropods, and diseases on corn growth and grain yield were evaluated in 1998 and 1999. Data on plant height, leaf area index, leaf dry matter, stem dry matter, and ear dry matter. Grain yields at DGPS locations were classified into four distinct clusters. In 1998 a dry season, clusters were strongly influenced by elevation and soil texture. Grain yields were higher at high elevations where water use was high and soil texture was heavy compared with low elevations. Grain yields at low elevations also were reduced by common smut that preferred dry conditions

Jaynes *et al.* (2003) observed that Cluster Analysis of Spacio-temporal Corn Yield Patterns in an Iowa Field. This study to determine cluster analysis used to decipher the temporal and spatial patterns of corn yield within a field. Non-hierarchal cluster analysis was applied to 6 years of corn yield data collected for 224 yield plots on a regular grid on the southern half of a 32-ha field. We were able to group the yield observations into five temporal yield patterns or clusters. The clusters were not randomly distributed across the field but instead formed contiguous areas roughly equivalent to landscape positions. Cluster membership was determined primarily by yield differences in years with growing season precipitation greater than the 40-years average. A multiple discriminate analysis was used to predict the spatial occurrence of the clusters from easily determined field attributes: soil electrical conductivity, elevation, slope, plan and profile curvature. The multiple discriminate functions were unable to distinguish between the two clusters located on the lowest portions of the landscape. Because of similar temporal yield patterns in these two clusters, they were combined and the multiple discriminate analyses repeated for four clusters. The result achieved 76 and 80% success rates in classifying the yield plots in to the correct yield clusters.

Hardeman (2003) noticed that use of cluster analysis for identification and classification of farming systems in Qing yang county, central north China. The result examined that the Cluster analysis by the average linkage method was used for the identification and classification of farming systems in Gansu on the Loess Plateau of central north China. Data from 26 townships describing socio-agricultural parameters of farm households were standardized and a Pearson correlation matrix was constructed. Parameters with the highest correlation coefficients and largest standards of deviation were selected for cluster analysis. These were: proportion of slope land, proportion of valley land, proportion of autumn sown crops per household, autumn sown crop yield, grain yield (kg/ha), number of draught animals per unit area and number of sheep per household. The resulting Dendrogram identified 3 major farming systems, topography being the principal component for discerning the major groups. Subsequent farm surveys confirmed differences between farming system groups I and III.

METHODOLOGY

Cluster analysis

Cluster analysis is usually done in an attempt to combine cases into groups when the group membership is not known prior to the analysis. Cluster analysis is a technique for grouping individual or objects into unknown groups. In biology, cluster analysis has been used for decades in the area of taxonomy, where living things are classified into arbitrary groups on the basis of their characteristics group. The classification proceeds from the most general to the most specific in steps. The most general classification was kingdom followed by phylum, subphylum, and class etc. (Richard A Johnson and Dean W. Wichern, 2002)

Clustering Methods: The commonly used methods of clustering fall into two general categories. (i) Hierarchical (ii) Non hierarchical. Hierarchical clustering techniques proceed by either a series of mergers or a series of successive divisions.

Agglomerative vs Divisive Clustering: Once the measures of association as well as the method for determining the distances between clusters have been considered.

Agglomerative Clustering (Leaves to trunk)

- Start with all sample units in '*n*' clusters of size one.
- Further at each step of the algorithm, the pair of clusters with the shortest distance combined into a single cluster.
- The algorithm stops when all sample units are combined into a single cluster of size *n*.

Divisive Clustering (Trunk to leaves)

- Start with **all** sample units in a single cluster of size *n*.
- Further at each step of the algorithm, clusters are partitioned into a pair of daughter clusters, selected to maximize the distance between each daughter.
- The algorithm stops when sample units are partitioned into '*n*' clusters of size one.

Measures of Association for Continuous Variables

The following standard notations are used all along

X_{ik} = Response for variable *k* in sample unit *i* (the number of individual species *k* at site *i*)

n = Number of sample units

p = Number of variables

Johnson and Wichern (1998) list four different measures of association (similarity) that are frequently used with continuous variables in cluster analysis.

Euclidean Distance: Euclidean distance is the most commonly used. For instance, in two dimensions, one can plot the observations in a scatter diagram, and measure the distances between the pairs of points. Generally the following equation can be used:

$$d(X_i, X_j) = \sqrt{\sum_{k-1}^{p}(X_{ik} - X_{jk})^2}$$

For each variable *k*, take the difference between the observations for sites *i* and *j*. These differences are then squared, and summed over *p* variables. This gives us the sum of the squared difference between the measurements for each variable. Finally, take the square-root of the result.

Measuring Association d_{12} Between Clusters 1 and 2

After determining the measurement of association between the subjects, the next thing to look at measuring the association between the clusters that may contain two or more members. There are multiple approaches that one can take. Methods for measuring association between clusters are called linkage methods.

Notations used:

$X_1, X_2, ..., X_k$ = Observations from cluster 1

$Y_1, Y_2, ... , Y_1$ = Observations from cluster 2

D(**x,y**) = Distance between a subject with observation vector **X** and a subject with observation vector **Y**

Table 1.1: Linkage Methods or Measuring Association d_{12} Between Clusters 1 and 2

Centroid Method	$d_{12} = d(\overline{x}, \overline{y})$	This involves finding the mean vector location for each of the clusters and taking the distance between these two centroids.
Single Linkage	$d_{12} = \min_{i,j} d(X_i, Y_j)$	This is the distance between the closest members of the two clusters.
Complete Linkage	$d_{12} = \max_{i,j} d(X_i, Y_j)$	This is the distance between the farthest apart members.
Average Linkage	$d_{12} = \frac{1}{kl}\sum_{i-1}^{k}\sum_{j-1}^{l} d(X_i, Y_j)$	This method involves looking at the distances between all pairs and averages all of these distances. This is also called UPGMA - Unweighted Pair Group Mean Averaging.

Ward's Method

It is an alternative approach for performing cluster analysis. Basically, it looks at cluster analysis as an analysis of variance problem, instead of using distance metrics or measures of association.

This method involves an agglomerative clustering algorithm. It will start out at the leaves and work its way to the trunk, so to speak. It looks for groups of leaves that it forms into branches, the branches into limbs and eventually into the trunk. Ward's method starts out with *n* clusters of size 1 and continues until all the observations are included into one cluster.

This method is most appropriate technique for analysing quantitative variables, and not for binary variables. Based on the notation that clusters of multivariate observations should be approximately elliptical in shape, it is generally assumed that the data from each of the clusters will be realized in the pattern of multivariate distribution. Therefore, it follows elliptical shape when plotted in a *p*-dimensional scatter plot.

Notations used are as follows: X_{ijk} denote the value for variable '*k*' in observation '*j*' belonging to cluster '*i*'.

Further, under this method the following concepts are defined:

1. **Error Sum of Square (ESS)**

$$ESS = \sum_i \sum_j \sum_k (X_{ijk} - \overline{X}_{i,k})^2$$

Summing over all variables and all of the units within each cluster and comparing the individual observations for each variable against the cluster means for that variable. If the Error Sum of Squares is small, then it suggests that data are closer to their cluster means, implying that a cluster of like units.

2. Total Sum of Squares (TSS)

$$ESS = \sum_i \sum_j \sum_k (X_{ijk} - \bar{X}_{i,k})^2$$

The total sum of squares is defined in the similar way of ESS. Further, comparing the individual observations for each variable against the grand mean.

3. R-square (R^2)

$$R^2 = \frac{TSS - ESS}{TSS}$$

The R^2 value is interpreted as the proportion of variation explained by a particular clustering of the observations.

Using Ward's Method indicate with all sample units in n clusters of size 1 each. In the first step of the algorithm, (n – 1) cluster are formed, one of size two and the remaining of size one. The error sum of squares and R^2 values are then computed. The pair of sample units that yield the smallest error sum of squares, or equivalently, the largest R^2 value will form the first cluster.

In the second step of the algorithm, (n – 2) clusters are formed from that (n – 1) clusters defined in step 2. This may include two clusters of size two or a single cluster of size three comprising of two items clustered in step 1. Again, the value of R^2 is maximized. Thus, at each step of the algorithm clusters or observations are combined in such a way to minimize the results of error from the squares or alternatively maximize the R^2 value. The algorithm stops when all sample units are combined into a single large cluster of size n.

DENDROGRAM

The Dendrogram is a graphical representation of the results of hierarchical cluster analysis. This appears in the form of tree like plot where, each step of hierarchical clustering is represented as a fusion of two branches of the tree resolving into a single one. The branches represent clusters obtained on each step of hierarchical clustering.

RESULTS AND DISCUSSION

Table 1.1 Eigen values of the Correlation Matrix for Crop Area during 1985-1995

In the present study the Cluster analysis was carried out based on area, production, and productivity of different agricultural crops which are predominantly grown in the districts of Karnataka and calculated for two different periods 1985-1995 and 1995-2005 separately. The Table-4.12. Indicates that the results of the Eigen values of the Correlation matrix for the area of 18 different agricultural crops are predominantly grown in the district for the period 1985-1995.

The table shows that the Eigen value and their corresponding proportion of variation as well as their cumulative proportion for 20 initial crop clusters. It can be seen that first, second and third, initial clusters have the Eigen values 13.450, 3.280, and 1.354 respectively which are having Eigen values greater than one, hence the wards method of clustering resulting only with three final clusters for this period of study. Further, the result depicts that first 17 districts resulting with cumulative value to one, during the period 1985-1995.

Table 4.3.2. Eigen values of the Correlation Matrix for Crop Area during 1995-2005

Table 4.3.2. indicate that the results of the Eigen values from the Correlation matrix for the area of 18 different agricultural crops grown in the districts for the period of 1995-2005. The table shows that the Eigen value and their corresponding proportion of variation as well as their cumulative proportion for 27 initial clusters. It can be seen that first, second and third, initial clusters have the Eigen values 20.6120, 3.9690 and 1.5690 respectively which are having Eigen values greater than one. Hence, the ward's method of clustering resulting only with three final clusters for period of study 1995-2005. However, the result depicts that first 15 districts resulting with cumulative value to one, during the period 1985-1995.

Table 4.3.1. Eigen values of the matrix for crop Area during 1985-1995 (Ward's method of clustering)

Eigen values of the correlation matrix				
Districts	*Eigen value*	*Difference*	*proportion*	*cumulative*
1	13.4500	6.2140	0.6725	0.6725
2	3.2800	2.0560	0.1640	0.8367
3	1.3540	1.0120	0.0677	0.9042
4	0.8380	0.6150	0.0419	0.9461
5	0.5140	0.4070	0.0257	0.9718
6	0.1440	0.1020	0.0072	0.9790
7	0.1100	0.0990	0.0055	0.9845
8	0.0780	0.0620	0.0039	0.9884
9	0.0600	0.0490	0.0030	0.9914
10	0.0500	0.3500	0.0025	0.9939
11	0.0400	0.0280	0.0020	0.9959
12	0.0300	0.0200	0.0015	0.9969
13	0.0160	0.0150	0.0001	0.9979
14	0.0140	0.0110	0.0008	0.9987
15	0.0100	0.0100	0.0008	0.9994
16	0.0070	0.0070	0.0005	0.9999
17	0.0020		0.0001	1.0000

Table 4.3.2: Eigen values of the matrix for crop Area during 1995-2005 (Ward's method of clustering)

Eigen values of the correlation matrix				
Districts	*Eigen value*	*Difference*	*proportion*	*cumulative*
1	20.6120	8.1620	0.7634	0.7634
2	3.9690	0.0930	0.1470	0.9804
3	1.5690	0.7140	0.0570	0.9674
4	0.4000	0.1260	0.0148	0.9822
5	0.1540	0.0980	0.0077	0.9899
6	0.1430	0.0640	0.0053	0.9952
7	0.0460	0.0210	0.0017	0.9962
8	0.0270	0.0140	0.0010	0.9979
9	0.0160	0.0090	0.0006	0.9985
10	0.0140	0.0080	0.0005	0.9990
11	0.0110	0.0070	0.0004	0.9994
12	0.0080	0.0040	0.0003	0.9996
13	0.0050	0.0030	0.0002	0.9998
14	0.0030	0.0020	0.0001	0.9999
15	0.0020		0.00001	1.0000

Figure-2. Crop Cluster based on Area during 1985-1995

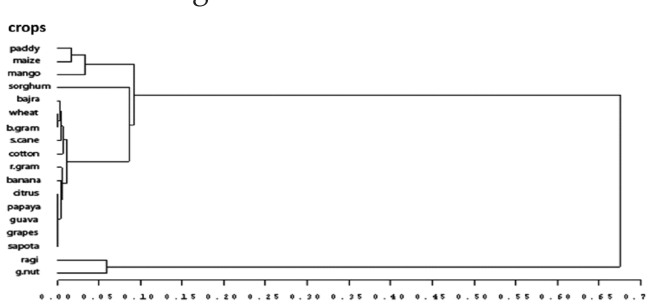
The Dendrogram (Fig.-2) showed that the cluster analysis was done based on the cultivated area of different crops across all the districts of Karnataka for the period of 1985-1995 at 10% semi partial R^2. From the findings, it can be seen that only three final major clusters were observed.

The first cluster formed by cotton, paddy, ground nut and ragi. This indicates only these crops have similarity in area of cultivation across all the districts of Karnataka. The second cluster mainly formed by sorghum alone. This indicates that this crop has similarity in area of cultivation across all the districts of Karnataka. Further, the third cluster has the combination of the rest of the crops.

Figure 3. Crop Cluster based on Area during 1995-2005

The Dendrogram (Fig. 3) showed that the cluster analysis was done based on the area of cultivation of different crops across all the districts of Karnataka for the period 1995-2005 at 10% semi partial R^2. From the result, it evident that only three major clusters was formed.

The first cluster formed by paddy maize and mango. This indicated that these crops have similarity in area of cultivation across all the districts of Karnataka. The second cluster mainly formed by ragi and ground nut. This indicates that two crops have similarity in area of cultivation across all the districts of Karnataka. Third cluster formed by remaining other crops. Based on this Dendrogram, it evident that the cotton production has shifted over the years during the period 1985-1995, when compared to the area during 1995-2005



Dendrogram method of crop cluster based on area during 1995-2005

Dendrogram method of crop cluster based on area during 1985-1995

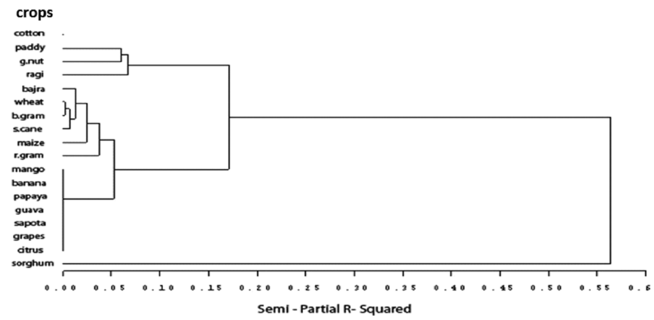


Figure 6. Crop Cluster Based on Productivity during 1985-1995

The Dendrogram (Fig-6) showed that the cluster analysis done based on the productivity of different crops across all the districts of Karnataka for the period of 1985-1995 at 10 percent semi partial R^2. Three final clusters formed, the first cluster formed by papaya crop. This indicated that papaya alone showing similarity in productivity across all the districts of Karnataka. The second cluster mainly formed by sapota, mango, grapes, sugarcane and banana. This indicates that these crops have similarity in productivity across all the districts of Karnataka and third cluster formed by remaining other crops. This showed all other fruit crops have similarity in productivity across all the districts of Karnataka under study.

Figure-7. Cluster Based Crop Productivity during 1995-2005

The Dendrogram (Fig-7.) showed that the cluster analysis done based on the productivity of different crops across all the districts of Karnataka for the period 1985-1995 at 10 percent semi partial R^2. Three final clusters were formed; the first cluster by papaya and sugarcane indicating similarity in productivity of these crops across all the districts of Karnataka. The second cluster formed by banana. This indicates, that was the only crop showing similarity in productivity across all the districts of Karnataka. The third cluster formed by considering remaining other crops. This indicated that remaining other crops have similarity in productivity of crops across all the districts of Karnataka. Based on the cluster

Dendrogram, it can be concluded that Banana was introduced as a major crop across all the districts showing similarity in productivity during the period 1995-2005.

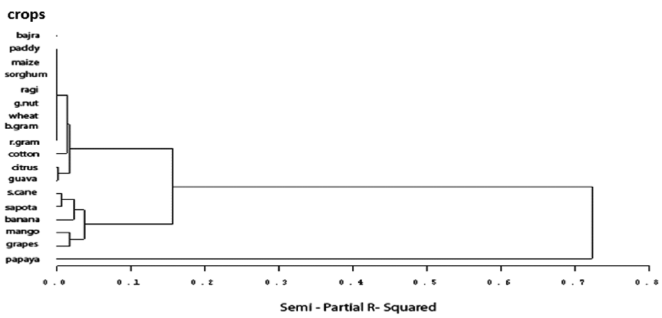


Figure 7. Dendrogram method of crop cluster based on productivity during 1995-2005

Table.4.3.7. showed that for first period of crop area between 1985 and 1995, various crops such as sorghum, cotton, paddy, ground nut and ragi showed similarity in area across the Karnataka. For the second period 1995-2005, crops like paddy, maize, mango, ragi and groundnut had similarity in area across all the districts of Karnataka. When the performance of crop clusters based on area between two periods was compared, it indicated that cotton crop has shifted by introduction of maize and mango over the years in the second period.

Table. 4.3.7. showed that for the period of 1985-1995, sugarcane and sorghum had similarity in production across all the districts of Karnataka. For the period of 1995-2005, only sugarcane showed similarity in production across all the districts of Karnataka. When comparison of the performances of crop clusters based on production between two periods was observed, it indicated that sorghum production shifted over the period.

Table. 4.3.7. showed that for the first period (1985-1995) papaya, sapota, mango, sugarcane, banana and grapes had similarity in productivity across all the districts of Karnataka. For second period 1995-2005, the crops like papaya and sugarcane showed similarity in productivity across all the districts of Karnataka. When the performances of crop clusters were compared based on productivity between two periods, it indicated that mango and banana production are shifted over the years.

Table 4.3.7. Overall performance of the cluster analysis

Aspects	*Study period (year)*	*First cluster*	*Second cluster*	*Third cluster*
Area	1985-1995	Cotton, Paddy, ground nut and Ragi (4 crops)	Sorghum (1 crop)	Other remaining crops (13 crops)
	1995-2005	Paddy, Maize and Mango (3 crops)	Ragi and Ground nut (2 crops)	Other remaining crops (12 crops)
Production	1985-1995	Sugarcane (1 crop)	Sorghum (1 crop)	Other remaining crops (16 crops)
	1995-2005	Sugarcane (1 crop)	Other remaining crops (17 crops)	
Productivity	1985-1995	Papaya (1 crop)	Sapota, Mango, sugarcane, Banana and Grapes (5crops)	Other remaining crops (12 crops)
	1995-2005	Papaya and Sugarcane (2 crops)	Banana (1 crop)	Other remaining crops (15 crops)

CONCLUSION

It showed that for first period of crop area between 1985 and 1995, various crops such as sorghum, cotton, paddy, ground nut and ragi showed similarity in area across the Karnataka. For the second period 1995-2005, crops like paddy, maize, mango, ragi and groundnut had similarity in area across all the districts of Karnataka. When the performance of crop clusters based on area between two periods was compared, it indicated that cotton crop has shifted by introduction of maize and mango over the years in the second period and also. showed that for first period 1985-1995, sugarcane and sorghum had similarity in production across the Karnataka. For second period 1995-2005, only sugarcane showed similarity in production across the Karnataka. When production between two periods was observed the result revealed that sorghum production was shifted over the period.

From all the tables we concluded that for the first period 1985-1995, papaya, sapota, mango, sugarcane, banana and grapes had similarity in production across all the districts of Karnataka. For the second period 1995-2005, the papaya and sugarcane showed similarity in production across the Karnataka. Based on productivity between two periods, it clearly indicated that mango and banana crops production are shifted over the years. Based on productivity major clusters were formed by the horticulture crops even though they are grown in less considerable area but productivity of these crops are high, so they formed an separate clusters.

REFERENCES

ANA IGLESIAS, CYNTHIA ROSENZWEIG AND DAVID PEREIRA., 2000, Agricultural impacts of climate change in Spain: developing tools for a spatial analysis, *Global Environmental Change,* **10**(1): 69-80.

BEHERA, U.K., SHARMAI, A.R., AND MAHAPATRA.C., 2007, Crop Diversification for Efficient Resource Management in India: Problems, Prospects and Policy, *Journal of Sustainable Agriculture,* **30**(3):234-244.

BRETT.F., EDWARD., SMITH.L AND HOWARD.O.ENGLAND., 1987, Regression and Cluster Analysis of Environmental Responses of Hybrid and Pure line Winter Wheat Cultivars, *Crop Science Society Of America*, **27**:659-664.

CALI SKI.T, AND HARABASZ, J., 1992, Dendrite Method for Cluster Analysis, *Communications in Statistics - Simulation and Computation*, **3**(1), 1–27.

CHANDRASEKHAR.C.P., 1993, Agrarian Change and Occupational Diversification: Non-Agricultural Employment and Rural Development in West Bengal, *Journal of Peasant Studies*, **20** (2) : 205–270.

KUTNER, NACHTSHEIM. AND NETER., 2003, Applied Linear Regression Models, 4th edition, McGraw-Hill Irwin., **pp**120-137.

LOREN W. TAUR., 1990, How diversified is your states agriculture, American journal of agriculture economics., 49(2):121-129.

MACHADO,S. E.BYNUM, JR., ARCHER, T. L., LASCANO, R. J., WILSON, J., BORDOVSKY, E., SEGARRA, K., BRONSON, D. M., NESMITH AND W. XU**.,** 2002, Spatial and Temporal Variability of Corn Growth and Grain Yield, *Crop Science Society of America,* **42**:1564-1576 .

MAHARAJAN GIRISH., 2004, Crop Diversification in the Farm sectors of Kangra District of Himachal Pradesh, *Indian Journal of Agricultural Economics*.**17**:199-217.

MO HUIDONG GU SHILIANG., 1987 Cluster Analysis for Agronomic Characters of Barley Varieties in Jiangsu-Zhejiang Shanghai Area *Scientia Agriculture Sinica*, **30**(7).143-157.

PANDEY .V.K. AND K.C.SHARMA., 1996, the crop diversification and self-sufficiency in food grains, *Indian Journal of Agricultural Economics*, **51**(4): 644-51.