



Received : August, 2010; Accepted : October, 2010

Research  
Paper

## Application of cluster analysis for spatial classification

S.R. TORANE, B.K. NAIK, J.M. TALATHI AND P.J. KSHIRSAGAR

See end of the article for authors' affiliations

Correspondence to :

**S.R. TORNE**

Department of Agricultural Economics, College of Agriculture, Dr. B.S. Konkan Krishi Vidyapeeth, Dapoli, RATNAGIRI (M.S.) INDIA

### ABSTRACT

Cluster analysis is a multivariate statistical procedure that starts with data set containing information about a sample of entities and attempts to reorganize these entities into relatively homogeneous groups or groups of highly similar entities, called as "clusters". Study was undertaken in north Konkan region of Maharashtra state. The variables were based on the set of 23 characteristics. The analysis has classified all 30 Tahsils of North Konkan region into five distinct clusters. The accuracy of clusters formed and appropriateness of the method adopted was tested by Discriminant analysis. It was concluded from the study that cluster analysis technique is useful in spatial classification, which can be further used for making development policy in general agriculture in particular for deciding the region specific cropping as well farming systems.

Torane, S.R., Naik, B.K., Talathi, J.M. and Kshirsagar, P.J. (2011). Application of cluster analysis for spatial classification, *Internat. Res. J. agric. Eco. & Stat.*, 2 (1) : 11-14.

**Key words :** Cluster analysis, Hierarchical agglomerative, Discriminant analysis, North Konkan

## INTRODUCTION

Cluster analysis is generic name for wide variety of procedures that can be used to create a classification. This procedure more empirically formed clusters or groups of highly similar entities. More specifically, a clustering method is multivariate statistical procedure that starts with a data set containing information about a sample of entities attempts to reorganize these entities into relatively homogeneous groups, called as "clusters". Cluster analysis (CA) is a classification method that is used to arrange a set of cases into clusters. The aim is to establish a set of clusters such that cases within a cluster are more similar to each other than they are in other clusters. Present study was undertaken with an objective to delineate the north Konkan region into different sub-regions on the basis of agro-climatic parameters.

## MATERIALS AND METHODS

The north Konkan region of Maharashtra state was selected purposively for the study. The north Konkan region is comprised of three districts namely, Thane, Raigad and Mumbai. Mumbai district was excluded from the study. The data on 23 characteristics such as area under various

crops such as mango, cashewnut, coconut, arecanut, paddy, *Nagli*, other cereals, pigeonpea, *Mung*, other pulses, chilli, other vegetables, groundnut, livestock such as cow, buffalo, bullock, poultry birds and also data regarding forest area, rainfall etc. were collected for the year 2007-08. Among the different methods of cluster analysis, Hierarchical agglomerative method was used in the present study. The district level data were obtained from District Socio-economic review and District statistical abstract for Raigad and Thane districts and also from office of Joint Directors of Agriculture, Thane and Raigad. A SPSS statistical software (statistical package for social sciences) was used to carry out the analysis. The data for the different variables used for each Tahsil were standardized between range of 0 to 1. The Pearson's correlation coefficient was used to form a proximity matrix. The complete linkage method was used for cluster making. A correctness or validation of clusters formed was assessed by a discriminant analysis.

## RESULTS AND DISCUSSION

The proximity matrix of cluster classification based on Pearson's correlation coefficient is presented in Table 1.