5.1 Introduction

Cluster Analysis, one of the Multivariate data analyses, and encompasses a number of different algorithms and methods for grouping objects of similar kind into respective categories and often pictorially represented in a tree structure called a dendrogram. The root of the tree consists of a single cluster containing all observations, and leaves correspond to individual observations. There are two modes of clustering; ‘R’ mode and ‘Q’ mode. For ‘R’ mode clustering, putting weight on groupings of taxa, should be in rows. It is also possible to find groupings of variables or associations (Q mode), by entering taxa in columns. The most straightforward way of computing distances between objects in a multi-dimensional space is to compute Euclidean distances.

Cluster analysis has been widely applied to explore the spatial and temporal similarities among communities or groups in a community in various fields of science and arts.

The Multivariate analysis of datas (63 species X 23 stations X environmental parameters) of Vellar Estuary are Clustered in Q mode, using unweighted paired group and Euclidean distance measure using PAST Software. The Cluster Analysis was carried out using the distribution of 63 foraminifera species in 23 stations (Tables 5.1 a-c) with
their environmental parameters for all the three collections. The resulted
dendograms are given (Fig. 5.1 a-c) and are explained in the following
pages.

5.1.1 May

The cluster analysis (dendrogram) based on the distribution of
foraminifera collected from 23 sampling stations during May 2007 is given
in fig. 5.2 a. The cluster analysis grouped the stations based on the
availability of species and environmental variables into two major
categories (α and β). The cluster α includes the sampling stations 3 to 11
and the cluster β includes the stations 1, 2, 12 to 23.

In general, group α supports more number of individuals and
species of foraminifera than group β. The group α can further be divided
into αa and αb where αa includes stations 3, 4, 6, 7, 8, 9, 10 and 11 and αb
includes only station 5. In the second groupings, αa is with lesser number
of individuals and species of foraminifera than αb. The sub cluster αb is
confined only to station 5, where the population and diversity are highest
in the season because of the congenial environmental conditions viz.,
salinity value: 31.4 ppt; CaCO₃: 14.6% and sediment fines- 48%.

The αa is further sub grouped into αa-1 and αa-2. The sub cluster
αa-1 includes the stations 8 to 11, where the living foraminiferal
population and diversity are moderate aided with the moderate salinity of
bottom water and moderate CaCO₃ content of sediments but with slightly
lessier sediment fines. The stations 3, 4, 6 & 7 of the cluster αa-2 provide a
congenial conditions (salinity values -> 30.2 ppt; CaCO$_3$ - 13.8 to 14.8% and sediment fines- 40-45%) for higher population and diversity.

The cluster $\beta$ is sub grouped into $\beta$a and $\beta$b, in which $\beta$a is grouped with stations 1 & 2 where the foraminiferal population and diversity are very low because of the sandy substrate. The sub cluster $\beta$b is further classified into $\beta$b-1 and $\beta$b-2. Under the $\beta$b-1, there are two divisions’ $\beta$b-1-1 and $\beta$b-1-2, the first of which is represented with the stations 21-23 where the population is almost absent due to very low salinity values of bottom water and low CaCO$_3$ content of the bottom sediments. The sub cluster $\beta$b-1-2 is further classified into $\beta$b-1-2-1 confined to the stations 18-20 where the population and diversity are very low because of the uncongenial conditions (low salinity values), while $\beta$b-1-2-2 representing the stations 16 & 17 is with low population because of lesser salinity, dissolved oxygen and sediment fines of the substrate.

The sub cluster $\beta$b-2 represents the stations 12 to 15, where the population is very moderate because of the moderate salinity values(24.0 – 26.5 ppt) and moderate CaCO$_3$ content (9.8 -11.8%).

It may be concluded that the cluster analysis grouped the stations which are accommodative and congenial to support higher foraminiferal population and diversity into the major cluster $\alpha$ and the sampling stations which are not so congenial for higher population and diversity into the cluster $\beta$. 
5.1.2 September

The cluster analysis (dendrogram) based on the distribution of foraminifera collected from 23 sampling stations during September 2007 is given in fig. 5.2 b. The cluster analysis grouped the stations based on the availability of species and environmental variables into two major categories (α and β). The cluster α includes the sampling stations 3 to 11 and the cluster β includes the stations 1, 2, 12 to 23.

The cluster α is divided into αa and αb in which αa is sub grouped into αa-1 and αa-2. Sub groups comprising of stations 8 to 11 are under αa-1 and stations 3 & 4 are under αa-2. The stations of αa-1 indicate the area of moderate population and diversity because of moderate salinity values (24.4 - 27.9 ppt), CaCO₃ content (10.6 - 11.5%) and lesser sediment fines, while the stations of αa-2 also encountered with moderate population and diversity because of the lesser sediment fines.

The sub cluster αb comprises stations 5-7, where the living foraminiferal population and diversity are higher because of the congenial environmental conditions (salinity > 29.5ppt; CaCO₃ 13.0-14.2%; sediment fines 40 -45 %) prevailed in these stations.

The cluster β is sub grouped into βa and βb, in which βa is grouped with stations 1 & 2 where the foraminiferal population and diversity are very low because of the uncongenial sandy substrate. The sub cluster βb is further classified into βb-1 and βb-2. Under the βb-1, there are two divisions’ βb-1-1 and βb-1-2, the first of which is again divided into βb-1-
1-1 represented with the stations 19-21 where the population and diversity are very low due to lesser salinity values of bottom water and βb-1-1-2 where the living and total population is absent, represented in stations 22 & 23, because of very low salinity values.

The sub cluster βb-2 represents the stations 12 to 15, where the population is very moderate because of the moderate salinity values (19.2 – 22.1ppt) and moderate CaCO\textsubscript{3} content (9.8 - 10.2%).

Hence, It may be concluded that the cluster analysis grouped the stations which are accommodative and congenial to support higher foraminiferal population and diversity into the major cluster α and the sampling stations which are not so congenial for higher population and diversity into the cluster β.

5.1.3 January

The cluster analysis dendrogram based on the distribution of foraminifera collected from 23 sampling stations during January 2008 is given in fig. 5.2 c. All the 23 stations were based on the availability of species and the environmental parameters grouped largely into two categories α and β. The cluster α comprises the sampling stations 3 to 11 and the cluster β comprising the stations 1, 2, 12 to 23.

The cluster α is divided into αa and αb in which αa sub grouped into αa-1 and αa-2. These sub groups comprising of stations 5 to 7 under αa-1 and stations 3 & 4 under αa-2. The stations of the cluster αa-1 provide an congenial conditions for higher population and diversity (salinity values
The stations of \( \alpha \)-2 also encountered with moderately higher population and diversity because of the optimum conditions as mentioned above but with slightly lesser sediment fines.

The sub cluster \( \alpha \)b is classified into \( \alpha \)b-1 represented by stations 10-11 and \( \alpha \)b-2 consists of stations 8-9, where the population and diversity are moderate because of the following environmental conditions. Salinity values 26.8-28.2 ppt; CaCO\(_3\) 11.6 to 12.7% and sediment fines 38-42%.

The cluster \( \beta \) is sub grouped into \( \beta \)a and \( \beta \)b, in which \( \beta \)a is grouped with stations 1 & 2 where the foraminiferal population and diversity are very low because of the uncongenial sandy substrate. The sub cluster \( \beta \)b is further classified into \( \beta \)b-1 and \( \beta \)b-2. Under the \( \beta \)b-1, there are two divisions’ \( \beta \)b-1-1 and \( \beta \)b-1-2, the first of which is represented with the stations 20-23 where the population are almost absent due to low salinity values of bottom water and low CaCO\(_3\) content of the bottom sediments. The sub cluster \( \beta \)b-1-2 where the living and total population is very low represented in stations 16 - 18.

The sub cluster \( \beta \)b-2 consists of stations 12 to 15, where the population is very moderate because of the moderate salinity values (21.8 – 26.8) and moderate CaCO\(_3\) content (9.8 -12.0%).

Hence, as mentioned for earlier months, it may here also be concluded that the cluster analysis grouped the stations which are
accommodative and congenial to support higher foraminiferal population and diversity into the major cluster $\alpha$ and the sampling stations which are not so congenial for higher population and diversity into the cluster $\beta$.

The cluster analysis, taking 23 stations as variables, gave rise to different dendrograms for different collections. From the grouping of different clusters it may inferred

(i) Higher populated stations and lesser populated stations form different clusters in all the dendrograms.

(ii) Immediate sub clusters, represent grouping of highly (stations 3-7), moderately (stations 8-11), less (stations 12-15) and very less (stations 16-20) populated stations and also grouping stations (21-23) with no living forms.

(iii) The clusters representing moderate population are dominated by cosmopolitan species.

(iv) Species diversity values are directly proportional to number of individuals and species and thus higher the species richness and individuals of species higher the diversity values.

(v) The distribution of clusters mainly depends on the salinity of the bottom water, Calcium carbonate and sediment fines (silt + clay) of the substrate.