International Journal of Agriculture S c i e n c e s
ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 7, Issue 9, 2015, pp.-693-697.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000217

FISH PRODUCTION ANALYSIS OF CACHAR DISTRICT (ASSAM) USING FUZZY C-MEAN CLUSTERING

KUMAR AVINASH[1], KUMAR DEEPAK[2*], SARKAR SUDIPTO[3]

[1,3] Department of Agricultural Engineering, Assam University, Silchar, Assam, 788 011
[2]Department of Hydrology, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, 247667
* Correspondence Author, E-mail: deepak.civil.iitdelhi@gmail.com, deepak01iitkgp@gmail.com

Received: July 03, 2015; Revised: October 28, 2015; Accepted: October 29, 2015

Abstract- Along with agricultural practices, fish farming is one of the main occupations of the north eastern state of India. In the present study, Cachar district of Assam has been selected for studying an optimal pattern of fish farming. Cachar district is one of the poorer areas of Assam. Tea plantation along with fish farming is a source of livelihood of the people of this region. A comprehensive survey for the fish farm size and the respective production has been done in the present study. The collected data is then studied using cluster analysis. Cluster analysis has been done using Fuzzy C-mean Clustering. During this analysis, the data have been studied for single fuzzy c-mean cluster centre, double fuzzy c-mean cluster centre and triple fuzzy c-mean cluster centre. The results suggest that single fuzzy c-mean cluster centre analysis is the best analysis as far as cluster analysis is concerned. Thus, the annual production of the study area can be clustered in a single group and the cluster centre is located at annual production of 900 kilograms per hectare for farm size of 0.54 hectare.

.**Keywords:** Cachar, Fish farm size, Fish production, Fuzzy, Cluster analysis, c-mean

Citation: Kumar Avinash, *et al.* (2015) Fish Production Analysis of Cachar District (Assam) Using Fuzzy *C*-Mean Clustring, International Journal of Agriculture Sciences, ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 7, Issue 9, pp.-693-697.

Introduction

India has the second largest demographical statistics and also has a high demand of consumption of natural resources for their survival and developments. In the present scenario's most critical challenge for developing countries is to fulfil the need of foods, water, energy, and clean environment. Modern global society has become increasingly aware of the dimensions of the stock of resources available to it, and of the various limits and constraints to their use. After the land based resources, water based resources are the most important among all the developing countries. The water-based resources are the fish farming, corals, geological minerals and the energy from the water bodies.

In India, after the independence, the fishery is developed as the complimentary of the agriculture. Therefore, presently fish production has a significant contribution in the Indian economy and culture. According to the FAO fishery yearbook 2012 [1], India is the fourth largest fish producer on the globe with 4.5 percentage annual growth rate. At present, the total 9.06 million metric tonnes of fishes produces which earn foreign exchange of US$ 3.51 billion (2012–13). It has been estimated that about 1.2 million ha of potential brackish water area available in India is suitable for farming. In addition to this, about 9.0 million ha of salt affected areas are also available. The patterns of marketing and consumption in fish and fish-derived products have developed over centuries and are affected by established practice and cultural tradition as well as availability. Therefore, fishery refers as the most sustainable resources for fulfilments of the food security as well as rural employments and empowerments of the second front people of the societies.

Systematic fisheries research that can develop and test hypothesis relating alternative management and regulation structures with fishery performance requires the integration of resource biology and ecology with the economic factors that determine fishers' behaviour in space and time. The complexity of fisheries management is increased by the major uncertainties inherent in natural systems, as well as by a range of other biological, social, political, and economic factors, requiring a precautionary approach to fisheries management (FAO, 1996) [2].In many parts of Asia, small-scale, low input aquaculture technologies are seen as an important tool for improving food security, especially in areas where there is a shortage of fresh fish. In Bangladesh, several projects have been undertaken to assist farmer adopt fish culture in small water bodies in and around their homesteads. Keys to successful implementation of research agenda in the fish farming must be part of an integrated approach that examines how changes in water use will impact all areas of operation. Therefore, for better understanding of the fish cultivation at a place, it is important to have a survey of that place. The outcome of this type of survey sometimes suggests very important information about the economic status of farmers in that locality. Furthermore, the followings are themes are identified as important elements of a research agenda on Indian fish farming systems at regional levels 1). Generation of eco-hydrological knowledge on the fish farming system, and institutional mechanisms for its sustainable use and protection from pollution. 2). Emerging technological options in water system managements.

Fish farming is the second important source of livelihood of Cachar district. More than 165 kinds of fishes are found in Assam. All the species are not cultivated commercially, Some species like; Ilish, chital, kandhuli, balisonda, puthi, mirika, bhangone, nara, rau, common carp, grass carp, silver carp, big head carp, singorah, arii, barali, magur, thilandmagur, singhi,goroi, cheng, kuchia, shol, sal, japanikawai, kholihona are cultivated commercially in Cachar.

Further, in 2007 Hortle [3] reported a regional study giving overview of production levels, consumption levels of fisheries. Descriptive statistics were used in methodologies as the data's were collected from over 20 different surveys. Jahan *et al.* (2010) [4] collected data from 225 farmers in 2001-05 for the Development of Sustainable Aquaculture Project (DSAP) to calculate the impact of aquaculture projects on fish consumption (using a before/after, with/without trial), calculated in terms of annual per capita fish. They also made calculations for the consumption of other foodstuffs, and record the types of fish that are consumed. Allison [5] collected the data of the relationship between export trade and national fish protein supply (not micro-nutrients) from FAO for the year of 1976-2007. Jahan [4] conducted the micro-studies on pro-aquaculture projects to improve incomes and consumption for participant households on their respected countries. In separate two experiments, Bene [6,7] conducted micro-level studies revealing the extent to which poorer groups participate in and benefit from fisheries or aquaculture. They used a Gini decomposition exercise and collected the data from their own survey. They collected data from 43 fishing camps by conducting surveys, interviews and group discussions. It includes records of peoples' incomes and expenditures during the previous year. They used statistical tests (an analysis of variance and pair wise multiple comparison) to determine correlations on the role of fisheries particularly regarding income generation. Despite the vast aquatic resources, Assam has not been able to produce ample fish to cater to the needs of its ever increasing population. Assam's share of the total inland fish production in India is 6.55 %. The size of a fish farm plays a very vital role in deciding the income of a farmer. In this study, an optimal fish farm size has been determined on the basis of total fish production in the farm. Similar type of study can also be obtained from literature referenced from [8-12].

For this study, an exhaustive data collection of the fish farm size and its respective fish production has been done. After doing an unbiased survey, the collected data for this study has been comprehensively studied using cluster analysis [13-15]. Cluster analysis is generally done to sub group objects of similar kind into respective categories. In Cachar district, the farm size varies from less than one hectare to three hectare. In accordance with farm size, the total fish production also varies. In the present study, the distribution of farm sizes and its respective fish production has been studied using Fuzzy-C mean clustering.

Study Area Descriptions

The Cachar district is situated in the southern part of the Indian state of Assam. The total geographical area of the district is 3,786 Sq. Km. The district lies between 92° 24' E and 93° 15' E longitude and 24° 22' N and 25° 8' N latitude which is 35 meters above mean sea level. The total population of Cachar district is 17, 36,319 as per Census 2011. The average density is 459 per sq. km. The location of Cachar district is shown in [Fig-1]. Total 21,516 hectares of land are used for fish farming in this district. Fish farming is the second important source of livelihood of cachar district. As per the district fisheries development officer, the total population of fishermen is about 1.5 Lakh. However, they able to produce only 21,500 tonnes of fish per year from the different water bodies, including the Barak river, whereas the annual demand are estimated at 36,000 tonnes.

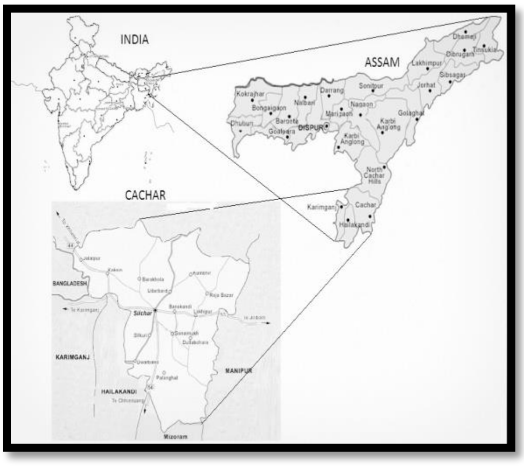

Fig-1-Cachar District Map

Among the 15 agro-climatic regions of the country, categorized/identified on the basis of homogeneity in agro- characteristics, Cachar falls in the Barak Valley zone. The agro climatic conditions of the district are conducive for various agricultural activities like development of the plantation crops *viz*., tea, rubber, cashew, coffee, areca nut, coconut and also aromatic plant like Patchouli. The area is characterised by a tropical monsoon climate, having three distinct seasons *viz*., summer (March-May), Rainy season (June-September) and winter. Towards the mid-April rain clouds starts covering the skyline. The Cachar district receives an average annual rainfall of 3000 mm in a year. Cachar is frequently inundated due to excessive rainfall and flooding by the river Barak.

Methodology

Data Collection

A survey was conducted in five blocks of Cachar district to collect the information related to the fish farming pattern. The following steps were undertaken for the data collection. First of all five main fish cultured blocks, which are nearby Assam University, Silchar, were identified with the help of district fishery officer. There are 15 blocks in Cachar district namely Silchar, Salchapra, Tapang, Udarbond, Sonai, Borkhola, Kalain, Katigorah, Banskandi, Binnakandi, Lakhipur, Rajabazar, Narsingpur, Borjalenga, Palonghat. Fish farming is carried out almost in all the blocks. However, Tapang, Narsingpur, Sonai, Borjalenga and Palonghat were identified as five extensive fish cultured blocks, nearby Assam University Silchar. A set of questionnaire was formed and distributed among the farmers to collect the data related to fish farming. The questionnaire contains the information related to the livelihood of farmer, technical and economical aspect of the fish farming pattern.

A principle of Participatory Rural Appraisal (PRA) was adopted during the visit in different blocks. The highest fish producers as per the questionnaire and with the consultation of the Gram-Mukhya in the selected blocks were also identified. Participatory Rural Appraisal is an approach used by non –governmental organizations and other agencies involved in rural development. The approach aims to incorporate the knowledge and opinions of rural people in the planning and management of development projects and programs funded by the GOI or any private institutions

Governing Equation and Approaches

Fuzzy C-mean clustering

Fuzzy c-means clustering is a method to group the data into one or more than one groups [16-19]. These groups are also known as clusters. The method of c-mean clustering has been developed by Dunn in 1973 [20] and further it has been modified by Bezdek in 1981 [21]. C-mean clustering is also used in pattern recognition. In fuzzy C-mean clustering, an algorithm is developed for searching a point in the space of a data set such that the distance of that point can be minimized from other points of that group. Thus, the objective function is to minimize the summation of distance from the cluster centre to all the points in that group.

The objective function used in fuzzy, c-mean clustering is mentioned in [Eq-1]. The aim of this algorithm is to minimize this objective function (Dembele and Kastner 2003) [22].

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| X_i - c_j \right\|^2, \ 1 \le m < \infty \qquad \text{[Eq-1]}$$

Where, J_m is the objective function; X_i is the ithdata point; c_j is the centre of the cluster; u_{ij} is the degree of membership of X_i in the jth cluster; m is any real number greater than 1.

Fuzzy partitioning or grouping in c-mean clustering is carried out through an iterative optimization of the objective function shown in [Eq-1], with the update of membership function u_{ij} and the cluster centers c_j shown in [Eq-2],[Eq-3] and [Eq-4] below.

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{X_i - c_j}{X_i - c_k} \right)^{\frac{2}{m-1}}} \qquad \text{[Eq-2]}$$

$$C_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} . x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \quad \text{[Eq-3]}$$

$$\max_{ij}\left(\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right) < \varepsilon \quad \text{[Eq-4]}$$

In the present study, for using fuzzy c-mean clustering, MATLAB has been used. The graphical user interface, which is a MATLAB tool, has been used for developing clusters.

Results and Discussion

In this section, the results obtained from fuzzy, c-mean clustering are discussed. After surveying the farm sizes of different farmers in Cachar district, it has been found that farmers in this region have varying sizes of farms on which fish farming is practiced. Often the farmers in this region are poor and uneducated too. The economic backwardness of the farmers may be due to their poor educational background. [Fig-2] depicts the educational status of all the farmers, which has been surveyed for this study. From this figure, it is clear that 48%, 47% and 5% of the farmers are below 5th grade, between 5th to 12th grades and graduate respectively. Most of the farmers in this area use conventional fish farming system and the production of this area usually depend upon the farm sizes.

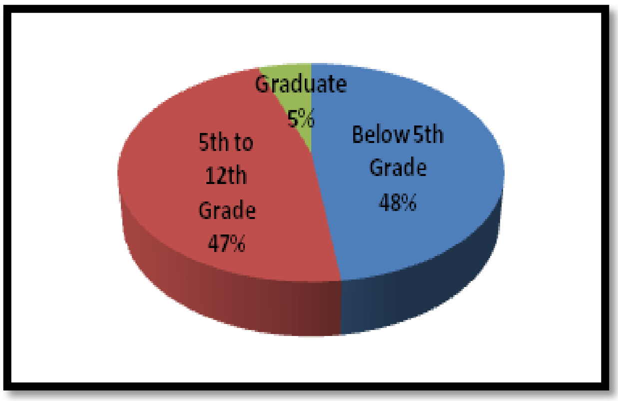

Fig-2- Educational status of survived farmers of Cachar

The annual production of farmers for different farm sizes of the surveyed farmers has been shown in [Fig-3]. By carefully analysing this figure, it can be observed that the farm sizes are very scattered and it is very difficult to exactly know the fish production of the area which maximum farmer share commonly for different sizes of the fish farm. For example, different farmers having farm size of 0.5 hectare shows a very diverse annual fish production. Similarly, for another farm sizes also, the annual production is very diverse. Cluster analysis is a method, which can able to group the data into most probable groups and by using this, one can able to know the cluster centre where all the data points belonging to a group is supposed to be converged. For getting a better picture of annual fish production from the study area for various farm sizes, a cluster analysis has been done.

A fuzzy C-mean clustering is adopted in this study. Three different sets of analysis have been performed in this study. In the first analysis set, abbreviated as 'Analysis-I', fuzzy, c-mean clustering has been used to generate single cluster of all the data sets. In the second type of analysis, abbreviated as 'Analysis-II', the same algorithm has been used to divide the whole data set into two clusters. Similarly, 'Analysis-III' has been done to divide the data set into three clusters. During these analyses, the centre of each cluster has been obtained. The cluster centre will indicate the point where a group of data is most likely to be converged.

In [Fig-4], the results obtained from Analysis-I have been presented. A point marked as 'C1' shown in [Fig-4] is the centre of the cluster. Thus, a farmer having farm size of 0.54 hectare and whose annual production is 900 kilogram per hectare per year is the most probable cluster centred production which maximum farmers seems to be converged.

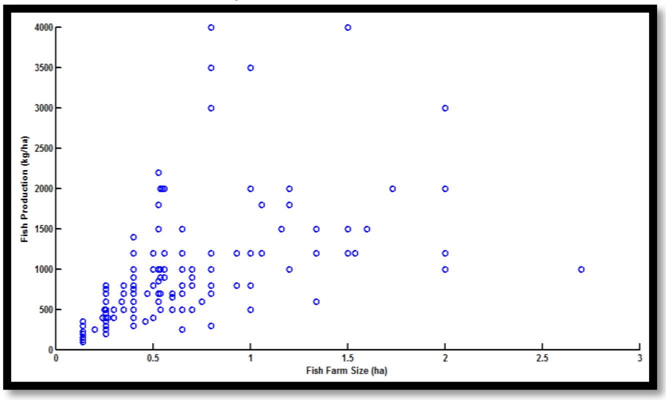

Fig-3- Farm production for various farm sizes of Cachar

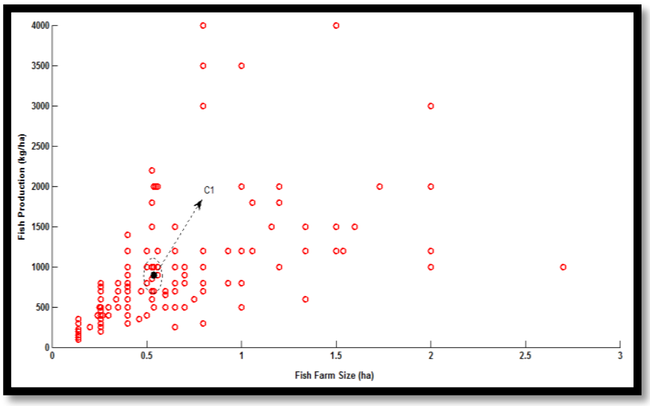

Fig-4- Single centred Fuzzy c-mean clustering

Similarly, fuzzy, c-clustering is adopted for double centre clustering to obtain two different points where the production of this area can be grouped. [Fig-5] shows two centres (namely M1 and M2) when the analysis has been done to separate the whole data set in two different cluster sets. [Fig-4] reveals that if the whole annual fish production will be clustered in two parts, one cluster centred having annual fish production will be located at 678.34 kg/ha with farm size of 0.51 hectares and the other cluster centre is located at 2338.65 kg/ha annual fish production having farm size of 1.02 hectares.

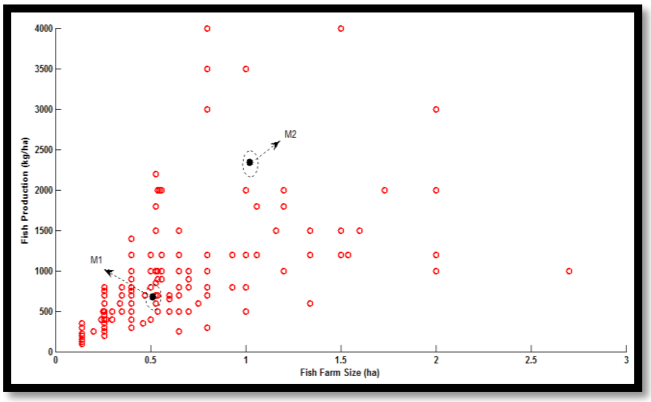

Fig- 5- Double centred Fuzzy c-mean clustering

The fuzzy C-clustering algorithm has also been used to divide the whole data set into three clusters. Thus, three cluster centres have been obtained by this process and the same has been shown in [Fig-6]. The first cluster centre has been located at D1 having annual production of 515 kg/ha for farm size of 0.4 ha. Similarly, the second and third cluster centre is located at D2 and D3 with annual production of 1249.27 kg/ha and 3323.82 kg/ha for farm size of 0.8 ha and 1.08 ha respectively. The production for its respective farm size for different analysis has also been shown in [Table-1].

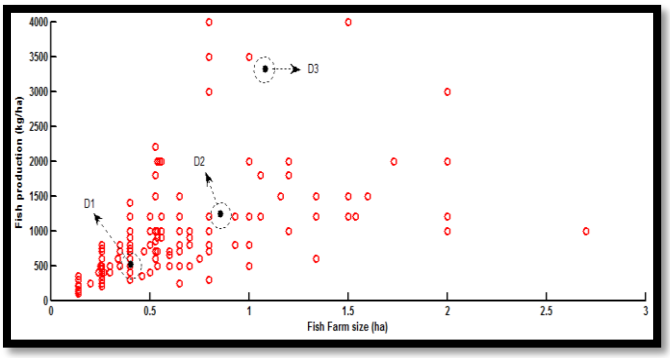

Fig- 6- Triple centred Fuzzy C-mean clustering

Table- 1- *Annual fish production for various analyses set*

Analysis Set	Number of Cluster	Fish Farm Size (ha)	Fish Production (kg/ha)
Analysis-I	Single Cluster centre	0.54	900
Analysis-II	Two Cluster centre	0.51	678.35
		1.02	2338.66
Analysis-III	Three Cluster centre	0.40	515.30
		0.86	1249.27
		1.08	3323.8

In the fuzzy c-mean clustering, the cluster centre is that centre which is at the minimum distance from a group of data points. Thus, in this type of analysis, an objective function governs the cluster centre. The aim of the objective function is to minimize the distance from the centre of the cluster. The numeric values of the objective function for Analysis-I, Analysis-II and Analysis-III has been shown in [Table-2]. The iteration count in the table suggests the minimum iteration in which objective function has been achieved. From [Table- 2], it can be stated that Analysis-I having single cluster centre has the minimum objective function. The objective function value for analysis-II and Analysis-III is 24373255.96 and 10167719.72 respectively. Thus, in analysis II and analysis-III, the cluster centres are not exactly converging. Thus, in the present study, analysis-I is the best analysis as far as cluster analysis is concerned. Thus, the annual production of the study area can be clustered in a single group and the cluster centre is located at annual production of 900 kilogram per hectare for farm size of 0.54 hectare.

Table -2- Iteration count and Objective function value for various number of cluster centres.

Analysis Set	Number of cluster	Iteration count	Objective Function value
Analysis-I	1	6	8.06×10^{-29}
Analysis-II	2	51	24373255.96
Analysis-III	3	56	10167719.72

Conclusion

In India, after the independence, Government of India developed fisheries as the complimentary of the agricultural practices. In the 21st century in India, fish production has a significant contribution. In the northeast regions of India such as Assam, Manipur, Mizoram, Fish is one of the good source of livelihood. Most farmers in this region has a small farm sizes and they depend on these farms, which also has pond, for there earning. In the present study, a survey in Cachar district has been done and data for fish farm size and production from these farms has been collected. After the collection of data, soft computing technique has been used to correlate the data and it has been tried to group the data in such a manner that farm size versus its production can get clustered optimally to give an elaborate picture of Cachar district. Fuzzy C-mean clustering has been used as a soft computing technique in this study. Thus, in the present study, analysis-I (single cluster centre) is the best analysis as far as cluster analysis is concerned. Thus, the annual production of the study area can be clustered in a single group and the cluster centre is located at annual production of 900 kilogram per hectare for farm size of 0.54 hectare.

References

[1] FAO yearbook Fishery statistics (2012) *Fishery and Aquaculture Statistics*, ISSN:2070-6057.
[2] FAO (1996) *Manual on the production and use of live food for aquaculture. FAO Fisheries Technical Paper*, 361,ISBN 92-5-103934-8.
[3] Hortle K.G. (2007) *Consumption and the yield of fish and other aquatic animals from the Lower Mekong Basin. MRC Technical Paper No. 16, Mekong River* Commission,Vientiane, pp.87.
[4] Jahan K. M., Ahmed M., Belton B. (2010) *Aquaculture Research*, 41, 481-495.
[5] Allison E.H. (2011) *Aquaculture, fisheries, poverty and food security, World Fish Centre Working Paper* 2011-65
[6] Bene C., Neiland A., Jolley T., Ovie S., Sule O., Ladu B., Mindjimba K., Belal E., Tiotsop F., Baba M., Dara L., Zakara A. and Quensiere J. (2003) *Journal of Asian and African Studies,* 38(1), pp. 17-51. http://dx.doi.org/10.1177/002190960303800102
[7] Bene C., Steel E., Kambala B., Gordon A. (2009) *Fish asthe "bank in the water"– Evidence from chronic-poorcommunities in Congo, Food Policy*, pp.108-118. http://dx.doi.org/10.1016/j.foodpol.2008.07.001
[8] Weinstein J. (2015) *Recent Challenges, Upgrades and Insights in the Monitoring of California's Marine Recreational Fisheries. In 145th Annual Meeting of the American Fisheries Society*.
[9] Fisheries N. O. A. A. (2014) *NOAA sees big increase in Bering Sea pollock survey abundance estimates: NOAA Fisheries*.
[10] Fock H. O., Kloppmann M. H., and Probst W. N. (2014) *Journal of Sea Research*, 85, 325-335.
[11] Foster J. (2015) *A Replication Approach to Controlled Selection in an Intercept Survey of Marine Recreational Fishing Trips. In 145th Annual Meeting of the American Fisheries Society, Afs.*
[12] Clarke M. E. (2015) *Experiences Using Autonomous Technologies to Survey Fish Populations. In 145th Annual Meeting of the American Fisheries Society, Afs.*
[13] Tan P. N., Steinbach M., and Kumar V. (2013) *Data Mining Cluster Analysis: Basic Concepts and Algorithms.*
[14] Zhou F., De la Torre F., and Hodgins J. K. (2013) *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(3), 582-596.
[15] Zoccatelli D., Parajka J., Gaal L., Blöschl G. and Borga M. (2015) *A process flood typology along an Alpine transect: classification based on cluster analysis. In EGU General Assembly Conference Abstracts* (Vol. 17, p. 15144).
[16] Selvakumar J., Lakshmi A., and Arivoli T. (2012) *Brain tumor segmentation and its area calculation in brain MR images using K-mean clustering and Fuzzy C-mean algorithm. In Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on* (pp. 186-190), IEEE.
[17] Kannan S. R., Ramathilagam S., and Chung P. C. (2012) *Expert Systems with Applications*, 39(7), 6292-6300.
[18] Sudha K. R., Raju Y. B., and Sekhar A. C. (2012) *International Journal*

of Electrical Power & Energy Systems, 37(1), 58-66.
[19] Tsai D. M., and Lin C. C. (2011) *Pattern recognition*, 44(8), 1750-1760.
[20] Dunn J. C. (1973) *Cybern. Syst*., 3, 32–57.
[21] Bezdek J. C. (1981) *Kluwer Academic*, pp. 256.
[22] Dembele D. and Kastner P. (2003) *Bioinformatics,* 19(8), 973-980.