GLACIER FACIES MAPPING USING FCM AND EVALUATION OF ITS ACCURACY

Submitted in partial fulfillment of the requirements for the award of degree of

Master of Technology

In

WATER RESOURCE ENGINEERING

CIVIL ENGINEERING

Submitted by

KRITIZA SHARMA

(2014PCW5106)

Under the Supervision of

Dr. Gunwant Sharma
Professor and Head
Department of Civil Engineering
MNIT Jaipur

CIVIL ENGINEERING DEPARTMENT

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

JUNE 2016

A

DISSERTATION REPORT

ON

GLACIER FACIES MAPPING USING FCM AND EVALUATION OF ITS ACCURACY

Submitted in partial fulfillment of the requirements for the award of degree of

Master of Technology

In

WATER RESOURCE ENGINEERING

Submitted by

KRITIZA SHARMA

(2014PCW5106)

Under the Supervision of

Dr. Gunwant Sharma

Professor and Head

Department of Civil Engineering

MNIT Jaipur

DEPARTMENT OF CIVIL ENGINEERING

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

JUNE 2016

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

DEPARTMENT OF CIVIL ENGINEERING

JAIPUR 302017

DECLARATION

I hereby certify that the work which is being presented in the dessertation report entitled **"GLACIER FACIES MAPPING USING FCM AND EVALUATION OF ITS ACCURACY"**, in partial fulfillment of the requirements for the award of the Degree of Master of Technology and submitted in the Department of Civil Engineering of the Malaviya National Institute of Technology Jaipur is an authentic record of my own work carried out during a period from August 2015 to June 2016 under the supervisions of Dr. Gunwant Sharma, Professor and Head, Department of Civil Engineering, Malaviya National Institute of Technology Jaipur and Dr. M. K. Arora, Director, PEC University of Technology, Chandigarh, India.

The matter presented in the report has not been submitted by me for the award of any degree of this or any other Institute.

(**KRITIZA SHARMA**)
Student ID: **2014PCW5106**

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Dr. Gunwant Sharma
Professor and Head
Department of Civil Engineering
MNIT Jaipur

ACKNOWLEDGEMENT

I owe a great thanks to a many people who helped and supported me during the project.

I would like to express gratitude and deep regards to my mentors Dr. Gunwant Sharma, Professor and Head, Department of Civil Engineering, MNIT Jaipur for his exemplary guidance, monitoring and constant encouragement throughout the project. His supervision and willingness to share their vast knowledge has helped me to complete the assigned task.

I would also like to thank Dr. M.K Arora, Director, PEC University of Technology, Chandigarh and Dr. Aparna Shukla, Scientist, Wadia Institute of Himalayan Geology, Dehradun for their guidance and support to complete this work.

I would also like to thanks Prof. A. K. Vyas, DPGC Convener, Department of Civil Engineering, MNIT Jaipur for extending every possible help and encouragement.

(Kritiza Sharma)

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Image classification techniques convert remote sensing data into informational data. Various conventional hard classification techniques are used widely to classify images and produce thematic maps. However, these techniques do not perform accurately when image contains mixed pixels due to loss of information. Thus, soft classification techniques can be used effectively which assigns a pixel multiple classes. Fuzzy based soft classifiers (FCM & PCM) are gaining wide popularity as fuzzy set theory takes into account the uncertainties present in nature. The output of FCM is dependent upon numerous factors. In the present, study three factors viz. fuzzy exponent ‘m’, A-norm and training data size are considered and their influence on FCM output has been studied. Fuzzy Error Matrix with MIN-PROD operator has been used for accuracy assessment and three parameters viz. overall accuracy, kappa coefficient and RMSE are determined for evaluation of accuracy. Results show that Mahalanobis norm outperformed the other two giving maximum OA and kappa values of 74.032 and 0.638 respectively. Accuracy values show a slight increase with increase in ‘m’ upto1.5 and then a sharp decrease with increase in ‘m’. With training data size, different trends are observed for different norms. OA/kappa measures outperformed RMSE measure for accuracy assessment giving RMSE of 1.0623 as compared to 2.907. Since the spatial distribution of various classes within a pixel cannot be determined through soft classification techniques so mapping is limited to facies area calculation only.

CHAPTER 1
INTRODUCTION

Himalayas is the origin of many glaciers and vital river systems of India. These glaciers form crucial fresh water sources and feed major perennial river systems of India like Ganga, Sutlej, Chenab, Brahmaputra, and Indus. Today, vast industrialization and urbanization has led to global climate change and glaciers are also responding to it. Climate change is affecting various glacial components like snow, ice, debris, debris mixed ice etc. ultimately causing the retreat of the glaciers. Various dams and water storage structures are built on the glacial fed rivers which are contributing to irrigation, water supply and also hydropower systems in our country. Climate change is affecting the availability of water in Himalayan rivers and ultimately all the above-mentioned provisions. Thus, it becomes imperative to study and monitor glaciers and their components in all aspects.

Because of the inaccessibility and harsh climatic conditions of glacier areas, remote sensing techniques can be used as a vital tool for the surveying and monitoring purpose. As different geographic features on Earth differ in physical, chemical or biological properties from each other; thus giving different spectral response in different bands. This property is used as basis of distinction among different geographical features in remote sensing. Spectral response of a class within a pixel is represented by a numeric value termed as digital number (DN).

One of the most important applications of remote sensing is image classification which categorizes all the pixels of an image into several land cover classes, generally called as clusters on the basis of their spectral response. Conventional classification techniques like maximum likelihood classification, parallelepiped method, and minimum distance method and nearest neighborhood assign a single class to a single pixel thus referred as per pixel or hard classification. All these methods are based on the assumptions that classes are mutually exclusive (no overlap), exhaustive (all classes present in the landscape are accounted for and none have been omitted) and hierarchical. But this is not the case in real life scenarios.

Land cover classes are rarely separated by hard or crisp boundaries; in fact there always exists fuzziness in between them. In such cases hard classification doesn't give accurate results as a pixel may consist of more than one class at the boundary and classification assigns one class per pixel; thus resulting in either overestimating or underestimating of classes. Normally, such pixels with more than one class are referred as mixed pixels.

Pan Sharpening (Panchromatic sharpening) is one of the techniques used to take care of mixed pixels in image. In this technique a high-resolution panchromatic image is merged with multispectral images of lower resolution to obtain a single high-resolution image. The spatial resolution of output image is same as that of panchromatic image. Thus spatial resolution of multispectral image is sharpened or increased. But this method has various disadvantages of its own.

Thus new techniques are being introduced which deals with mixed pixels more precisely. Such techniques are called soft classification techniques as each pixel is assigned the proportion or fraction of a number of classes laying within pixel area on the ground rather than only one class i.e. most predominant class as in hard classification. Thus, DN value of a pixel is dependent on individual spectral response of all its constituent classes. All these techniques make use of spectral response of a pixel for image classification.

Fraction images equal to number of thematic classes are obtained as output of soft classification. Fraction image of a particular class consists of same size (number of rows and columns) as that of original image where each pixel represents the fraction of respective class within the pixel with sum of fractions of all the classes within a pixel corresponds to one. Various soft classification techniques include linear mixing models, artificial neural networks, Fuzzy c means, probabilistic c means, support vending machines and many others.

However major drawback of all these soft classification techniques is that they do not provide the spatial distribution of constituent classes (how the classes are aligned) within a pixel thus limiting its use to prepare land use/ land cover map out of fraction images which shows only the proportion of pixel area covered by particular land cover class(Atkinson, 1997).

Image classification is incomplete without its accuracy assessment. It tells how accurately the image has been classified. Various methods like cross-entropy method, Euclidean and L1 distance method, correlation coefficient measures were widely used for accuracy assessment of soft classification but the biggest disadvantage of these methods was that they provided only correlational solutions not actual accuracy values as provided by conventional error matrix in hard classification accuracy assessment. Binaghi (1999) gave the concept of fuzzy error matrix (FERM) to assess the accuracy of soft classification techniques. FERM is similar to conventional error matrix used for accuracy assessment of hard classifier except that elements of FERM can be real numbers as well as opposed to only integer values in latter.

1.1 Problem related to Glacier Areas

Snow covered areas of glaciers give very high reflectance in visible range of spectrum thus resulting in saturation of bands for most of the widely used sensors like Landsat TM and MSS due to their low radiometric resolution (8 bits).Various moderate resolution sensors like AWiFS on-board IRS-P6 and MODIS onboard Terra and Aqua satellites have high radiometric resolution (10 bits) thus preventing saturation. Only disadvantage related to these sensors is their low spatial resolution which results in large number of mixed pixels in the image. Thus, data from indigenous sensor AWiFS has been used in this study and soft classification techniques are employed.

1.2 Objectives of the Study

The present study focusses on the role of different parameters like training data size, fuzzy exponent 'm' and various A-norms on Fuzzy C means classification output. Objectives of study are listed below:

1. Identification of various facies of glacier and collect appropriate training data for supervised FCM classification.
2. Implementation of FCM algorithm for glacier facies mapping in MATLAB and identify the role of previously mentioned parameters in FCM.
3. Evaluation of fraction images obtained from FCM classification by preparing FERM algorithm in MATLAB.
4. Obtain the optimum combination of all the parameters corresponding to maximum accuracy.

1.3 Thesis Outline

Present thesis has been divided in six chapters.

Chapter 1outlines the introduction and main objectives of present study. The applicability remote sensing for monitoring of glaciers has been described in detail in this chapter. Various techniques to take care of mixed pixels in case of coarse resolution data are presented briefly.

Chapter 2 presents a comprehensive literature review on the concepts used in various soft classification techniques; their mathematical background, limitations and advantages. A brief literature review on various accuracy assessment measures has been present in detail. Need for accuracy assessment, various assessment parameters are presented comprehensively.

Chapter 3 is allocated for the description of the study area and data used for the complete work. Methodology used for present study has been discussed in detail in this chapter.

Chapter 4 illustrates various results obtained for glacier mapping. Variation of FCM with 'm' value, training data size and A-norms has been illustrated through graphs.These results are discussed in brief in relation to objectives in the present chapter.

Chapter 5 concludes the research work and recommends the future work.

Chapter 6 contains references from which we acquired a lot of knowledge for this project work.

CHAPTER 2

LITERATURE REVIEW

A remote sensing image usually consists of both pure as well as mixed pixels. Mixed pixels generally occur at the boundaries of different geographical features, or along the boundary line of linear features like rivers, sea coast etc. Major cause of mixed pixels in an image is coarser spatial resolution of sensor. Coarser the resolution more will be the number of mixed pixels in an image. Thus if image data of moderate resolution sensors like MODIS with spatial resolution of 250m, 500m, 1000m or AWiFS (Advanced Wide Field Sensor) with spatial resolution of 56m are used; it becomes imperative to take mixed pixels into account while classifying the image.

Fisher (1997) gave four land cover scenarios which causes mixed pixel problem. When size of geographical feature is smaller than size of a pixel, it consists of more than one class within it.
In case when two classes are separated by hard boundary and pixel lies at boundary of two classes, it contains both the classes. Third case arises when one class gradually transforms into other one like forest area gradually transforming into shrubs/grasses. Finally, when width of a feature is less than the size of a pixel, it consists of multiple classes. Usually this case arises in case of small water streams.

To overcome the problems of mixed pixels various soft classification techniques are employed.All these techniques are known as soft classification techniques. Instead of assigning each pixel only one class, these techniques assign multiple classes to a pixel. Each class has a certain proportion [0,1] in a pixel depending on its areal extent in pixel area. More is the areal extent, more will be the proportion of corresponding class.A brief overview of these techniques is discussed in present chapter.

One of the methods to solve the mixed pixel problems is Linear Mixing Model (LMM). This method attempts to model the spectral response from a mixture of classes within a mixed pixel and the approach is generically termed as ‘mixture modeling’. It is based on the principal that DN of a mixed pixel is linear sum of mean of spectral response of constituent classes weighted over fraction of constituent class in that particular pixel.

Linear mixture model assumes that spectral response of a pixel varies linearly with the proportion of its constituent classes. Image pixels representing pure class spectral response i.e. pure pixels are identified and their spectral response is termed as ‘end-member spectra’.

Mathematically, spectral response of a pixel can be defined as linear sum of components described below:

$$x_i = \sum_{j=1}^{c} M_{ij}\left(f_j\right) + e_i$$

Where

x_i = spectral response of a mixed pixel in ith band

M_{ij}=endmember spectral response of jth class in the pixel for ith band

f_j=fraction of jth class in pixel for ith band

e_i=error term for ith band

j = 1, 2, 3 ... n (Number of classes assumed)

i = 1, 2, 3 ... m (Number of Spectral bands for the sensor system)

Matrix M_{ij} represents end member spectra where rows represent spectral response of number of constituent j classes and column represents the respective ith spectral band of pixel. The error term (e_i) is included because spectral response of a pixel will never vary linearly with its individual response of constituent classes due to atmospheric interaction of reflected radiations. Two constraints are added to above model. First, the sum of the proportions of all the classes for any pixel equates to one. Also, the proportion values of classes must be non-negative.

Both constraints are described below with mathematical equations:

Constraint 1: $\sum_{j=1}^{n} f_j = 1$

Constraint 2: $f_j \geq 1$

For any multispectral image, spectral response of each pixel of this image can be defined as a linear combination of a finite set of components as

$$x_1 = M_{11}f_1 + M_{12}f_2 + M_{13}f_3 + \cdots + e_1$$

$$x_2 = M_{21}f_1 + M_{22}f_2 + M_{23}f_3 + \cdots + e_2$$

$$x_3 = M_{31}f_1 + M_{32}f_2 + M_{33}f_3 + \cdots + e_3$$

In above equations, number of unknowns includes $f_1, f_2, f_3\ and f_4$. Since number of unknowns and number of equations are equal thus these unknowns can be calculated easily. Three possible cases arise here as listed below:

Case 1: n = c

Where,

n= Number of bands

c=number of classes

In this case number of unknowns can easily be found out.

Case 2: n > c

In this case these equations will result in infinite solutions. In order to find class fractions sum of squares of error is minimized. This method is known as Constrained Least Squares Method (CLSM). It estimates the proportion of each component within a pixel by minimizing the sum of squares of the errors.

Case 3: n < c

In this case all number of unknowns cannot be determined.

One of the major limitations of this method is the assumption of linear variation of spectral response of constituent classes within a pixel which is quite rare in practical scenarios.For example in vegetation region radiations interaction with environment can make spectral response non-linear. Moreover, sum of square of errors approach poses problems associated with outliers.

Maximum likelihood classification, widely used for hard classification can also be used to produce class proportions in a pixel. MLC involves a large computational work as it calculates the probability of each class belonging to a particular pixel. But since in hard classification only class with maximumprobabilty is considered, rest of the computation becomes unavailable to user. The class probabilities calculated can be used to define the proportion of class in a pixel. Foody et al. (1992)

The Maximum Likelihood Classification tool considers variance and covariance of the class signatures when assigning each pixel to one of the classes. With the assumption of normal distribution of training data, a class can be characterized by the mean vector and the covariance matrix. Using these characteristics for each pixel value, the statistical probability is computed for each class to determine the membership of the pixel to the class.

The estimated probability density function for class $w_i$is computed using the equation:

$$p\left({}^{x}/_{w_i}\right) = \frac{1}{\sigma_i (2\pi)^{\frac{1}{2}}} e\left[\frac{-1}{2} \frac{(x-u_i)^2}{\sigma_i{}^2}\right]$$

where,

e= base of natural logarithm,

x=pixel DN value,

u_i=estimated mean of training data of particular class,

$\sigma_i{}^2$=estimated variance of all the brightness values in a class

Therefore, we need to store only the mean and variance of each training class to compute the probability function.

Therefore, to classify a pixel in the multispectral remote sensing dataset with an unknown measurement vector X, a maximum likelihood decision rule computes the product for each class and assigns the pattern to the class having the largest product in case of hard classification. This product for any class can also be related with probability of occurrence of respective class in a pixel. Thus it gives the fraction of all the classes constituting a pixel.

The maximum likelihood method has an advantage from the viewpoint of probability, but there are few cases which need to be addressed:

Case1: Training data should be sufficient enough to calculate mean vector and the variance-covariance matrix.

Case2: The inverse matrix of the variance-covariance matrix becomes highly unstable when high correlation exists between two bands. This case arises when ground truth data is very homogeneous.

Case3: When training data is not normally distributed (true in most cases), the maximum likelihood method may give inaccurate results.

Thus the use of likelihood functions as soft classifier is valid only if the classes of interest are high separable and distributions overlap only slightly.(Boucher, 2009)(Tatem, 2001)

Fuzzy based soft classification techniques are gaining wide popularity as they precisely represent the uncertainties present in nature. Fuzzy set theory eliminates the output in the form of 0 or 1. Instead, it takes into account partial membership values [0,1].Each member of fuzzy set has a membership value ; 0 represents null membership and 1 represents full membership. Fuzzy set theory introduces uncertainty in the form of degree of membership and eliminates hard and crisp boundaries. It illustrates a scenario where one individual pixel does not belong to a single class but it belongs to all the classes with different degree of membership.

Fuzzy c-means is an iterative process based on fuzzy set theory which starts by assigning pixels to classes randomly. Next, every pixel is given a membership degree for each class. With every iteration, degree of membership and center of cluster is updated. Once the center of clusters has been determined a pixel that lies close to the center of a cluster will have a higher degree of membership than another pixel which lies far away than the center of the cluster. The iteration aims at minimizing the objective function that represents the distance of a pixel from the center of the class weighed by membership value of class in that particular pixel. Dunn (1973) ,Bezdek et al., (1984).

FCM aims at minimizing the following Objective function:

$$J_m = \sum_{k=1}^{n} . \sum_{i=1}^{c} (u_{ik}) \quad ^{m} (d_{ij})^2$$

where,

n= total number of pixels,

c= number of classes,

u_{ik}= fuzzy membership value of kth pixel in ith class,

d_{ij}=distance of pixel i from center of cluster j,

m= fuzzy exponent; ranges from [1, ∞].

Further, distance d_{ij} is represents distance between pixel(X_i) and cluster center of jth class x_j and is given by,

$d_{ij}{}^2 = (X_i - x_j)^2 = (X_i - x_j)^T$ A (X_i - x_j)

Where A is the weight matrix. Out of various A-norms, three A-norms are widely used each induced by specific weight matrix. These A-norms are Euclidean, Mahalanobis and Diagonal norms. Formulations for each norm are as under (Bezdek, 1981):

A= I	Euclidean Norm
A= $D_j{}^{-1}$	Diagonal Norm
A= $C_j{}^{-1}$	Mahalanobis Norm

Where I is the identity matrix,D_j is the diagonal matrix with its diagonal elements as eigen values of variance covariance matrix for each class, C_j is given by:

$$C_j = \sum_{i=1}^{n}((X_i - x_j)(X_i - x_j)^T)$$

The typeof A-norm in the distance measurementdetermines the shape of the clusters. The norm influences the clustering criterion by changing the measure of dissimilarity. The Euclidean norm induces hyperspherical clusters (surfaces of constant membership are hyperspheres). Both the diagonal and the Mahalanobis norm generate hyperellipsoidal clusters. With the diagonal norm, the axes of the hyperellipsoids are parallel to the coordinate axes, while with the Mahalanobis norm the orientation of the hyperellipsoid is arbitrary.

Shapes of clusters formed for each of A-norm has been shown in fig. on next page.

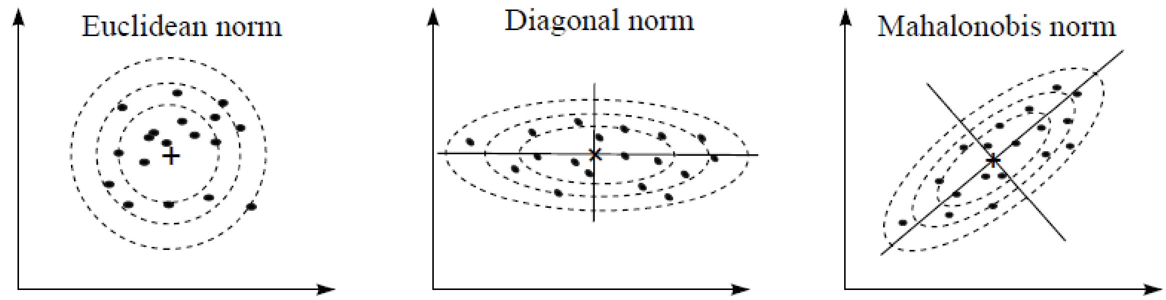


Fig. 1 Different A-norms used in clustering

Membership value is calculated using following equation:

$u_{ij} = \sum_{k=1}^{c}(.\frac{x_{ij}^2}{x_{ik}^2})\hat{}(-1/(m-1)$i ϵ (1,…,.n) , j ϵ (1,….,c)

where,

$$x_{ik}^2 = \sum_{j=1}^{c} x_{ij}^2$$

And $x_{ij}^2 = |d_i - c_j|^2$

Center of cluster is calculated by following equation:

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m}$$

The iterative process terminates when either number of user specifies maximum number of iterations are achieved or when objective function gets minimized to predefined specific value. The output of soft classification largely depends on fuzzy exponent ‘m’ value which represents degree of fuzziness in FCM .The value of ‘m’ varies from [1,∞] where 1 represents no fuzziness at all (hard classification) and as its value increases beyond 1 fuzziness increases accordingly . Increase in fuzziness represents almost equal share of all the classes within a pixel and as ‘m’ value gets close to 1 pixel belongs more to one class than others. However, till now not much work has been done to find out optimum value of ‘m’.

Fuzzy c-means generally produces accurate class composition estimates when all classes have been defined and included in the training phase of the classification. However, the presence of untrained classes would degrade the estimation accuracy of the fuzzy c-means. Moreover FCM classification accuracy largely depends upon the value of fuzzy exponent taken and there is no such criterion to choose m value. Therefore a counterpart of this technique, namely Possibilistic C (Krishnapuram, 1996) (Krishnapuram, A possibilistic approach to clustering, 1993)-Means is used for robustness towards the untrained classes (Krishnapuram and Keller, 1993; Krishnapuram and Keller, 1996).

PCM is similar to FCM except for few modifications which applied to overcome the limitation of FCM when all the classes have not been trained. In PCM, membership of a pixel in class doesn't depend on number of classes whether classes are trained or not. Thus presence of untrained classes can degrade accuracy of FCM but not of PCM.

One of the primary constraint in FCM stating that the sum of all the membership values within a pixel sums up to one is removed in PCM.

Here,

$$\sum_{j=1}^{c} u_{ij} \neq 1$$

Membership matrix is calculated using following equation:

$$u_{ik} = \frac{1}{1 + \left(\frac{{d_{ik}}^2}{n}\right)}.$$

where,

n = parameter that defines the distance at which membership value to class equals 0.5

When classes have not been defined exhaustively then PCM will give accurate results as compared to FCM.

Image classification through neural network techniques has been gaining wide popularity because its ability to solve nonlinear mixture problems (Carpenter et al., 1999). Data classification in remote sensing has been solved using a variety of feed-forward neural network. All these models are differentiated on the basis of arrangement of neurons in network and the type of activation function used.

Recently, neural network method has been applied to remote sensing data for classification problems.Three-layer back propagation algorithm is used to classify Landsat TM data (McClelland et al.)..These methods don’t use any prior assumption about probability distribution of dataset thus this method will give accurate results as compared to other methods if prior assumption about data is not met. Despite of its merits there is a major drawback of neural networks that they are insufficient in explaining the physical system being modeled in the entire process.That’s why they are referred as ‘black box’ methods.(Liang, 2008)

Various researchers have studied all these methods and compared the results to find out which method is better in which conditions. A brief review of it has been presented in this chapter below.

Foody (1995)concluded thataccuracy assessment techniques used for conventional hard classification cannot be used for evaluating soft classifiers as pixels have partial class membership. Here, entropy measure is used for evaluating accuracy which illustrates how closely the output of soft classification represents true class fractions. Cross entropy is calculated for both soft classified data and reference data from probability distribution of classes within a pixel.

Foody (1996) has evaluated the performance of FCM and fuzzy neural networks for land-cover mapping. He concluded that in case of mixed pixels where a pixel contains more than one class; hard classification might not be accurate methods for image classification thus either it has to be softened to get more accurate results or fuzzy classification approach can be used.Outputs from fuzzy c means classification and fuzzy artificial neural network came out to be more accurate than conventional hard classifications. Although posterior probabilities obtained from softening of hard classification were not exactly same as that of FCM output. He studied the effect of different values of fuzzy exponent and concluded that classification output becomes most accurate when m is taken as 2.0.

Zhang et al. (1997)used fuzzy approach for suburban land cover mapping and for comparative analysis hard and fuzzy classifications were tested using their respective evaluation techniques .concluded that fuzzy c means has outperformed conventional hard cover classification. Kappa coefficient values were found to be almost doubled when fuzzy approach has been used.

Bastin (1997)performed linear mixture modeling and artificial neural networks to classify Envisat MERIS and Landsat ETM image data set. He compared these soft classification outputs with hard classification. Linear mixture model and artificial neural networks gave better results than hard classifiers. To produce the test data, Landsat image of study area was classified using Maximum likelihood classification and then resampled. He concluded that there no significant difference between LMM and artificial neural network methods in case of certain land cover classes. However, ANN outperformed LMM where pixels having high degree of mixture are to be classified.

Zang and Foody (2001)used fully fuzzy classification approach (which takes into account fuzziness of individual pixel in all three stages of supervised classification i.e. training, allocation and testing) and compared the results with partially fuzzy classification approaches (fuzziness is accounted in only one or two stages of classification). Results showed that fully fuzzy approaches have performed better than the partial one. Two algorithms were used in fully fuzzy mode one being supervised fuzzy c means and other one is artificial neural networks. It is further suggested that derivation of fuzzy ground data and classified data is the most critical factor for fully-fuzzy classifications.

Ibrahim et al. (2005)compared the results of soft classification techniques when they are used in fully fuzzy mode .Uncertainties in class allocation resulted due to the presence of mixed pixel has been included in training, allocation and testing of data set. Three soft classification techniques – maximum likelihood classification (softened), fuzzy c means classifier and possibilistic c means classifiers were used and results showed an increase in classification accuracies with fully fuzzy mode. FERM has been used for accuracy assessment. Possibilistic c means gave the highest accuracy out of these three approaches.

Okeke et al. (2005) proposeda new and efficient method for determining the optimum value of fuzzy exponent for fuzzy c means classification. In this method, fuzzy classification output has been used to predict the original data .For different values of fuzzy exponent, the difference between original data and predicted data has been calculated. Fuzzy exponent value which gives the minimum difference is considered the optimum value. Class members values obtained from optimum 'm' value were found to be very close to true class proportion values. This method has been suitable for fuzzy c means algorithm and it's all extensions as well.

Jain et al. (2008) used ANN to study the decreasing trend of water flow in wetland area using LISS-II and LISS-III data. They concluded that correlation coefficient and determination coefficient can be improved by using ANN technique though ANN is not suitable for base flow simulation due to very large spatial scale used.

Out of all these soft classification tools, fuzzy c means has been used in the present study. Various previous studies showed that fuzzy c means can be used as an effective tool to take into account the uncertainties involved in ground truth data. However, it has limitations also. Output of FCM varies greatly with fuzzy exponent value 'm'. No formulation or method has been discovered till yet to find 'm' exponent. Different studies have suggested different values of 'm'. In this present study also value of 'm' has been varied from 1.1 to 6 in order to find the optimum value corresponding to maximum accuracy.

Results of any classification technique are validated through various accuracy assessment measures. Generally, in accuracy assessment classified data obtained by using algorithms is compared to reference data (which itself can be inaccurate) at selected sample points class by class. Use of a conventional error matrix (confusion matrix) is widely acceptable method for accuracy assessment of hard classification. Typical error matrix is a square matrix where columns represent reference data and rows represents classified data values.

Various measures like Overall Accuracy, User's accuracy, producer's accuracy, kappa coefficient are used to evaluate accuracy. Overall accuracy (OA) is the simplest statistic which represents the ratio of sum of diagonal elements to sum of total elements (sum of all rows/columns), thus representing number of correctly classified pixels out of entire sample pixels. Additionally, accuracies of individual classes can be evaluated through users and producers accuracies. User's accuracy is obtained by dividing total number of correct pixels in a particular class to sum of pixels of that class obtained from classified data. It is called so as it represents how accurately user has classified the pixels. If no. of correct pixel of a class is divided by number of class pixels in reference data then it represents producers accuracy as producer of the image is interested to know how well image is classified.

Major drawback of OA is that it doesn't consider off-diagonal elements of error matrix; thus there's another statistics called kappa statistic which is widely used for evaluation of accuracy. Generally, OA is given by following formulation:

$$\text{Overall Accuracy (OA)} = \frac{\text{sum of diagonal elements of error matrix}}{\text{sum of all the elements of matrix}}$$

Following table depicts typical error matrix and various accuracy measures

Classified data ↓ Reference Data →

Class	A	B	C	Row Total
A	25	5	5	35
B	4	20	3	27
C	7	9	30	43
Column Total	36	34	38	108

Table 1: Sample Error Matrix

Overall Accuracy:

OA= (25+20+30) / 108 = 85 %

User's and producer's accuracy is represented by following table:

Class	User's Accuracy (%)	Producer's Accuracy (%)
A	25/35= 71.4	25/36=70
B	20/27=74.07	20/34=60
C	30/43=70	30/38=79

Table 2: Sample User's and Producer's Accuracy

Kappa Statistics:

$$K = \frac{N \sum_{k=1}^{r} x_{kk} - \sum_{k=1}^{r} (x_{k+} \times x_{+k})}{N^2 - \sum_{k=1}^{r} (x_{k+} \times x_{+k})}$$

Where, r represents number of rows in error matrix, x_{kk} represents diagonal elements with kth row and kth column, x_{k+} and x_{+k} represents marginal totals of row and column, N represents total number of observations (Bishop et al., 1975).

Confusion matrix method has been used as a standard method for accuracy assessment of hard classifiers where each pixel is assigned only one class. However for soft classifiers which give fraction images as output; standardized accuracy assessment method has not been established yet. Various accuracy assessment procedures were put forward like Entropy, Cross Entropy ,Euclidean and the L1- distance, Correlation Coefficients, fuzzy error matrix , among these fuzzy error matrix is most acceptable method. FERM (Fuzzy Error Matrix) represents the generalized form of conventional error matrix except for the fact that it is based on fuzzy set theory. The layout of FERM is similar to traditional error matrix with columns representing reference data (R_n) and rows representing classified data (C_m). Since FERM elements represent class fractions so they can be real numbers.

Various fuzzy set operators are used to construct FERM out of classified and reference data sets. Once FERM is prepared, it should exhibit following two characteristics

a). Diagonalization: If classified data and reference data match perfectly then matrix should be a diagonal matrix.

b).Marginal Sums: Sum of grades of classified and reference data should be equal to marginal sums.

Agreement –disagreement measures are required to conform to equation (1) where A and D represents agreement and disagreement respectively.s_{nk}'andr_{nl}' denote underestimation or overestimation errors in nth pixel i.e. disagreement . s_{nk}andr_{nl} represents correctly matched class fraction in nth pixel i.e. agreement.

$$C(s_{nk}, r_{nl}) = A(s_{nk}, r_{nl}) \quad \text{if } k=l \qquad \text{or} \qquad D(s_{nk}', r_{nl}') \quad \text{if } k \neq l \tag{1}$$
$$s_{nk}' = s_{nk} - \min(s_{nk}, r_{nl})$$
$$r_{nl}' = r_{nl} - \min(s_{nk}, r_{nl})$$

Various operators used to establish FERM need to satisfy all these properties enlisted in the table given above. A brief overview of these operators is discussed in next article.

Minimum operator (MIN) is the classic fuzzy set intersection operator which measures the maximum overlap between the classes in classified data and reference data. It satisfies all the basic properties enlisted in table 3; but it does not satisfy marginal sums property of FERM. Since it always estimates maximum overlap between the classes thus marginal sum value may exceed actual reference/ classified fraction values. Moreover, it produces non-nulloff-diagonal values in case of perfect match. Thus use of MIN operator is limited for to obtain FERM.
Similarity Index (SI) operator is a modification over MIN operator. MIN operator values are normalized by sum of grade values. Thus overestimation of marginal sums is reduced to a certain extent but still it doesn't satisfy it properly. Moreover, homogeneity property and diagonalization are also violated.

Product operator (PROD) is a probabilistic operator which measures the joint probability of a pixel in reference and classified data set belonging to a class; given that pixels are classified independently. If we consider a random point in a space consisting of n pixels then joint probability that point belongs to k class in classified data and l class in reference data will be the product of s_{nk} and r_{nl} ; given that class fractions s_{nk} and r_{nl} are obtained independently.
LEAST operator (Pontius &Connors, 2006) measures minimum overlap between class fractions of classified and reference data thus underestimating the marginal sums.

Table showing these four basic operators and their mathematical formulations is given below:

Operator	Agreement	Disagreement	Soft Confusion
MIN-PROD	$min\ (s_{nk}, r_{nl})$	$\frac{(s_{nk}{}' \times r_{nl}{}')}{\sum_i r_{ni}{}'}$	Constrained expected
MIN-MIN	$min\ (s_{nk}, r_{nl})$	$min\ ((s_{nk}{}', r_{nl}{}')$	Constrained maximum
MIN-LEAST	$min\ (s_{nk}, r_{nl})$	$min\ ((s_{nk}{}', r_{nl}{}' - \sum_i r_{ni}{}', 0)$	Constrained minimum

Table 3: Basic Operators

Neither of the four basic operators satisfy diagonalization characteristic, thus basic operators are rarely used. Thus there came the concept of composite operators which includes two operators at a time. Since MIN operator is the only operator which satisfies all the basic properties of agreement and disagreement thus in composite operators MIN operator is widely used for agreement i.e. diagonal values of FERM.

MIN-MIN, MIN-LEAST and MIN- PROD are the three composite operators. MIN-MIN operator uses MIN operator for calculating agreement first and then disagreement is measured based on overestimations and underestimations resulted from former measures. MIN-LEAST uses MIN for agreement and LEAST for disagreement values. Similarly, MIN-PROD
(Pontius and Cheuk, 2006) uses MIN for agreement and normalized PROD for disagreement.

Table showing all these composite operators is given below:

Operator	Agreement	Disagreement	Soft Confusion
MIN-PROD	$min\ (s_{nk}, r_{nl})$	$\frac{(s_{nk}' \times r_{nl}')}{\sum_i r_{ni}'}$	Constrained expected
MIN-MIN	$min\ (s_{nk}, r_{nl})$	$min\ ((s_{nk}', r_{nl}')$	Constrained maximum
MIN-LEAST	$min\ (s_{nk}, r_{nl})$	$min\ ((s_{nk}', r_{nl}' - \sum_i r_{ni}', 0)$	Constrained minimum

Table 4: Composite operators

MIN-PROD operator has been used in the present study for the assessment of soft classification. It uses MIN operator for diagonal elements and PROD operator for off-diagonal elements thus combining fuzzy and probabilistic approaches. Firstly, maximum overlap between corresponding classes in classified and reference data is calculated using MIN operator following the calculation of expected overlap between residual class fractions using PROD operator. It neither underestimates the values as in LEAST operator nor overestimates as in MIN operator, thus giving a mid-value between two extreme values as given by MIN and LEAST operators.

A brief review of the studies on various accuracy assessment measures of soft classification techniques has been presented in this chapter below.

Binaghi et al. (1999)proposed a new method for soft classification accuracy assessment which is the extended form of conventional error method used for accuracy assessment of hard classifiers. Partial membership of classes is taken into account while forming fuzzy based error matrix. Fuzzy Error Matrix similar to classical error matrix has been proposed where each element of matrix represents partial membership of class and it can have fractional value unlike in conventional error matrix. Various accuracy parameters like overall accuracy, producer's accuracy, user's accuracy, kappa coefficient etc. are calculated.

Pontius et al. (2006)proposed the concept of MIN- PROD composite operator for the preparation of fuzzy error matrix. Conventional methods use Boolean operators to compute error matrix in case of hard classification and multiplication operator like minimum operator in case of soft classification. Various limitations associated with conventional methods of computation of error matrix are described in this study. A composite operator is being proposed which eliminates most of the errors associated with basic operators.

Silván-Cárdenas et al. (2008)proposed a more general soft classification accuracy assessment method that takes into consideration the uncertainty in class distribution within a pixel. Major drawbacks for previous accuracy assessment methods based on fuzzy error matrix have been identified. A new fuzzy error matrix called as subpixel confusion–uncertainty matrix (SCM) in which each elements represents confusion intervals in the form of a center value ± maximum error was developed which accounts for uncertainty in class membership in a pixel. Various accuracy assessment parameters like overall accuracy, user's and producer's accuracy can be calculated from SCM.

Jain et al. (2008) used ANN to study the decreasing trend of water flow in wetland area using LISS-II and LISS-III data. They concluded that correlation coefficient and determination coefficient can be improved by using ANN technique though ANN is not suitable for base flow simulation due to very large spatial scale used.

Dwedi et al. (2012)stated that soft classification based on fuzzy sets like FCM and PCM has been used as a standard method to compare the performance of other soft classifiers. Accuracy of FCM and PCM classifier has been evaluated through FCM, SCM and kappa coefficients. Out of composite operators MIN-LEAST and MIN-MIN, former one has given better results with highest classification accuracy and highest kappa coefficient. The optimized value of fuzzy exponent 'm' came out to be 4.

Prasad M.S. et al. (2015)proposed new method for assessment of soft classifier’s accuracy. Two measures were proposed: fuzzy similarity measures and fuzzy certainty measures. Firstly synthetic data is prepared and membership values are calculated using FCM algorithm on this data. Fuzzy similarity and fuzzy certainty measures are evaluated for each class using membership values and reference dataset. Results showed that both these measures calculated values differently and gave different results. Fuzzy certainty measures used in this study were similar to that used to measure goodness of fit in various statistical models.

CHAPTER 3

STUDY AREA, DATA USED AND METHODOLOGY

3.1 Study Area

Study area for the present work is a part of Chenab Basin situated in Lahaul and Spiti district of Himachal Pradesh. It extends from latitude 31.74º N to 33.20º N and longitude 77.33º E to 78.23º E. It forms a part of Greater Himalayas. Two main rivers Chandra and Bhaga originate from Chenab basin which further meets at Tandi to form Chenab river system. There are around 200 glaciers in Chenab basin of varying sizes. Present study area forms a part of Samudratapu glacier; second largest glacier in upper Chandra basin after Bara SigriGlacier. Fig. 2shows the AWiFS FCC (standard) image of the study area.

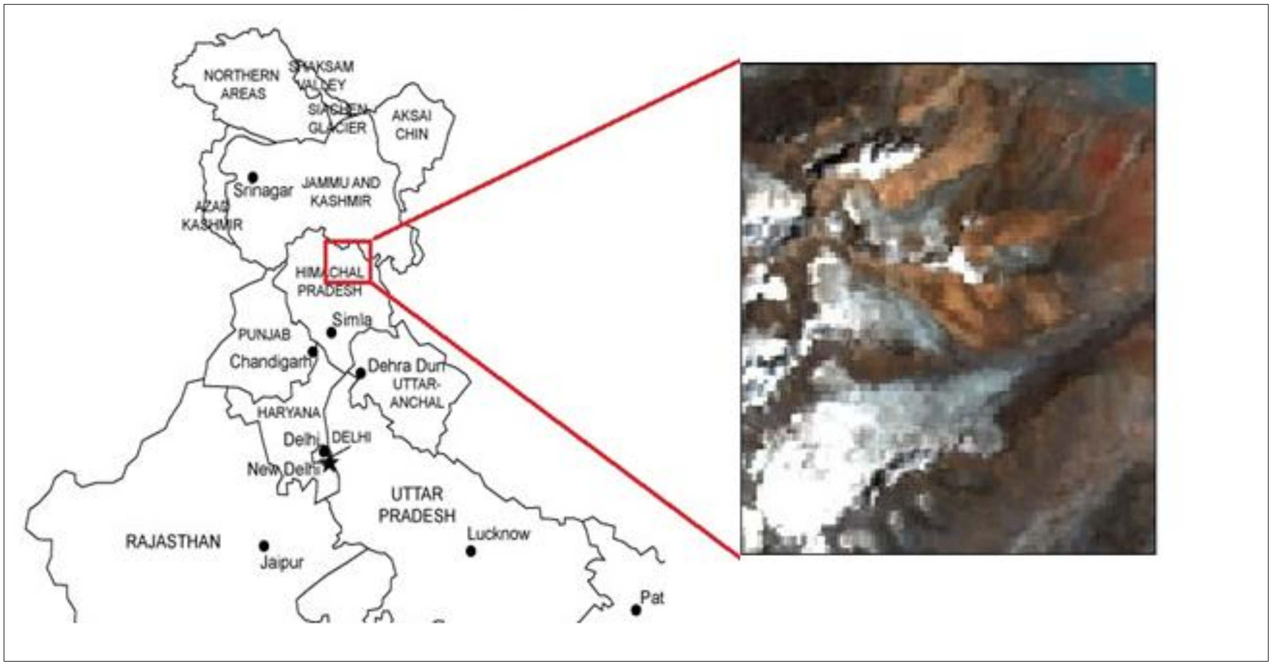


Fig. 2 Study area of Present work

3.2 Data Used

AWiFS and ASTER satellite images are used for soft classification and preparation of reference map/reference fraction images respectively. AWiFS image shown in Fig2(date of acquisition: 7^{th} September, 2005) has been used for soft classification purpose. AWiFS sensor operates in four spectral bands (band 1→ 0.52-0.59μm, band 2 →0.62-0.68μm, band 3→0.77-0.86μm and band 4 →1.5-1.77μm) with a moderate spatial resolution of 56m and radiometric resolution of 10 bits. Because of its high radiometric resolution, saturation of bands over snow is reduced to a greater extent thus making it suitable for glacier studies.

ASTER is high-resolution sensor on board terra satellite with a spatial resolution of 15 m. ASTER (date of acquisition: 8th September, 2005) data has been used to prepare the reference fraction images which are further used to evaluate accuracy of soft classification of AWiFS image.

3.3 Methodology

Firstly, a comprehensive study of soft classification techniques has been conducted. Principles involved in each method, their mathematical background and advantages and limitations have been studied thoroughly. Fuzzy c means classification technique and its algorithm has been studied in detail. Since study area chosen is a glacier area thus general literature about various glaciers, different facies of glaciers have been studied. A thorough review of spectral response of glacier features has been done in order to identify training data for supervised FCM classification. A flowchart of complete methodology is shown in figure 1.

A flowchart of entire methodology used in present study is given below:

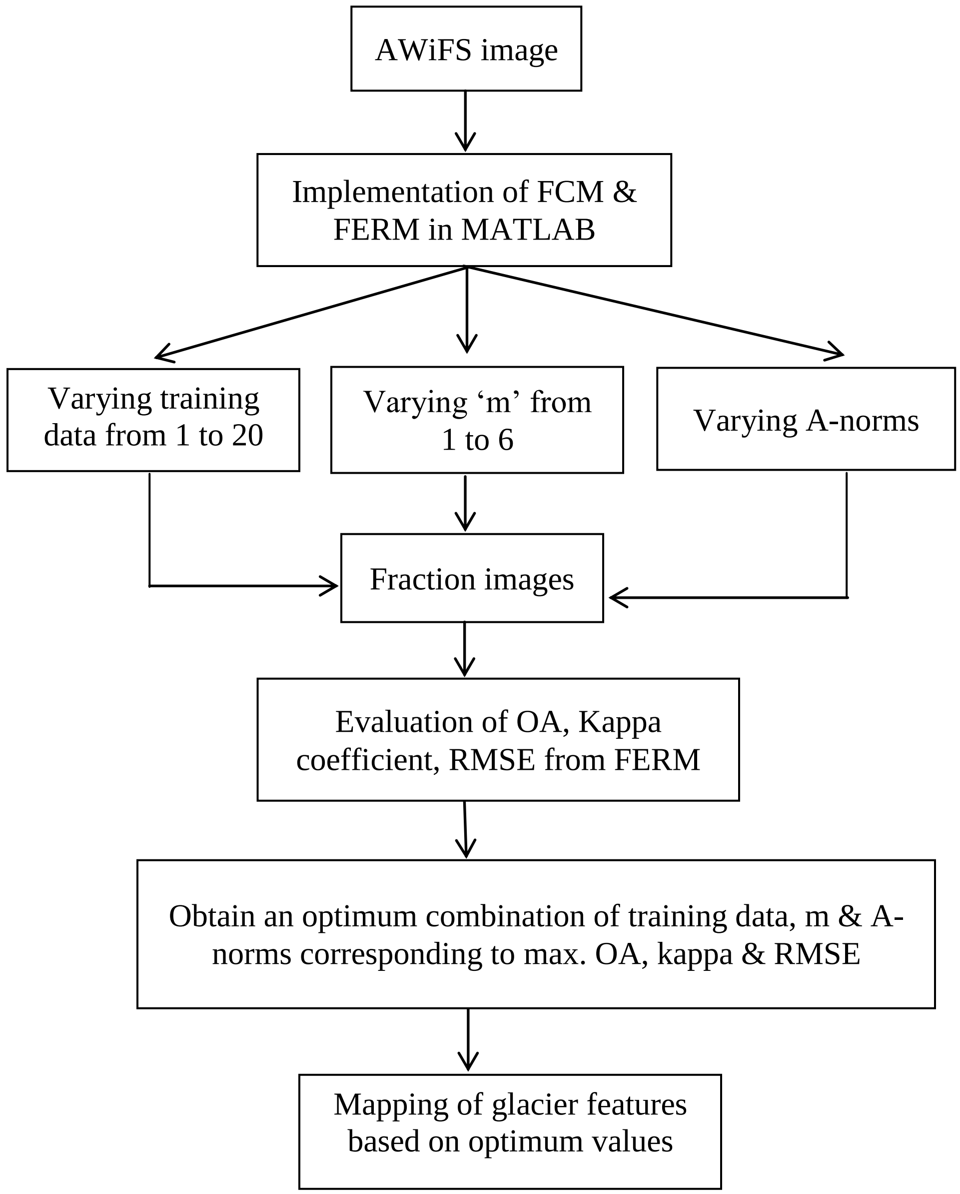


Fig. 3 Flowchart of methodology

Seven classes have been taken into account viz. snow, ice, ice mixed debris, valley-rock, lake, debris, shadow. Input AWiFS image has both pure and mixed pixels, but for training data only pure pixels of a particular class have been selected. Earlier eight classes were chosen; eighth one being moraine. Moraine is a part of debris (rock fragments) which lies on the boundary of glacier. The chemical composition of debris and moraine is same; only difference being lower spectral response of moraine than debris as moraine lies in direct contact with glacier snow/ice.

A considerable part of debris area in input image was under shadow hence showing lower spectral response similar to moraine. Thus it was difficult to choose training data which completely separates these two classes. When FCM was implemented for classification, very low accuracy values were obtained for both debris and moraine. Training data was improved again and again but it couldn't help. Finally, these two classes were merged together to form a single class 'debris'.

Another major problem faced during classification is with water bodies' fraction image. There are three water streams in input image merging to form a large lake. While selecting the training data, training pixels from these streams are also included. But spectral response of these pixels is different from main lake pixels. Due to the presence of debris underneath, their spectral response in SWIR band is much higher than lake pixels as shown in fig. 5.Very low accuracy values are obtained for water bodies' fraction images. Water stream pixels show higher membership value in class 'debris'; resulting in very low accuracy values. Thus instead of taking water bodies as a class, only lake pixels are considered neglecting water streams pixels to form new class 'lake'. Water stream pixels are considered to be a part of debris.

Following figure shows the spectral response of water streams pixels and lake pixels.

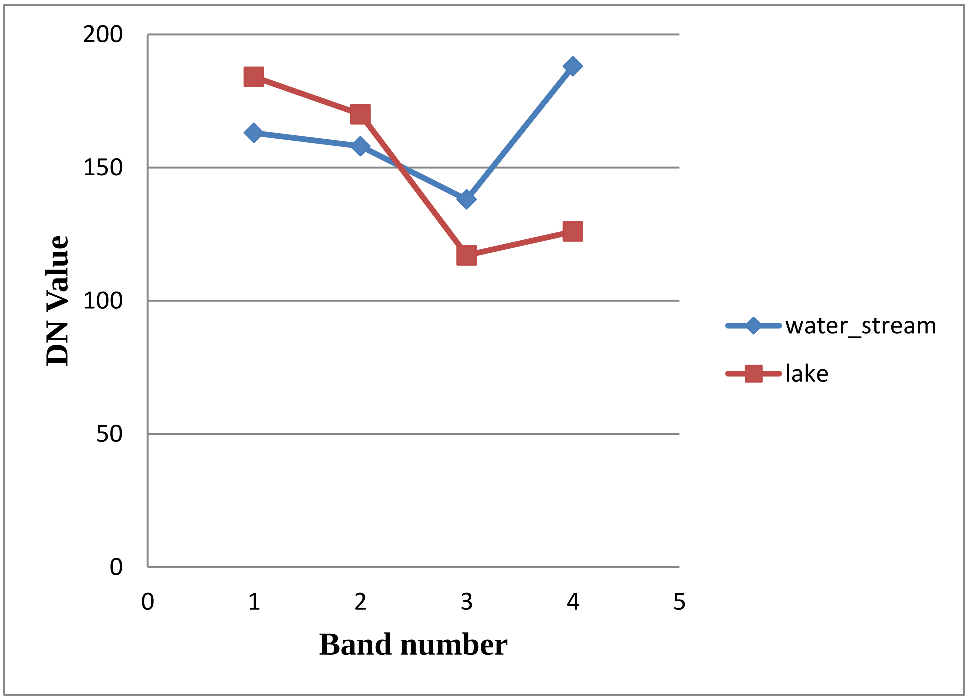


Fig. 4 Spectral response of water stream pixels and lake pixels

Finally, seven classes are considered for soft classification and training data has been selected for each of the class.While selecting the training data, it has been made sure that AOI's (Area of Interest) are well placed and are uniformly scattered thought the class. Signature files of each of class have been used to train FCM algorithm. Matlab software has been used to develop the algorithm of Fuzzy c means in supervised mode. Also FERM algorithm for accuracy assessment of fraction images with MIN-PROD operator has been developed in Matlab.

Mean spectral response for each of the class obtained from training data has been shown in fig. 6 on next page.

Following figure shows the mean spectral response of all the seven classes chosen:

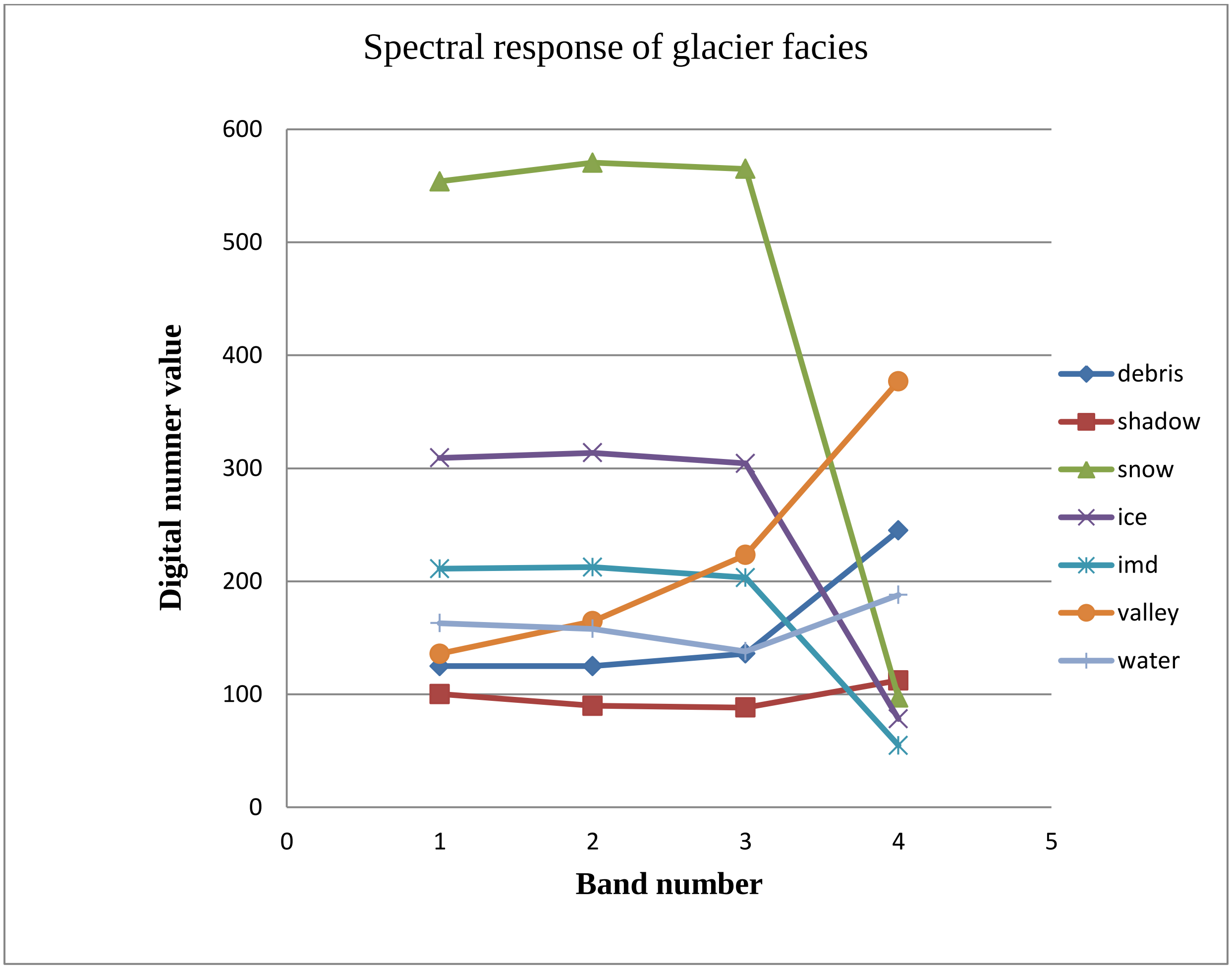


Fig. 5 Mean spectral response of glacier facies

Once training data has been selected , it is used as an input to FCM algorithm .Main objective of this study is to evaluate the performance of FCM algorithm with three varying parameters i.e. training data size, fuzzy exponent ‘m’ and A-norms. Training data size is varied from 3 percent up to 21 percent at an interval of 3 percent and ‘m’ value has been varied from 1.1 to 7 at an interval of 0.2 .Accuracy assessment of output fraction images has been carried out using FERM to evaluate overall accuracy, kappa coefficient and RMSE.Keeping training data size constant; fuzzy exponent ‘m’ value is varied from 1 to 7 for all 3 A-norms. Fraction images are obtained for all three A-norms with varying ‘m’ values. Three values of ’m’ corresponding to each of A-norm have been selected based on maximum accuracy. Out of these three values, ‘m’ value corresponding to maximum accuracy has been selected. Similar procedure has been followed for rest of training data sizes and ultimately an optimum combination of training data size, fuzzy exponent ‘m’ and A-norms has been obtained corresponding to maximum accuracy.

CHAPTER 4

RESULTS AND DISCUSSIONS

Results obtained for supervised Fuzzy C-means classification of AWiFS image has been discussed in detail in present chapter.

Firstly , classification of image is carried out for training data size of 21 percent for each of A-norms with varying fuzzy exponent value 'm' from 1 to 7.Overall accuracy and Kappa coefficients came out to be maximum for mahalononbis distance followed by diagonal distance and then least for Euclidean distance. For constant training data size and A-norms, value of 'm' shows an increasing trend up to 1.4 and then a decreasing trend. Maximum values of OA and kappa coefficient obtained are 74 percent and 0.63 respectively.

Figure below shows the variation of 'm' values for three A-norms with training data size of 21 percent.

Euclidean distance

m	1.1	1.3	1.5	1.7	1.9	2	2.2	2.4	2.6
OA	66.80	67.267	66.417	64.163	61.165	59.505	56.206	53.078	50.269
Kappa	0.537	0.544	0.535	0.509	0.475	0.456	0.419	0.385	0.355
RMSE	0.24712	0.23412	0.22285	0.21426	0.20931	0.2082	0.20821	0.21036	0.21375

2.8	3	3.2	3.4	3.6	3.8	4	4.4	5	6
47.779	45.594	43.712	42.059	40.622	39.358	38.249	36.399	34.318	31.99
0.329	0.306	0.287	0.27	0.2704	0.242	0.231	0.222	0.21	0.19
0.21771	0.22183	0.22586	0.22966	0.23319	0.23643	0.23938	0.24449	0.25054	0.25764

Diagonal distance

m	1.1	1.3	1.5	1.7	1.9	2	2.2	2.4	2.6
OA	57.802	58.851	59.05	58.358	57.061	56.245	54.446	52.55	50.651
kappa	0.455	0.468	0.47	0.461	0.445	0.435	0.413	0.391	0.368
RMSE	0.30386	0.28852	0.27264	0.25812	0.24661	0.24211	0.23544	0.2314	0.22939

2.8	3	3.2	3.4	3.6	3.8	4	4.4	5	6
50.651	47.162	45.59	44.146	42.836	41.646	40.562	38.704	36.49	33.868
0.345	0.3292	0.311	0.295	0.281	0.268	0.257	0.237	0.199	0.178
0.22939	0.22932	0.23044	0.23197	0.23374	0.23561	0.23751	0.24122	0.24622	0.25286

Mahalanobis distance

m	1.1	1.3	1.5	1.7	1.9	2	2.2	2.4	2.6	2.8
OA	73.119	73.773	72.918	70.388	66.874	64.997	61.212	57.706	54.559	51.805
kappa	0.627	0.638	0.63	0.6	0.558	0.536	0.49	0.449	0.413	0.381
RMSE	0.22632	0.2075	0.19473	0.18811	0.18649	0.18709	0.19018	0.19476	0.19995	0.2052

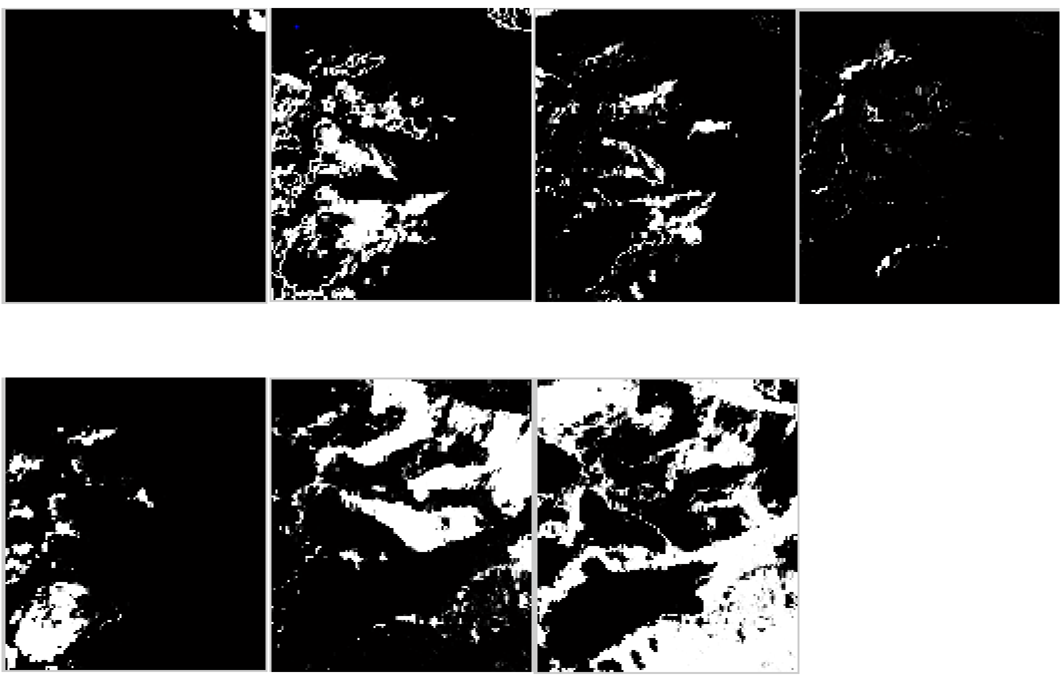
3	3.2	3.4	3.6	3.8	4	4.4	5	5.5	6
49.409	47.349	45.553	43.975	42.588	41.361	39.308	36.954	35.457	34.255
0.354	0.331	0.311	0.294	0.279	0.266	0.244	0.219	0.203	0.191
0.21024	0.21493	0.21932	0.22313	0.22666	0.22986	0.23537	0.24193	0.24624	0.24978

Table 5: Accuracy assessment parameters corresponding to varying ‘m’ values for Euclidean, Diagonal and Mahalanobis distances and training data size of 21 percent. Numerals in red represent maximum OA & kappa and min RMSE

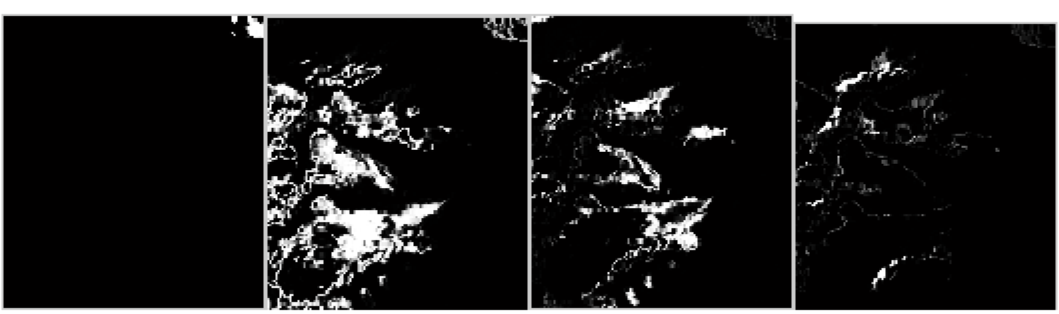
It is clear from above table that for each A-norm, OA/kappa values decrease with increase in ‘m’ value. As the value of ‘m’ increases, fuzziness among the classes also increases, thus showing lower accuracy values at higher ‘m’.

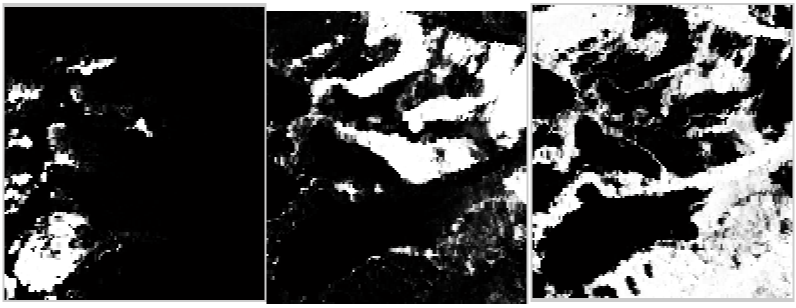
Fraction images obtained corresponding to varying ‘m’ values for mahalononbisdistance are shown below in fig 7.

‘m’=1.1

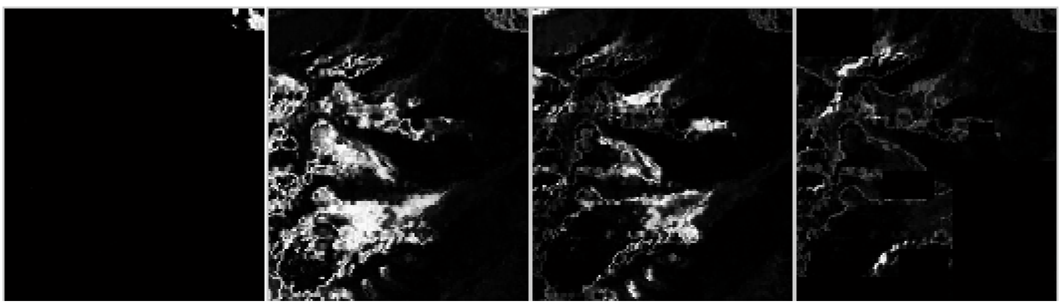


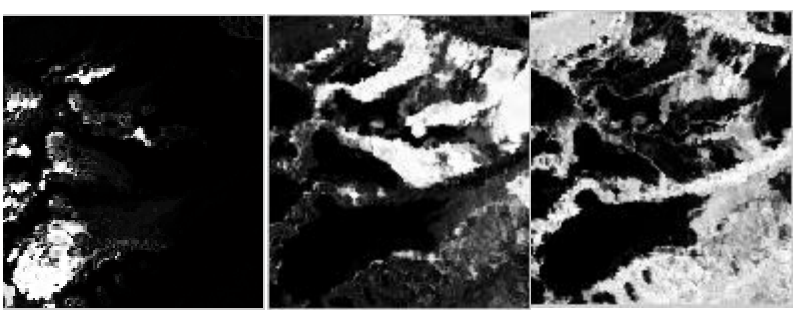
‘m’=1.3



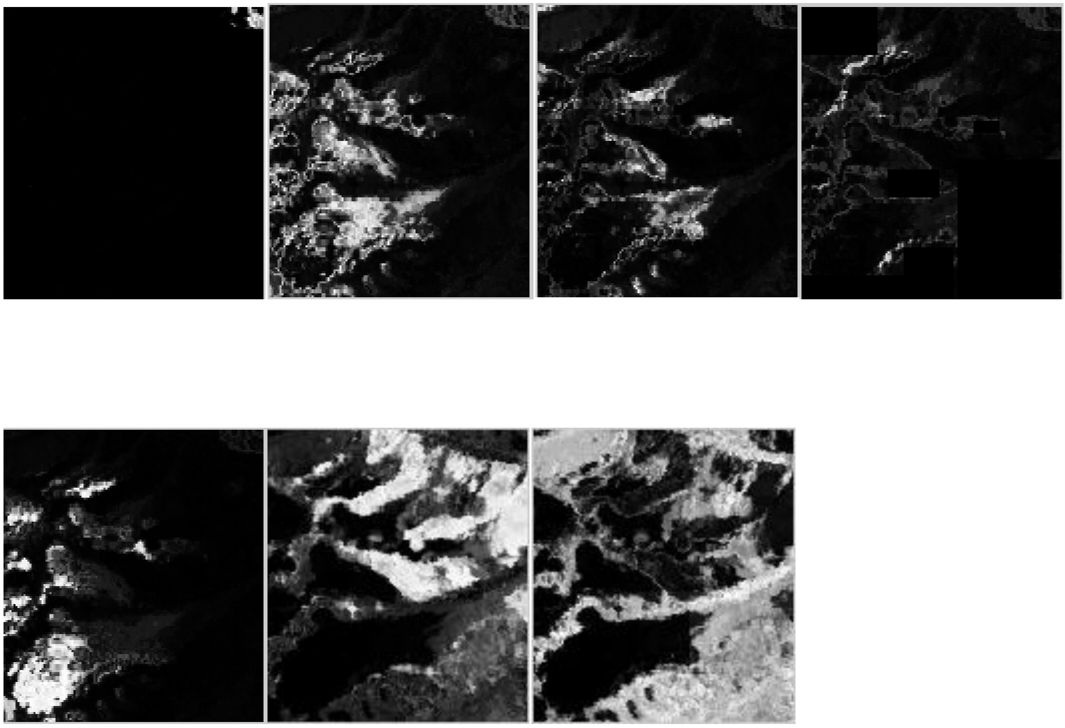


'm'=1.7

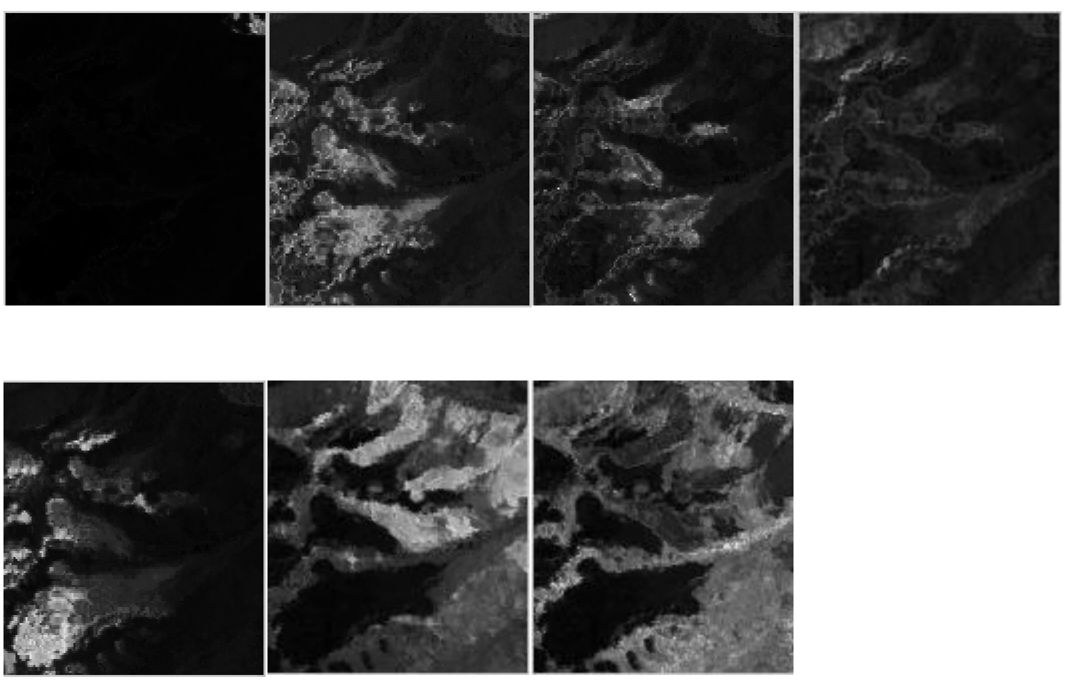




'm'=2.0



'm'=3.0



‘m’=4.0

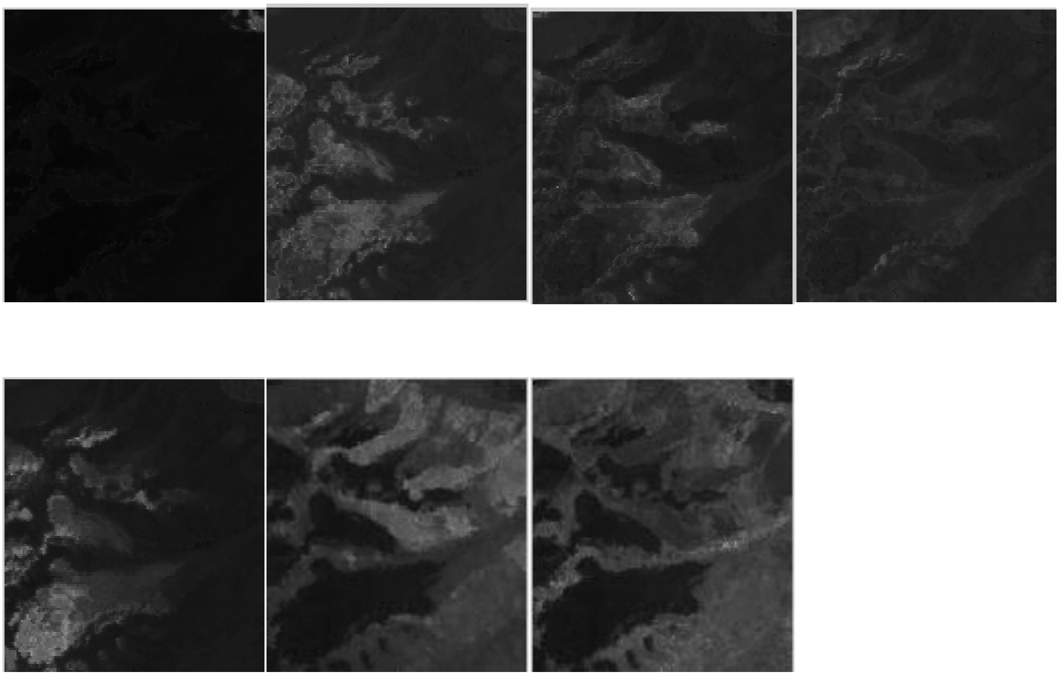
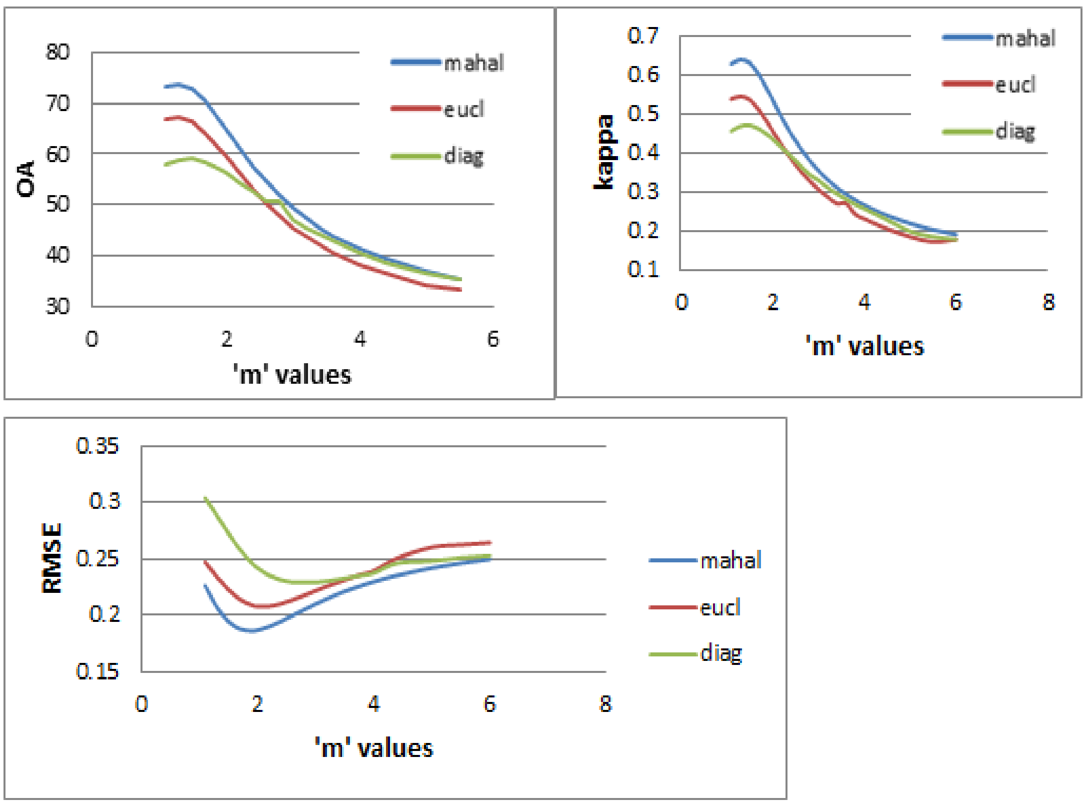


Fig. 6 Fraction images of lake, ice, imd, shadow, snow, rock-valley and debris respectively with varying ‘m’ values for mahalononbis distance

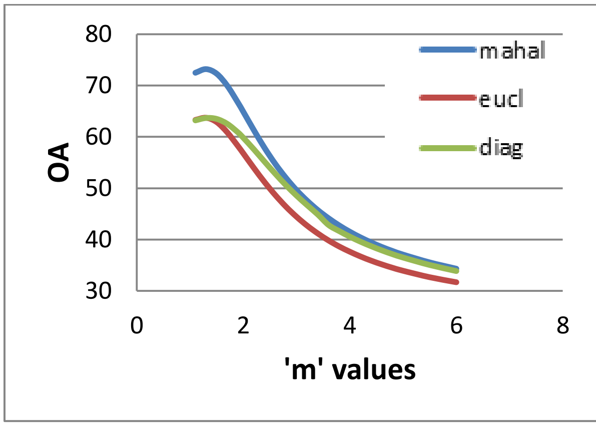
From above shown fraction images, it is clear that with the increase in ‘m’ value, fraction images are getting darker. Bright white pixels hardly exist in fraction images for ‘m’ value greater than 3. This can be explained by the fact that as ‘m’ value increases, fuzziness among the classes also increases resulting in almost similar membership values for all the classes within a pixel, thus giving darker and greyish fraction images. Bright white pixels representing higher membership values are present only in fraction images with lower ‘m’ values.

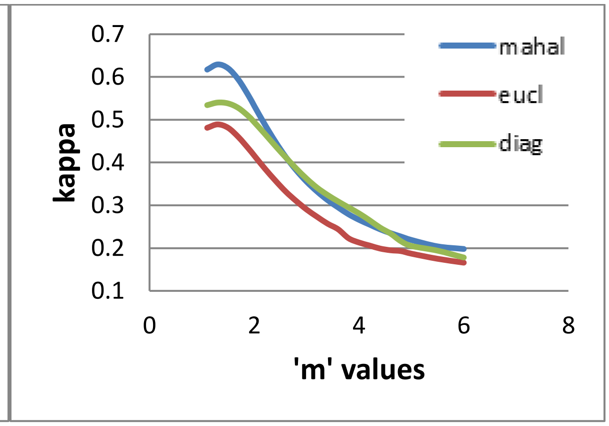
For all the values of training data size (3 percent- 21 percent), almost similar trends have been observed with mahalononbis distance showing maximum accuracy values followed by diagonal distance and then Euclidean distance showing least accuracy. Variation of OA, kappa and RMSE for each A-norm with varying values of training data size (3 percent- 21 percent) has been shown in following fig. 7

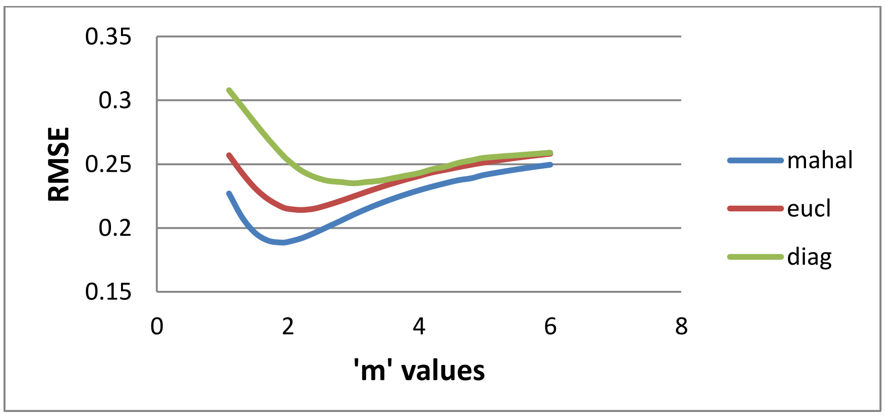
Training data=21 %



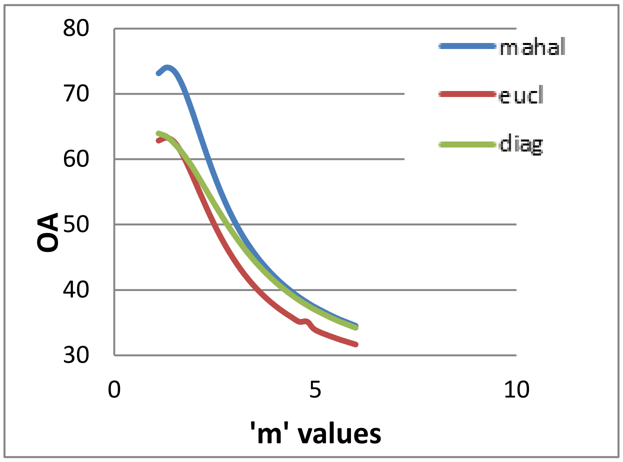
Training data=18%

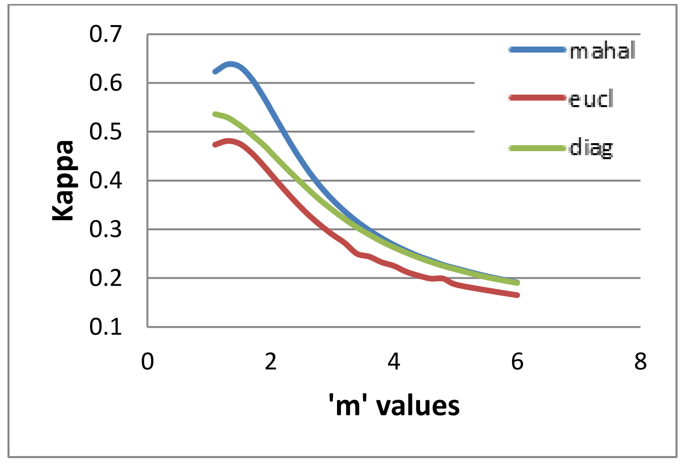


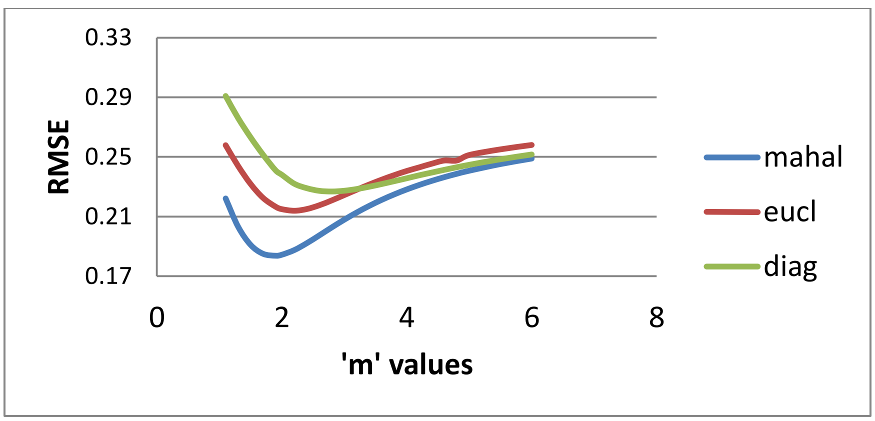




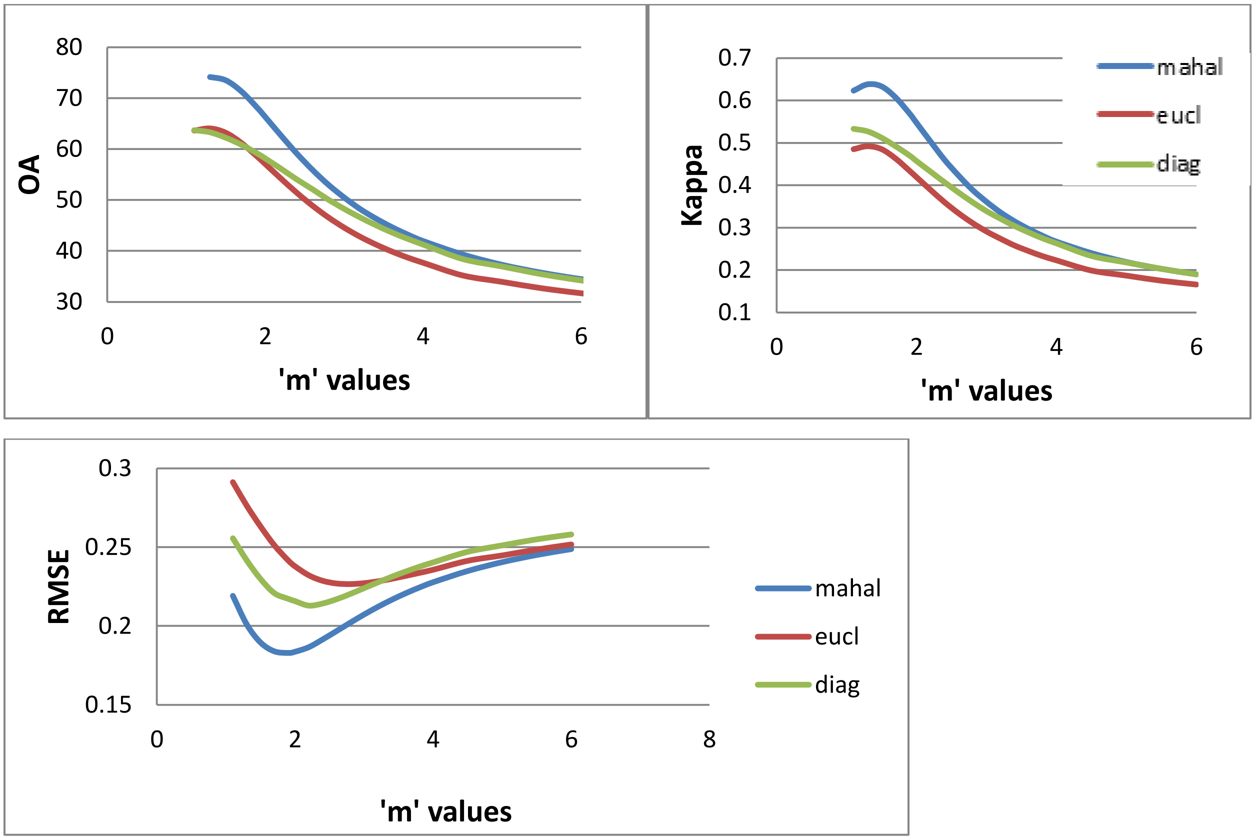
Training data =15%



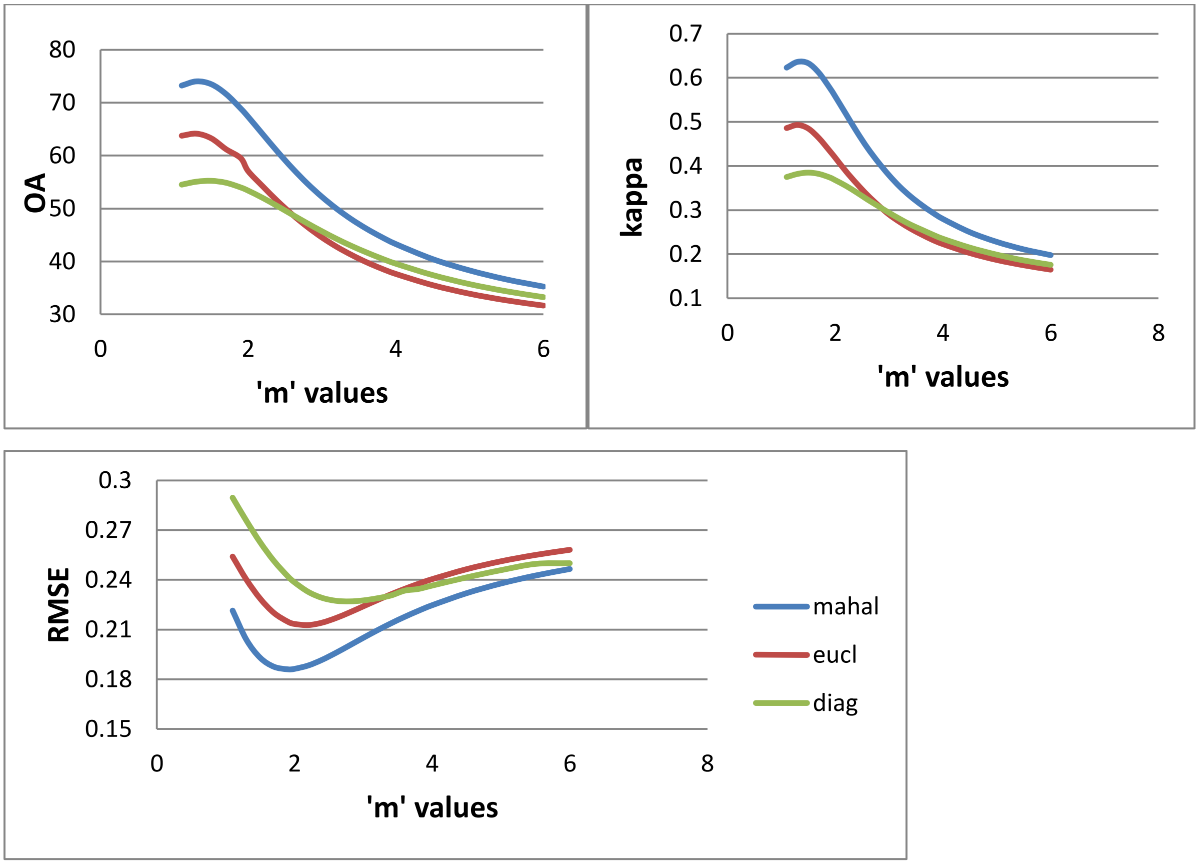




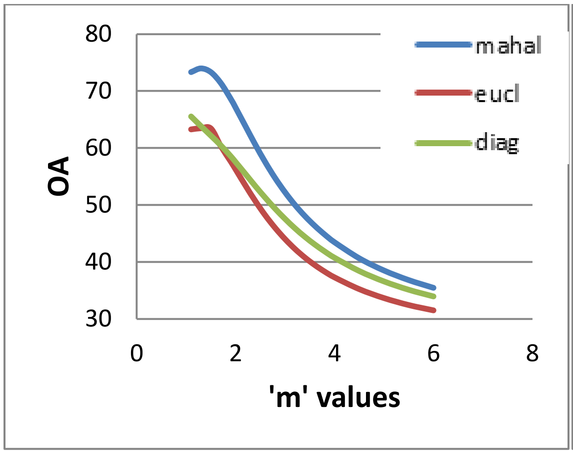
Training data=12%

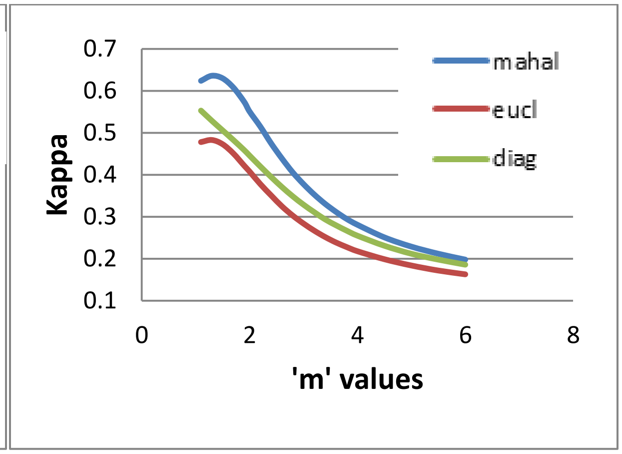


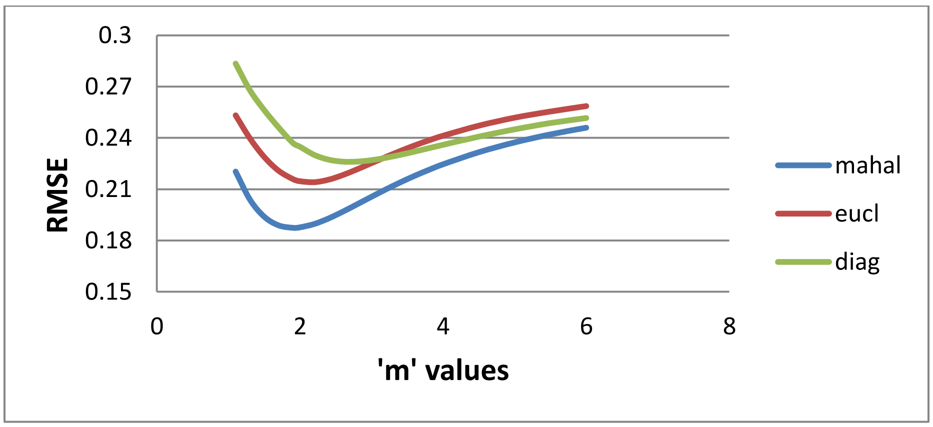
Training data= 9%



Training data = 6%







Training data=3%

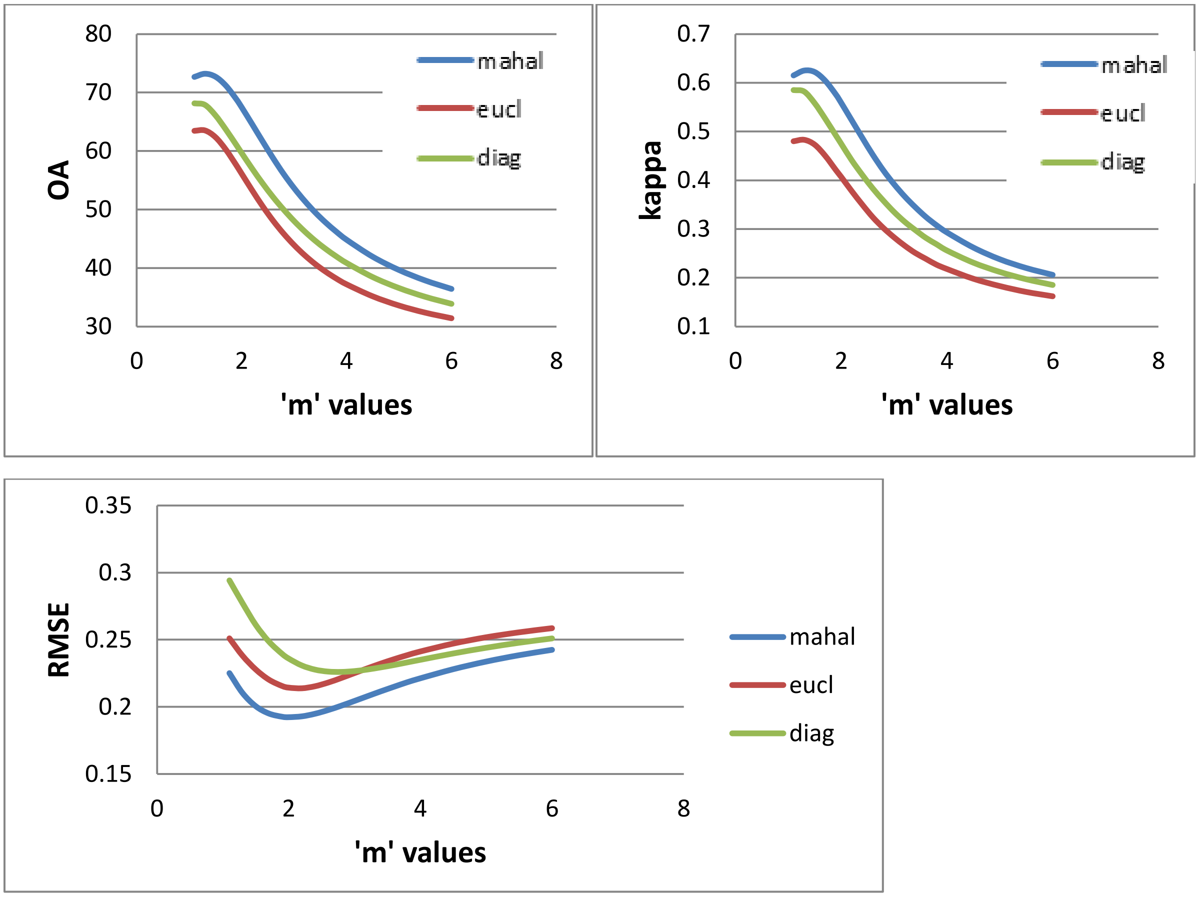


Fig. 7 Variation of OA, kappa & RMSE corresponding to A-norms with ‘m’ for varying training data sizes.

Form above graphs, it has been concluded that kappa coefficient and OA varies in similar manner for each norm. Both these accuracy parameters show a slight increase first up to ‘m’ value of 1.4 then decreases continuously with increasing ‘m’ value. RMSE values on the other hand decrease first and then show increasing trend. Maximum OA value and kappa coefficient is obtained for mahalononbis A-norm with training data size of 21 percentand ‘m’ value of 1.3.Maximum value for OA and kappa obtained are 74.032 and 0.682 respectively. Minimum RMSE value of 0.18287 has been observed for training data size of 12 percent with mahalononbis A-norm and ‘m’ value of 1.9.

For each training data size, Mahalanobis norm has outperformed the other two. It can be explained by the fact that Mahalanobisnormenables the FCM algorithm to form hyper-ellipsoidal clusters with any possible random orientation thus taking into account most of the diversities present in a class. Euclidean distance on the other hand being the simplest and most basic norm forms hyperspherical clusters only. It considers minimum variations in a class, thus showing least accuracy values. Diagonal norm on the other hand can form hyperellipsoidal clusters same as Mahalanobis but their orientation is fixed; making it somewhat rigid than Mahalanobis. Thus, diagonal norm shows accuracy values in between Mahalanobis and Euclidean.

For representing the variation of FCM with training data size, maximum value of OA/ kappa and minimum value of RMSE obtained for each of A-norm with each training data size are considered.

Variation of maximum values of OA/ kappa and minimum RMSE with training data size has been shown below in fig. 8 on next page.

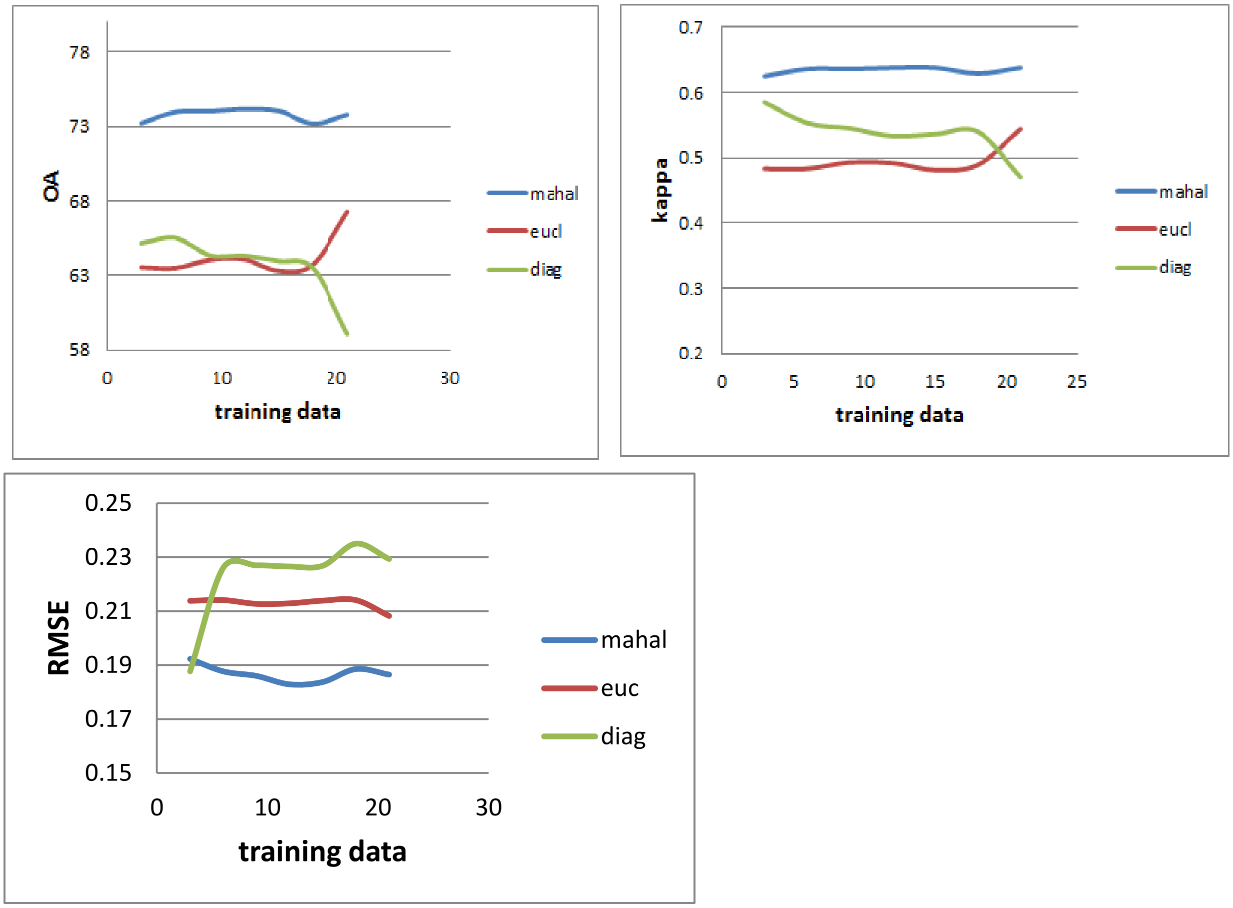


Fig. 8 Variation of maximum OA, maximum Kappa & minimum RMSE corresponding to three A-norms with training data.

From the above graphs, it has been observed that different A-norms show different trends with variation in training data size.

Mahalanobis and Euclidean norm doesn't vary much with training data size, though beyond 21 percent, Euclidean norm is showing a significant increase in OA/kappa.

While diagonal distance is showing completely different behavior than other two norms .It shows highest accuracy values at low training data size(around 6 %) and a further decrease in accuracy with increase in training data size . At 21 percent, accuracy reduces greatly for diagonal norm.

RMSE trends are showing similar results too with Mahalanobis and Euclidean norm showing very little variation with training data size and diagonal norm showing an increase in RMSE with increase in training data size.

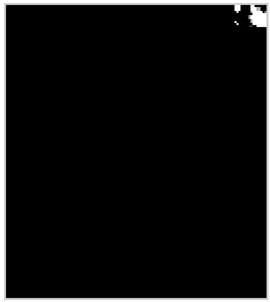
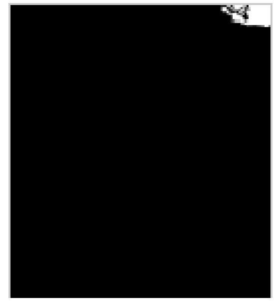
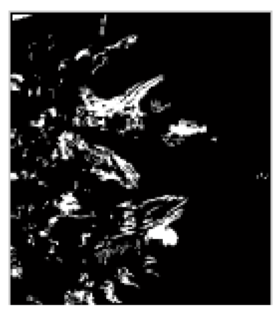
Taking all these parameters into account, maximum OA value is observed for following combination of parameters shown in table:

Parameters	Training data size (%)	A-norm	‘m’ value	Critical value (max. for OA& kappa; min for RMSE)
OA	15	Mahalanobis	1.3	74.032
Kappa	15	Mahalanobis	1.3	0.638
RMSE	12	Mahalanobis	1.9	0.18287

Table 6: Optimal values of parameters taken under consideration

Maximum values of accuracy assessment parameters (OA & kappa) have been observed forMahalanobis norm, ‘m’ value of 1.3 and training data size of 15 percent. On the other hand, minimum RMSE value has been observed for training data size of 12 percent, Mahalanonbis norm and ‘m’ value of 1.9.

Fraction images corresponding to this optimum combination of OA/kappa and reference data have been shown in fig. 9.

Class	Classified Images	Reference Images
Lake		
Ice		
IMD		
Shadow		

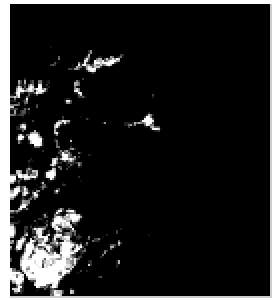
Class	Classified Images	Reference Images
Snow		
Valley		
Debris		

Fig. 9 Comparison of fraction images corresponding to maximum accuracy and reference images

Thematic land cover mapping cannot be done by soft classification techniques as they do not provide any information about spatial distribution of the classes. Hence, only area corresponding to each class can be determined .Taking into account the optimum combination of parameters, areal extent of classes has been calculated.

Classes		OA & kappa (training data size – 15 %, m- 1.3, A-norm – Mahalanonbis)	RMSE (training data size – 12 %, m- 1.9, A-norm – Mahalanonbis)	Reference fraction images
		Area (Km^2)	Area (Km^2)	Area(Km^2)
1.	Lake	4.51	3.903	9.912
2.	Ice	162.74	147.815	155.862
3.	IMD	71.30	81.088	86.772
4.	Shadow	36.28	41.253	18.333
5.	Snow	71.42	105.014	65.734
6.	Valley	261.51	298.170	272.594
7.	Debris	455.08	385.630	444.715

Table 8: Area covered by glacier facies

Two sets of facies area are obtained one based on maximum OA/kappa and other based on min RMSE. In order to determine which set is more accurate, RMSE between each of the set and reference area is calculated. Results show that area based on OA/kappa is more accurate (RMSE = 1.063) than that from RMSE (RMSE= 2.907).

CHAPTER 5

CONCLUSION

From the result and analysis in the study the following conclusion can be drawn as below

- Fuzzy c means is one of the widely accepted soft classification technique. However, its output is dependent on number of factors including training data size, fuzzy exponent 'm' and type of A-norms used.
- There doesn't exist any formulation or rule to get the optimum values of these parameters in order to have most efficient results. Different studies propose different values of 'm' or A-norm. Present study shows that with increase in 'm' value, there's a decrease in accuracy values. OA decreases to 20 percent when 'm' value increases to 7.Similarly, kappa coefficient also reduces with an increase in 'm' value. OA and kappa follow the same trend.
- RMSE first shows a decreasing trend up to average 'm' value around 2 and it increases with further increase in 'm' value. Out of three A-norms, mahalononbis norm outperformed the other two with Euclidean norm giving least accuracy and diagonal norm giving accuracy in between the other two. With training data size, each A-norm behaves differently.
- Mahalanobis norm shows very little variation with training data size, making it almost insensitive to this parameter. Euclidean norm also shows similar trends as that of Mahalanobis except that beyond 18 percent data size, it shows significant increase in accuracy.
- Diagonal distance however shows maximum accuracy at low data size and with further increase in size, its accuracy decreases considerably, especially beyond 18 percent. Once the optimum combination for parameters has been obtained mapping is done based on it.
- Area obtained based on OA/kappa is more accurate than that of RMSE thus making the former one more precise accuracy assessment parameter than RMSE.

CHAPTER 6

REFERENCES

1. Atkinson P.M " Mapping sub-pixel boundaries from remotely sensed images", Innovations in GIS IV, 166-180 (1996).

2. Bastin L "Comparison of Fuzzy C-Means classification ,linear mixture modeling and MLC probabilities as tools for unmixing coarse pixels", International Journal of Remote Sensing, 3629-3648 (1997).

3. Benedictsson J. A "Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data", IEEE Transactions on Geoscience and remote sensing, 540-551 (1990).

4. Bezdek J. E "FCM: The fuzzy c-means clustering algorithm", Computer and Geosciences, 191-203 (1984).

5. Binaghi E " A Fuzzy set based accuracy assessment of soft classification", Pattern Recognition Letters, 935-948, (1999).

6. BoucherA "Sub-pixel mapping of coarse satellite remote sensing images with stochastic simulations from training images", Mathematical Geosciences, 265-290, (2009).

7. CarpenterG " A neural network method for mixture estimation for vegetation mapping", Remote Sensing of Environment", 138-152 (1999).

8. Dunn J "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters", Journal of Cybernetics, 32-57 (1973).

9. DwediR, Kumar A and Ghosh S. K "Study of Fuzzy Based Classifier Parameter Using Fuzzy Matrix", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-3 (2012).

10. Fisher P "The pixel: a snare or a dilusion", International Journal of Remote Sensing, 679-685 (1997).

11. Foody G. M "Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data", ISPRS Journal of Photogrammetry and Remote Sensing, 2-12 (1995).

12. Foody G.M "Approaches for the production and evaluation of fuzzy land cover classifications from", International Journal of Remote Sensing, 1317-1340 (1996).

13. Foody G. M "Fully fuzzy supervised classification of suburban land cover using remote sensing Imagery", International Journal of Remote Sensing, 2731-2738 (1998).

14. Foody G. M and Arora M. K "Incorporating mixed pixels in training, allocating and testing stages of supervised classification", Pattern Recognition Letters, 1389-1398, (1996).

15. Foody G. M " fully fuzzy supervised classification of sub urban land cover from remotely sensed imagery :statistical and artificial neural network approach", International Journal of Remote Sensing, 615-628 (2001).

16. Jain S. K , Sarkar A, V garg "Impact of declining trend of flow in Harike wetland,India", water resource management", 409-421 (2008).

17. Krishnapuram R. A "A possibilistic approach to clustering", IEEE Transactions on Fuzzy Systems", 98-110 (1993).

18. Krishnapuram R. A "The possibilistic c-means algorithm: insights and recommendations" , IEEE Transactions on Fuzzy Systems, 385-393 (1996).

19. LiangS. S "Emerging issues in land remote sensing", Advances in Land Remote Sensing: Systems, Modelling, Inversion and Application,, 485-494 (2008).

20. Okeke F. K "Methods of fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analysis" International Journal of Remote sensing, 153-176 (2006).

21. Pontius R. G and Cheuk M. L "A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions", International Journal of Geographical Information science, 1-30 (2006).

22. Prasad M.S and Arora M. K "Assessing the quality of fuzzy land cover classification by similarity and certainty measures",Journal of Geomatics(2015).

23. Silván-Cárdenas J. L"Sub-pixel confusion–uncertainty matrix for assessing soft classifications", Remote Sensing of Environment, 1081-1095 (2008).

24. Lillesand T. M "Remote sensing and image interpretation. geological magazine", 248-249 (1994).

25. Tatem A. L "Super-resolution target identification from remotely sensed image using a Hopfield neural network", IEEE Transactions on Geoscience and Remote Sensing, 781-796 (2001).

26. Thornton M. A "A linearised pixel-swapping method for mapping rural linear land cover features from fine spatial resolution remotely sensed imagery", Computers and Geosciences, 1261-1272 (2007).

27. Woodcock C. E "The factor of scale in remote sensing", Remote Sensing of Environment, 311-332 (1987).