Exploring measures of similarity and dissimilarity for fuzzy classifier: from data quality to distance quality

Sayan Mukhopadhaya

March, 2016

IIRS Supervisor

Dr. Anil Kumar

INDIAN INSTITUTE OF REMOTE SENSING
Indian Space Research Organisation
Department of Space, Government of India

ITC Supervisor

Dr. Ir. Alfred Stein

UNIVERSITY OF TWENTE.
ITC
FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Exploring measures of similarity and dissimilarity for fuzzy classifier: from data quality to distance quality

SAYAN MUKHOPADHAYA
Enschede, the Netherlands, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

SUPERVISORS:

ITC Supervisor	:	Prof. Dr. Ir. A. Stein
IIRS Supervisor	:	Dr. Anil Kumar

THESIS ASSESSMENT BOARD:

Chairperson (ITC)	:	Dr. A. A. Voinov
External Examiner	:	Dr. S. K. Ghosh

OBSERVERS:

ITC Observers	:	Dr. N.A.S. Hamm
		Dr. V. A. Tolpekin
IIRS Observers	:	Dr. S. K. Srivastav
		Dr. Sameer Saran

INDIAN INSTITUTE OF REMOTE SENSING
Indian Space Research Organisation
Department of Space, Government of India

UNIVERSITY OF TWENTE.
ITC FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

Dedicated to my family and my teachers……..

ABSTRACT

Remote sensing images are predominantly affected by the presence of mixed pixels. Soft classifiers have the advantage to handle the mixed pixels due to the shortcomings of hard classifiers. The fuzzy based classifiers have shown to be robust and accurate when classifying land use and land cover maps. In the literature, the fuzzy *c*- means classifier has been studied with Euclidean, Mahalanobis and diagonal Mahalanobis norms. In this study, the fuzzy *c*- means classifier has been studied with nine other similarity and dissimilarity measures: Manhattan distance, chessboard distance, Bray-Curtis distance, Canberra, Cosine distance, correlation distance, mean absolute difference, median absolute difference and normalised squared Euclidean distance. Both single and composite modes were used with a varying weighted constant (m) at different α-cuts. Formosat-2 image and Landsat-8 image of 8m and 30m spatial resolution were used to implement the weighted norms respectively. Formosat-2 image of finer resolution was used as the reference image for the accuracy assessment of Landsat-8 image of coarser resolution. The results showed that the best single and composite norms were obtained by optimizing the weighted constant (m). This helps in controlling the degree of fuzziness at various α-cuts. The two best single norms obtained were combined to study the effect of composite norms on the datasets used. An image to image accuracy check was done to assess the accuracy of the classified images. Fuzzy Error Matrix (FERM) was used to measure the accuracy assessment outcomes for Landsat-8 dataset with respect to Formosat-2 dataset. Cosine norm was found to be the best single norm among all the norms with an overall accuracy of 75.24%, followed by the Euclidean norm. These two norms were combined to form the composite norm which showed an overall accuracy of 69.80%. The accuracy of the classification was also measured in the case of an untrained class (wheat), which resulted in a decrease in the overall accuracy in comparison to the trained case. To conclude FCM classifier with Cosine norm performed better than the conventional Euclidean norm. But, due to the incapability of FCM classifier to handle noise properly, the classification accuracy was around 75%.

Keywords: Fuzzy *c*-Means Classifier, Classification, Similarity and Dissimilarity measures, Distance, Fuzzy Error Matrix

ACKNOWLEDGEMENTS

First and foremost, I thank GOD for showering blessings all throughout my life. I thank my parents, family and my beloved brother *Supratik* for giving me the constant support, encouragement, motivation and most importantly for tolerating me till today.

I want to express my heartily thankfulness to all my teachers from IIRS and ITC for helping me with their vast knowledge and also for technically supporting me throughout this course.

Dr. Anil Kumar, my supervisor from IIRS, who is unique as a guide and also as a teacher. He has provided constant support and motivation towards completing this research work. I highly admire his thoughts towards life and work, technical knowledge, constant support, positive criticism and most importantly providing guidance and time when needed the most. Dear Sir, thank you for helping me throughout and it has been a privilege for me to work under your guidance, which I will cherish all throughout my life. I am grateful to you.

Prof. Dr. Alfred Stein, my supervisor from ITC is a person with clarity in thoughts and ideas. His in-depth knowledge and constant guidance were a boon for me. He has been a strong guiding force all throughout my research work. I have been honoured and feel blessed to have a supervisor like him. His promptness and clarity in emails helped me to shape my research work properly. Thank you sir, for providing new ideas and helping all throughout.

I also like to thank Dr. S. K. Srivastav, Group Head, Remote Sensing and Geoinformatics group, for solving all our problems and guiding with necessary advice when required. I also like to thank Dr. Sameer Saran, Head of Department, Geoinformatics Department, for being a person who can be approached anytime with all types of problems and also for providing necessary resources and advices to successfully accomplish my research work.

I would also like to thank Dr. Nicholas Hamm for his support all throughout this course and especially during the stay at Netherlands. I would also thank Dr. Valentyn Tolpekin for his polite and helpful nature, and also for providing necessary advice during our mid-term assessment.

At last but not the least, big thanks to all my friends and I take this opportunity to appreciate them individually, *Aparna, Solanki, Amit, Nitin, Anchit, Gokul .P, Gokul .G, Vinit, Hati, Nandi, Soham, Aliandro, Surojit, Arsh, Arnab, Nikki, Sreeja, Ram sir* and *Pruthvi* for their constant encouragement.

I would also like to thank all my ITC friends: Ipsit, Maral, Yalda, Grace, Melkamu, Maurice, Gunjan, Kavita, Rushikesh, JR and many others for their time and making our stay happy and joyful.

I would like to thank especially *Shenbaga Rajan* (my roomie), *Atreya (Basu et al.)* and *Rajasweta* for helping me all throughout and providing enormous encouragement throughout the research work phase.

I also express my deep gratitude towards all the staff members of the GID department and CMA, IIRS, especially *Mr. Pankaj Aggarwal* and *Ms. Sangeeta* for providing all kinds of support whenever required.

Sayan Mukhopadhaya

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLOGY USED IN THE REPORT

X_j	Vector pixel value
I	Identity matrix
A	Weight matrix
$\|\| \ \|\|_A^2$	Square of norm of A
d_{ij}	Square of distance between the sampled point and cluster centre
μ_{ij}	Class membership values of a pixel *i* belonging to *j*
Y	Subset of X
x_i	Pixel spectral response
$X = \{x_1, x_2, \ldots, x_n\}$	Set of n random sample points
c	Total no. of clusters
$J_m(U,V)$	Objective Function
m	Weighted constant ($1 < m < \infty$)
n	Total no. of pixels
U	n * *c* membership matrix
V	Mean vector for class *i*
v_i	Cluster centre
λ	Composite weighting component ($0 < \lambda < 1$)
$D(X_j, V_i)$	similarity and dissimilarity measures
$\mu(x)$	membership value of a sample point *x*

1. INTRODUCTION

1.1. Background

Remotely sensed image data are classified applying a classifier to generate user-defined labels (Mather and Tso, 2009). A Land Use/Land Cover (LULC) map is required for land use planning, preparing land cover maps, to check the health of the crops, etc. Thematic maps have a wide application among the end products of remote sensing. Spatial variations in phenomenon like geology, land surface elevation, soil type, vegetation, etc. are also displayed in a thematic map (Tyagi et al., 2015). In the digital domain, thematic maps are created by assigning labels to each pixel in an image and, this process is known as Digital Image Classification (Harikumar, 2014). Many factors affect the classification of remotely sensed image data into a thematic map such as the approach for image processing and classification, the quality and selection of remotely sensed data, the topography of the terrain, etc. These factors also affect the accuracy of the classification (Lo and Choi, 2004).

Many previous works also show that image classification algorithms have been developed, which show a significant confidence in extraction of information and generation of thematic maps (Gong et al.,1992; Kontoes et al.,1996; San Miguel-Ayanz et al.,1997; Foody, 1996; Stuckens et al., 2000; Franklin et al., 2002; Otukei et al., 2010; Landgrebe, 2003; Gallego, 2004; Richards and Jia, 2006; Tso and Mather, 2001). But, classifying a remote sensing image into a thematic map is a big challenge as there are many factors, which may be involved like: landscape complexity, specification of the data used, the algorithms used for image processing and classification, etc. and these factors may affect the success of classification (Foody et al., 1997; Stehman, 1997).

The term classification as defined by Chambers Twentieth Century Dictionary is the ‘act of forming into a class as per rank or order of persons or things’. The procedure to classify all pixels in an image into land cover classes is the main objective of an image classification technique (Lillesand et al., 1994). Classifications can be either one-to-one classification or one-to-many classification. One to one classification can be called as hard classification and a one to many classification can be called as soft classification technique (Mather and Tso, 2009). The probability that a pixel belongs to a class is equal to 0 or 1 in hard classification i.e. a pixel belongs to one particular class. In soft classification, a pixel can be assigned to more than one class with a value between 0 and 1 (Mather and Tso, 2009) (Figure 1.1). “Soft classifiers provide for each pixel a measure of the degree of similarity for every class” (Choodarathnakara et al., 2012).

However, heterogeneity of classes within a pixel may occur. This is commonly defined as a mixed pixel (Harikumar, 2014). The presence of mixed pixels is the cause of different problems in mapping and monitoring of land cover. The most severe effect of mixed pixels is in the mapping of diverse landscape using images of coarser resolution (Foody, 2002). The fuzzy set approach has been found quite suitable for solving the mixed pixel problem (Kumar et al., 2006a).

Fuzzy set theory introduced by Zadeh (1965) uses the concept of uncertainty in the definition of a set by removing the crisp boundary concept into a function of the degree of membership or non-membership (Binaghi et al., 1999) (Figure 1.1). Fuzzy logic using fuzzy set theory provides important tools for data mining and to determine the data quality and has been proven to have the ability to present uncertain data that contain vagueness, uncertainty and incompleteness (Stein, 2010). This is especially observed if the databases are complex. Classifiers based on fuzzy set theory like the Fuzzy *c*-Means classifier (FCM) (Bezdek et al., 1984) has been studied with weighted norms such as Euclidean norm, Mahalanobis norm and diagonal Mahalanobis norm for solving mixed pixel problems in remote sensing images (Kumar et al., 2006b). Earlier, other measures of similarity and dissimilarity measures such as the correlation, Canberra, Cosine distance, etc. have not been studied with FCM classifier. In this work, these measures were studied with FCM classifier. Common statistical analyses have been used in the past to calculate similarities for a fuzzy set like works done by Lopatka and Pedzisz (2000) and also by Besag et al. (1986). However, these analyses have been heuristic and are rather general. Therefore, it is important to consider the analysis of vague and ambiguous data with a degree of membership. Also, to determine the distance between the fuzzy sets, α-cuts have been used to get a better accurate distance between the fuzzy sets and also to avoid or check the overlap between the cluster centres (Dilo, 2006).

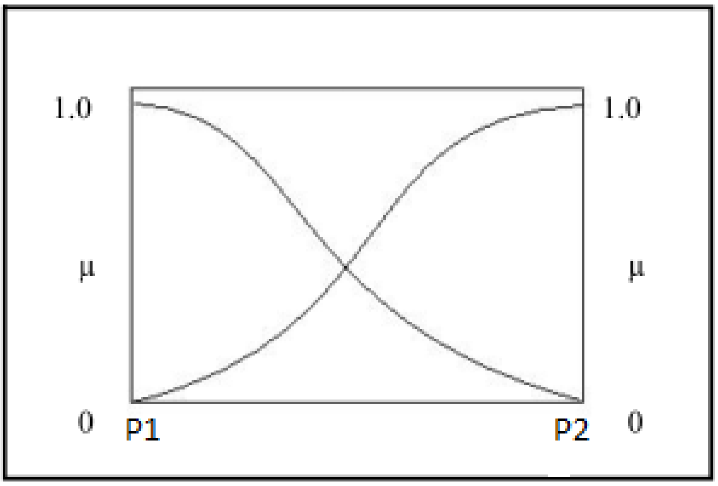


Figure 1.1 Fuzzy membership concept (Zadeh, 1965)

Similarity and dissimilarity are concepts that have been used before by researchers to build automated systems that assist humans in solving classification issues (Goshtasby, 2012). Measure for similarity and dissimilarity can be metric, non-metric, independent and dependent. Metric measures do not deal with all the topologies that are required for fuzzy classifiers (Dilo, 2006). Non-metric measures are quite effective for comparing images captured by different sensors (Pekalska et al., 2006). Independent measures are independent of the scale of the data or the rotational or translational of axes (Le Maitre, 1982). Dependent measures largely depend upon the class that has to be classified (Cheplygina et al., 2012). These measures are used, for example, to analyse the correspondence of images stored in a database to an observed image from a camera or sensor.

Measures of similarity can also be used to locate an object of interest (where the model of the object is given as a template) in an observed image, by finding the most appropriate place in the image where the template can fit. Measures of similarity can provide solutions when the templates and saved images and the observed image should neither have rotational nor scaling differences, and hence both the images match completely (Goshtasby, 2012). This shows the dependency between them. The dissimilarity measure between two datasets can be considered as a distance between them which quantifies their independency.

The works by Binaghi et al. (1999), Zhang and Foody (1998), Congalton (1991) and Martin et al. (1989) demonstrated that the accuracy of classified images can be evaluated by various ways. The conventional method of error matrix is not to be used as it assigns a pixel to a single class, which is hard classification. The Fuzzy Error Matrix (FERM) introduced by Binaghi et al. (1999) can be used to evaluate the accuracy of soft classified images. Though it is quite captivating, it is not regarded as a standard method to calculate the accuracy of soft classified images. In this work, the soft classified images were evaluated using an image to image accuracy by considering reference image of finer resolution than the classified image of coarser resolution.

1.2. Motivation and Problem Statement

Remotely sensed images of coarser resolution are used for diverse purposes. These images, when classified, give an erroneous result due to the presence of mixed pixels (Figure 1.2). Thus, soft classification methods are chosen over hard classification methods to handle the mixed pixels. Early works by researchers have studied fuzzy classifiers with Euclidean, Mahalanobis and diagonal Mahalanobis norms and are able to handle the mixed pixel. Fuzzy based classifiers with Euclidean norm cannot handle complex environment (Wan-zhi et al., 2013). In this work, different other measures of similarity and dissimilarity for fuzzy classifiers is explored not just for the data quality but also for the distance quality. This will assist researchers in decision making on which norm is most accurate with a higher distance quality. Previous works has showed that Fuzzy *c*-Means (FCM) classifier have been studied with three weighted norms (Euclidean, Mahalanobis and diagonal Mahalanobis) only. Throughout this work, it has been tried to incorporate the other various similarity and dissimilarity norms in FCM classifiers along with the α-cuts. A comparative study has been taken into account to find out the best possible single or composite measures for both similarity and dissimilarity norms on the virtue of their output data quality results, as all the distance norms have not been studied extensively and the conventional process of using Euclidean norm in FCM lacks the handling capacity of complex environment.

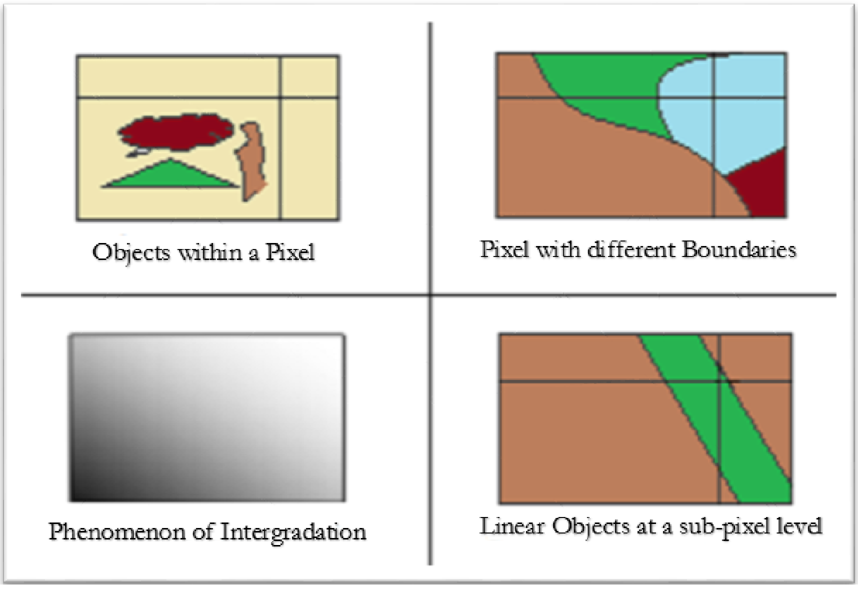


Figure 1.2 Causes of Mixed Pixels (Fisher, 1997)

1.3. Research Objectives

The main objective of this proposed research work was to study the behaviour of similarity and dissimilarity measures with a Fuzzy *c*-Means (FCM) approach.

The main objective was reached by defining the following sub-objectives:

- To develop an objective function for the fuzzy *c*-means classifier with similarity and dissimilarity measures.
- To optimize parameters of FCM classifier with similarity and dissimilarity measures
- To study FCM objective function with single or composite, similarity and dissimilarity measures using the α-cuts.
- To evaluate the performance of the proposed FCM classifier in the case of untrained classes.

1.4. Research Questions

The following are the research questions identified from the research objectives for the proposed work:

- How can similarity and dissimilarity measures be incorporated into the FCM classifier approach?
- How single or composite, similarity and dissimilarity measures work with different α-cuts along with FCM objective function?
- What will be the effect of using composite distance norms on FCM as compared to single distance norm?

1.5. Innovation Aimed At

The innovations intended in this study are:

- To study similarity and dissimilarity measures as single or composite distance norm with FCM classifier.
- To find out the best distance norm to solve the mixed pixel problem in an image.
- To make an optimal combination of two distance norms with FCM approach.

1.6. Research Approach

To answer the research questions and research objectives of this work, an objective function for the Fuzzy *c*- Means (FCM) classifier has been developed to handle the mixed pixel problem along with similarity and dissimilarity measures. The images of Formosat-2 and Landsat-8 satellites were geometrically corrected and geo-registered, and simulated images containing classes same as the remotely sensed images has been used. Supervised classification approach has been applied while incorporating various distance norms for similarity and dissimilarity measures using FCM classifier. Norms considered were Manhattan, chessboard, Bray-Curtis, Canberra, Euclidean, Mahalanobis, diagonal Mahalanobis, median-absolute-difference, mean-absolute-difference and normalized-squared-Euclidean for dissimilarity measures. Cosine and correlation norms were used for similarity measures. A certain combination of norms has been used to form a composite measure for evaluating its performances with respect to the single best norm. The classification has been conducted by using FCM objective function by incorporating the aforesaid norms at different α-cuts. The accuracy assessment has been done for both single and composite distance norms.

1.7. Thesis structure

The whole thesis has been organised into a total of six chapters. ***Chapter one*** includes the background information of the research work along with the important facets of the topic, the motivation and problem statement, research questions and the approach taken for the research. ***Chapter two*** describes the details of the related work that has been done in the past by various researchers. ***Chapter three*** includes the information of the study area chosen and the materials used along with the details of the methodology adopted. ***Chapter four*** describes the details of the classification techniques. ***Chapter five*** shows the results obtained along with the discussion of the results. Finally, the conclusion of the research work with recommendations leading to future research has been mentioned in ***chapter six***.

2. LITERATURE REVIEW

This chapter has different sections giving an introduction (section 2.1) to the previous research works on land cover classification method (section 2.2); Fuzzy *c*- Means (FCM) Classification on remote sensing images (section 2.3); different similarity and dissimilarity measures (section 2.4) and also about the usages of α- cuts (section 2.5).

2.1. Introduction

In this chapter, an overall view has been given of the various works done by researchers on the extraction of land cover followed by different norms which have been used in soft classification techniques on the basis of similarity or dissimilarity criterion.

Boyd et al.(2006); Foody et al. (2006) and Li et al. (2011) showed that there is a need to have information about all the classes in the training set exhaustively, to determine a specific class by using supervised classification. This, however, may result in a considerable error (Foody et al., 2006). Hence, supervised classification or hard classification is inappropriate for extracting a specific class (Foody et al., 2006). A problem like the occurrence of mixed pixels will be encountered as well by this conventional approach of classification (Upadhyay et al., 2013). Kumar et al.(2006b) showed that mixed pixels are found on the boundary of two or more classes in an image due to the pixel size compatibility with the class size.

The mixed pixel problem can be solved by the fuzzy set theory, by using a membership function along with α-cuts and quantifying the degree of belongingness of a pixel to a class (Dilo, 2006). Foody (2000) showed that Fuzzy *c*-Means classifier can be used to solve the mixed pixel problem. This has been recognized in the past as well: “Fuzzy set theory provides a useful technique to allow a pixel to be a member of more than one category or class with graded membership” (Shankar et al., 2006). Lee et al.(1996); Wang et al.(2005) and Upadhyay et al.(2014) showed that norms like Euclidean, Mahalanobis and diagonal Mahalanobis have been incorporated with FCM classifier. Tyagi et al. (2015) show that a fuzzy classifier along with similarity and dissimilarity measures can be used to solve the mixed pixel problem. Lee et al.,(2009) showed that if a similarity measure of a data-set has been found, it can also represent the dissimilarity, as a high level of similarity of data shows a low level of dissimilarity measure. The measure of similarity can be calculated based on the distance between the data used (Lee et al., 2009). There is a relationship between distance and similarity measures and the combination of similarity measure and distance measure shows the totality of information (Xuecheng, 1994).

2.2. Land Cover Classification

The main purpose of image classification is to classify every pixel either on the basis of one to one classification (hard classification) or one to many classification (soft classification) (Mather and Tso, 2009). There are many classification methods to classify a remotely sensed image into different land cover types. According to Swain and Davis (1979) these methods can be categorized into:

a. Methods based on whether a process of training is needed or not, i.e. supervised and unsupervised classification respectively.
b. Methods based on the usage and requirement of any parametric model (i.e. parametric and non-parametric).

There are many algorithms developed for classifying images. Amid the prevailing algorithms, the most widespread are the maximum likelihood classifier (MLC), support vector machine (SVM), decision tree classifiers and neural network classifiers. Maximum Likelihood Classifier (MLC) algorithm is a supervised statistical approach for thematic mapping using pixel based information. MLC follows Gaussian rule approach and it becomes unreliable when the class size is small (Gopinath, 1998), but works fine for a large class size though there is a high degree of computation. Despite its limitations, as it follows a normal distribution function for the signature of the classes (Swain and Davis, 1979), it is a common and widely used classification algorithm (Wang, 1990 and Hansen et al.,1996).

Neural network classifiers (NNC) avoid some of the problems that are faced in MLC by choosing a non-parametric approach. They do not follow a Gaussian rule approach. Neural networks have an advantage of high computation rate due to the presence of huge parallel networks, which resulted in the development of various other types of neural networks (Lippmann, 1987) such as: the most commonly used network in the classification of remote sensing images is the Multi-Layer Perceptron (MLP) (Paola and Schowengerdt, 1995; Atkinson and Tatnall, 1997). Artificial neural network (ANN) however may be very complex, as the learning rate can be high for the data of higher dimensionality. Large sets of training data are required for generalization as the data structure becomes complex on increasing the data dimensionality (Ablin and Sulochana, 2013).

Decision tree classifier (DTC) uses a different approach for land cover classification. Safavian and Landgrebe (1991) showed that a decision tree breaks a complex problem of classification into several stages of simple processes of decision making. There are univariate and multivariate decision trees, determined on the basis of the amount of variables used at each stage (Friedl and Brodley, 1997). At a global scale, land cover classification is done using univariate decision trees (De Fries et al., 1998; Hansen et al., 2000). Multivariate decision trees are generally more compact than univariate decision trees and are also sometimes more accurate than univariate decision trees (Brodley and Utgoff, 1995). The hierarchical method provides an advantage that it is easily interpreted than ANN, as the tree structure can be observed as a white box (Roosta et al., 2012). Another advantage is that it needs less complex training in comparison to ANN, but decision frames need to be framed for decision trees and they become complex when there is a large number of decision rules (Mather and Tso, 2009).

The Support Vector Machine (SVM) classifier is based on learning classification technique. It is used to allocate the labels as they were originally found in linear binary classifier (Mather and Tso, 2009). Construction of a separating hyperplane based on the properties of the training samples is the core operation of SVM. SVM has a large variety of applications. Osuna and Freund (1997) applied SVMs for human face detection along with digital image classification. Mukherjee et al. (1997) and Pal et al. (2005) used SVM for classifying remote sensing images. Huang et al. (2002) have showed that SVM gives higher accuracies than other classifiers like MLC, NNC and DTC. However, SVMs can be a time consuming process as shown by Patra and Bruzzone (2011).

Hard classifiers are poor in accounting information within mixed pixels and an analyzer has to adopt different methods like soft classifiers to handle mixed pixels. Soft classifiers result in different proportions of belongingness of classes within a single pixel. Presently, there are various classifiers like fuzzy classifiers, artificial neural network (ANN), etc. which can be used as soft classifiers. Fuzzy set theory classification takes heterogeneity and imprecise nature of the real world into account. It can also be used as supervised classification. The next sections provide a literature review on Fuzzy *c*- Means classification and the various distance measures that have been studied in this study.

2.3. Fuzzy *c*- Means Classification

Fuzzy *c*- Means (FCM) is a popular fuzzy clustering method that has been used for various applications for solving problems in the domain of remote sensing data. FCM is used with either supervised or unsupervised modes. Bezdek et al.(1984) showed that distance norms can be incorporated into FCM for clustering purpose with an unsupervised mode.

Various other works show that FCM can be used to classify remotely sensed data. The work by Zhu (1997) shows how fuzzy logic can be used along with similarity algorithms to find out the uncertainty in a remotely sensed image. Thus, provides the areas where accuracy is high. Other works also show that fuzzy logic and fuzzy set theory can be used to classify remotely sensed images (Ji, 2003; Shalan et al., 2003). The aforesaid works showed how mixed pixels are handled at the allocation stage for class identification within a pixel. This is represented in the form of membership value of a class related to the class composition of the pixel. FCM approach with a supervised mode can also be used to classify remote sensing images (Wang, 1990).

Foody (1996) and Bastin (1997) had evaluated the execution of Fuzzy *c*- Means (FCM) classifier and concluded that FCM provides a better approximation of sub-pixel land cover classes and thus can easily map the real world scenario.

Zhang and Foody (1998) applied FCM classification algorithm for classifying and mapping real life scenario. It was inferred that the obtained outputs were advantageously accurate while applying fuzzy classification and evaluation methods over conventional hard classification or partially fuzzy methods. Ibrahim et al. (2005) concluded that to produce accurate and proper land cover classification the concept of mixed pixels (which shows variability in the allocation of class) should be incorporated at all stages of the classifying process of remotely sensed images. Dwivedi et al. (2012) carried out a comparison of FCM (Fuzzy *c*-Means) and PCM (Possibilistic *c*-Means) and conducted an accuracy assessment by using FERM, SCM and Fuzzy Kappa Coefficient; norms considered were namely Euclidean, Mahalanobis and diagonal Mahalanobis only.

2.4. Measures Of Similarity and Dissimilarity

Zwick et al. (1987) studied and compared nineteen measures of similarity and dissimilarity with the different fuzzy sets. These measures were both geometric and set-theoretic, and were compared on the basis of their behavioral performances. It was concluded that distance measures could be evaluated on one's interest and the best distance measure should be chosen on the basis of high correlations for the particular situation. Deer et al. (1996); Takahashi et al. (2011) and Charulatha et al. (2013) had done a comparative study on FCM classifier with various distance metrics like Mahalanobis, Euclidean, Manhattan, Canberra, Tchebychev and Cosine. The results showed that the different distance metrics work differently with the variation of weighting exponent "*m*" and it was concluded that there is a need of exhaustive exploration of the distance metrics for different kind of datasets on various clustering algorithms.

Das (2013) analyzed how pattern recognition technique can be used with Fuzzy *c*-Means (FCM) classifier. In this work, the data analyzed was in the form of numerical vectors with predefined clusters. Besides, Euclidean other distances like Canberra and Hamming were also used in FCM classifier to get the variation in the outputs of membership values of the objects in the different clusters. The results showed that Euclidean produced the fastest and the most expected outputs whereas the outputs with Canberra were slowest and the least expected. Kouser et al. (2013) had applied K-means clustering algorithm with distances measures like Euclidean, Manhattan and Chebyshev. The experimental results showed that the overall accuracy of Chebyshev distance and Euclidean distance are comparable, whereas Chebyshev distance had the highest number of iterations.

Dik et al. (2014) showed how fuzzy clustering results improve when a weighting factor is introduced in the inter-object distances. The distances considered were Euclidean, Manhattan, Spearman and Chebyshev incorporated with FCM and were tested on three datasets. The results showed that there was a significant improvement in the accuracy when weighted distances were considered over unweighted distances. Sinwar et al. (2014) studied two distance metrics, Euclidean and Manhattan, incorporated with simple K-Means clustering algorithm on two real and one synthetic dataset. The results of the experiments performed showed that Euclidean approach has better outcomes than Manhattan approach on the basis of number of iterations for calculating the centroid of the datasets used during the overall clustering process.

2.5. Fuzzy α- Cuts

Reznik et al. (1994) demonstrated the method of α-cut border mapping. This method was implemented along with a proportional–integral–derivative controller (PID controller). The results showed that the method of α -cut border mapping is quicker than defuzzification of fuzzy output set. Thus, it was as good as, or comparable to real-time control applications. Kainz (2007) and Ponce-Cruz et al. (2010) explained the concept of α-cut vividly and described a fuzzy set being composed of crisp sets by using the concept of α -cuts. It was also explained that α -cut concept can be used to know all the elements which belong to a fuzzy set and also possess some degree of membership. Xexéo (1997) explained that the concept of α-cut is important as it could be used to deduce fuzzy functions from crisp sets. He also described the difference between the concepts of α-cut and threshold level. Dunyak et al. (1997); Abebe et al. (2000); Wong et al. (2001) and Yang et al., (2009) studied the concept of α-cut with classifiers based on fuzzy set theory and explained the usage of α-cut while analyzing the uncertainty in the model parameters by showcasing the advantages and drawbacks.

Kreinovich (2013) extended his ideas to fuzzy mathematics and fuzzy data processing from fuzzy logic and made some important proofs for α-cuts, such as:

- The membership function and α-cut representations are not same from the algorithmic point of view.
- Prevailing of a *c*-membership function for which computation of α-cuts are not possible and vice-versa is also true.
- In general, computation of fuzzy data processing is not possible for membership functions, but exceptions are there for α-cuts.

Other authors have shown that α-cuts can be used for solving various problems like;

- Lee et al.(2015) showed the usage of α-cut as a filter in proxy caching mechanism for wireless services. This mechanism was demonstrated to monitor the traffic flow and thus guaranteeing exact and faster streaming of services while buffer caching. The results of the work showed that the given mechanism has better performance than other caching techniques like S-caching, I-caching and C-caching mechanisms.

3. STUDY AREA, MATERIAL USED AND METHODOLOGY

This chapter explains the study area with the reasons for choosing the study area and the materials used for completing the work along with the methodology. The explanation for using a simulated image and specifications of the sensors from which the datasets are acquired are also explained and described.

3.1. Study Area

The study area selected for this project work was Haridwar, Uttarakhand and is shown in Figure 3.1. The district shares its boundaries by Dehradun in the north, Pauri Garhwal in the east while, west and south are bounded by districts of Uttar Pradesh. The central latitude and longitude of the district are 29.956° N and 78.170° E respectively. The coverage of the area is 2.664 km x 2.192 km in the east to west and north to south direction respectively. The land is fertile with river Ganga flowing through the district and agriculture remains the mainstay of the district. Five classes are considered: Water, Riverine Sand, Wheat Crop, Forest, and Fallow Land.

The main reason for selecting the study area was the presence of diversity in terms of land use classes, such as vegetation type (wheat), riverine sand, forest, fallow land and water. Due to the diversity of land use and land cover classes, there is also the presence of mixed pixels at the boundaries of the classes and this will help to examine the capacity of FCM classifier with different similarity and dissimilarity measures for classification. Field ground truth data of study area was available as the field visit was conducted on 16th March, 2015. Datasets from the sensors FORMOSAT-2 and LANDSAT-8 were also available of the same time frame to check the image to image accuracy of the classifier

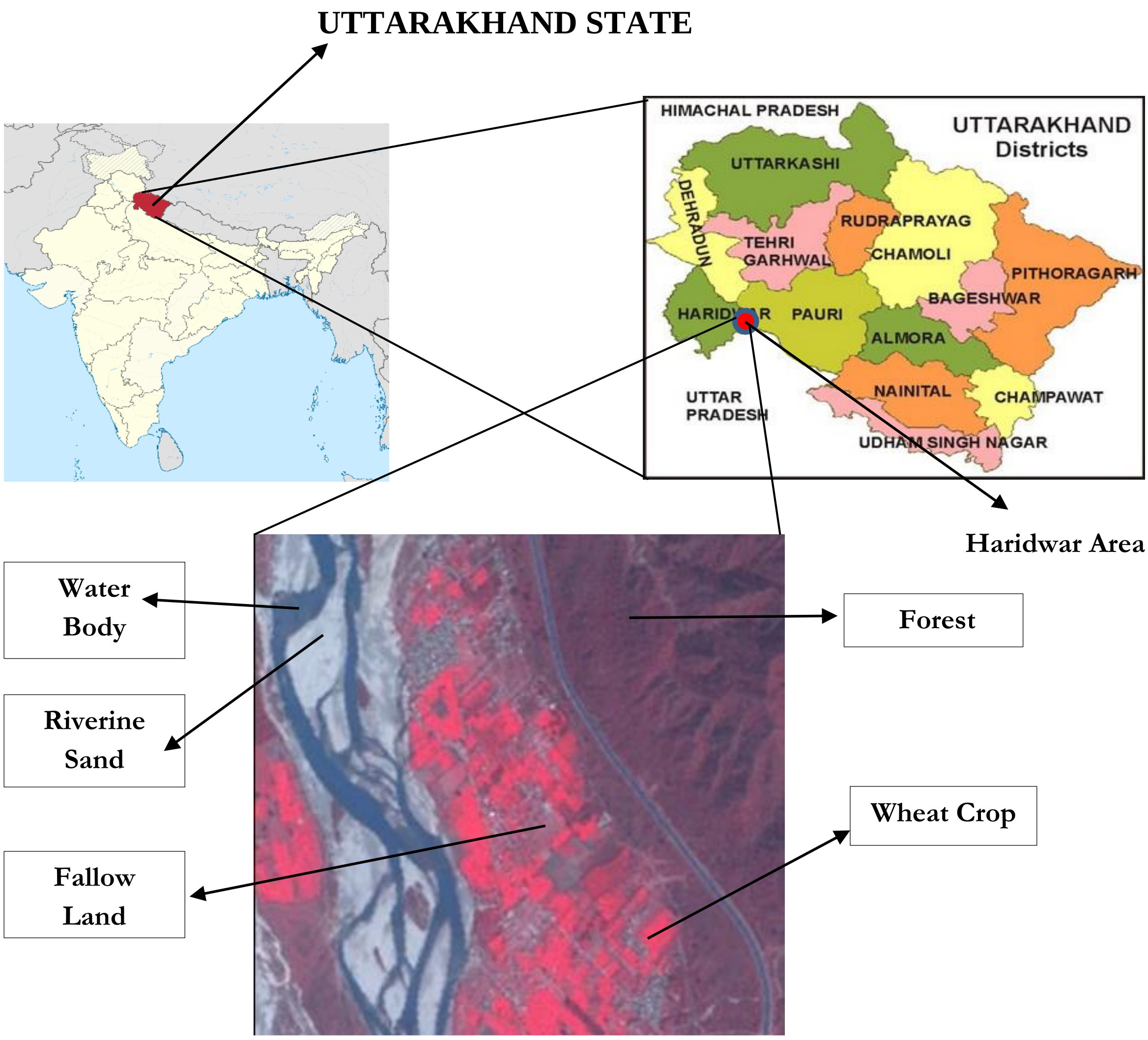


Figure 3.1: Image of the data is of Haridwar area, Uttarakhand, India

3.2. Materials used

In any research work the suitable use of remotely sensed data is necessary depending on the usability of the proper algorithms. These data may vary in spectral, spatial and temporal attributes. In this research work, multispectral images of 8m and 30m resolution of FORMOSAT-2 and LANDSAT-8 satellites were used. The formosat-2 satellite was developed by National Space Organisation (NSPO), Taiwan and was launched on May 21, 2004. The main aim of the FORMOSAT-2 mission has been to capture remotely sensed data on land and oceans of the earth with a daily revisit (Corporation, 2013). The landsat-8 satellite was developed and launched by National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS) on February 11, 2013. It is the eighth satellite in the satellite program of the Landsat. The main aim of the LANDSAT-8 mission is to provide optimum resolution images to segregate land use and land cover features to track down the usability of land and water (Corporation, 2015). The soft fractional outputs of finer resolution FORMOSAT-2 images were used to validate the soft fractional outputs of LANDSAT-8. Table 3.1 shows the specifications of the satellite data used:

Specification	FORMOSAT-2	LANDSAT-8
Spatial Resolution (m)	8m	30m
Spectral Resolution	• B1: 0.45 - 0.52 μm (Blue) • B2: 0.52 - 0.60 μm (Green) • B3: 0.63 - 0.69 μm (Red) • B4: 0.76 - 0.90 μm (Near-infrared)	• B1: 0.450 - 0.515 μm (Blue) • B2: 0.525 - 0.600 μm (Green) • B3: 0.630 - 0.680 μm (Red) • B4: 0.845 - 0.885 μm (Near-infrared)
Sensor Footprint	24 km x 24 km	185 km x 170 km
Return interval	Daily	After every 16 days

Table 3.1: FORMOSAT and LANDSAT satellite specification

3.2.1. The simulated image

In this research work, simulated images of multi-spectral data of Formosat-2 (4 bands) and Landsat-8 (7 bands) has been taken to study the performances of all the norms i.e. Euclidean, Mahalanobis, diagonal Mahalanobis, Cosine, correlation, Canberra, Manhattan, chessboard, Bray-Curtis, mean absolute difference, median absolute difference and normalized squared Euclidean with FCM classifier. Simulated FORMOSAT-2 and LANDSAT-8 images contain five classes: water body, wheat, forest, fallow land and riverine-sand. In these simulated images, we have intentionally mixed classes in a specific ratio and also have created an intra-class variation. Based on these controlled conditions the ability of handling the mixed pixel problem and detecting the intra-class pixel value variation were tested on the simulated image. Details of the simulated images are explained in figures A-1 and A-2 (Appendix A).

3.3. Methodology

The main objective of this work was to develop an objective function for the Fuzzy *c*-Means classifier with similarity and dissimilarity measures, by incorporating the concept of α-cuts. This section of the chapter describes the steps taken to accomplish the objectives of section 1.3.

The flow chart of the methodology adopted and developed has been presented in Figure 3.2.

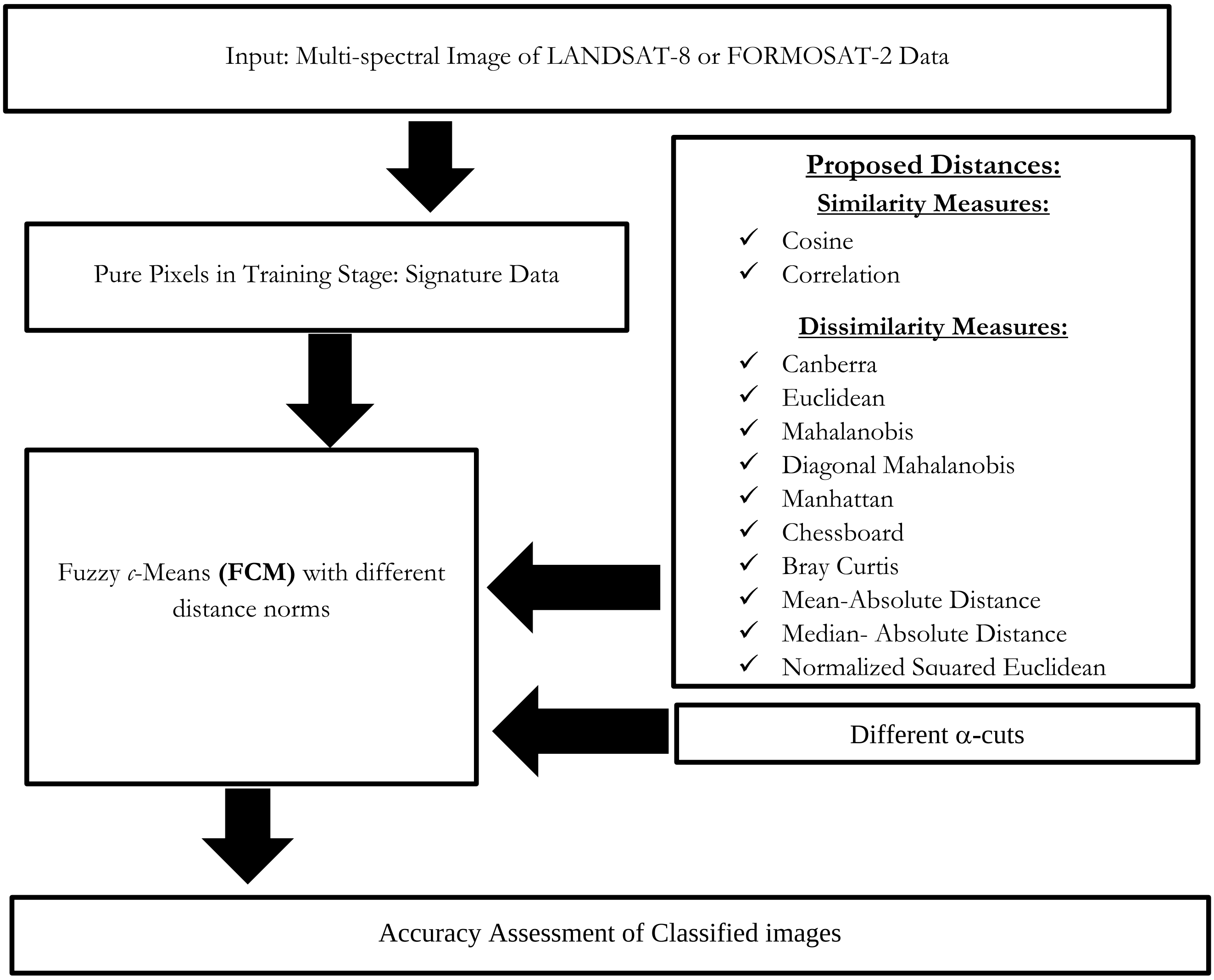


Figure 3.2. Research Methodology for this research work

3.4. Reference Dataset preparation

The outputs of FCM classifier are soft classified outputs. Hence, for the calculation of accuracy of the outputs, there is a need for soft reference data. The outputs of the classifier were soft outputs for each of the concerned class. In this research work, the classified soft outputs of Formosat-2 having finer resolution were used as the reference images for evaluating the image to image accuracy of the classified Landsat-8 images. The soft ground data were unable to be acquired due to the following reasons (Chawla, 2010):

- To locate a subpixel class on the ground is not possible.
- It is also not possible to accurately measure the stretch of a class at a sub-pixel level on the ground.
- Due to inaccessibility in some areas, the ground data was very difficult to collect in a soft mode.
- There may be presence of an error in the ground data, hence standard accuracy assessment can be termed as a degree of agreement but not the true value that is present on ground (Foody, 2002).

In this research work, outputs of the soft classification were of the type of fractional images for each considered class. The fractional images of Formosat-2 having finer resolution were used as the reference data (images) for assessing the accuracy of Landsat-8 fractional images. The images from Formosat-2 and Landsat-8 satellites were acquired of nearly same time frame. Hence, occurrence of errors due to temporal changes in the datasets were avoided. Kloditz et al. (1998) suggested a method using multi-resolution concept so that the estimation of accuracy after classification is possible for an image of low resolution by means of an image of finer resolution, where pixels of finer resolution for an area play a part to the pixels of low resolution of that same area during the assessment. It has also been observed that the pattern of the low-resolution image is conserved and there was also no damage to the inherent information of the image.

3.5. Sub-pixel classification algorithms

Supervised FCM classifier was used to generate the results for the sub-pixel classification. Three approaches namely fuzzy *c*-means (FCM), FCM with single measure and FCM with composite measures were applied.

3.5.1. Fuzzy *c*-Means (FCM)

There are many fuzzy based clustering algorithms. The outputs of all the sub-pixel classifications are in the form of fractional images for each concerned class. The optimization of the parameter is regarding the optimization of the weighted-constant (*m*) for each of the similarity and dissimilarity measures. This optimization is done on the simulated image by considering each norm with a fixed m-value and then checking the behaviour of the norm for classifying the following:

1. Pure pixel area (intra-class variation as well as membership value must be tending to one and hence the pixel DN-value should nearly 255 on an 8-bit scale)

if the 1st condition is satisfied, then the behaviour of the similarity measure was checked on;

a) Areas where there is a mixture of two classes, membership values must be tending to 0.5 for each class within a pixel (the DN-values should be nearly 127.5 for each class on an 8-bit scale)

b) Areas where there is a mixture of three classes, membership values must be tending to 0.3, 0.3 and 0.4 for each class within a pixel (the DN-value should be 76.5, 76.5 and 102 respectively on an 8-bit scale)

The flowchart for optimization of the weight constant 'm' is shown in figure A-3 (Appendix A.). This optimization of the weighted-constant (*m*) was done for both single as well as composite norms.

3.5.2. FCM with similarity measures

Mainly two types of measures were considered: similarity measures and dissimilarity measures. In this research work, two similarity measures were used: Cosine norm and correlation norm and ten dissimilarity measures were used: Bray-Curtis norm, Canberra norm, chessboard norm, diagonal Mahalanobis norm, Euclidean norm, Mahalanobis norm, Manhattan norm, mean absolute difference norm, median absolute difference norm and normalized-squared-Euclidean norm. Following the implementation of the similarity and dissimilarity measures, the optimization of the weighted constant '*m*' was achieved for each measure. The best single measure was selected based on the minimum difference with the expected output using the simulated image for the optimized '*m*'-value.

3.5.3. FCM with composite similarity measures

The composite similarity and dissimilarity measures were obtained from the best possible single measures. In the composite measures, the weight factor λ varies in between 0.1 to 0.9 with an interval of 0.1. For the composite measures, the optimization of '*m*' and λ were also necessary and this was accomplished in the same manner as in figure A-3 (Appendix A.). The untrained case of outputs were also verified by not using the signature data of one class in the FCM classifier (Byju, 2015), here we have considered the wheat field as the untrained class.

The membership values produced in a pixel by a class is represented in the form of fractional images, which are the classified outputs of a soft classifier (Harikumar, 2014). The total number of fractional images produced is equal to the number of concerned classes. Selecting the training samples was very important for all the approaches as it helped to determine the quality of classification. Hence, the mean of the membership grade of all the samples collected was measured for each of the concerned class.

3.5.4. FCM with α-cuts

The concept of α-cut is to create a threshold for the membership value of a pixel in the concerned class. The outputs obtained from both the single or composite use of similarity and dissimilarity measures were checked by the α-cuts from 0.5 to 0.9 with an interval of 0.1. The value of α-cut was restricted from 0.5 to 0.9 because if the value of α is below 0.5, then there will be an overlap of degree of membership of a class for a pixel and if the value of α is 1, then it represents the centre of the cluster of the concerned class (Yang et al., 2009). The outputs obtained at different α-cuts for both single and composite measures were evaluated for their accuracy to obtain the best α-cut value.

3.6. Accuracy assessment

Accuracy assessment is one of the most important aspect for diagnosing the quality of the outputs after classification. Image to image accuracy assessment was performed by taking FORMOSAT-2 data as the reference dataset for LANDSAT-8 data. To generate kappa statistics and overall accuracy fuzzy error matrix (FERM) and sub-pixel confusion uncertainty matrix (SCM) were used.

4. MEASURES OF SIMILARITY WITH FUZZY CLASSIFIERS

This chapter emphases on the fuzzy classification algorithm which includes developed Fuzzy *c*-Means (FCM) algorithm incorporating a total of twelve similarity and dissimilarity measures (similarity measures – Cosine and correlation; dissimilarity measures – Canberra, Bray-Curtis, chessboard, Manhattan, mean absolute difference, median absolute difference, normalised squared-Euclidean, Euclidean, diagonal Mahalanobis and Mahalanobis) in a single mode or composite mode along with fuzzy α- cuts. These measures along with the soft classifier (FCM) generate fuzzy outputs as fractional images.

4.1. Fuzzy *c*-Means Clustering Algorithm

Clustering is a method of grouping of pixels which has spectral similarity in multispectral space (Richards and Jia, 2006). Clustering segregates the pixels into multiple clusters on the basis of the similar properties (Fig. 4.1). There are a few common clustering techniques used for remotely sensed data such as, the iterative optimization clustering algorithm (Ball and Hall, 1965), single pass clustering algorithm, hierarchical clustering technique and clustering technique based on histogram peak selection (Letts, 1978; Richards and Jia, 2006). Furthermore, clustering can be segregated based on "hard" and "soft" methods of clustering (Jafar and Sivakumar, 2013). In hard clustering, a pixel of an input data is allocated to a particular cluster but in soft clustering (fuzzy clustering) a pixel is allocated a fuzzy membership value, with respect to each cluster (class), which shows the degree of belongingness of a pixel for a specific class (Zadeh, 1965).

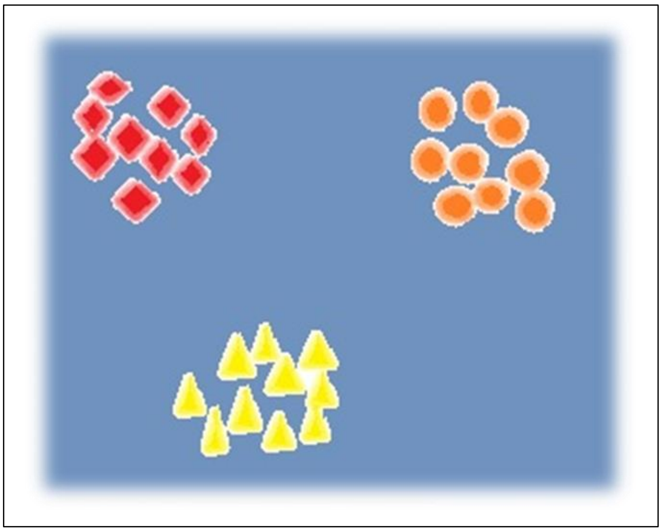


Fig. 4.1 Clustering

The Fuzzy *c*-Means (FCM) classifier (Bezdek, 1981) is a widely considered soft clustering technique (Jafar and Sivakumar, 2013). FCM provides membership value ranging from 0 to 1 to each pixel of the sample data for the different clusters (classes) (Bezdek et al., 1984).

In the concept of fuzzy membership, a pixel can be partially associated with many land cover classes. Thus, an idea of membership vector comes up with the value ranging from 0 to 1 for a sample of each class. Hence, a pixel can be associated with a class up to a certain level and may be associated with another class with another level and this level of association is shown by fuzzy membership values. In spectral space, the fuzzy membership value is the highest (closer to 1) for a point to a class, which lies next to the cluster centre of that class. In fuzzy membership values, there are no sharp partitions of the clusters for the spectral space. The main advantage of fuzzy membership value is that there is no loss of information, unlike hard partitioning technique, during determining the membership of a pixel (Wang, 1990). The concepts of hard partitioning technique and fuzzy membership value in spectral space is shown in figure 4.2.

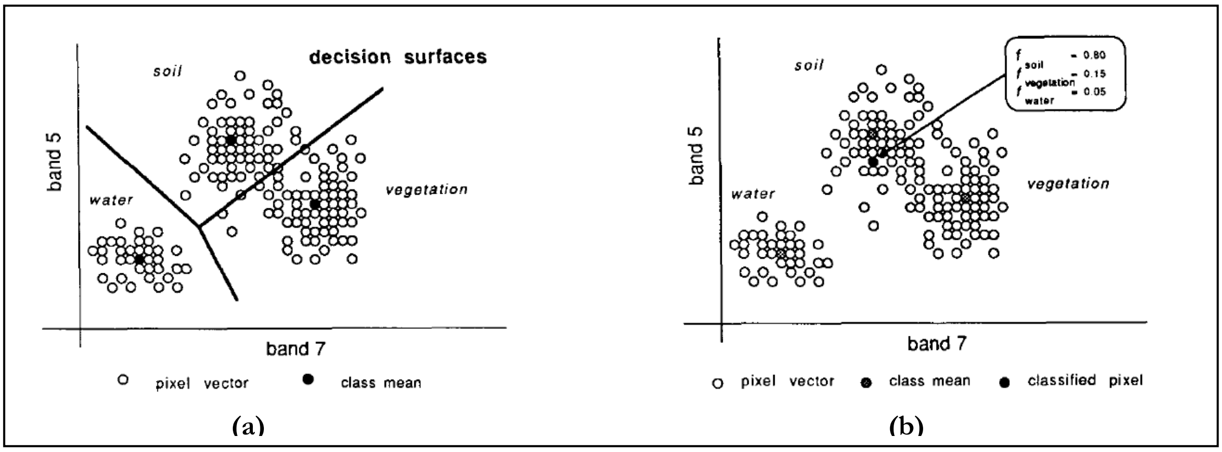


Figure 4.2 (a) Hard partitioning and (b) fuzzy membership partitioning of spectral space (Wang, 1990)

In hard partitioning technique (Figure 4.2. a) the spectral space is partitioned by crisp boundaries, thus the possibility of a pixel belonging to more than one class is omitted, whereas in fuzzy partitioning technique (Figure 4.2. b) membership values are assigned to a pixel which helps to depict the belongingness of a pixel to more than one class. Thus, fuzzy partitioning technique of spectral space can depict a real world situation better than hard partitioning technique and also helps to produce outputs close to ground information as there is no loss of information unlike hard partitioning (Wang, 1990). A fuzzy set is better described by a function of membership values that is associated with each sample data (pixel) ranging from 0 to 1. Let us consider a set of classes, represented by Y, in a spectral space X, then the fuzzy set is described as follows in equation 4.1 (Gehler and Scholkopf, 2009).

$$Y = \{f(x, \mu(x)) \mid x \in X\} \qquad 4.1$$

Here the membership value is represented by $\mu(x)$ and the sample pixels in the spectral space X is represented by x (Zadeh, 1965). Each pixel in the spectral space has a membership of value ranging from zero to one. The membership values close to unity represent a higher degree of similarity between the pixel and the concerned cluster (Bezdek et al., 1984).

Fuzzy clustering algorithm is considered as another possible way of clustering apart from an unsupervised classification of the data using *k*-means. Fuzzy clustering technique is a clustering type which allows one pixel to belong to more than one clusters with a certain membership value for each cluster present in the spectral space. FCM algorithm, which was proposed by Dunn (1974) and later generalized by Bezdek (1981), is one of the most commonly used fuzzy clustering technique. In the concept of supervised classification using FCM, each pixel belongs to some cluster or other clusters with a certain membership value respectively and the sum of the membership values has to be unity. In FCM algorithm the spectral space (dataset) $X = \{x_1, x_2 \ldots, x_n\}$ is partitioned into c number of fuzzy subsets. A fuzzy partitioning of the spectral space X into c-partitions may be represented by (c × n) form of matrix U, where all entries are in the form of μ_{ij} representing the membership value of a pixel for a class (Mather and Tso, 2009). But the U matrix is subject to some constraints stated in equations 4.2a and 4.2b (Mather and Tso, 2009):

$$\mu_{ij} \in [0, 1] \tag{4.2a}$$

and

$$\sum_{j=1}^{c} \mu_{ij} = 1 \text{ for all } i \tag{4.2b}$$

In FCM, the criterion for clustering can be attained by reducing the least-square error objective function (Mather and Tso, 2009) stated in equation 4.3 with certain constraints mentioned in equations 4.4a, 4.4b and 4.4c (Mather and Tso, 2009):

$$J_m(U,V) = \sum_{j=1}^{n}\sum_{i=1}^{c} \mu_{ij}{}^{m} D(X_j, V_i) \tag{4.3}$$

with certain constraints,

$$\sum_{j=1}^{c} \mu_{ij} = 1 \text{ for all } i \tag{4.4a}$$

$$\sum_{i=1}^{n} \mu_{ij} > 0 \text{ for all } j \tag{4.4b}$$

$$0 \leq \mu_{ij} \leq 1 \text{ for all } i, j \tag{4.4c}$$

where, n denotes the sum of the number of pixels present, c denotes the total number of classes, μ_{ij} the fuzzy membership value of the ith pixel for class j, m is the weighing exponent $1<m<\infty$, which determines the degree of fuzziness, X_j is the vector pixel value, V_i is the mean vector of a class and $D(X_j, V_i)$ is a similarity or dissimilarity measures as described in Eqn. (4.8) to Eqn. (4.20) and Eqn. (4.22). The matrix μ_{ij} of class membership is mentioned in equation 4.5 wherein d_{ik}^2 is calculated by equation 4.6 (Dwivedi et al., 2012):

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}}} \quad , i = 1, \ldots\ldots c, j = 1, \ldots\ldots n \tag{4.5}$$

where,

$$d_{ik}^2 = \sum_{j=1}^{c} d_{ij}^2 \tag{4.6}$$

Weighted constant (*m*): The degree of fuzziness is controlled by the value of *m* and it is also known as the fuzzifier. As the value of *m* is changed from near to unity (1) to infinity (∞), there is a corresponding change of FCM from a hard classifier to a complete fuzzy classifier. Cannon et al. (1986) has studied the effects of *m* on FCM and suggested that the value of weighted constant *m* should range in between 1.3 to 1.8. Zimmermann (2001) asserts in his book that the value of *m* should be 2, but there was lack of theoretical reasoning for selecting the value. Pal and Bezdek (1995) has suggested that the value of the weighted constant *m* should lie in the interval of 1.5 to 2.5 and the value of *m* equals to 2.0, which is the mean and midpoint of the interval, was a preferred choice.

4.2. Similarity and Dissimilarity Measures

Considering two sets of measurements $X = \{x_1, x_2, \ldots\ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots\ldots, y_n\}$, the similarity and dissimilarity between the two sets is a measure of quantifiable dependency or independency between the sets. Measurements of any two objects or phenomena can be represented by *X* and *Y*. A similarity measure S is to be considered as a metric if it shows increasing sequences of value of dependency corresponding to the values in the sequence. The following properties are satisfied by a metric similarity S for all orders of *X* and *Y* (Theodoridis and Koutroumbas, 2009; Goshtasby, 2012):

i) The range is limited: $S(X, Y) \leq S_0$, where S_0 is some arbitrarily large number.
ii) Symmetric: $S(X, Y) = S(Y, X)$
iii) Reflexivity: $S(X, Y) = S_0$, only when $X = Y$
iv) Triangle Inequality: $S(X, Y)\, S(Y, Z) \leq [Z(X, Y) + S(Y, Z)]\, S(X, Z)$.

Between the sequences *X* and *Y*, the largest possible similarity is S_0.

A dissimilarity measure D is to be considered as a metric if it shows increasing sequences of value of independency corresponding to the values in the sequence. The following properties are satisfied by a metric dissimilarity D for all orders of X and Y (Duda et al., 2001; Theodoridis and Koutroumbas, 2009; Goshtasby, 2012):

i) Non- negativity: $D(X, Y) \geq 0$.

ii) Symmetric: $D(X, Y) = D(Y, X)$

iii) Reflexivity: $D(X, Y) = 0$, only when $X = Y$

iv) Triangle Inequality: $D(X, Y) + D(Y, Z) \geq D(X, Z)$.

Besides having the desirable properties of a metric, a similarity measure can be effective though it may be not metric. Similarity measures have a value ranging from zero to unity, whereas dissimilarity measures have a value ranging from zero to infinity (∞), but this value can be normalized to a value ranging from zero to unity. The relationship between similarity (S) and normalized dissimilarity (D) can be shown by the equation 4.7:

$$S(X,Y) = 1 - D(X,Y) \tag{4.7}$$

In few of situations, a dissimilarity measure is converted into similarity measure so that it makes the computation easier for further procedures.

There are a lot of applications and usages of similarity or dissimilarity measures like it helps in distinguishing one object from another; the objects can be grouped on the basis of similarity and dissimilarity; a new object can be classified into a group based on the behavior as per the similarity or dissimilarity measures; thus further actions and decisions can be planned based on the prediction and structural information of the data. In this study, a total of twelve similarity and dissimilarity measures have been studied with Fuzzy *c*-Means (FCM) classifier in single or composite mode. The following sections describes the mathematical functions of similarity and dissimilarity measures.

Manhattan: The Manhattan metric estimates the distance based on the sum of the differences between the values of the concerned variables at any location. It is known as city block metric or taxicab metric. It is also used to compare images and is also one of the oldest dissimilarity measures. If, we define vector pixel value like $X_j = (X_{j1}, X_{j2}, X_{j3}, \ldots\ldots\ldots\ldots, X_{jb})$ and the mean values as $V_i = (V_{i1}, V_{i2}, V_{i3}, \ldots\ldots\ldots\ldots, V_{ib})$, then the Manhattan distance can be described like in equation 4.8 (Hasnat et al., 2013):

$$D(X_j, V_i) = Abs(X_{j1} - V_{i1}) + Abs(X_{j2} - V_{i2}) + \ldots\ldots\ldots + Abs(X_{jb} - V_{ib}) \tag{4.8}$$

where, b shows the total amount of bands in the image.

The following figure 4.3a depicts the Manhattan Distance.

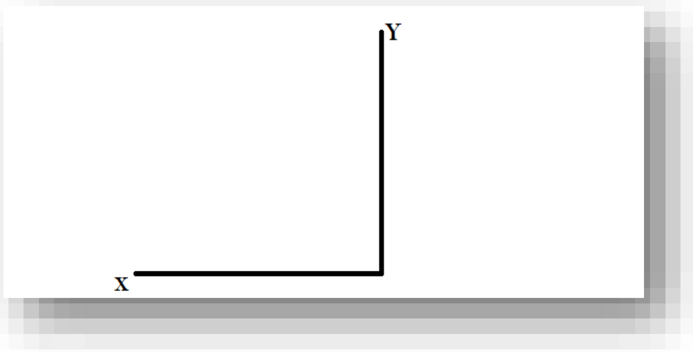


Figure 4.3a. The Manhattan distance between two points X and Y on a grid.

Bray-Curtis: The Bray-Curtis dissimilarity measure is named after J. Roger Bray and John T. Curtis (Bray and Curtis, 1957). It is a statistical approach, which is used for quantifying the compositional dissimilarity among two objects of different types. This quantitative approach is based on the number of counts at each object. It is a non-metric dissimilarity approach which is used for many applications and results are robust and reliable. Bray-Curtis dissimilarity is a modified way of the Manhattan dissimilarity measure, where the total summation of the differences among the variables is standardized with respect to the total summation of the object variables. Equation 4.9 shows the general equation of Bray-Curtis dissimilarity (Schulz, 2007):

$$d^{BCD}(i,j) = \frac{\sum_{k=0}^{n-1}|y_{i,k} - y_{j,k}|}{\sum_{k=0}^{n-1}|y_{i,k} + y_{j,k}|} \qquad (4.9)$$

In equation 4.9, d^{BCD} is the Bray-Curtis dissimilarity measure between two objects i and j, k is the variable index and n depicts the total amount of variables in y. The outcomes of Bray-Curtis dissimilarity range from zero to unity, where zero defines that the two objects have the similar composition and represent exactly same coordinates and unity defines that the two objects do not have any similarity. If both the objects are at zero coordinates, then Bray-Curtis dissimilarity measure is not defined (Bloom, 1981). The Bray-Curtis dissimilarity is not a distance as it does not satisfy the triangle inequality.

Chessboard: Chessboard is defined as a metric of greatest differences for two vectors along any dimensional coordinates in a vector space (Abello et al., 2002). It is also called Chebyshev (Tchebychev) distance after the name of Pafnuty Chebyshev. In the game of chess, the least moves required by a king to move from a square on a chessboard to another is same as the Chebyshev distance between the square centers, with a side length of one unit dimension in a 2-dimensional space (Heijden et al., 2004). It is depicted by the equation 4.10 (Moore, 2002; Balu, 2015):

$$D(X_j, V_i) = Max[Abs(X_{j1} - V_{i1}), Abs(X_{j2} - V_{i2}), \ldots\ldots\ldots\ldots, Abs(X_{jb} - V_{ib})] \quad (4.10)$$

where, b shows the total amount of bands in the image.

The following figure 4.3b depicts the difference between chessboard distance and Euclidean distance (Balu, 2015):

√2	1	√2
1		1
√2	1	√2

1	1	1
1		1
1	1	1

Figure 4.3b. Euclidean distance (left-hand side) vs Chessboard Distance(right-hand side)(Moore, 2002).

Canberra: Canberra distance was introduced by Lance and Williams (1966) and later it was refined in 1967 (Lance and Williams, 1967). It is a numerical measurement of the distance between two points in a vector space. It has been used for various purposes like a metric for comparison of ranked lists (Jurman et al., 2009) and also in computer security by using intrusion detection (Emran and Ye, 2001). It is similar to Manhattan distance metric and it is mathematically defined as the absolute difference among the variables of the objects concerned with respect to the summation of the absolute value of the variables before it is summed. Equation 4.11 shows the working of Canberra distance (Johnson and Wichern, 1998; Emran and Ye, 2001):

$$D(X_j, V_i) = \frac{Abs(X_{j1} - V_{i1})}{Abs[X_{j1}] + Abs[V_{i1}]} + \frac{Abs(X_{j2} - V_{i2})}{Abs[X_{j2}] + Abs[V_{i2}]} + \ldots\ldots \quad (4.11)$$
$$\ldots\ldots + \frac{Abs(X_{jb} - V_{ib})}{Abs[X_{jb}] + Abs[V_{ib}]}$$

where, b shows the total amount of bands in the image.

Mean Absolute Difference: The mean absolute difference is a statistical measurement of dispersion which is equal to the average value of the absolute difference of two independent numbers acquired from a probability distribution. Mathematically, it can be defined as the summation of the absolute differences between the variables of two independent objects with identical distribution of same order and type divided by the total number of variables. The mean absolute difference is generally depicted by Δ or as MD. Equation 4.12 shows the mathematical working of mean absolute difference (Vassiliadis et al., 1998):

$$D(X_j, V_i) = \frac{1}{b}[Abs(X_{j1} - V_{i1}) + Abs(X_{j2} - V_{i2}) + \ldots\ldots\ldots + Abs(X_{jb} - V_{ib})] \quad (4.12)$$

where, b shows the total amount of bands in the image.

Median Absolute Difference: Manhattan dissimilarity measure produces an exaggerated value for the distance measure when salt and pepper or impulse noise is present in the image of fixed size with n number of pixels. Manhattan dissimilarity measure calculates the summation of the absolute difference of the intensity of the corresponding pixels of two different images. The median absolute differences (MAD) may be used instead of the average of absolute differences so that the effect of the noises is reduced on the dissimilarity measure. Although, salt and pepper noise has a considerable effect on Manhattan norm, but it has minimal effect on MAD (Sari et al., 2012). MAD is mathematically defined as finding out the differences between the absolute intensities of the corresponding pixels of two images and then taking the median of the orderly data as the dissimilarity measure. Equation 4.13 (Scollar et al., 1984) shows the mathematical working of MAD:

$$D(X_j, V_i) = Median[Abs(X_{j1} - V_{i1}), Abs(X_{j2} - V_{i2}), \ldots\ldots, Abs(X_{jb} - V_{ib})] \quad (4.13)$$

where, b shows the total amount of bands in the image.

Normalised Squared Euclidean: Normalised squared Euclidean calculates the normalised squared Euclidean distance amidst two vectors. It normalises the measure with respect to the contrast of the image. Normalised squared Euclidean requires normalization of the intensities of the pixels before calculating the summation of squared differences among the pixels of two images. Equation 4.14 (Wolfram, 2010) shows the mathematical formula:

$$D(X_j, V_i) = \frac{\begin{array}{l} Abs\{X_{j_1} + \frac{1}{b}(-X_{j1} - X_{j2} .. - X_{jb}) - \mathrm{V}_{i1} + \frac{1}{b}(V_{i1} + V_{i2} .. + V_{ib})\}^2 + .. \\ + Abs\{X_{jb} + \frac{1}{b}(-X_{j1} - X_{j2} .. - X_{jb}) - \mathrm{V}_{ib} + \frac{1}{b}(V_{i1} + V_{i2} .. + V_{ib})\}^2 \end{array}}{\begin{array}{l} 2[Abs\{X_{j_1} + \frac{1}{b}(-X_{j1} - X_{j2} .. - X_{jb})\}^2 + .. \\ + Abs\{X_{jb} + \frac{1}{b}(-X_{j1} - X_{j2} .. - X_{jb})\}^2 + \\ Abs\{\mathrm{V}_{i1} + \frac{1}{b}(-V_{i1} - V_{i2} .. - V_{ib})\}^2 + .. + Abs\{\mathrm{V}_{ib} + \frac{1}{b}(-V_{i1} - V_{i2} .. - V_{ib})\}^2] \end{array}} \quad (4.14)$$

where, b shows the total amount of bands in the image.

Cosine: Cosine similarity measure calculates the Cosine of the angle between two vectors present in an inner product space. The value of the Cosine of the angle ranges from -1 to 1. The Cosine measure at zero degree angle is 1 and it decreases at any angle other than zero. Thus, vectors of similar orientation have a Cosine similarity of 1, vectors at a right angle have a Cosine similarity of 0 and vectors which are exactly opposite to each other have a Cosine similarity of -1. But, generally Cosine similarity is used in positive space, so the values are bounded from 0 to 1. Cosine similarity is used for high dimensional positive spaces. Cosine similarity gives a measurement of similarity about two vectors with respect to each other (Singhal, 2001).

This technique is used for the calculation of cohesion among the clusters in the field of data mining (Tan et al., 2005). Equation 4.15 (Ye, 2011) shows the mathematical formula of Cosine similarity:

$$D(X_j, V_i) = 1 - \frac{X_{j1}V_{i1} + X_{j2}V_{i2} + .. + X_{jb}V_{ib}}{\sqrt{Abs[X_{j1}]^2 + .. + Abs[X_{jb}]^2}\sqrt{Abs[V_{i1}]^2 + .. + Abs[V_{ib}]^2}} \tag{4.15}$$

where, b shows the total amount of bands in the image.

Correlation: Correlation similarity is a measure of finding the correlation between two vectors. It uses a standardized angular separation method by centring the coordinates towards its mean vector value. The correlation output is within the range of -1 to 1. The correlation output is normalised for a positive vector space, hence the output ranges from 0 to 1. It is a similarity measure rather than a distance measure. The similarity between two vectors is computed by using the Pearson-r correlation (Sarwar, 2001).Equation 4.16 (Zhang et al., 2008) shows the correlation mathematical formula:

$$D(X_j, V_i) = 1 - \frac{\begin{array}{c}[\{X_{j1} + \frac{1}{b}(-X_{j1} - X_{j2}.. - X_{jb})\}\{V_{i1} + \frac{1}{b}(-V_{i1} - V_{i2}.. - V_{ib})\} + .. \\ + \{X_b + \frac{1}{b}(-X_{j1} - X_{j2}.. - X_{jb})\}\{V_b + \frac{1}{b}(-V_{i1} - V_{i2}.. - V_{ib})\}]\end{array}}{\begin{array}{c}\sqrt{Abs[X_{j1} + \frac{1}{b}(-X_{j1}... - X_{jb})]^2 + .. + Abs[X_b + \frac{1}{b}(-X_{j1}... - X_{jb})]^2} \\ \sqrt{Abs[V_{i1} + \frac{1}{b}(-V_{i1} - V_{i2}.. - V_{ib})]^2 + .. + Abs[V_b + \frac{1}{b}(-V_{i1} - V_{i2}.. - V_{ib})]^2}\end{array}} \tag{4.16}$$

Euclidean: Euclidean distance is the normal distance between two objects in a metric space. The norm associated is known as the Euclidean norm. Bezdek et al. (1984) introduced this norm with FCM classifier in the form an identity matrix. Equation 4.17 shows the mathematical form of Euclidean norm used for FCM:

$$D(X_j, V_i) = I, \text{ where } I \text{ is the identity matrix} \tag{4.17}$$

Diagonal Mahalanobis Norm: Diagonal Mahalanobis norm is the diagonal matrix D_j consisting of diagonal elements which are the eigenvalues of the variance-covariance matrix C_j shown in equation 4.20. Equation 4.18 (Bezdek et al., 1984), shows the mathematical form of diagonal norm:

$$D(X_j, V_i) = D_j^{-1} \tag{4.18}$$

Mahalanobis Norm: Mahalanobis distance was introduced by Mahalanobis (1936). It measures the distance amidst a point and a distribution. The distance tends to zero as the point tends to move towards the mean of the distribution and vice versa. Bezdek et al. (1984) used this distance in the form a variance-covariance matrix C_j for FCM. Equation 4.19 shows the mathematical formulation used for FCM:

$$D(X_j, V_i) = C_j^{-1} \tag{4.19}$$

$$C_j = \sum_{i=0}^{N} (x_i - v_j)\ (x_i - v_j)^T \tag{4.20}$$

where,

$$v_j = \sum_{j=1}^{N} x_j / N \tag{4.21}$$

4.3. Composite Measure

Composite measure can be generated by using any of the two measures (similarity or dissimilarity) in combination by choosing a weighting component λ. By using a combination of two among the twelve similarity or dissimilarity measures (details in section 4.2), a composite measure can be created as mentioned in Eqn. (4.22).

$$D_c = \lambda D_a + (1 - \lambda) D_b \tag{4.22}$$

where,

D_c is Composite measure and λ is a weighting component, $0 \leq \lambda \leq 1$, D_a and D_b can be any similarity or dissimilarity measure.

4.4. α - cuts

If A is a fuzzy subset of universal set X, then the α-cut set of the fuzzy set A, will be written as A[α] and is defined as $\{x \in X \mid A(x) \geq \alpha\}$, for $0 < \alpha \leq 1$. The α equals to 0 cut, or A[0], should be defined separately because $\{x \in X \mid A(x) \geq 0\}$ is always the whole universal set X (Buckley and Eslami, 2002).

4.5. Accuracy Assessment

Assessment is a very important step to quantify the results of the outputs and to compare them with other techniques of classification (Okeke and Karnieli, 2006). The error matrix or confusion matrix or contingency table is one of the ways to showcase the accuracy of results obtained through a classification. The error matrix produces the settlement of accuracy assessment between the data that are classified and the data that are used as a reference along with wrongly classified outputs. Several statistical processes such as the Kappa coefficient, user's accuracy, producer's accuracy and overall accuracy have been introduced on the basis of the error matrix. These processes are used to sum up all the statistics about accuracy assessment.

In the sole case of hard classification, the error matrix is used for the accuracy assessment as in hard classification a single pixel belongs to a single class and, not when a pixel may belong to two or more classes (Silván-Cárdenas and Wang, 2008). In the case of soft classification, other methods like fuzzy error matrix (FERM), sub-pixel confusion uncertainty matrix (SCM), etc. were introduced for assessing the accuracy (Congalton, 1991; Binaghi et al., 1999; Jr and Cheuk, 2006). Fuzzy Error Matrix was introduced for measuring the accuracy of soft classifiers. The following section describes the methods used for accuracy assessment of soft classified outputs.

4.5.1. Fuzzy Error Matrix (FERM)

The error matrix is a square matrix in the form of rows and columns, where the rows depict the classified data (pixels) and the columns depict the elements with respect to the referenced data (pixels). The diagonal elements of an error matrix represent the pixels that are classified correctly and the elements in the off-diagonal position show the wrongly classified pixels. But, in the case of FERM, both the referenced data and the classified data are in the form of a fuzzy set, having membership values ranging in between 0 and 1. FERM is created on the basis of the MIN operator which offers a maximum overlap among the classified and the referenced data at a sub-pixel level. Equation 4.23 (Binaghi et al., 1999) shows the mathematical formulation for FERM operator:

$$\mu_{C_m \cap R_n}(x) = \min(\mu_{C_m}(x),\ \mu_{R_n}(x) \tag{4.23}$$

Where, R_n depicts the membership value from the referenced data, in the form of a set, which is allotted to class n, C_m depicts the membership value from the classified data, also in the form of a set, which are allotted to class m and the membership value of a pixel with respect to the classes is shown by μ. The overall accuracy is the primitive form of statistics gathered from an accuracy assessment. In error matrix, the overall accuracy is measured by summing up the diagonal components of the matrix and dividing the total by the sum of the sample components in the matrix. In the case of FERM, the overall accuracy is measured by calculating the sum of the diagonal components divided by the total membership value of the referenced data. Equation 4.24 (Kumar, 2007) shows the mathematical formulation:

$$OA_{FERM} = \frac{\sum_{i=1}^{c} M(i,j)}{\sum_{i=1}^{c} R_j} \tag{4.24}$$

Here, OA depicts the overall accuracy, M (i, j) depicts the element of the m^{th} class of the soft classified result and n^{th} class of the soft reference record, c depicts the total number of classes and R_j depicts the total summation of the membership value of n class in the soft reference data.

4.5.2. Subpixel confusion uncertainty matrix (SCM)

Determination of the true overlap in between classes which are on the basis of fractional land-cover is challenging. This kind of situation is known as sub-pixel area allocation problem (Silván-Cárdenas and Wang, 2008). The spatial distribution of the classes determines the minimum or maximum overlap of the classes in a pixel. This kind of problem gives rise to solutions such as a unique solution or no solution. For, unique solution, there is a chance of overestimation or underestimation of classes and hence, sub-pixel confusion matrix can be uniquely defined. For the case of no solution, as there is a lack of unique solution, hence the solutions are depicted by confusion intervals. SCM has confusion intervals, which are shown as central value ± maximum error. The confusion matrix produced for a soft classifier output satisfies the following (Silván-Cárdenas and Wang, 2008):

- *Property of Diagonalization*: If the data that are considered equals the classified data, then the matrix is a diagonal matrix.
- *Property of Marginal Sums*: The total summation of the marginal equals the total values both from the data that are assessed and the classified data.

Several operators were introduced for measuring the relationship between pixel and class in sub-pixel classifications. MIN operator provides the maximum overlap possible among the data that are classified and the data that are assessed. This process may result in an overestimation of the actual agreement and disagreement at a sub-pixel level, thus results in larger marginal sums. The Similarity Index (SI) produces a sub-pixel overlap normalization and is also a modification of the MIN operator. The PROD operator produces the expected overlap possible among the data that are classified and the data that are assessed. The LEAST operator produces the minimal sub-pixel overlap possible among the two concerned classes (Silván-Cárdenas and Wang, 2008).

Several other composite operators like MIN-PROD, MIN-MIN and MIN-LEAST were introduced as the basic operators were unable to fulfil the property of diagonalization. The MIN-MIN operator operates by assigning the components at the diagonal and then the off-diagonal components. The MIN-LEAST operator operates by using the MIN operator for the components at the diagonal and the LEAST operator for the components at the off-diagonal positions. The MIN-PROD operator operates by using the MIN operator for the components on the diagonal of the matrix and the normalized PROD operator for the components at the off-diagonal positions. To determine the minimum and maximum overlapping at a sub-pixel level operators like MIN-MIN and MIN-LEAST were put forth respectively. MIN-PROD operator is used when utmost a class has been overestimated or underestimated (Silván-Cárdenas and Wang, 2008).

5. RESULTS AND DISCUSSION

This chapter describes the classified outputs achieved by using different classification methods and the analysis of them. The following section 5.1 presents the results of FCM classification on the simulated image of the FORMOSAT-2 dataset and LANDSAT-8 dataset, which was used for getting the optimized value of the weighted constant '*m*'. The next sections of 5.2 and 5.4 show the results of the optimized 'm' with the FORMOSAT-2 dataset for single measure and composite measure respectively. The sections of 5.3 and 5.5 contain the results of FCM classification of the LANDSAT-8 dataset for single measure and composite measure respectively and the sections 5.6 and 5.7 depict the results of both FORMOSAT-2 and LANDSAT-8 datasets by using α-cuts for single measure and composite measure respectively. The results of untrained class is shown in section 5.8 and followed by the discussion of the results in section 5.9.

5.1. Identification of best measure and estimation of the parameter

The behavioural characteristics of the developed FCM were studied in details using simulated image. This simulated image was developed to estimate the parameters and also to check the accuracy of the FCM classification. The simulated image was developed according to the study area selected, containing all the classes present in the study area and a within the class variation was incorporated to check the capability of the FCM classifier to detect variation at an intra-class level. The simulated image has been generated for the Formosat-2 image as well as for the Landsat-8 image. The membership grade of a pixel with respect to a class in a fractional image ranges from 0 to 1. In order to eliminate the cumbersome process of handling decimal digits between 0 and 1, the membership grades were up-scaled to 8-bit values ranging from 0 to 255. In FCM, the membership grade of zero for a pixel denotes that there is no belongingness of the pixel to a concerned class and the membership grade of 255 for a pixel denotes that the pixel completely belongs to the concerned class. In this research work, the fractional images from Formosat-2 dataset has been used as the referenced images to calculate the accuracy of Landsat-8 dataset.

The best parameter value of weighted constant "*m*" was estimated for the developed FCM algorithm with the simulated image (details in section 3.2.1), as the input values of the image and the values of the expected outputs were known. This optimized parameter of weighted constant "*m*" was also used to check the effect of change of the degree of fuzziness on the accuracy. The weighted-constant or fuzzifier was optimized within the value ranging from 1.10 to 3.00. By, using the optimization of parameter technique, the best two norms out of all twelve norms were chosen to form a composite measure (details in section 4.3). The optimization of the parameter value of weighed constant "*m*" was also estimated on the supervised FCM algorithm with this composite measure. Lastly, with this optimized parameter value of weighted constant "*m*" for the best similarity or dissimilarity measure (single or composite) was implemented into the supervised FCM classification algorithm.

The simulated image was used for optimization of the weighted constant "*m*" parameter and also to find out the best similarity or dissimilarity measures for both Formosat-2 and Landsat-8 datasets. This optimized parameter of "*m*" along with the best similarity or dissimilarity measure was used for the image to image accuracy assessment for the image of coarser resolution, Landsat-8. The aforementioned method was also used to optimize the weighted-constant "*m*" parameter to find the best similarity or dissimilarity measure for the Landsat-8 image. Accuracy assessment techniques like FERM (details in section 4.5.1) and SCM (details in section 4.5.2) were used to measure the accuracy of the classified images.

5.1.1. Fuzzifier or Weighted Constant (*m*)

Parameter estimation for Formosat-2 using simulated image

Here, the study was executed with Formosat-2 simulated image (details in section 3.2.1) with five different classes. Firstly, the FCM algorithm was implemented with five different classes namely Fallow-Land, Forest, Riverine-Sand, Water and Wheat to optimize the weighted constant "*m*" for the FCM algorithm on various similarity and dissimilarity measures (mentioned in section 4.2) by following the well-defined method (details in section 3.6.1). After executing this method, a comparative exploration was done on the effect of the fuzzifier "m" on each similarity measures incorporated to the FCM algorithm.

At first, we have implemented the FCM classification algorithm for all the similarity measures on the simulated image of the Formosat-2 dataset. The value for the weighted-constant or fuzzifier "*m*" was carefully chosen on the basis of the results obtained in the classification. In the results, the criteria for optimality was based on the classification of the pure pixels, whose value should reach the target value of 255 and the intra-class variation should be least for that concerned class. Along with the aforementioned criteria, the mixed pixel should also be classified according to the target values (details in section 3.6.1). The results obtained showed that, for the Formosat-2 simulated image the optimal value of "*m*" was achieved at m equals to 2.7 for Cosine norm, which was the best measure according to the criteria defined in section 3.6.1. Figure A-4 (Appendix A) shows the outputs of FCM algorithm with a simulated image for Cosine norm with m equals to 2.7. Table 5.1 shows the results of all the similarity measures while handling the pure pixel classes and also its behaviour for within the class variation. Table 5.2 shows the results of all the similarity measures while handling the mixed pixels of two different classes. Table 5.3 shows the results of all the similarity measures while handling the mixed pixels of three different classes.

Norms with *m*-value	Water	Wheat	Forest	Riverine-Sand	Fallow-Land	Total Variation of pure pixel class
Canberra (1.9)	247-239 = 8	249 – 242 = 7	246- 238 = 8	253-252 = 1	246 – 237 = 9	33
Cosine (2.7)	253-252 = 1	254 – 253 = 1	253-252 = 1	254 – 254 = 0	253-252 = 1	4
Euclidean (2.5)	253-250=3	254-253=1	253-250=3	254-253=1	252-248=4	12
Chessboard (1.9)	252-249=3	253-252 = 1	251-248=3	253-251=2	251-246=5	14
Mean absolute distance (1.9)	248-240=8	251-246=5	248-241=7	253-251=2	246-237=9	31
Diagonal Norm(2.5)	253-250=3	254-253=1	253-250=3	254-253=1	252-248=4	12
Median absolute distance (1.9)	250-246=4	252-250=2	250-246=4	253-251=2	249-243=6	18
Manhattan (1.9)	248-240=8	251-246=5	248-241=7	253-251=2	246-237=9	31
Bray-Curtis (1.9)	247-240=7	241-247=4	248-240=8	253-252=1	246-237=9	29
Correlation (2.5)	0-0	0-0	0-0	0-0	0-0	-
Mahalanobis (2.5)	124-122	252-249	124-122	254-253	237-201	-
Normalised Squared Euclidean (2.7)	255	255	254	0	254	-

Table 5.1. The similarity measures for handling the pure pixel classes and also its behaviour within the class variation (Membership value was calculated on an 8-bit scale i.e., the target values for a class were 255 and 254 (with a variation of 1 within the class)).

Norms with *m*- value	Riverine Sand – Forest	Riverine Sand – Fallow Land	Water – Wheat
Canberra (1.9)	68 – 46	29-76	35-54
Cosine (2.7)	69-33	46-154	14-20
Euclidean (2.5)	50-52	18-94	30-29
Chessboard (1.9)	53-54	26-105	28-28
Mean absolute distance (1.9)	51-52	20-77	39-40
Diagonal Norm(2.5)	50-52	18-94	30-29
Median absolute distance (1.9)	49-50	22-97	37-37
Manhattan (1.9)	51-52	20-77	39-40
Bray-Curtis (1.9)	65-47	28-74	37-44
Correlation(2.5)	19-26	12-210	10-39
Mahalanobis(2.5)	32-34	3-18	57-49
Normalised Squared Euclidean (2.7)	24-24	15-211	16-16

Table 5.2. The similarity measures for handling the mixed pixel containing two classes (Membership value was calculated on an 8-bit scale i.e., the target value for each class was 127.5 respectively).

Norms with m- value	Water-Forest-Riverine Sand	Water-Riverine Sand-Wheat	Riverine Sand - Fallow Land - Wheat
Canberra (1.9)	52-52-49	49-49-48	47-60-47
Cosine (2.7)	32-23-78	27-45-23	48-69-23
Euclidean (2.5)	61-53-35	44-33-49	31-65-47
Chessboard (1.9)	50-55-41	33-40-48	40-55-45
Mean absolute distance (1.9)	59-58-36	55-36-47	34-62-45
Diagonal Norm(2.5)	61-53-35	44-33-49	31-65-47
Median absolute distance (1.9)	58-50-39	45-34-49	34-63-46
Manhattan (1.9)	59-58-36	55-36-47	34-62-45
Bray-Curtis (1.9)	55-53-48	50-48-48	45-58-47
Correlation(2.5)	60-16-51	6-6-15	3-7-8
Mahalanobis(2.5)	24-24-13	27-17-151	14-28-161
Normalised Squared Euclidean (2.7)	41-20-61	13-15-9	20-42-10

Table 5.3. The similarity measures for handling the mixed pixel containing three classes (Membership value was calculated on an 8-bit scale i.e., the target values for each class were 76.5, 76.5 and 102 respectively).

Here the mixed pixels are simulated with two types of variations, one with the composition of 50:50 (as shown in figures A-1 and A-2, Appendix-A) between two different classes and another with the arrangement of 30:30:40 (as shown in figures A-1 and A-2, Appendix-A) among three different classes. The target membership value expected for a pixel belonging completely to a class must be close to 255 (on an 8-bit scale) and the target membership value for a pixel of the mixed pixels of two different classes must be close to 127.5 (on an 8-bit scale) i.e., 50% of the full membership value of a pixel belonging to a concerned class and the target membership value for a pixel of the mixed pixels of three different classes must be close to 76.5, 76.5 and 102 (on an 8-bit scale) i.e., 30%, 30% and 40% of the full membership value of a pixel belonging to a concerned class respectively.

The results obtained for Formosat-2 simulated image as shown in table 5.1, depict that Cosine norm at m equals to 2.7 shows the best result among all the similarity measures for handling the pure pixels in an image and also can detect the intra-class variation properly. The results shown in table 5.2 and table 5.3 show that the measures were unable to handle the mixed pixels properly. This can be due to the inefficiency of FCM classifier to handle noise. Here, the mixture of two or more classes creates noise for the other concerned class during classification and hence, the developed FCM algorithm cannot handle the mixed pixels properly. A similar analysis was done on the simulated image of Landsat-8, which resulted in Cosine norm with m equals to 2.5 showed the best result while handling the pure pixels in an image and also while detecting the intra-class variation.

5.1.2. Optimization of weighting component (λ) for composite measure

In section 4.3 it has been discussed that for composite measure a weighting component (λ) is required, which provides weight λ to a norm D_a and 1 – λ to another norm D_b. For, a composite measure it was essential to optimize both the parameters of λ and m. The values considered for λ was ranging from 0.10 to 0.90. However, the classification may result in misclassified outputs when the weight set for a norm D_a is greater than D_b. This kind of misclassification arises if the performance of D_a is better than D_b and with a larger value of weighting component (λ) to D_b in a composite situation will result in a measure with inferior results. Figure A-5 (Appendix A) shows that fallow-land and forest classes have misclassification in the results. The two best norms obtained from the results shown in table 5.1, table 5.2 and table 5.3 were Cosine and Euclidean. These two norms were used to make the composite norm. The results obtained after optimization of both parameters m and λ on the simulated images show that the composite norm of Cosine and Euclidean were optimized at "*m*" equals to 2.5. However, there was no significant change observed while changing the value of λ from 0.10 to 0.99. Table 5.4, table 5.5 and table 5.6 shows the comparison between the results of the best single norm and the results of the composite norm. In figure A-5 (Appendix A), it was observed that the fallow land was misclassified as forest and also forest was misclassified with water and fallow land.

Norms with m- value	Water	Wheat	Forest	Riverine-Sand	Fallow-Land	Total Variation of pure pixel class
Cosine (2.7)	253-252 = 1	254 – 253 = 1	253-252 = 1	254 – 254 = 0	253-252 = 1	4
Euclidean + Cosine (2.5)	253-250=3	254-253=1	253-250=3	254-253=1	252-248=4	12

Table 5.4 The comparative results of the best single similarity measures and the composite measure while handling the pure pixel classes and also its behaviour within the class variation (Membership value was calculated on an 8-bit scale i.e., the target values for a class is 255 and 254 (with variation of 1 within the class))

Norms with m- value	Riverine Sand – Forest	Riverine Sand – Fallow Land	Water – Wheat
Cosine (2.7)	69-33	46-154	14-20
Euclidean + Cosine (2.5)	50-52	18-94	30-29

Table 5.5 The comparative results of the best single similarity measures and the composite measure while handling the mixed pixel containing two classes (Membership value was calculated on an 8-bit scale i.e., the target values for each class was 127.5 respectively).

Norms with m- value	Water-Forest-Riverine Sand	Water-Riverine Sand-Wheat	Riverine Sand - Fallow Land - Wheat
Cosine (2.7)	32-23-78	27-45-23	48-69-23
Euclidean + Cosine (2.5)	61-53-35	40-33-49	31-65-47

Table 5.6 The comparative results of the best single similarity measures and the composite measure while handling the mixed pixel containing three classes (Membership value was calculated on an 8-bit scale i.e., the target values for each class were 76.5, 76.5, 102 respectively).

5.2. FCM classification Results for Formosat-2 Dataset using single similarity measure

FCM classifier was applied with a supervised approach for classification of Formosat-2 data. For this process, a total of 20 training pixels were carefully chosen from each of the land cover class. The training sites were selected at various locations spread well over the Formosat-2 image. In FCM classifier using supervised approach Cosine norm was considered (details in section 5.1.1). The weighted-constant or fuzzifier for Cosine was optimized at *m* equals to 2.7. The results of FCM classification on Formosat-2 data has been shown in figure 5.1.

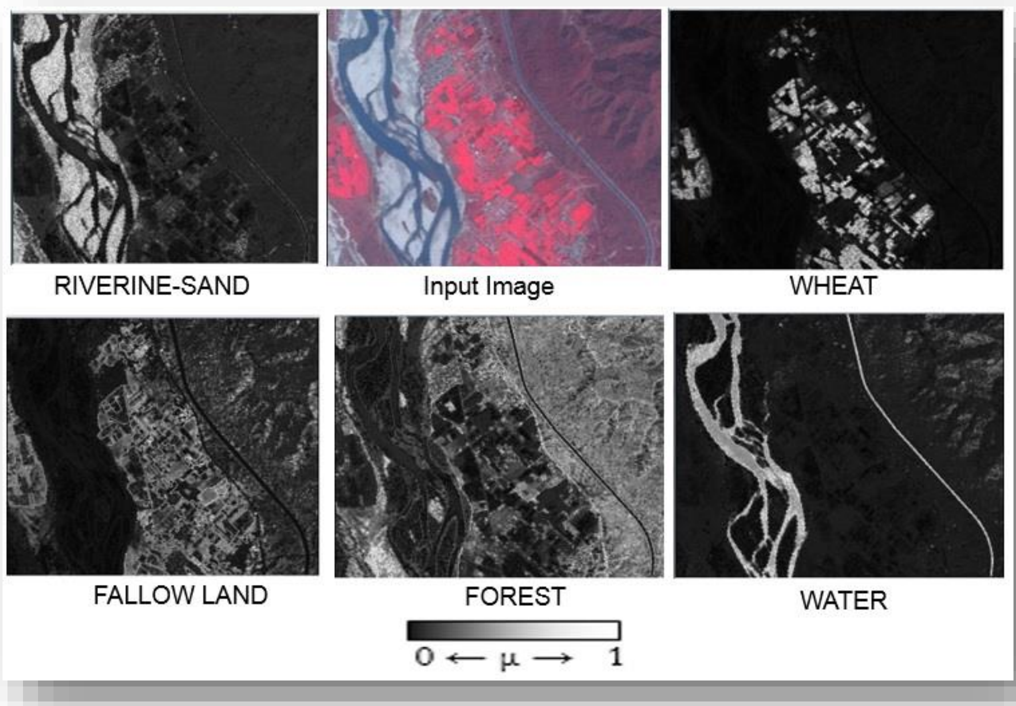


Figure. 5.1 Fractional images of FCM classification with Cosine norm at *m* equals to 2.7 for Formosat-2 data.

The membership value of all the pixels in the fractional images is ranging from 0 to 1. These fractional images were used as the reference data for measuring the accuracy of landsat-8 classified fractional images. With the purpose of getting the resolution (10m) of Formosata-2 data (reference data) on a scale with Landsat-8 data (30m), the method of *mean-aggregation* was implemented. As the method of *mean-aggregation* was applied on the reference data of finer resolution, the sensor's *point spread function* (PSF) was overlooked.

5.3. FCM classification Results for Landsat-8 Dataset using single similarity measure

FCM classifier was applied with a supervised approach for classification of Landsat-8 data. For this process, a total of 20 training pixels were carefully chosen from each of the land cover class. The training sites were selected at various locations spread well over the Landsat-8 image. In FCM classifier using supervised approach Cosine norm was considered (details section 5.1.1). The results of FCM classification on Landsat-8 data has been shown in figure 5.2.

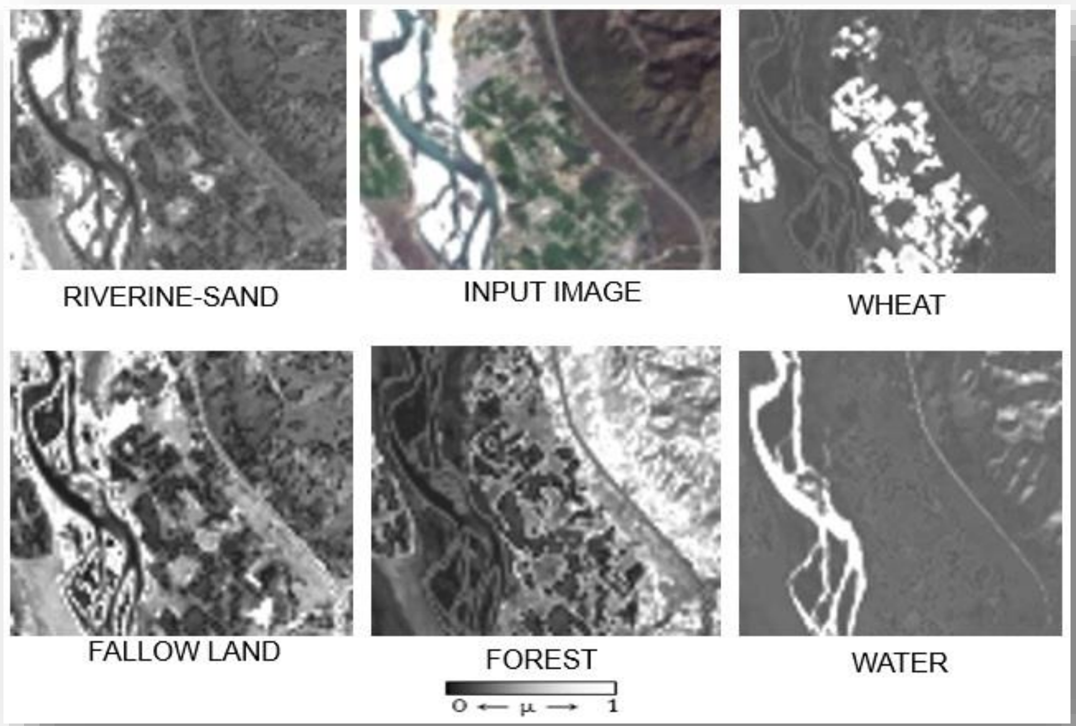


Figure. 5.2 Fractional images of FCM classification with Cosine norm at *m* equals to 2.5 for Landsat-8 data.

The membership value of all the pixels in the fractional images was ranging from 0 to 1. In this study, mean-aggregation method was used to maintain the scale ratio of resolutions of the reference data and the assessed data. The accuracy was assessed by the fuzzy based techniques like FERM and SCM (details in section 4.5). The method of mean-aggregation was also followed by FERM and SCM so that the referenced data and the assessed data are at the same scale (Binaghi et al., 1999 and Silván-Cárdenas et al., 2008). As the referenced data and the assessed data were brought to the same scale, the following fuzzy accuracy operators like, FERM and SCM were used to measure the accuracy of the fractional images of the Landsat-8 dataset. 500 sample points (pixels) were selected randomly as the test sites to carry out the accuracy assessment. The fuzzy user's accuracy, producer's accuracy, kappa coefficient and overall accuracy were computed for all the fuzzy accuracy operators using the error matrices. Table 5.7 shows the detailed statistics of the accuracy assessment.

Accuracy Assessment Operators	FERM	SCM
User's Accuracy (%)		
Riverine Sand	86.05	86.88 ± 3.77
Fallow Land	60.90	62.45 ± 5.52
Forest	81.23	82.35 ± 4.15
Water	81.72	82.87 ± 2.97
Wheat	63.75	65.22 ± 5.26
Producer's Accuracy (%)		
Riverine Sand	66.30	67.78 ± 5.39
Fallow Land	85.92	86.60 ± 2.50
Forest	80.14	81.18 ± 2.75
Water	50.86	53.50 ± 9.36
Wheat	77.41	78.72 ± 3.68
Overall Accuracy (%)	73.96	75.24 ± 4.72
Fuzzy Kappa value		0.68 ± 0.06

Table 5.7 Details of accuracy assessment for classification results of Landsat-8 data using single measure

Rendering to the table 5.7, the overall accuracy of the fuzzy classification of Landsat-8 data was found to be ranging in between 73% to 76% (73.96% in FERM, 75.24% in SCM) and SCM shows an uncertainty of ±4.72% indicating a range of uncertainty over the different land cover classes. The results obtained from the fuzzy kappa statistics for this classification ranges in between 0.62 to 0.74 for SCM fuzzy accuracy operators. The mean fuzzy kappa value as obtained for SCM was 0.68 and the uncertainty associated with SCM kappa (± 0.06) was found to be low, indicating low error sources in the calculation of uncertainty (Silván-Cárdenas et al., 2008). From the statistics obtained in table 5.7, it can be inferred that the overall performance of the developed FCM classification method for Landsat-8 dataset was unaffected by the uncertainty.

The statistics of the classification of the Landsat-8 data by FCM approach for single measure has been shown in figure A-6 (Appendix –A). The results show better user's accuracy than producer's accuracy for all the fuzzy tools used for accuracy assessment. Thus, it can be inferred that omission error has occurred more than the commission error for the concerned land cover classes.

5.4. FCM classification Results for Formosat-2 Dataset using composite similarity measure

Composite similarity measures were tried to incorporate the characteristics of both the norms that are used in forming the composite similarity measure. For, this study a combination of the two best norms obtained from section 5.1 namely, Cosine and Euclidean were considered. The optimized value of the fuzzifier (*m)* as obtained (details in section 5.1.2) was equal to 2.5 for the composite similarity measure. However, there was no significant change in the classified outputs while optimizing the value of weighting constant (λ). Thus, the value of weighting component was set at λ equals to 0.5 (mean value of the range of λ [0.10, 0.90]) for the classification, which signifies the contribution of both the norms of an equal distribution of 50%. The result of supervised FCM classification using the composite measure has been shown in figure 5.3.

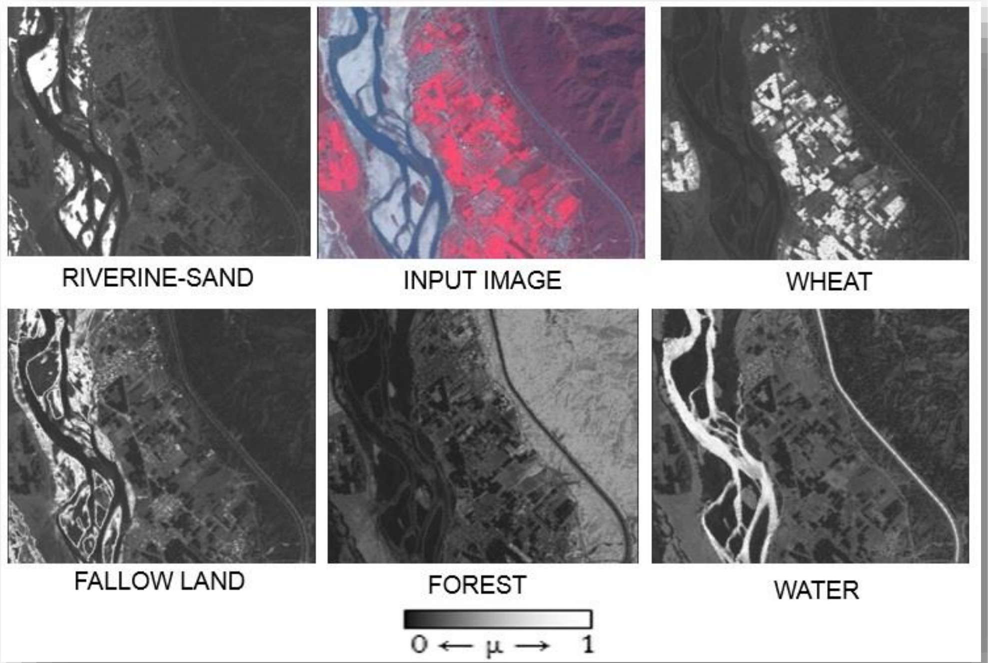


Figure. 5.3 Fractional images of FCM classification with composite measure at *m* equals to 2.5 and λ equals to 0.5 for Formosat-2 data.

The membership value of the pixels distributed all over the fractional images ranges from 0 to 1. These fractional images were used as the reference data for measuring the accuracy of Landsat-8 classified images in the section 5.5.

5.5. FCM classification Results for Landsat-8 Dataset using composite similarity measure

In this study, supervised FCM classification approach was used for classification of Landsat-8 dataset. A total of 20 training sites were carefully selected for each class. The training sites were chosen from various locations spreading all over the image. The two best norms obtained in section 5.1.1 were Euclidean and Cosine with an optimized fuzzifier (*m)* value equals to 2.5. . However, there was no significant change in the classified outputs while optimizing the value of weighting constant (λ). Thus, the value of weighting component was set at λ equals to 0.5 (mean value of the range of λ [0.10, 0.90]) for the classification, which signifies the contribution of both the norms of an equal distribution of 50%. The results of FCM classification on Landsat-8 data has been shown in figure 5.4.

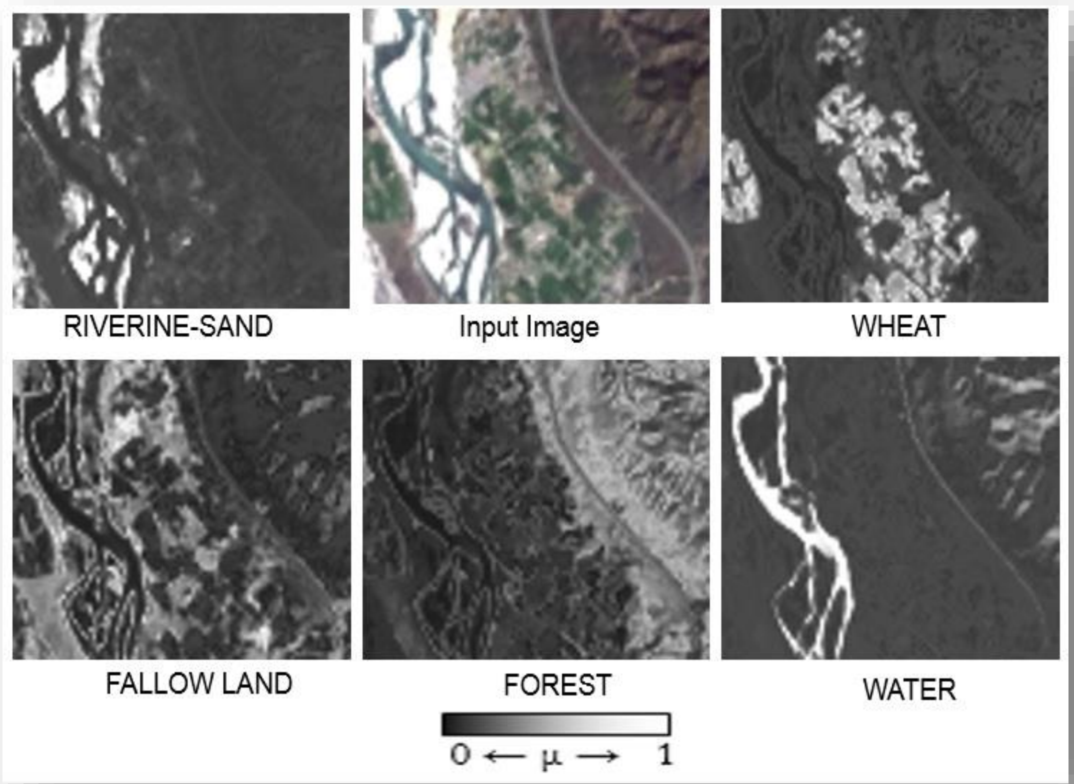


Figure. 5.4 Fractional images of FCM classification with composite measure at *m* equals to 2.5 and λ equals to 0.5 for Landsat-8 data.

The membership values of the pixels present in the fractional images range in between 0 and 1. The accuracy assessment of these classified images was measured using various fuzzy based accuracy tools namely, FERM and SCM (details in section 4.5). 500 random sample points were carefully chosen as the test sites for carrying out the accuracy assessment. Different accuracy statistics were calculated like user's accuracy, producer's accuracy, overall accuracy and kappa coefficient value for all the accuracy tools by means of error matrices. Table 5.8 shows the details of the statistics.

Accuracy Assessment Operators	FERM	SCM
User's Accuracy (%)		
Riverine Sand	66.67	68.24 ± 6.25
Fallow Land	44.84	46.26 ± 4.25
Forest	86.30	86.94 ± 2.59
Water	76.83	77.76 ± 1.68
Wheat	64.31	65.92 ± 5.29
Producer's Accuracy (%)		
Riverine Sand	68.08	69.67 ± 5.93
Fallow Land	77.72	78.51 ± 2.20
Forest	75.28	76.46 ± 3.35
Water	42.30	43.72 ± 5.17
Wheat	78.99	80.11 ± 2.99
Overall Accuracy (%)	68.57	69.80 ± 4.21
Fuzzy Kappa value		0.61 ± 0.06

Table 5.8 Details of accuracy assessment for classification results of Landsat-8 data using composite measure.

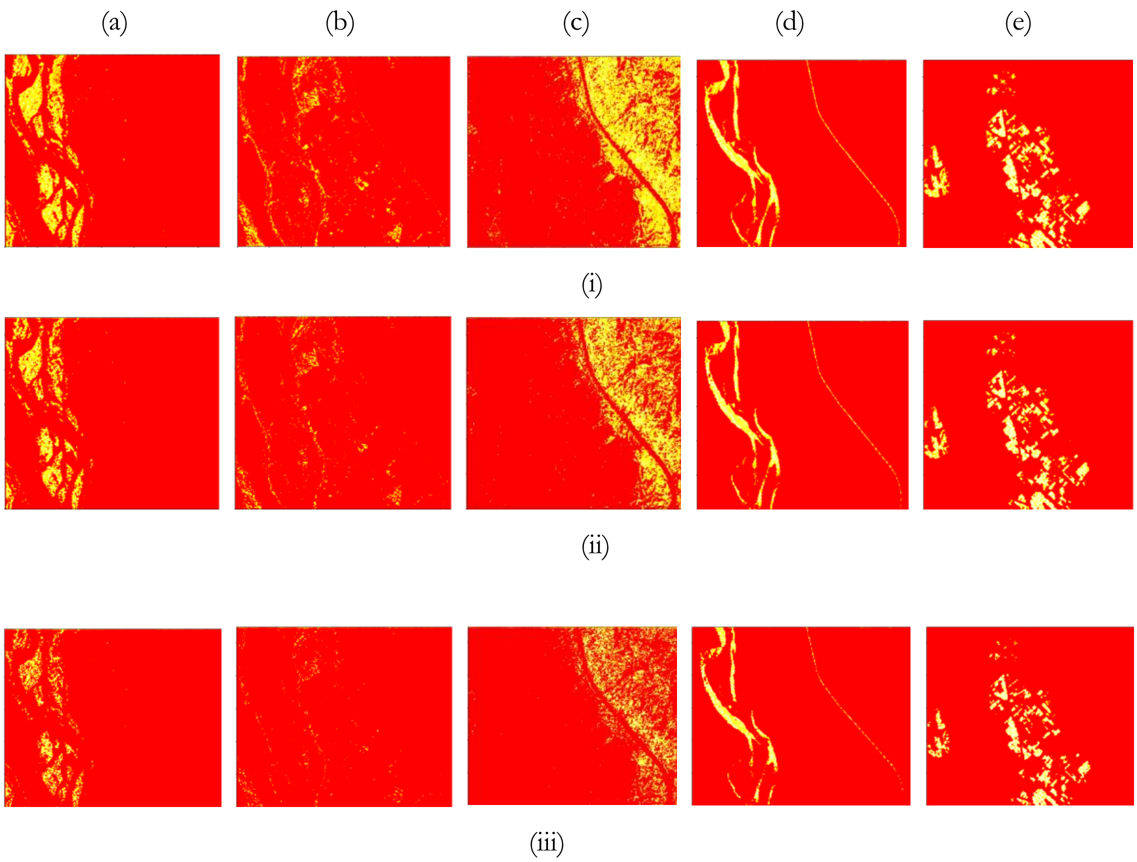
Rendering to the results observed in the table 5.8, the overall accuracy of the classified images of Landsat-8 was found to be in the range from 65% to 74% (68.57% in FERM, 69.80% in SCM) with SCM showing an uncertainty of ± 4.21%, which indicates uncertainty over the land cover classes present in the dataset. The results obtained for the fuzzy kappa statistics ranges from 0.55 to 0.66 for the SCM fuzzy accuracy assessment tool. The kappa value for SCM shows an uncertainty of ±0.06 from the mean value of 0.61. From the statistics obtained in table 5.8, it can be inferred that the overall performance of the developed FCM using composite measure for Landsat-8 dataset was unaffected by the uncertainty. More distinctly the performances of the developed FCM classifier with the various accuracy operators has been shown in figure A-7 (Appendix A). It has been also observed that results for producer's accuracy was better than user's accuracy for all the fuzzy based accuracy assessment. Hence, it can be inferred that the error of commission was more than the omission error for the concerned classes.

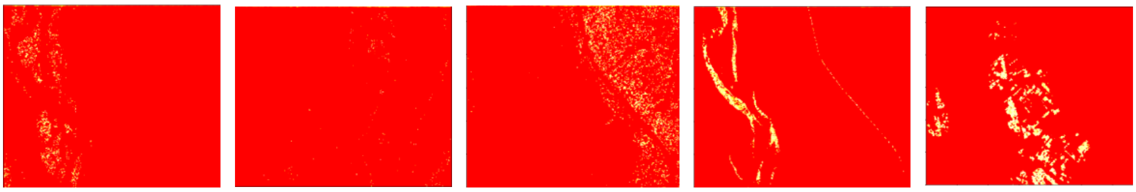
5.6. Results of FCM Classifier using α-cuts with single norm

Though fuzzy clustering techniques with cluster cores have good clustering characteristics, however, there can be difficulties in cluster cores produced by FCM in case non-spherical shape clusters. Like, the cluster cores of two overlapping clusters (line structure) cannot be determined by FCM (Yang et al., 2009). So, to describe the general core of the clusters of any shape, α-cut has been incorporated in the FCM algorithm. The cluster cores generated by FCM was such that if the distance between the pixel and the cluster center of the concerned class was less than a defined threshold (α-cut value), then that pixel would be belonging to that class with membership grade value of 1. In this study, the α-cut value was taken in the range from 0.5 to 0.9 with an interval of 0.1, as suggested by Yang et al. (2009).

The α-cut FCM algorithm was implemented on the results obtained from sections 5.2 and 5.3. The generated fractional images of the α-cut FCM algorithm has been shown in figure 5.5 for each α-cut starting from 0.5 to 0.9 with an interval of 0.1 for Formosat-2 data using Cosine norm at m equals to 2.7.

Figure 5.5. Generated fractional images for Cosine norm at optimized m value of 2.7 of Formosat-2 data for (i) α-cut = 0.5 (ii) α-cut = 0.6 (iii) α-cut = 0.7 (iv) α-cut = 0.8 (v) α-cut = 0.9 for all the classes (a) Riverine-Sand (b) Fallow-Land (c) Forest (d) Water (e) Wheat.





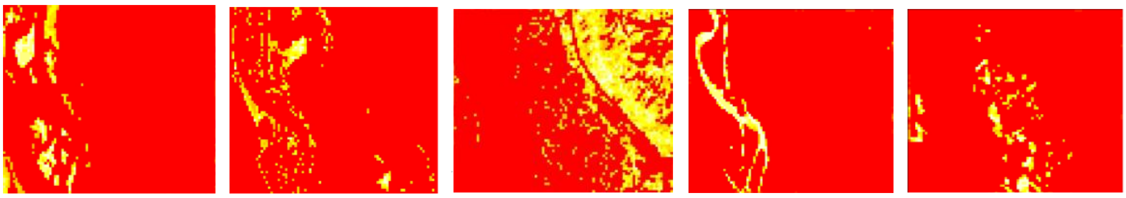
(iv)

(v)

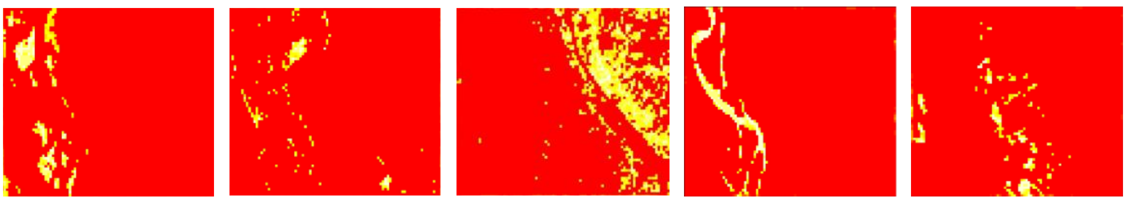
The membership grade of all the pixels in these fractional images ranges from 0 to 1. These fractional images were used as the reference data for measuring the accuracy of the results obtained from the Landsat-8 image. From the images in figure 5.5, it has been observed that as the α-cut value was increased from 0.5 towards 0.9 with an interval of 0.1, the membership grades of the pixels which were less than the threshold value were removed. The generated fractional images of the α-cut FCM algorithm has been shown in figure 5.6 for each α-cut starting from 0.5 to 0.9 with an interval of 0.1 for Landsat-8 data using Cosine norm at m equals to 2.5.

Figure 5.6. Generated fractional images for Cosine norm at optimized *m* value of 2.5 of Landsat-8 data for (i) α-cut = 0.5 (ii) α-cut = 0.6 (iii) α-cut = 0.7 (iv) α-cut = 0.8 (v) α-cut = 0.9 for all the classes (a) Riverine-Sand (b) Fallow-Land (c) Forest (d) Water (e) Wheat.

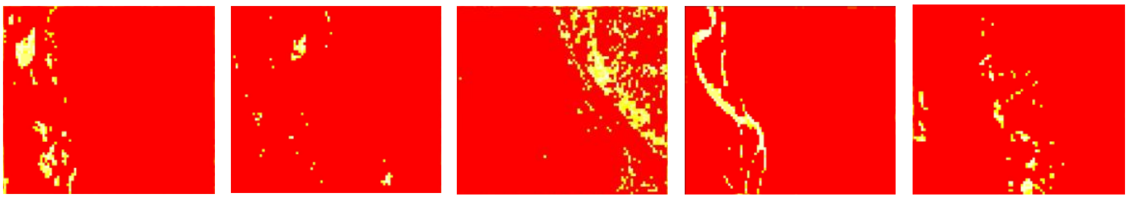
(a) (b) (c) (d) (e)



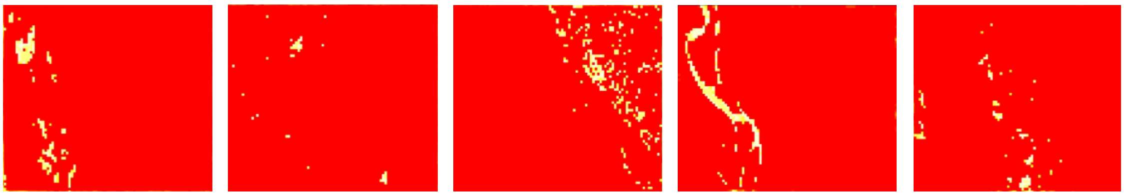
(i)



(ii)

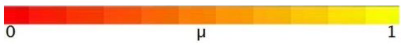


(iii)



(iv)

(v)



The membership grade of all the pixels in these fractional images ranges from 0 to 1. From the figures, it has been observed that as the α-cut value was increased from 0.5 towards 0.9 with an interval of 0.1, the membership grades of the pixels which were less than the threshold value were removed. The accuracy assessment was done on these fractional images with the fractional images from the Formosat-2 dataset as the referenced images. Fuzzy accuracy operators like FERM and SCM were used to measure the accuracy of the classified images. Total of 500 sample points (pixels) were chosen randomly as the test sites for carrying out the accuracy assessment. The fuzzy user's accuracy, producer's accuracy, kappa coefficient and overall accuracy for all the accuracy operators were calculated for all the fuzzy accuracy operators using the error matrices for an α-cut value equal to 0.5, 0.6, 0.7, 0.8 and 0.9 respectively (Table B-1 to Table B-5, Appendix B).

5.7. Results of FCM Classifier using α-cuts with composite norm

Fuzzy clustering methods have good clustering features, but while handling non-spherical shape cluster, it may show some errors. For instance, FCM cannot determine cluster cores for two overlapping line structured clusters (Yang et al., 2009). Thus, the concept of α-cut was incorporated in this study to determine the cluster core for any shape. In this study, the α-cut FCM algorithm was implemented on the classified images of the Formosat-2 dataset using the composite measure as obtained in section 5.4 with α-cut values of 0.5, 0.6, 0.7, 0.8 and 0.9. The fractional images of the α-cut FCM algorithm have been shown in figure 5.7 for each α-cut value ranging from 0.5 to 0.9 with an interval of 0.1.

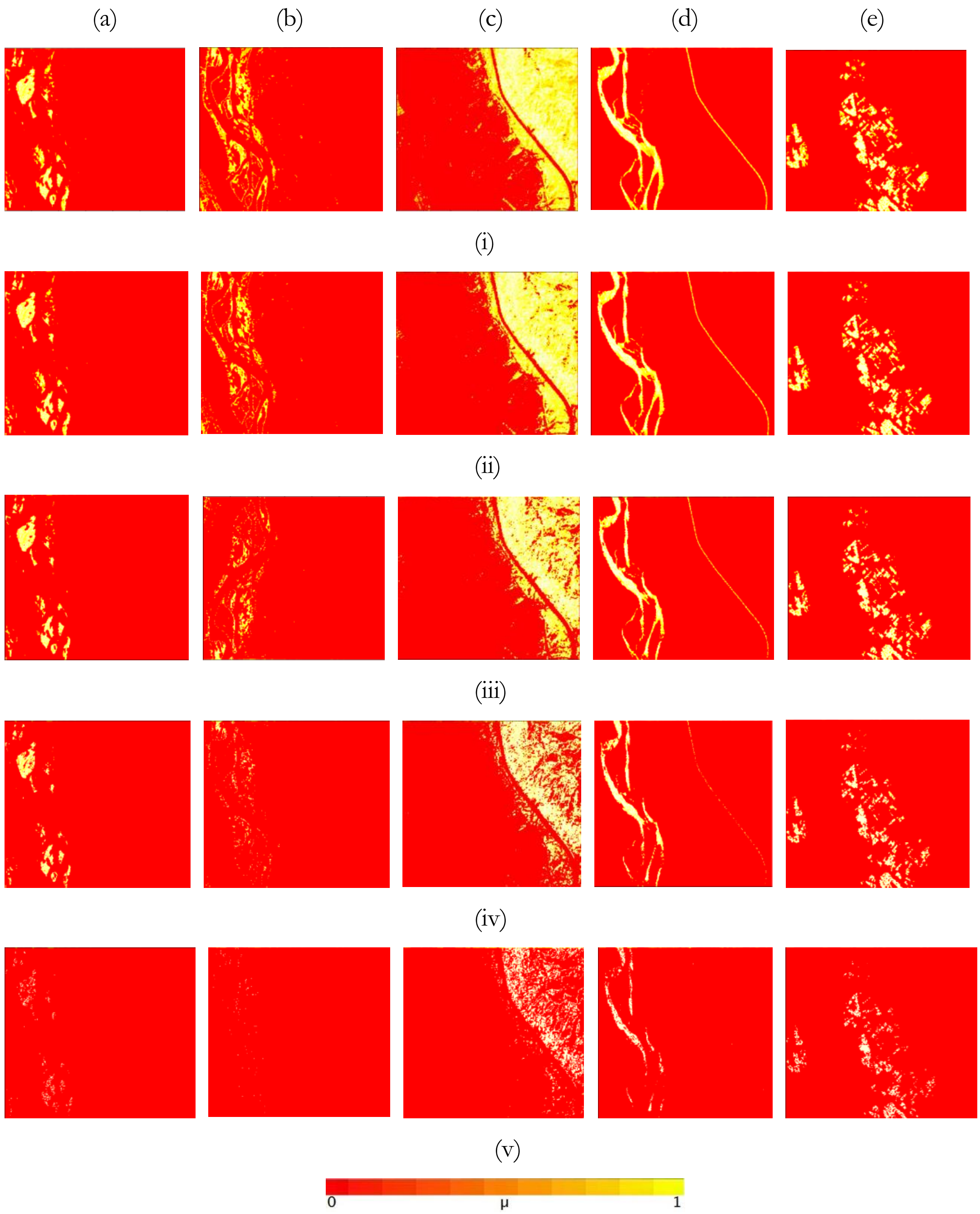


Figure 5.7 Generated fractional images for composite measure of Cosine and Euclidean norms at optimized *m* value of 2.5 and λ value of 0.5 of Formosat-2 data for (i) α-cut = 0.5 (ii) α-cut = 0.6 (iii) α-cut = 0.7 (iv) α-cut = 0.8 (v) α-cut = 0.9 for all the classes (a) Riverine-Sand (b) Fallow-Land (c) Forest (d) Water (e) Wheat.

The membership values of all the points (pixels) in these fractional images lies in between 0 and 1. These resultant fractional images were used as the reference images for the α-cut FCM classified images of Landsat-8 data. It was observed that on increasing the α-cut (threshold) value from 0.5 to 0.9 with an interval of 0.1, the pixels containing membership value less than the threshold were removed.

The α-cut FCM algorithm was applied on the classified images of the Landsat-8 dataset (section 5.5) with α-cut values ranging from 0.5 to 0.9 with an interval of 0.1. The fractional images of the α-cut FCM algorithm have been shown in figure 5.8 for each α-cut value.

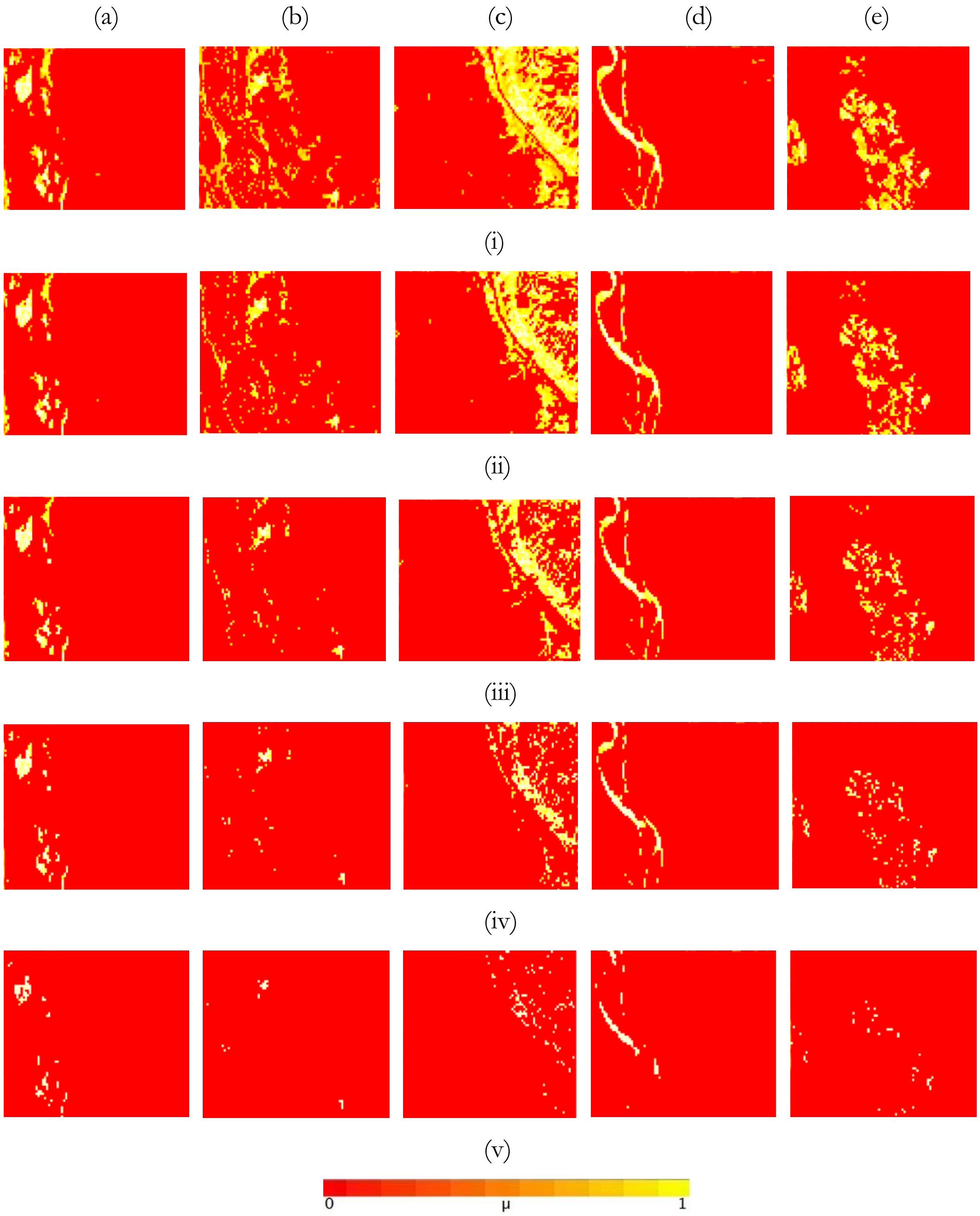


Figure 5.8 Generated fractional images for composite measure of Cosine and Euclidean norms at optimized m value of 2.5 and λ value of 0.5 of Landsat-8 data for (i) α-cut = 0.5 (ii) α-cut = 0.6 (iii) α-cut = 0.7 (iv) α-cut = 0.8 (v) α-cut = 0.9 for all the classes (a) Riverine-Sand (b) Fallow-Land (c) Forest (d) Water (e) Wheat.

The membership value of all the pixels in these fractional images (figures 5.5 to 5.8) was within the range of 0 and 1. From the results obtained, it was observed as there was an increment in the value of α-cut (threshold) the pixels below the threshold value were discarded. The accuracy assessment was performed on these fractional images using the fractional images of the Formosat-2 dataset as the reference images. Various fuzzy accuracy operators were used like FERM and SCM for calculating the accuracy. 500 random sample points were chosen as the test sites for performing the accuracy assessments. The fuzzy user's accuracy, producer's accuracy, overall accuracy and kappa coefficient were calculated for all the fuzzy accuracy assessing tools for α-cut values from 0.5 to 0.9 with an interval of 0.1 respectively (Table C-1 to Table C-5, Appendix C).

5.8. Untrained Classes

During the training stage of a classifier, some classes were ignored resulting in the untrained class. The untrained classes depict higher degree of membership for classes which are spectrally different and hence, resulting in a drop in the accuracy of the classification (Foody, 2000). In this work, for the developed FCM classifier, the mean values of wheat for Formosat-2 and Landsat-8 datasets were not considered for training. Figure 5.9 compares the overall accuracy results of both single measure as well as composite measure for both trained (tables 5.7 and 5.8) and untrained cases respectively.

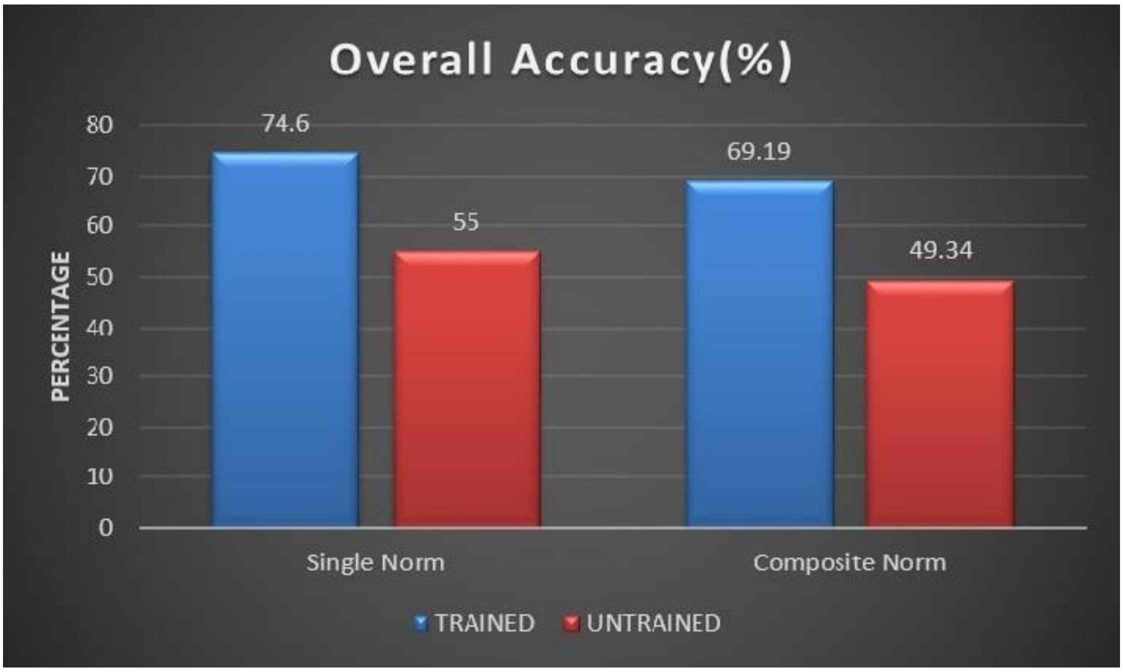


Figure 5.9 Details of accuracy assessment in trained and untrained case of classification results for Landsat-8 data with a single norm and composite norm respectively.

The results obtained after accuracy assessment in figure 5.9 showed that the overall accuracy for trained classes was more than the untrained classes ranging from 49% to 55% for single and composite measures respectively. These results showed that the removal of a class (wheat) for untrained class reduced the overall accuracy, however, the trend was constant like for trained dataset.

5.9. Discussion of Results

This section converses about the different results achieved by the classification algorithm used. In this study single and composite, similarity or dissimilarity measures were integrated into the FCM objective function to handle mixed pixels in a remote sensing data. The main objective of this research was to study the behaviour of similarity and dissimilarity measures while handling the mixed pixels.

The foremost focus of this research work was on the optimization of different parameters for the different similarity and dissimilarity measures used in the FCM classifier. Setting optimal values for various parameters was necessary for their proper performance. These values might get changed by varying the datasets. Optimal values of *m* were achieved on the basis of the working of the particular norm while handling the pure and mixed pixels (details in section 3.6.1). For FCM using single measures and FCM using composite measures the optimization of the fuzzifier (*m*) was computed within the range from 1.1 to 3.0. On the basis of the statistics from Table 5.1, Table 5.2 and Table 5.3, Cosine norm was found to be the best among all the measures with an optimized "*m*" value of 2.7 for Formosat-2 data and "*m*" value of 2.5 for Landsat-8 dataset. The composite measure was formed by using the two best norms (Euclidean norm and Cosine norm) obtained from section 5.1. The optimized value of "*m*" for the composite measure was at 2.5 for both Formosat-2 dataset and Landsat-8 dataset.

FCM classification has resulted in an overall accuracy (using SCM operator) of 75.24% and 69.80% for Landsat-8 image with Formosat-2 image as the referenced image while using single measure and composite measure respectively. The average user's accuracy using the same operator for Landsat-8 image for single and composite measures were 75.95% and 69.02% respectively and the average producer's accuracy were 73.56% and 69.69% respectively. From these results, it was evident that there was an overall decline in the accuracy while using composite measures instead of single measure. Izakian and Pedrycz (2014) showed the usage of weighted composite measures, however, the results were unsatisfactory. The performance of the composite measures depends on the type of single measures taken into account as a combination. Taking two best norms in a combination would give a deterioration in the results with respect to single best norms, due to weighting component (details in section 4.3).

The classification was tested on the various α-cut values starting from 0.5 up to 0.9 with an interval of 0.1 with the both single measure and composite measures with '*m*' value of 2.5 and 2.7. The range of α-cut was chosen from 0.5 to 0.9 because the FCM clustering algorithm is sensitive to noise and thus the FCM clustering outputs do not show any result with a membership value of 1 (Yang et al., 2009). When a suitable value of α-cut was chosen, noise and outliers would be outside the cluster cores as α-cut was a threshold value and any pixel with a membership value less than the threshold would not be considered. That's the reason we get either a very high accuracy or an accuracy of zero, as the α-cut value was moved towards 0.9 (Table C-1 to C-5, Appendix C). The concept of α-cut was used for identification of the noisy points. For example, if the α-cut value was set at 0.6 in the developed FCM algorithm, then the pixels having membership value greater than or equal to 0.6 would belong to the cluster core such that all the membership values of these pixels would become 1 for that concerned cluster and zero for the remaining clusters. However, if the pixels had membership value below 0.6, then the membership of the pixels remains same. As the pixels had a small membership value for the cluster, these pixels would be considered as noise and were removed from those clusters. Figure 5.5 to 5.8 show the clusters of the classes without the noisy pixels namely Riverine-Sand, Fallow-Land, Forest, Water and Wheat at different α-cut values of 0.5, 0.6, 0.7, 0.8 and 0.9 respectively for both single measure and composite measure incorporated in the FCM algorithm.

The classification was also verified on untrained classes where the FCM classifier was not trained about a class (in this study, wheat was untrained). There is an overall decrease in the accuracy in the untrained case in comparison with the trained case. The average overall accuracy for the single measure in case of trained and untrained was 74.6% and 55% respectively and for composite measures 69.19% and 49.34% respectively. However, figure 5.9 showed the overall trend of the accuracy was same for both trained and untrained with respect to single and composite measures respectively. This trend also explained that the incapability of FCM classifier to handle noise properly. On the removal of a class from the training samples, there was fall in the accuracy of nearly 20% in both single and composite measures. This showed that for the classification, the class (wheat) removed was noise for the other classes in the training samples, hence, there was a dip in the overall accuracy from trained case to untrained case.

By considering the overall accuracy of the classification (tables 5.7 and 5.8), it can be established that FCM with the single measure (Cosine) performs better in classification than FCM with composite measures. FCM with α-cut also reduces the noise in the classified images (figures 5.5 to 5.8), which helps in the handling of the mixed pixel problem in a better way. In this study, all the similarity and dissimilarity measures were evaluated for images of both medium and coarser resolutions. However, the behaviour of the measures may differ with different datasets and these similarity or dissimilarity measures may also be evaluated with a large number of various datasets to get a robust conclusion. Kumar et al. (2007) and Dutta (2009) stated in their works that FCM with Euclidean norm performs better than diagonal Mahalanobis and Mahalanobis norms but, in this study, it was found that Cosine norm outperforms the Euclidean norm (section 5.1.1).

All research work has its qualities and drawbacks, thus, this research work has a few merits and demerits with respect to the prevailing techniques and hence, there is a need for analysis. The SWOT (strength, weakness, opportunities and threats) analysis has been taken into consideration to analyse the advantages and limitations of this work, shown in table 6.1.

Table 6.1. The SWOT analysis of this research work

External Factors / **Internal Factors**	**Opportunities**	**Threats**
	1. Soft classifications techniques are more prevalent to handle mixed pixels than hard classification techniques. 2. High-end computers are available for fast computation.	1. In depth understanding of all similarity and dissimilarity along with FCM classifier is required to apply and develop as per requirement. 2. Accuracy assessment tools and generation of soft classified data is still not globally accepted.
Strength 1. Mapping of the real world can be done accurately. 2. Long and in-depth background of mathematics and fuzzy set theory.	1. The developed FCM classifier can be used to map a region with vague boundaries and also in crop mapping. 2. By using high-end computers, the developed FCM classifier along with α-cut can be used in place other fuzzy based classifiers.	1. People with expertise can be taken into account for mapping the real world using the developed FCM classifier along with the α-cut concept. 2. The importance of the developed FCM can be upgraded by widespread field visit along with better accuracy tools.
Weakness 1. More computational time was taken for calculating the membership values during the incorporation of α-cut.	1. Advancement of computer science can lessen the overall time for computing along with the help of experts.	1. Applicability of the developed FCM is restricted due to the unavailability of standard assessment tools.

6. CONCLUSIONS AND RECOMMENDATIONS

In this chapter, the conclusions attained through a detailed study of the existing and the developed method is presented. Section 6.1 presents the conclusions on the basis of the research objectives and questions. Section 6.2 shows the answers to the research questions and section 6.3 presents the recommendations for future work.

6.1. Conclusions

The occurrence of mixed pixels in the remote sensing images is largely due to the mismatch of resolution of the images with respect to class size. Due to the presence of mixed pixels, there are chances of having an inaccuracy in the results obtained after classification. Sub-pixel classification using fuzzy based classifiers such as FCM are a solution to this kind of uncertainty in data. Thus, to solve the problem of mixed pixels a fuzzy based approach using different similarity and dissimilarity measures have been studied in this research work. The main objective of this work was to the study behaviour of similarity and dissimilarity measures with FCM while handling the mixed pixels. The comparative study of different norms used, Cosine norm with '*m*' value of 2.7 attained the highest overall accuracy during the classification. It was also witnessed that optimization of parameters like weighted constant '*m*' and weighting component λ played a major role in the overall performance of the FCM based classifier.

Various accuracy assessment methods are available to measure the accuracy of the classification. Fuzzy Error Matrix (FERM), Sub-Pixel Confusion Uncertainty Matrix (SCM) along with different operators like MIN-PROD, MIN-MIN and MIN-LEAST has been recommended for the measurement of accuracy. The Landsat-8 image of the coarser resolution was assessed with a Formosat-2 image of finer resolution. The overall accuracy was low for the image of coarser resolution. This may be due to the lack of adjacency of the information in the coarser image with respect to the ground truth information.

Among the different similarity and dissimilarity measures, Cosine and Euclidean norms has given the best overall performance. These two norms were combined to form a composite norm with λ value equals to 0.5 and '*m*' value of 2.5. The composite measure has lower overall accuracy in comparison with the single norm (table 5.7 and 5.8). The performance of composite measure depends on the individual performance of single norms selected to combine and form the composite measure. If the best single norm with higher performance is pooled with a norm having lower performance, the resultant norm will also have lower performance.

The concept of α-cut has also been incorporated into the FCM function to minimize the effect of noise in the FCM classifier. This has been done due to the incapability of FCM to handle the noise properly (Yang et al., 2009). In this study, the effect of untrained class on the accuracy of the classification was also carried out by dropping wheat as an untrained class which resulted in decrease in the overall accuracy in comparison to the trained case (figure 5.9).

To conclude FCM classifier with Cosine norm performed better than the conventional Euclidean norm. But, due to the incapability of FCM classifier to handle noise properly, the classification accuracy was around 75%.

6.2. Answers to the research questions

- How can similarity or dissimilarity measures be incorporated into the FCM classifier approach?

Answer: The answer to the above question is given in section 4.2, 4.3 and 4.4.

- How single or composite similarity measures work with different α-cuts along with FCM objective function?

Answer: The different α-cut values namely 0.5, 0.6, 0.7, 0.8 and 0.9 were incorporated with the results obtained by classifying the Landsat-8 as well as Formosat-2 images using single or composite measures (details in sections 5.6 and 5.7). The α-cut values help in removing the noise or outliers in the classified outputs (figures 5.5 to 5.8).

- What will be the effect of using composite measure on FCM as compared to single distance measure?

Answer: Composite measure was formed by using the two best single measures with a certain weighting component. Thus, on increasing the weighting component, there is a change in the classification outputs. In this study, Euclidean and Cosine norms were combined to form a composite measure. However, it was obtained that the composite measure has a low accuracy with respect to the single norm when incorporated with FCM algorithm.

6.3. Recommendations

The developed FCM algorithm using various similarity and dissimilarity while incorporating the concept of α-cuts handles the problem of mixed pixels properly. However, there are a few limitations for this research work due to the low accuracy. The results while handling mixed pixels can be improved by considering the following ideas:

- ✓ The performance of developed α-cut FCM classifier with different similarity and dissimilarity measures can be tested for a large heterogeneous area with high complexity in the land cover.
- ✓ The performance of α-cut in FCM classifier can be proposed further for handling the noise in FCM, as it can form cluster cores with membership grade of 1, unlike the FCM clustering algorithm.
- ✓ The accuracy tools, as well as the mode of generating the soft reference data for accuracy assessment for soft classifiers, are still actively research in the field of digital image processing.
- ✓ Instead of supervised FCM approach, this study can be tested with unsupervised FCM classification.
- ✓ The accuracy assessment was done with datasets of coarser resolution and medium resolution, however if the accuracy assessment is obtained with two datasets of finer resolutions. Then, the robustness of the methodology used in this study can be evaluated properly.
- ✓ If the method to calculate the area using the outputs of this research work is formed, then this methodology can be used to calculate the concerned area of interest e.g., vegetation fields, forest area, etc.
- ✓ The results obtained for the single norm in this research also gives an opportunity to further study other fuzzy based classification techniques which use the similarity or dissimilarity measures like Possibilistic *c*-Means (PCM) classifier, etc.

REFERENCES

Abebe, A., Guinot, V., & Solomatine, D. (2000). Fuzzy alpha-cut vs. Monte Carlo techniques in assessing uncertainty in model parameters. *Proc. 4th International Conference on Hydroinformatics, Iowa City, USA*, (July), 1–8. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.112.8621&rep=rep1&type=pdf

Abello, J., Pardalos, P. M., & Resende, M. G. C. (2002). *Handbook of Massive Data Sets*. (J. Abello, P. M. Pardalos, & M. G. C. Resende, Eds.) (Vol. 4). Boston, MA: Springer US. http://doi.org/10.1007/978-1-4615-0005-6

Ablin, R., & Sulochana, C. H. (2013). A Survey of Hyperspectral Image Classification in Remote Sensing. *International Journal of Advanced Research in Computer and Communication Engineering*, *2*(8), 2986–3000.

Atkinson, P. M., & Tatnall, A. R. (1997). Introduction: neural networks in remote sensing. *International Journal of Remote Sensing*, *18*(4), 699–709. http://doi.org/10.1080/014311697218700

Ball, G. H., & Hall, D. J. (1965). ISODATA, a novel method of data analysis and pattern classification. *Analysis*, (AD699616), 1–79. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=AD0699616

Balu, R. (2015). *Design and development of automatic appendicitis detection system using sonographic image mining. Shodhganga : a reservoir of Indian theses @ INFLIBNET*. Bharathiar University. Retrieved from http://shodhganga.inflibnet.ac.in/handle/10603/33597

Bastin, L. (1997). Comparison of fuzzy c-means classification, linear mixture modelling and MLC probabilities as tools for unmixing coarse pixels. *International Journal of Remote Sensing*, *18*(17), 3629–3648. http://doi.org/10.1080/014311697216847

Bencherif, F., Mouss, L. H., & Benaicha, S. (2013). Fuzzy Alpha-cuts to capture customer requirements in improving product development. In *2013 11th International Symposium on Programming and Systems (ISPS)* (pp. 1–6). IEEE. http://doi.org/10.1109/ISPS.2013.6581493

Besag, J., & Besag, J. (1986). On the Statistical Analisys of Dirty Pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, *43*(3), 259–302.

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Boston, MA: Springer US. http://doi.org/10.1007/978-1-4757-0450-1

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy *c*-means clustering algorithm. *Computers & Geosciences*, *10*(2–3), 191–203. http://doi.org/10.1016/0098-3004(84)90020-7

Binaghi, E., Brivio, P. a., Ghezzi, P., & Rampini, A. (1999). A fuzzy set-based accuracy assessment of soft classification. *Pattern Recognition Letters*, *20*, 935–948. http://doi.org/10.1016/S0167-8655(99)00061-6

Bloom, S. (1981). Similarity Indices in Community Studies: Potential Pitfalls . *Marine Ecology Progress Series*, *5*, 125–128. http://doi.org/10.3354/meps005125

Boyd, D. S., Sanchez-Hernandez, C., & Foody, G. M. (2006). Mapping a specific class for priority habitats monitoring from satellite sensor data, (August 2015). http://doi.org/10.1080/01431160600554348

Bray, J. R., & Curtis, J. T. (1957). An Ordination of the upland forest community of southern Wisconsin.pdf. *Ecology Monographs*. http://doi.org/10.2307/1942268

Brodley, C. E., & Utgoff, P. E. (1995). Multivariate decision trees. *Machine Learning*, *19*(1), 45–77. http://doi.org/10.1007/BF00994660

Buckley, J. J., & Eslami, E. (2002). Fuzzy Sets. In *An Introduction to Fuzzy Logic and Fuzzy Sets* (pp. 21–53). Physica-Verlag HD. http://doi.org/10.1007/978-3-7908-1799-7_3

Byju, A. P. (2015). *Non-Linear Separation of classes using a Kernel based Fuzzy c -Means (KFCM) Approach*. Enschede.

Cannon, R., Dave, J., Bezdek, J., & Trivedi, M. (1986). Segmentation of a Thematic Mapper Image Using the Fuzzy c-Means Clusterng Algorthm. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-24*(3), 400–408. http://doi.org/10.1109/TGRS.1986.289598

Charulatha, B. S., Rodrigues, P., Chitralekha, T., & Rajaraman, A. (2013). A Comparative study of different distance metrics that can be used in Fuzzy Clustering Algorithms. *International Journal of Emerging Trends and Technology in Computer Science (IJETTICS)*.

Chawla, S. (2010). *Possibilistic c - means - spatial contextual information based sub - pixel classification approach for multi - spectral data*. Enschede. Retrieved from http://www.itc.nl/library/papers_2010/msc/gfm/chawla.pdf

Cheplygina, V., Tax, D. M. J., & Loog, M. (2012). Class-dependent dissimilarity measures for multiple instance learning. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7626 LNCS*, 602–610. http://doi.org/10.1007/978-3-642-34166-3_66

Choodarathnakara, A. L., Kumar, T. A., Koliwad, S., & Patil, C. G. (2012). Soft Classification Techniques for RS Data 1. *International Journal of Computer Science Engineering and Technology*, *2*(11), 1468–1471. Retrieved from http://www.ijcset.net/docs/Volumes/volume2issue11/ijcset2012021101.pdf

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35–46. http://doi.org/10.1016/0034-4257(91)90048-B

Corporation, S. I. (2013). Formosat-2 Satellite Sensor. Retrieved February 11, 2016, from http://satimagingcorp.tw.cn/satellite-sensors/formosat-2.html

Corporation, S. I. (2015). LANDSAT 8 Satellite Sensor. Retrieved February 11, 2016, from http://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/landsat-8/

Das, S. (2013). Pattern Recognition using the Fuzzy c-means Technique. *International Journal of Energy, Information and Communications*, *4*(1), 1–14.

De Fries, R. S., Hansen, M., Townshend, J. R. G., & Sohlberg, R. (1998). Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, *19*(16), 3141–3168.

http://doi.org/10.1080/014311698214235

Deer, P., Eklund, P., & Norman. (1996). A Mahalanobis Distance Fuzzy Classifer. In Narasimhan & Jam (Eds.), *4th Australian and New Zealand Conference on Intelligent Information Systems (ANZIIS '96)* (pp. 220–224).

Dik, A., Bouroumi, A., & Ettouhami, A. (2014). Weighted distances for fuzzy clustering. *Applied Mathematical Sciences*, *8*(4), 147–156. http://doi.org/10.12988/ams.2014.311651

Dilo, A. (2006). *Representation of and reasoning with vagueness in spatial information: A system for handling vague objects. Geoinformatics* (Vol. PhD).

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern Classification. *New York: John Wiley, Section*. http://doi.org/10.1007/BF01237942

Dunn, J. C. (1974). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, *3*(3), 32–57. http://doi.org/10.1080/01969727308546046

Dunyak, J., & Wunsch, D. (1997). Training fuzzy number neural networks with alpha-cut refinements. In *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation* (Vol. 1, pp. 189–194). IEEE. http://doi.org/10.1109/ICSMC.1997.625747

Dutta, A. (2009). *Data Incorporating Spatial Contextual Information by using Markov Random Field Fuzzy c -Means Classification of Multispectral Data Incorporating Spatial Contextual Information by using Markov Random Field. University of Twente Faculty of Geo-Information and Earth Observation (ITC).* Enschede. Retrieved from http://www.itc.nl/library/papers_2009/msc/gfm/dutta.pdf

Dwivedi, R., Kumar, A., & K. Ghosh, S. (2012). Study of Fuzzy based Classifier Parameter Across Spatial Resolution. *International Journal of Computer Applications*, *50*(11), 17–24. http://doi.org/10.5120/7814-1074

Emran, S. M., & Ye, N. (2001). Robustness of Canberra Metric in Computer Intrusion Detection. In *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security United States Military Academy* (pp. 80–84).

Fisher, P. (1997). The pixel: A snare and a delusion. *International Journal of Remote Sensing*, *18*(3), 679–685. http://doi.org/10.1080/014311697219015

Foody, G. M. (1996). Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data. *International Journal of Remote Sensing*, *17*(7), 1317–1340. http://doi.org/10.1080/01431169608948706

Foody, G. M. (2000). Estimation of sub-pixel land cover composition in the presence of untrained classes. *Computers and Geosciences*, *26*(4), 469–478. http://doi.org/10.1016/S0098-3004(99)00125-9

Foody, G. M. (2002). The Role of Soft Classification Techniques in the Refinement of Estimates of Ground Control Point Location. *Photogrammetric Engineering & Remote Sensing*, *68*(9), 897–903.

Foody, G. M., & Arora, M. K. (1997). An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing*, *18*(4), 799–810. http://doi.org/10.1080/014311697218764

Foody, G. M., Mathur, A., Sanchez-Hernandez, C., & Boyd, D. S. (2006). Training set size requirements for the classification of a specific class. *Remote Sensing of Environment*, *104*(1), 1–14. http://doi.org/10.1016/j.rse.2006.03.004

Franklin, S. E., Peddle, D. R., Dechka, J. a., & Stenhouse, G. B. (2002). Evidential reasoning with Landsat TM, DEM and GIS data for landcover classification in support of grizzly bear habitat mapping. *International Journal of Remote Sensing*, *23*(August 2014), 4633–4652. http://doi.org/10.1080/01431160110113971

Friedl, M. a., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, *61*(3), 399–409. http://doi.org/10.1016/S0034-4257(97)00049-7

Gallego, F. J. (2004). Remote sensing and land cover area estimation. *International Journal of Remote Sensing*, *25*(15), 3019–3047. http://doi.org/10.1080/01431160310001619607

Gehler, P. V., & Scholkopf, B. (2009). An introduction to kernel learning algorithms. In G. Camps-Valls & L. Bruzzone (Eds.), *Kernel Methods for Remote Sensing Data Analysis* (pp. 25–45). Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/9780470748992

Gong, P., & Howarth, P. J. (1992). Frequency-based contextual classification and gray-level vector reduction for land-use identification. *Photogrammetric Engineering and Remote Sensing*, *58*(4), 423–437. Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=5155171

Gopinath, R. A. (1998). Maximum likelihood modeling with Gaussian distributions for classification. *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, *2*(914), 661–664. http://doi.org/10.1109/ICASSP.1998.675351

Goshtasby, A. A. (2012). Chapter 2 : Similarity and Dissimilarity Measures. In S. Singh & S. B. Kang (Eds.), *Image Registration Principles, Tools and methods* (pp. 7–66). Springer-Verlag London Limited. http://doi.org/10.1007/978-1-4471-2458-0

Hansen, M. C., Defries, R., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*(6), 1331–1364. http://doi.org/10.1080/014311600210209

Hansen, M. C., Dubayah, R., & DeFries, R. S. (1996). Classification trees: An alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, *17*(5), 1075–1081. http://doi.org/10.1080/01431169608949069

Harikumar, A. (2014). *Effect of discontinuity adaptive MRF models with Noise classifier*. Enschede.

Hasnat, A., Halder, S., Bhattacharjee, D., Nasipuri, M., & Basu, D. K. (2013). Comparative Study of Distance Metrics for Finding Skin Color Similarity of Two Color Facial Images. In *Computer Science & Information Technology (CS & IT)* (pp. 99–108). Academy & Industry Research Collaboration Center (AIRCC). http://doi.org/10.5121/csit.2013.3210

Heijden, F. van der, Duin, R. P. W., de Ridder, D., & Tax, D. M. J. (2004). *Classification, Parameter Estimation and State Estimation : An Engineering Approach Using MATLAB. John Willey & Sons.* Chichester, UK: John Wiley & Sons, Ltd. http://doi.org/10.1002/0470090154

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*(4), 725–749. http://doi.org/10.1080/01431160110040323

Ibrahim, M. a., Arora, M. K., & Ghosh, S. K. (2005). Estimating and accommodating uncertainty through the soft classification of remote sensing data. *International Journal of Remote Sensing*, *26*(14), 2995–3007. http://doi.org/10.1080/01431160500057806

Izakian, H., & Pedrycz, W. (2014). Agreement-based fuzzy C-means for clustering data with blocks of features. *Neurocomputing*, *127*, 266–280. http://doi.org/10.1016/j.neucom.2013.08.006

Jafar, O. A. M., & Sivakumar, R. (2013). A Comparative Study of Hard and Fuzzy Data Clustering Algorithms with Cluster Validity Indices. In *International Conference on Emerging Research in Computing, Information, Communication and Applications, ERCICA 2013* (pp. 775–782). Elsevier Publications.

Ji, M. (2003). Using fuzzy sets to improve cluster labelling in unsupervised classification. *International Journal of Remote Sensing*, *24*(4), 657–671. http://doi.org/10.1080/01431160210146226

Johnson, R. A., & Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis. Pearson Education International*. http://doi.org/10.1198/tech.2005.s319

Jr, R. G. P., & Cheuk, M. L. (2006). A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *International Journal of Geographical Information Science*, *20*(1), 1–30. http://doi.org/10.1080/13658810500391024

Jurman, G., Riccadonna, S., Visintainer, R., & Furlanello, C. (2009). Canberra distance on ranked lists. In T. Lu & C. Boutilier (Eds.), *Advances in Ranking* (pp. 22–27).

Kainz, W. (2007). Fuzzy Logic and GIS. Vienna: University of Vienna.

Kloditz, C., Van Boxtel, a, Carfagna, E., & Van Deursen, W. (1998). Estimating the accuracy of coarse scale classification using high scale information. *Photogrammetric Engineering and Remote Sensing*, *64*(2), 127–133.

Kontoes, C. C., & Rokos, D. (1996). The integration of spatial context information in an experimental knowledge-based system and the supervised relaxation algorithm—two successful approaches to improving SPOT-XS classification. *International Journal of Remote Sensing*, *17*(16), 3093–3106. http://doi.org/10.1080/01431169608949132

Kouser, K., & Sunita. (2013). A comparative study of K Means Algorithm by Different Distance Measures. *International Journal of Innovative Research in Computer and Communication Engineering*, *1*(9), 2443–2447.

Kreinovich, V. (2013). Membership Functions or alpha-Cuts ? Algorithmic (Constructivist) Analysis Justifies an Interval Approach. *Applied Mathematical Science*, *7*(5), 217–288. Retrieved from http://digitalcommons.utep.edu/cgi/viewcontent.cgi?article=1711&context=cs_techrep

Kumar, A. (2007). Investigation in Sub-pixel classification approaches for Land Use and Land Cover Mapping. IIT Roorkee. Retrieved from http://shodhbhagirathi.iitr.ac.in:8081/jspui/handle/123456789/1601

Kumar, A., Ghosh, S. K., & Dadhwal, V. K. (2006a). A comparison of the performance of fuzzy algorithm versus statistical algorithm based sub-pixel classifier for remote sensing data. In *Proceedings of mid-term symposium ISPRS* (pp. 8–11). Retrieved from http://www.isprs.org/proceedings/XXXVI/part7/PDF/153.pdf

Kumar, A., Ghosh, S. K., & Dadhwal, V. K. (2006b). Sub-pixel land cover mapping: SMIC system. In *ISPRS International Symposium on Geospatial Databases for Sustainable Development, Goa, India*. Retrieved from http://www.isprs.org/proceedings/xxxvi/part4/wg-iv-2-6.pdf

Kumar, A., Ghosh, S. K., & Dadhwal, V. K. (2007). Full fuzzy land cover mapping using remote sensing data based on fuzzy c-means and density estimation. *Canadian Journal of Remote Sensing*, *33*(2), 81–87.

Lance, G. N., & Williams, W. T. (1966). Computer Programs for Hierarchical Polythetic Classification (``Similarity Analyses''). *The Computer Journal*, *9*(1), 60–64. http://doi.org/10.1093/comjnl/9.1.60

Lance, G. N., & Williams, W. T. (1967). Mixed-Data Classificatory Programs I – Agglomerative Systems. *Australian Computer Journal*, *1*(1), 15–20.

Landgrebe, D. A. (2003). *Signal theory methods in multispectral remote sensing*. *Wiley*. http://doi.org/10.1002/0471723800

Le Maitre, R. . (1982). Cluster Analysis and Classification. In *Numerical Petrology : Statistical Interpretation of Geochemical Data* (N/A, pp. 163–177). Elsevier Science. http://doi.org/10.1016/B978-0-444-42098-5.50014-6

Lee, C., & Lee, G. K. (2015). An Alpha Cut (α-cut) Filter-Based Proxy Caching Control for Wireless Coverage Streaming Services. *Wireless Personal Communications*. http://doi.org/10.1007/s11277-015-3032-8

Lee, S. H., Pedrycz, W., & Sohn, G. (2009). Design of similarity and dissimilarity measures for fuzzy sets on the basis of distance measure. *International Journal of Fuzzy Systems*, *11*(2), 67–72.

Lee, Y., & Horng, S.-J. (1996). The Chessboard Distance Transform and the Medial Axis Transform are Interchangeable. In *Parallel Processing Symposium, 1996., Proceedings of IPPS '96, The 10th International* (pp. 424–428). Honolulu, HI: IEEE. http://doi.org/10.1109/IPPS.1996.508090

Letts, P. A. (1978). Unsupervised Classification in The Aries Image Analysis System. In *Proceedings of the 5th Canadian Symposium on Remote Sensing* (pp. 61–71).

Li, W., Guo, Q., & Elkan, C. (2011). A positive and unlabeled learning algorithm for one-class classification of remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(2), 717–725. http://doi.org/10.1109/TGRS.2010.2058578

Lillesand, T. M., & Kiefer, R. W. (1994). Remote sensing and image interpretation -- 3rd ed. *John Wiley and Sons, Inc., New York*, 524–647.

Lippmann, R. P. (1987). An Introduction ' to Computing with Neural Nets. *Ieee Assp Magazine*, (April).

Lo, C. P., & Choi, J. (2004). A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *International Journal of Remote Sensing*, *25*(14), 2687–2700. http://doi.org/10.1080/01431160310001618428

Lopatka, J., & Pedzisz, M. (2000). Automatic modulation classification using statistical moments and afuzzy classifier. *WCC 2000 - ICSP 2000. 2000 5th International Conference on Signal Processing Proceedings. 16th World Computer Congress 2000*, *3*. http://doi.org/10.1109/ICOSP.2000.893385

Mahalanobis, P. C. (1936). Mahalanobis distance. *Proceedings National Institute of Science of India*, *49*, 234–256.

Martin, L., & Howarth, P. (1989). Change-detection accuracy assessment using SPOT multispectral imagery of the rural-urban fringe☆. *Remote Sensing of Environment*, *30*, 55–66. http://doi.org/10.1016/0034-4257(89)90047-3

Mather, P., & Tso, B. (2009). *Classification methods for remotely sensed data* (Second). CRC Press - Taylor & Francis Group. Retrieved from https://www.crcpress.com/product/isbn/9781420090727

Moore, A. (2002). The case for approximate Distance Transforms. *Information Research*. http://doi.org/10.1.1.98.5245

Mukherjee, S., Osuna, E., & Girosi, F. (1997). Nonlinear prediction of chaotic time series using support vector machines. *IEEE Workshop on Neural Networks for Signal Processing VII*, 511. http://doi.org/10.1109/NNSP.1997.622433

Okeke, F., & Karnieli, a. (2006). Methods for fuzzy classification and accuracy assessment of historical aerial photographs for vegetation change analyses. Part I: Algorithm development. *International Journal of Remote Sensing*, *27*(1), 153–176. http://doi.org/10.1080/01431160500166540

Osuna, E., & Freund, R. (1997). An improved training algorithm for support vector machines. *Neural Networks for Signal Processing VII --- Proceedings of the 1997 {IEEE} Workshop*, 276–285. http://doi.org/10.1109/NNSP.1997.622408

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, *12*, S27–S31. http://doi.org/10.1016/j.jag.2009.11.002

Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, *26*(5), 1007–1011. http://doi.org/10.1080/01431160512331314083

Pal, N. R., & Bezdek, J. C. (1995). On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy Systems*, *3*(3), 370–379. http://doi.org/10.1109/91.413225

Paola, J. D., & Schowengerdt, R. A. (1995). A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of Remote Sensing*, *16*(16), 3033–3058. http://doi.org/10.1080/01431169508954607

Patra, S., & Bruzzone, L. (2011). A fast cluster-assumption based active-learning technique for classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(5), 1617–1626. http://doi.org/10.1109/TGRS.2010.2083673

Pekalska, E., Harol, a, Duin, R. P. W., Spillmann, B., & Bunke, H. (2006). Non-Euclidean or Non-metric Measures Can Be Informative. *Structural, Syntactic, and Statistical Pattern Recognition*, 871–880. http://doi.org/10.1007/11815921_96

Ponce-Cruz, P., & Ramirez-Figueroa, F. D. (2010). Fuzzy Logic. In *Intelligent Control Systems with*

LabVIEW™ (pp. 9–46). London: Springer London. http://doi.org/10.1007/978-1-84882-684-7_2

Reznik, L., & Stoica, A. (1994). Mapping alpha-cut borders: classification and PID realization. In *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference* (Vol. 3, pp. 1604–1607). IEEE. http://doi.org/10.1109/FUZZY.1994.343935

Richards, J. A., & Jia, X. (2006). *Remote Sensing Digital Image Analysis. New York*. Berlin/Heidelberg: Springer-Verlag. http://doi.org/10.1007/3-540-29711-1

Roosta, A., Setoodeh, P., & Jahanmiri, A. (2012). Artificial Neural Network Modeling of Surface Tension for Pure Organic Compounds. *Industrial & Engineering Chemistry Research*, *51*(1), 561–566. http://doi.org/10.1021/ie2017459

Safavian, R. S., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transaction Is on Systems, Man, and Cybernetics*, *21*(3), 660–674. http://doi.org/10.1109/21.97458

San Miguel-Ayanz, J., & Biging, G. S. (1997). Comparison of single-stage and multi-stage classification approaches for cover type mapping with TM and SPOT data. *Remote Sensing of Environment*, *59*(1), 92–104. http://doi.org/10.1016/s0034-4257(96)00109-5

Sari, S., Roslan, H., & Shimamura, T. (2012). Noise Estimation by Utilizing Mean Deviation of Smooth Region in Noisy Image. In *2012 Fourth International Conference on Computational Intelligence, Modelling and Simulation* (pp. 232–236). IEEE. http://doi.org/10.1109/CIMSim.2012.89

Sarwar, B. M. (2001). Correlation-based Similarity. Retrieved February 10, 2016, from http://www10.org/cdrom/papers/519/node13.html

Schulz, J. (2007). BRAY-CURTIS DISSIMILARITY. Retrieved February 10, 2016, from http://www.code10.info/index.php%3Foption%3Dcom_content%26view%3Darticle%26id%3D46:articlebray-curtis-dissim%26catid%3D38:cat_coding_algorithms_data-similarity%26Itemid%3D57

Scollar, I., Weidner, B., & Huang, T. S. (1984). Image enhancement using the median and the interquartile distance. *Computer Vision, Graphics, and Image Processing*, *25*(2), 236–251. http://doi.org/10.1016/0734-189X(84)90106-3

Shalan, M. a., Arora, M. K., & Ghosh, S. K. (2003). An evaluation of fuzzy classifications from IRS 1C LISS III imagery: A case study. *International Journal of Remote Sensing*, *24*(15), 3179–3186. http://doi.org/10.1080/0143116031000094791

Shankar, B. U., Meher, S. K., Ghosh, A., & Bruzzone, L. (2006). Remote Sensing Image Classification : A Neuro-fuzzy MCS Approach. *Imaging*, 128–139.

Silván-Cárdenas, J. L., & Wang, L. (2008). Sub-pixel confusion-uncertainty matrix for assessing soft classifications. *Remote Sensing of Environment*, *112*(3), 1081–1095. http://doi.org/10.1016/j.rse.2007.07.017

Singhal, A. (2001). Modern Information Retrieval: A Brief Overview. *Data Engineering Bulletin*, *24*(4), 35–43. http://doi.org/10.1.1.117.7676

Sinwar, D., & Kaushik, R. (2014). Study of Euclidean and Manhattan Distance Metrics using Simple K-Means Clustering. *International Journal for Research in Applied Science and Engineering Technology*

(IJRASET), *2*(5), 270–274.

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, *62*(1), 77–89. http://doi.org/10.1016/S0034-4257(97)00083-7

Stein, A. (2010). Fuzzy methods in image mining. *Studies in Fuzziness and Soft Computing*, *256*, 243–268. http://doi.org/10.1007/978-3-642-14755-5_10

Stuckens, J., Coppin, P. R., & Bauer, M. E. (2000). Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment*, *71*(3), 282–296. http://doi.org/10.1016/S0034-4257(99)00083-8

Swain, P. H., & Davis, S. M. (1979). *Remote Sensing: The Quantitative Approach*. *McGraw-Hill International Book Company*. New York, USA.

Takahashi, A., Endo, Y., & Miyamoto, S. (2011). Fuzzy c-means clustering for data with tolerance using cosine correlation. In *2011 IEEE International Conference on Granular Computing* (Vol. 11, pp. 619–624). IEEE. http://doi.org/10.1109/GRC.2011.6122668

Tan, P.-N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining. *Journal of School Psychology*, *19*, 51–56. http://doi.org/10.1016/0022-4405(81)90007-8

Theodoridis, S., & Koutroumbas, K. (2009). Clustering: Basic Concepts. In *Pattern Recognition* (pp. 595–625). Elsevier. http://doi.org/10.1016/B978-1-59749-272-0.50013-X

Tyagi, U., & Kumar, A. (2015). Similarity and Dissimilarity Measures with Fuzzy Classifiers : SMIC Tool. In *9th International Conference on ASEICT*. New Delhi.

Upadhyay, P., Ghosh, S. K., & Kumar, A. (2013). Ecological Informatics Moist deciduous forest identi fi cation using temporal MODIS data — A comparative study using fuzzy based classi fi ers. *Ecological Informatics*, *18*, 117–130. http://doi.org/10.1016/j.ecoinf.2013.07.002

Upadhyay, P., Ghosh, S. K., & Kumar, A. (2014). A Brief Review of Fuzzy Soft Classification and Assessment of Accuracy Methods for Identification of Single Land Cover, *2*(Mlc), 1–13.

Vassiliadis, S., Hakkennes, E. A., Wong, J. S. S. M., & Pechanek, G. G. (1998). The Sum-Absolute-Difference Motion Estimation Accelerator.

Wang, F. (1990). Fuzzy Supervised Classification of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(2), 194–201. http://doi.org/10.1109/36.46698

Wang, L., Zhang, Y., & Feng, J. (2005). On the Euclidean Distances of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(8), 1334–1339.

Wan-zhi, Z., & Zeng-cai, W. (2013). Rural Road Detection of Color Image in Complicated Environment. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, *6*(6), 161–168. Retrieved from http://dx.doi.org/10.14257/ijsip.2013.6.6.15

Wolfram. (2010). NormalizedSquaredEuclideanDistance. Retrieved February 10, 2016, from https://reference.wolfram.com/language/ref/NormalizedSquaredEuclideanDistance.html

Wong, K. W., Fung, C. C., Eren, H., & Gedeon, T. (2001). Determination of parameter d50c of hydrocyclones using improved multidimensional alpha-cut based fuzzy interpolation technique. In

IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188) (Vol. 3, pp. 1602–1606). IEEE. http://doi.org/10.1109/IMTC.2001.929474

Xexéo, G. (n.d.). Fuzzy logic. Rio de Janeiro: Computing Science Department and Systems and Computing Engineering Program, Federal University of Rio de Janeiro, Brazil.

Xuecheng, L. (1994). Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems*, *62*(3), 377. http://doi.org/10.1016/0165-0114(94)90124-4

Yang, M., Wu, K., Hsieh, J., & Yu, J. (2009). Alpha-Cut Implemented Fuzzy Clustering Algorithms and Switching Regressions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *38*(3), 588–603. http://doi.org/10.1109/TSMCB.2007.915537

Ye, J. (2011). Multicriteria decision-making method based on a cosine similarity measure between trapezoidal fuzzy numbers. *International Journal of Engineering, Science and Technology*, *3*(1), 272–278. http://doi.org/10.4314/ijest.v3i1.67654

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. http://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, J., & Foody, G. M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, *19*(14), 2721–2738. http://doi.org/10.1080/014311698214479

Zhang, M., Therneau, T., Mckenzie, M. A., Li, P., & Yang, P. (2008). A Fuzzy C-Means Algorithm Using a Correlation Metrics and Gene Ontology. *Nature Medicine*, 6–9.

Zhu, A. (1997). Measuring uncertainty in class assignment for natural resource maps under fuzzy logic. *Photogrammetric Engineering & Remote Sensing*, *63*(10), 1195–1202.

Zimmermann, H.-J. (2001). *Fuzzy Set Theory—and Its Applications. Kluwer, Boston, 2nd ed., 1993*. Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-010-0646-0

Zwick, R., Carlstein, E., & Budescu, D. V. (1987). Measures of similarity among fuzzy concepts: A comparative analysis. *International Journal of Approximate Reasoning*, *1*(2), 221–242. http://doi.org/10.1016/0888-613X(87)90015-6

APPENDIX A

Details of the simulated image for Formosat-2 dataset is given in Fig. A-1:

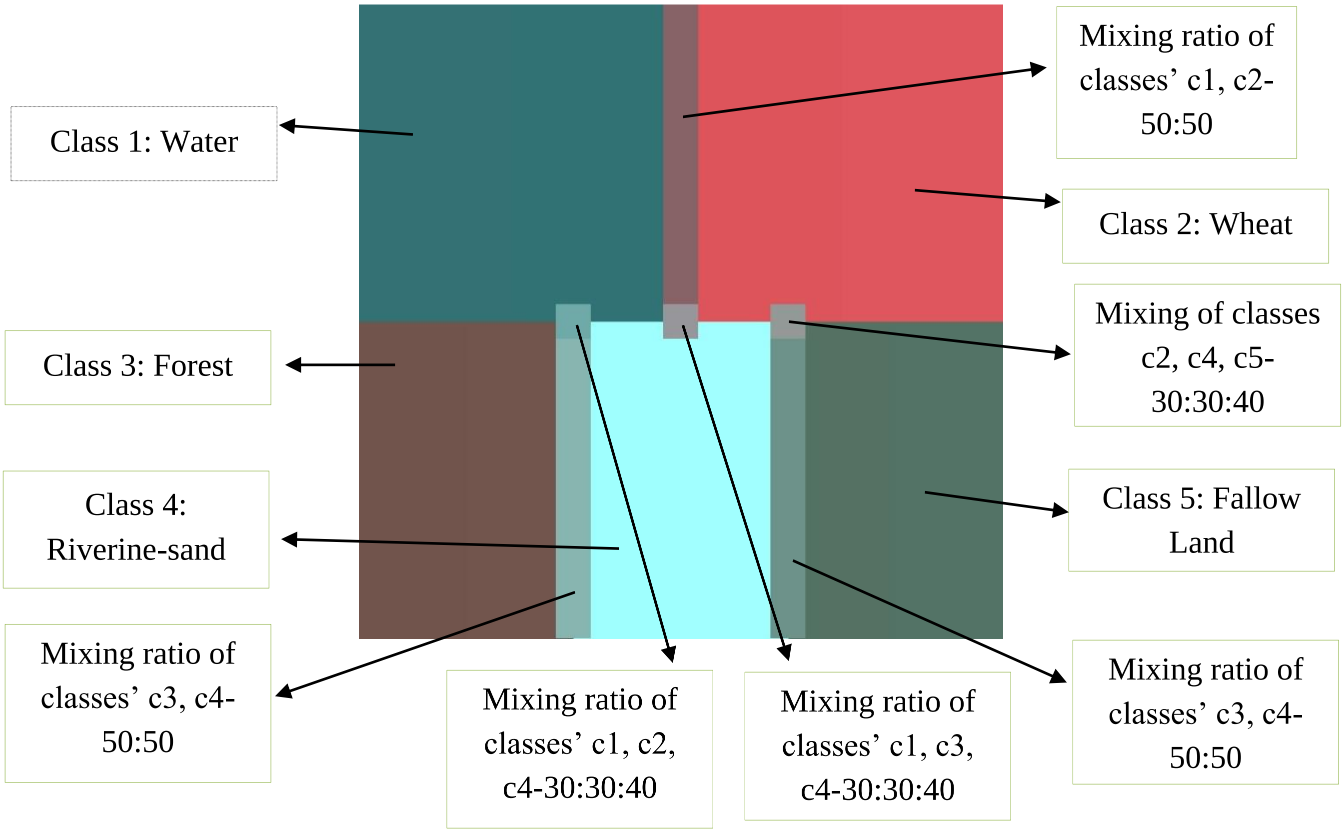


Figure A-1: Simulated Image Details of Formosat-2 dataset

Details of the simulated image for Landsat-8 dataset is given in Fig. A-2:

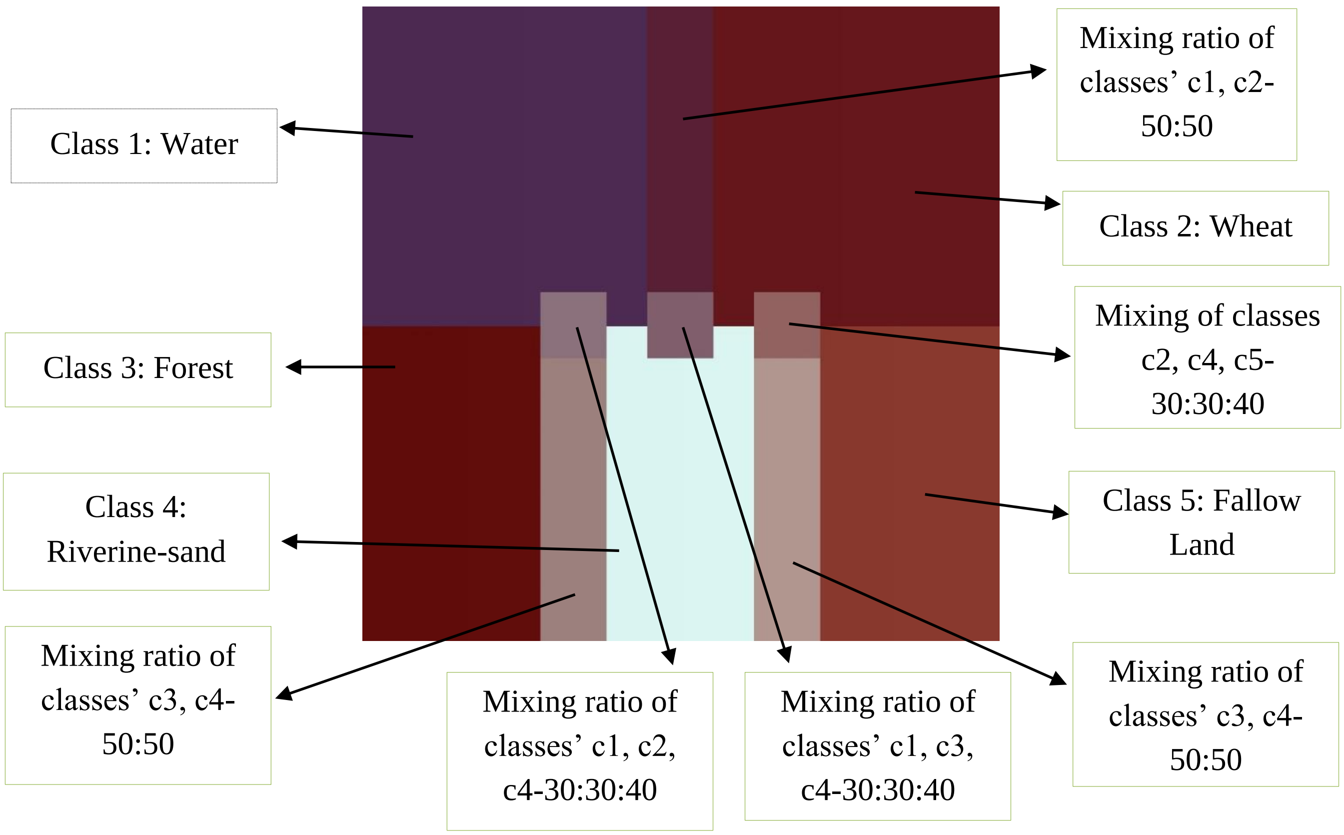


Figure A-2: Simulated Image Details of Landsat-8 dataset

The flowchart for optimization of the weight constant '*m*' is shown in figure A-3

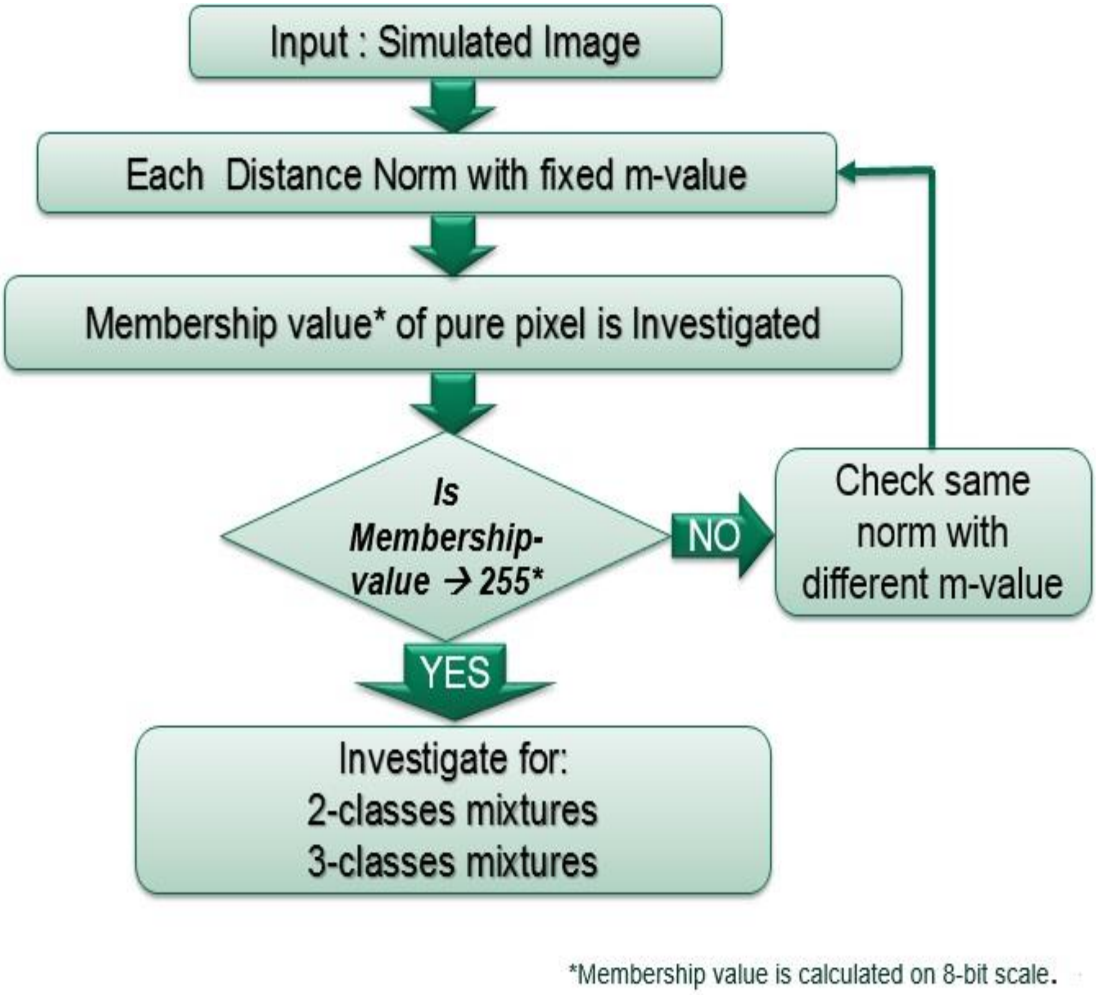


Figure A-3 Flow Chart for optimizing the parameter

The results of FCM algorithm with simulated image of Formosat-2 data for Cosine norm with *m* equals to 2.7 has been shown in figure A-4.

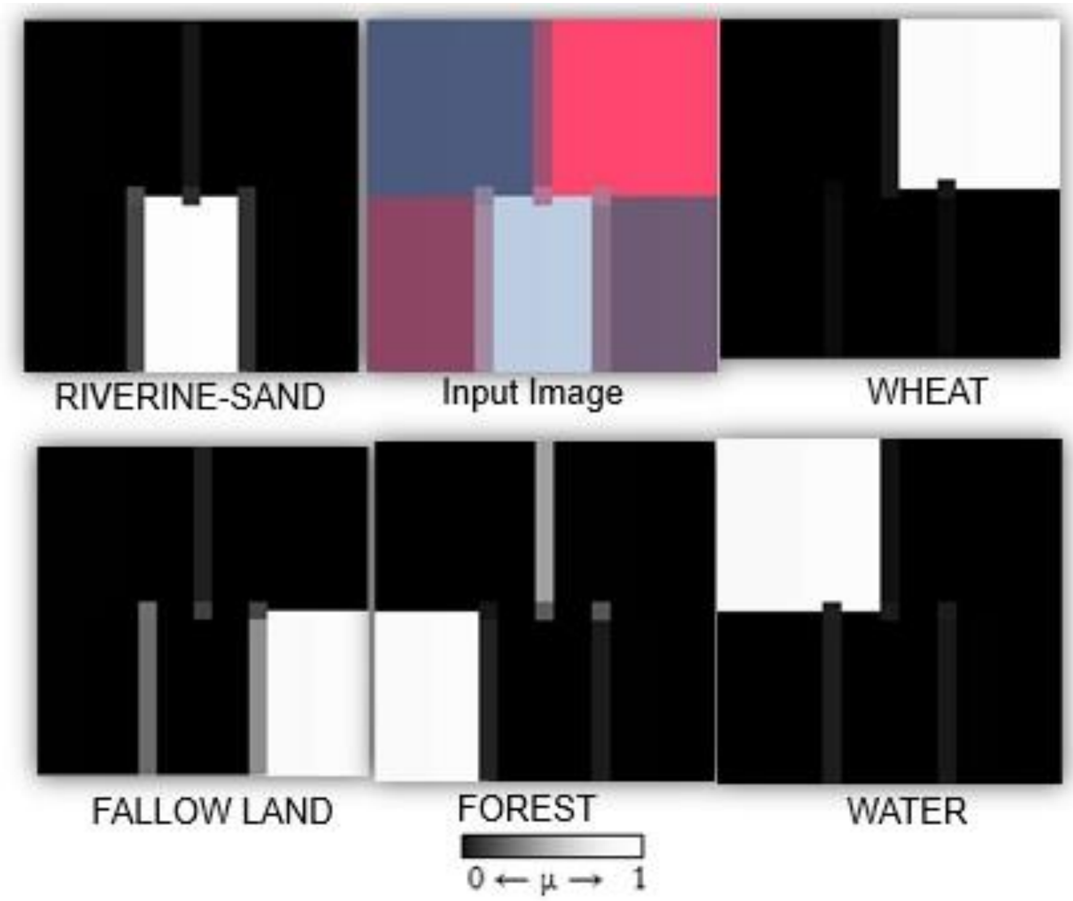


Figure. A-4. The result of simulated image using Cosine norm with *m* equals to 2.7

Figure A-5. The misclassified outputs in red circles while using composite measure with Euclidean and Cosine norms with *m* equals to 2.5 and λ equals to 0.5.

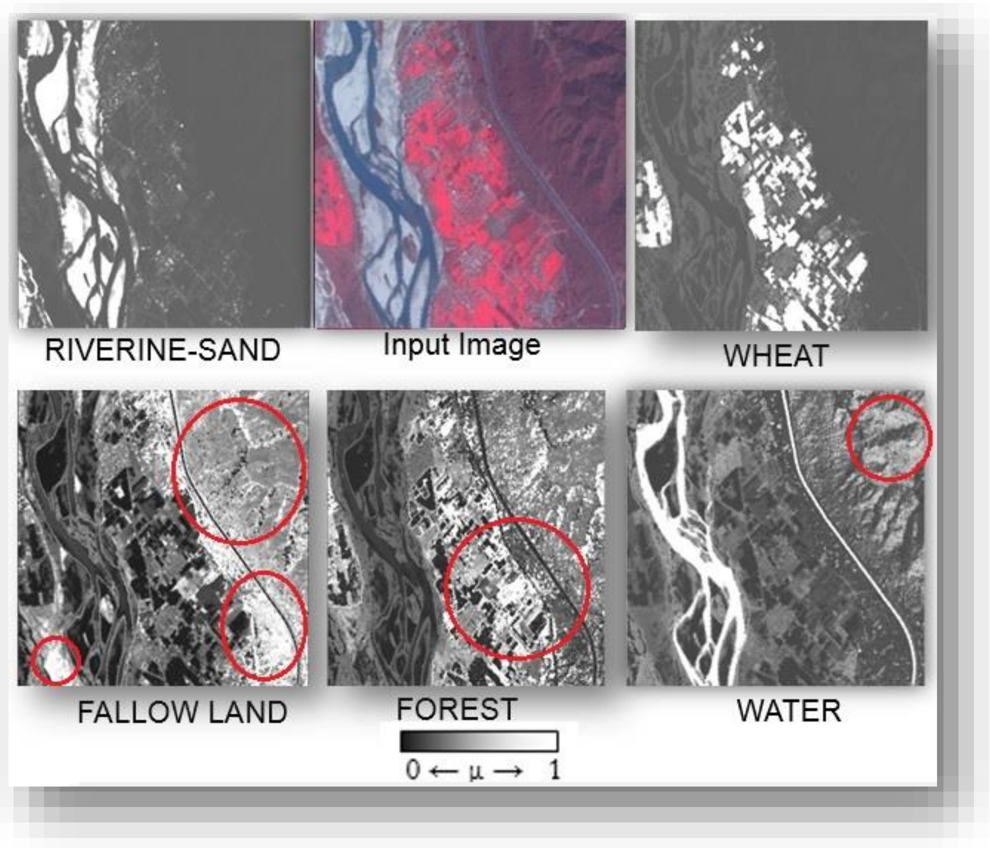


Figure A-6. The overall accuracy assessment for FCM using single measure in Landsat-8 data

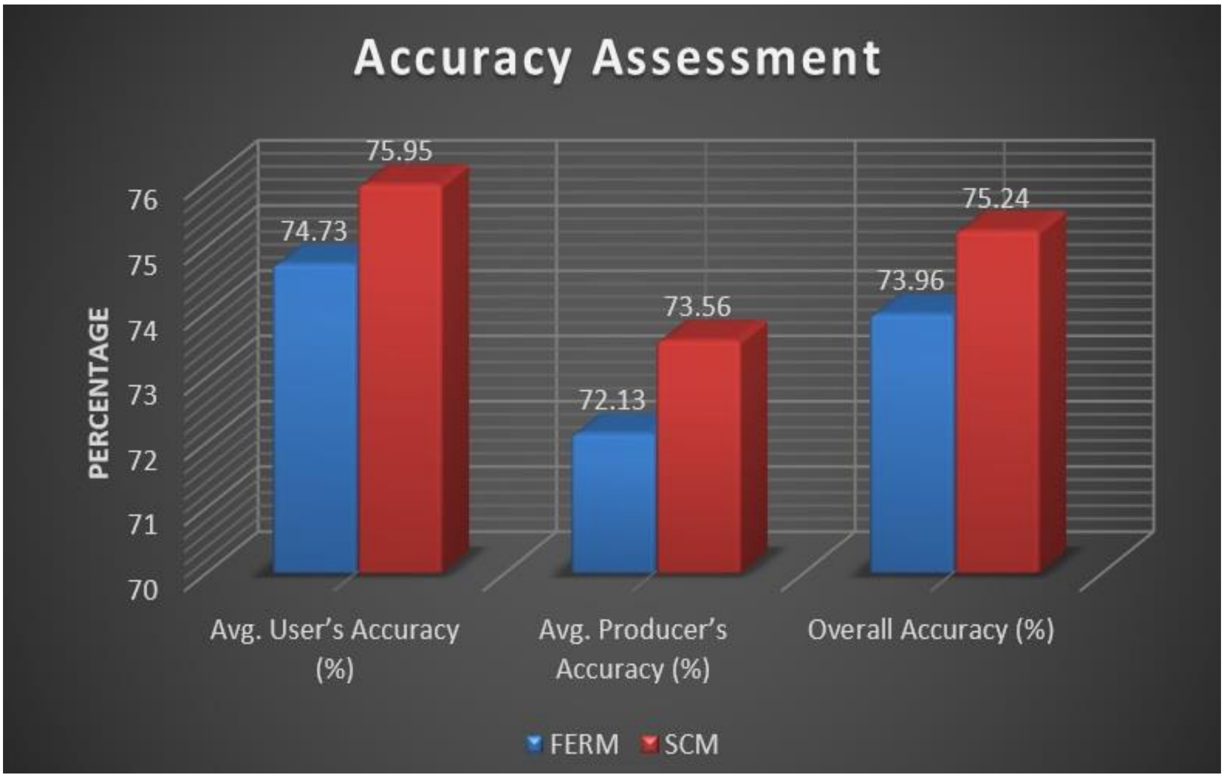
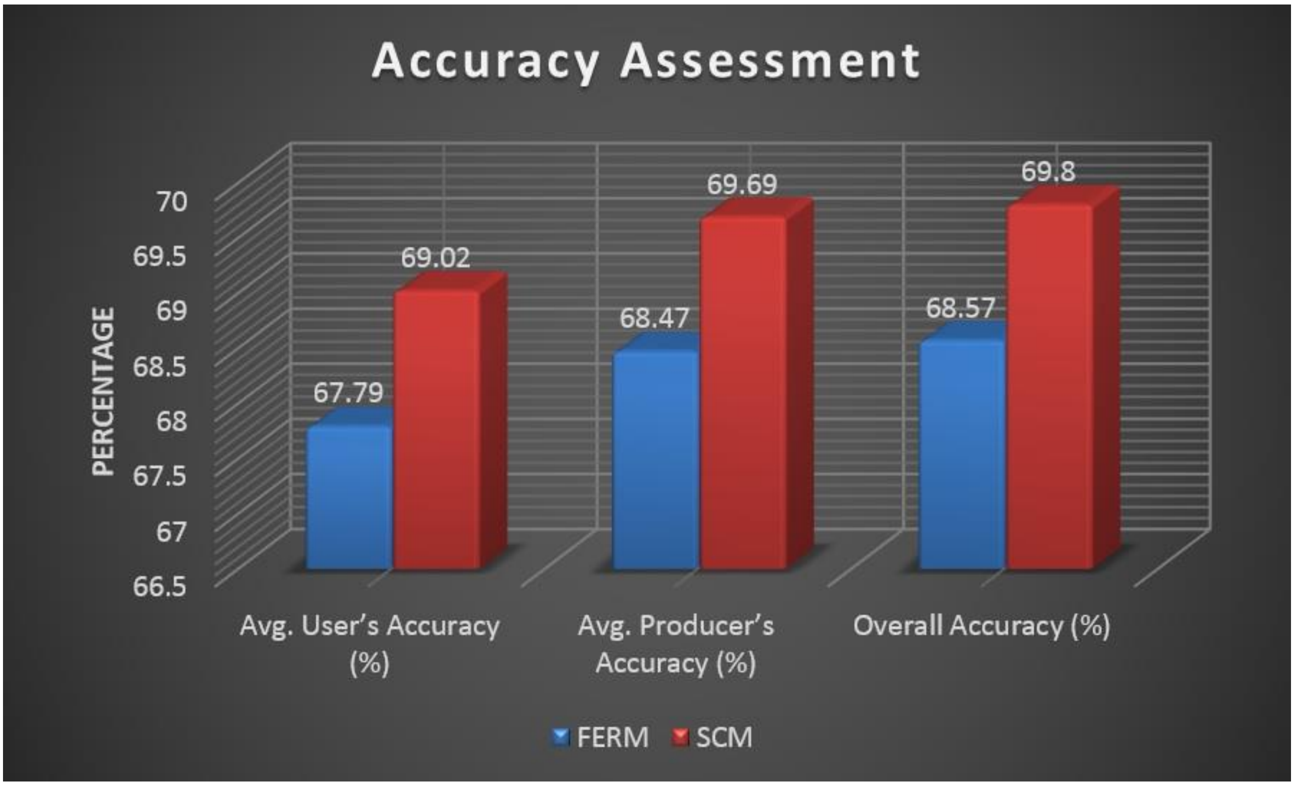


Figure A-7. Accuracy assessment for by FCM using composite measure in Landsat-8 data



APPENDIX B

Accuracy Assessment of classified outputs for Landsat imagery with Cosine norm at m equal to 2.5 and α-cut values 0.5, 0.6, 0.7, 0.8 and 0.9 has been shown in Table B-1, B-2, B-3, B-4 and B-5 respectively.

Table B-1 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.5

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	38.98	67.82 ± 6.92	67.44	60.90	74.75
Fallow Land	9.44	27.14 ± 9.52	23.50	17.62	36.66
Forest	36.48	85.88 ± 8.99	83.80	76.89	94.87
Water	23.60	61.97 ± 7.52	60.56	54.44	69.49
Wheat	23.99	83.33 ±11.10	79.20	72.23	94.43
Producer's Accuracy (%)					
Riverine Sand	28.02	59.48 ±10.73	56.94	48.75	70.21
Fallow Land	11.18	48.42 ±21.92	34.18	26.50	70.34
Forest	52.70	98.06 ± 1.80	97.60	96.26	99.86
Water	11.01	36.07 ±9.43	33.53	26.64	45.51
Wheat	30.12	79.34 ±10.25	77.67	69.10	89.59
Overall Accuracy (%)	31.59	74.16 ±10.14	71.81	64.03	84.30
Fuzzy Kappa value		0.64 ± 0.15	0.62	0.52	0.77

Table B-2 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.6

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	**36.07**	**67.95 ± 4.55**	**68.11**	**63.40**	**72.51**
Fallow Land	**3.22**	**10.66 ± 3.45**	**9.33**	**7.21**	**14.11**
Forest	**27.59**	**84.57 ± 6.81**	**82.57**	**77.76**	**91.38**
Water	**26.91**	**68.42 ± 7.01**	**66.82**	**61.41**	**75.43**
Wheat	**16.68**	**67.95 ±10.21**	**63.74**	**57.74**	**78.16**
Producer's Accuracy (%)					
Riverine Sand	**16.84**	**54.45 ± 9.14**	**52.01**	**45.32**	**63.59**
Fallow Land	**2.16**	**21.99 ±12.62**	**11.89**	**9.36**	**34.62**
Forest	**38.16**	**98.87 ± 1.01**	**98.45**	**97.85**	**99.88**
Water	**17.90**	**55.52 ± 7.29**	**54.04**	**48.23**	**62.82**
Wheat	**19.43**	**67.52 ± 8.07**	**65.90**	**59.76**	**75.89**
Overall Accuracy (%)	**22.92**	**72.20 ± 7.82**	**70.08**	**64.38**	**80.01**
Fuzzy Kappa value		**0.62 ± 0.11**	**0.60**	**0.53**	**0.72**

Table B-3 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.7

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	**19.53**	**94.46 ±5.54**	**91.32**	**88.93**	**100.0**
Fallow Land	**4.01**	**18.57 ± 4.76**	**16.91**	**13.81**	**23.33**
Forest	**19.23**	**93.05 ± 4.95**	**90.40**	**88.11**	**98.0**
Water	**38.46**	**84.37 ± 0.82**	**84.22**	**83.55**	**85.18**
Wheat	**21.74**	**80.85 ± 6.25**	**78.19**	**74.60**	**87.10**
Producer's Accuracy (%)					
Riverine Sand	**10.90**	**45.20 ± 7.86**	**42.24**	**37.34**	**53.05**
Fallow Land	**6.42**	**43.81 ±21.26**	**26.98**	**22.55**	**65.07**
Forest	**30.00**	**100.00**	**100.0**	**100.0**	**100.0**
Water	**16.73**	**80.49 ± 4.91**	**78.35**	**75.58**	**85.40**
Wheat	**20.39**	**98.51 ± 1.49**	**97.63**	**97.02**	**100.0**
Overall Accuracy (%)	**19.54**	**82.61 ± 5.58**	**80.40**	**77.04**	**88.20**
Fuzzy Kappa value		**0.76 ± 0.08**	**0.74**	**0.70**	**0.84**

Table B-4 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.8

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	**6.52**	**45.29 ± 3.52**	**45.29**	**41.76**	**48.81**
Fallow Land	**0.0**	**NaN**	**NaN**	**NaN**	**NaN**
Forest	**7.08**	**100.0**	**100.0**	**100.0**	**100.0**
Water	**19.13**	**83.86 ± 3.62**	**81.0**	**80.24**	**87.47**
Wheat	**14.24**	**87.05 ± 4.90**	**84.08**	**82.15**	**91.95**
Producer's Accuracy (%)					
Riverine Sand	**17.31**	**74.96 ± 5.03**	**70.93**	**69.93**	**80.00**
Fallow Land	**0.0**	**NaN**	**NaN**	**NaN**	**NaN**
Forest	**9.06**	**95.33 ± 3.40**	**92.52**	**91.94**	**98.73**
Water	**12.43**	**55.64 ± 4.20**	**55.07**	**51.44**	**59.84**
Wheat	**7.92**	**100.0**	**100.0**	**100.0**	**100.0**
Overall Accuracy (%)	**10.31**	**74.65 ± 3.86**	**73.27**	**70.80**	**78.51**
Fuzzy Kappa value		**0.65 ± 0.05**	**0.63**	**0.60**	**0.71**

Table B-5 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.9

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	0.0	0.0	0.0	0.0	0.0
Fallow Land	0.0	0.0	0.0	0.0	0.0
Forest	2.83	100.0	100.0	100.0	100.0
Water	6.49	100.0	100.0	100.0	100.0
Wheat	0.0	NaN	NaN	NaN	NaN
Producer's Accuracy (%)					
Riverine Sand	0.0	NaN	NaN	NaN	NaN
Fallow Land	NaN	NaN	NaN	NaN	NaN
Forest	3.02	100.0	100.0	100.0	100.0
Water	6.11	65.25 ± 5.37	61.83	59.88	59.88
Wheat	0.0	0.0	0.0	0.0	0.0
Overall Accuracy (%)	1.52	35.29	34.78	32.88	37.69
Fuzzy Kappa value		0.27	0.26	0.25	0.29

APPENDIX C

Accuracy Assessment of classified outputs for Landsat imagery with composite measures formed by Cosine norm and Euclidean norm at *m* equal to 2.5, weighting constant (λ) equal to 0.5 and α-cut values 0.5, 0.6, 0.7, 0.8 and 0.9 has been shown in Table C-1, C-2, C-3, C-4 and C-5 respectively.

Table C-1 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.5

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	**21.05**	**34.08 ± 5.69**	**32.68**	**28.39**	**39.76**
Fallow Land	**12.10**	**35.50 ± 8.19**	**33.80**	**27.31**	**43.69**
Forest	**56.65**	**93.95 ± 3.38**	**93.45**	**90.56**	**97.33**
Water	**46.42**	**74.14 ± 5.98**	**72.76**	**68.15**	**80.12**
Wheat	**26.26**	**76.33 ± 5.67**	**75.32**	**70.66**	**82.00**
Producer's Accuracy (%)					
Riverine Sand	**27.80**	**60.85 ± 8.31**	**59.77**	**52.54**	**69.16**
Fallow Land	**15.34**	**34.48 ± 7.29**	**32.04**	**27.19**	**41.77**
Forest	**57.31**	**97.74 ± 1.41**	**97.62**	**96.34**	**99.15**
Water	**28.19**	**60.01 ± 5.35**	**59.83**	**54.66**	**65.36**
Wheat	**24.19**	**61.37 ± 11.6**	**58.02**	**49.76**	**72.97**
Overall Accuracy (%)	**37.97**	**75.33 ± 6.38**	**74.37**	**68.95**	**81.71**
Fuzzy Kappa value		**0.64 ± 0.11**	**0.63**	**0.56**	**0.73**

Table C-2 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.6

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	**17.14**	**34.89 ± 5.98**	**33.59**	**28.91**	**40.87**
Fallow Land	**28.00**	**59.38 ± 7.55**	**59.42**	**51.83**	**66.92**
Forest	**53.65**	**93.23 ± 2.33**	**92.90**	**90.90**	**95.56**
Water	**40.94**	**68.58 ± 3.47**	**68.13**	**65.11**	**72.05**
Wheat	**18.60**	**79.10 ± 4.17**	**79.24**	**74.92**	**83.27**
Producer's Accuracy (%)					
Riverine Sand	**22.52**	**52.74 ± 5.21**	**53.02**	**47.53**	**57.95**
Fallow Land	**29.09**	**60.82 ± 6.53**	**59.31**	**54.30**	**67.35**
Forest	**52.91**	**99.92 ± 0.08**	**99.88**	**99.84**	**100.0**
Water	**19.70**	**49.32 ± 6.91**	**49.18**	**42.42**	**56.23**
Wheat	**12.35**	**55.53 ± 8.19**	**53.52**	**47.35**	**63.72**
Overall Accuracy (%)	**33.62**	**78.05 ± 4.76**	**77.75**	**73.30**	**82.81**
Fuzzy Kappa value		**0.66 ± 0.08**	**0.66**	**0.61**	**0.73**

Table C-3 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.7

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	13.24	24.31 ± 3.35	24.00	20.96	27.66
Fallow Land	5.97	20.24 ± 2.77	20.66	17.47	23.01
Forest	37.20	90.45 ± 2.58	89.90	87.88	93.03
Water	28.48	67.72 ± 4.46	66.52	63.27	72.18
Wheat	6.45	56.73 ± 5.86	54.68	50.87	62.59
Producer's Accuracy (%)					
Riverine Sand	17.23	52.73 ± 2.96	54.27	49.77	55.69
Fallow Land	4.57	25.83 ± 8.65	20.58	17.17	34.48
Forest	32.89	100.00	100.0	100.0	100.0
Water	14.69	40.86 ± 4.12	40.64	36.74	44.98
Wheat	3.81	27.37 ± 3.21	27.37	24.16	30.59
Overall Accuracy (%)	18.36	64.42 ± 4.63	63.96	59.78	69.05
Fuzzy Kappa value		0.50 ± 0.08	0.50	0.45	0.55

Table C-4 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.8

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	14.36	49.07 ± 3.57	47.52	45.50	52.65
Fallow Land	0.0	0.0	0.0	0.0	0.0
Forest	33.31	98.57 ± 0.97	97.95	97.60	99.54
Water	11.22	49.16 ± 5.97	46.32	43.19	55.14
Wheat	12.63	78.14 ± 6.11	84.08	72.03	84.25
Producer's Accuracy (%)					
Riverine Sand	21.17	50.14 ± 3.57	48.45	45.05	55.23
Fallow Land	0.0	0.0	0.0	0.0	0.0
Forest	27.59	100.0	100.0	100.0	100.0
Water	7.32	43.19 ± 5.97	41.72	39.12	47.26
Wheat	3.25	100.0	100.0	100.0	100.0
Overall Accuracy (%)	16.86	81.40 ± 3.14	73.27	78.26	84.54
Fuzzy Kappa value		0.64 ± 0.08	0.63	0.60	0.69

Table C-5 Details of accuracy assessment for classification results of Landsat-8 data at α-cut equals to 0.9

Accuracy Assessment Operators	FERM	SCM	MIN-PROD	MIN-MIN	MIN-LEAST
User's Accuracy (%)					
Riverine Sand	5.55	34.23 ± 0.87	34.14	33.36	35.10
Fallow Land	0.0	0.0	0.0	0.0	0.0
Forest	7.64	100.0	100.0	100.0	100.0
Water	28.50	100.0	100.0	100.0	100.0
Wheat	2.80	100.0	100.0	100.0	100.0
Producer's Accuracy (%)					
Riverine Sand	16.29	100.0	100.0	100.0	100.0
Fallow Land	0.0	NaN	NaN	NaN	NaN
Forest	3.76	100.0	100.0	100.0	100.0
Water	23.74	73.74 ± 1.49	73.15	72.25	75.23
Wheat	0.57	100.0	100.0	100.0	100.0
Overall Accuracy (%)	7.32	80.82 ± 1.19	80.35	79.63	82.01
Fuzzy Kappa value		0.66 ± 0.02	0.65	0.64	0.68

APPENDIX D

IMPLEMENTATION OF α-CUT FOR FUZZY *C*-MEANS (FCM) IN R

```
rm(list=ls(all=TRUE))

require(rgdal)
require(raster)
```

#taking input image as raster

```
img.file_1=raster ("C:/Users/mgi14-9417/Documents/outputs/landsat/img_real_subset_landsat_Cosine_2.5/clay.img")
img.file_2=raster ("C:/Users/mgi14-9417/Documents/outputs/landsat/img_real_subset_landsat_Cosine_2.5/fallow.img")
img.file_3=raster ("C:/Users/mgi14-9417/Documents/outputs/landsat/img_real_subset_landsat_Cosine_2.5/forest.img")
img.file_4=raster ("C:/Users/mgi14-9417/Documents/outputs/landsat/img_real_subset_landsat_Cosine_2.5/water.img")
img.file_5=raster ("C:/Users/mgi14-9417/Documents/outputs/landsat/img_real_subset_landsat_Cosine_2.5/wheat.img")
```

#Dimension of image files

```
dim(img.file_1)
dim(img.file_2)
dim(img.file_3)
dim(img.file_4)
dim(img.file_5)

row=74                    #Row value from dim
col=89                    #Col value from dim

row_subset=74
col_subset=89


n=row*col
```

#Value of images is taken with respect to 1, so its divided by 255 (on an 8-bit scale image)

```
value_1=(img.file_1)/255
value_2=(img.file_2)/255
value_3=(img.file_3)/255
value_4=(img.file_4)/255
value_5=(img.file_5)/255


value_11=value_1
value_22=value_2
value_33=value_3
value_44=value_4
value_55=value_5
```

#Loop for checking the value of image pixels with respect to the α-cut value (threshold)

```
for (a in seq(0.5, 0.9, 0.1))
{
 value_1=value_11
 value_2=value_22
 value_3=value_33
 value_4=value_44
 value_5=value_55

 for (row in 2:row_subset)
 {
  for (col in 2:col_subset)
  {
   if(value_1[row,col] >0)
   {
   if (value_1[row,col] < a)
   {
    value_1[row,col] = 0
   }
   }
   if (value_2[row,col] < a)
   {
    value_2[row,col] = 0
   }

   if (value_3[row,col] < a)
   {
    value_3[row,col] = 0
   }

   if (value_4[row,col] < a)
   {
```

```
      value_4[row,col] = 0
    }

    if (value_5[row,col] < a)
    {
      value_5[row,col] = 0
    }

  }
}
```

#To change value of the pixels from values with respect to 1 to a scale of 255

```
value_1=value_1*255
value_2=value_2*255
value_3=value_3*255
value_4=value_4*255
value_5=value_5*255
```

#Outputs

```
rf <- writeRaster(value_1, filename="clay.tif", format="GTiff", overwrite=TRUE)

rf <- writeRaster(value_2, filename="fallow.tif", format="GTiff", overwrite=TRUE)

rf <- writeRaster(value_3, filename="forest.tif", format="GTiff", overwrite=TRUE)

rf <- writeRaster(value_4, filename="water.tif", format="GTiff", overwrite=TRUE)

rf <- writeRaster(value_5, filename="wheat.tif", format="GTiff", overwrite=TRUE)

}
```

APPENDIX E

E.1 Publications

S.Mukhopadhaya, A.Kumar, A.Stein – "An effective approach of similarity and dissimilarity measures with alpha-cut"
(Draft prepared. To be submitted in a peer reviewed journal)