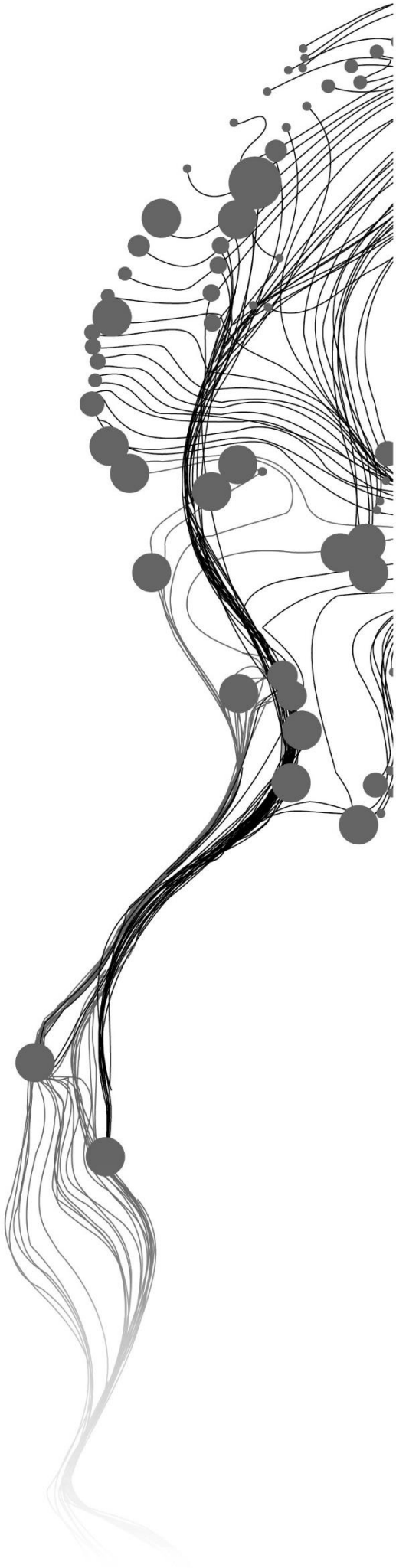
LOCAL CONVOLUTION INFORMATION FOR FUZZY BASED CLASSIFIER

ANJALI MADHU
March, 2019

SUPERVISORS:
Dr. Anil Kumar
Dr. Peng Jia



LOCAL CONVOLUTION INFORMATION FOR FUZZY BASED CLASSIFIER

ANJALI MADHU
Enschede, The Netherlands, March, 2019

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.
Specialization: Geoinformatics

SUPERVISORS:
Dr. Anil Kumar
Dr. Peng Jia

THESIS ASSESSMENT BOARD:
Prof.Dr. Ir. A. Stein(Chair)
Prof. S.K. Ghosh (External Examiner, Indian Institute of Technology (IIT), Roorkee)

ABSTRACT

Choosing a classifier for remote sensing image analysis is a difficult task. Effective classification algorithms with improved classification accuracy are essential in the presence of complexities such limited spatial resolution, fuzziness in ground substances, the presence of mixed pixels and vague pixels in the image, spectral variation etc. The mixed-pixel problem in land use/land cover classification can be tackled by using fuzzy classification algorithms in which the image pixels are assigned to more than a single class. The conventional fuzzy classification algorithms such as Fuzzy *c*-means (FCM) and Possibilistic *c*-means (PCM) use only the spectral information of the pixels for classification. In such classification, the information about the spatial distribution of pixels is ignored. The incorporation of spatial information in the classifiers can reduce probable ambiguities and produce more accurate classification results.

In this research, three FCM based spatial contextual classifiers from literature, FCM-S, FLICM and ADFLICM, were selected and their performance has been examined on remotely sensed image. These classifiers use spatial information within a defined local neighbourhood of each pixel for estimating the class membership of that pixel. The FCM-S, FLICM and ADFLICM classifiers incorporate spatial information by using a simple approach of modifying the FCM objective function to include a term that controls the effect from neighbouring pixels. A similar approach has been used to modify the PCM classifier to develop PCM based local spatial information classification algorithms PCM-S, PLICM and ADPLICM in this research.

Supervised classification with the algorithms developed was performed on Landsat-8 image (30m). For the validation of results, soft reference data was created from a finer resolution Formosat-2 image (8m) of the same study area captured around the same time. Various experiments were conducted to analyse the performance of FCM based and PCM based local spatial information classification algorithms developed. To better analyse the effect of incorporating spatial information into base classifiers, the performance of the classifiers were compared to the respective base classifiers (FCM and PCM). Fuzzy Error Matrix (FERM), Root Mean Square Error (RMSE) and Mean Membership Difference methods were used to analyse the performance of the classification algorithms quantitatively.

The results suggested that FCM based and PCM based local spatial information classifiers outperformed the conventional FCM and PCM based classifiers in terms of overall accuracy. However, the local spatial information algorithms may produce over smooth results which may cause loss of image details. Appropriate local spatial information classifiers should be chosen for classification based on the nature of the classes to be classified.

Keywords: *Fuzzy classification, Fuzzy c-means, Possibilistic c-means, local spatial information, spatio – contextual, local information*

ACKNOWLEDGEMENTS

The journey I went through to accomplish this research has been an erudite experience. I was able to learn, understand, and solve various complexities that I faced during the research. Overcoming the obstacles and arriving at the conclusions has brought me immense pleasure. I am grateful to the mighty force for being alive and for being able to pursue the quest with motivation.

This thesis will be incomplete without mentioning a few people who created a significant impact on me and who inspired me to face all the challenges with determination during the study. I take this opportunity to thank each and everyone who has directly or indirectly been involved or encouraged me in completing this research.

I want to express my deep sense of gratitude towards my supervisors Dr. Anil Kumar and Dr. Peng Jia for providing timely comments and suggestions to improve my work. Dr. Anil has been always approachable and patient enough to listen and clear my queries. Thank you for being an invaluable teacher and for being always there to help me out. I am grateful to Dr. Peng Jia for all the brainstorming discussions and comments which were necessary to conceptualise and organise the research in the right direction. Those were some valuable lessons for a beginner like me in the research field.

I want to thank all my teachers in IIRS and ITC for sharing their wisdom and knowledge with me. It has been a wonderful experience to be your student, and I will cherish everything that I have learned from you all. I would like to specifically mention Dr. Sameer Saran (Head of Geo-informatics Department, IIRS), for being the ever helpful person that he is. I also wish to thank every staff in IIRS and ITC for making my tenure memorable.

I wish to thank my friends without whom the completion of this research would have been impossible. You stood with me, helped me and motivated me when I was in agony and uncertainty. I am genuinely grateful for having you all in my life. Finally, I would like to thank my ever reassuring family for the moral support provided to me. You pushed me to pursue my research and had belief in me when I lacked the self-confidence I required. Thank you for always being there to encourage me to follow my dreams.

Anjali Madhu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

LULC: Land Use Land Cover
FCM: Fuzzy *c*-Means
PCM: Possibilistic *c*-Means
FCM-S: Fuzzy *c*-Means with Spatial constraint
FLICM: Fuzzy Local Information *c*-Means
ADFLICM: Adaptive Fuzzy Local Information *c*-Means
PCM-S: Possibilistic *c*-Means with Spatial constraint
PLICM: Possibilistic Local Information *c*-Means
ADPLICM: Adaptive Possibilistic Local Information *c*-Means
FERM: Fuzzy Error Matrix
OA: Overall Accuracy
PA: Producer's Accuracy
UA: User's Accuracy
RMSE: Root Mean Square Error

1. INTRODUCTION

1.1. Background

The advent of remote sensing has helped humankind in obtaining useful information by means of observing an object, a scene, a phenomenon without having any physical contact with it. Image interpretation or image analysis, which is the most crucial step in remote sensing, converts an image into useful information for wide range of users in applications such as land use/land cover mapping, geological and soil mapping, forestry, environmental assessment, disaster assessment, urban and regional planning, archaeology and many others (Lillesand, Kiefer, & Chipman, 2004; Sabins, 1997). Land use land cover mapping is maybe the most useful application derived from remote sensing as this information is vital for other applications such as assessment and monitoring of vegetation types, environment impact assessment, crop and yield estimation, natural resource management and monitoring, urban planning etc. (Ibrahim, 2004).

Image classification or pattern recognition in the context of remote sensing converts radiometric information contained in the image to thematic information by linking objects (pixels) in the study area to one or more elements in the user-defined label set (Tso & Mather, 2009). They allow identification of an area in the image as corresponding to a particular ground cover type. Classification procedure can be supervised or unsupervised. Supervised classification uses appropriate algorithms to label the pixels in an image as specific land cover classes that they represent on the ground. In this case, the land use/land cover classes are known beforehand, and the pixels are assigned to available classes based on their spectral properties which are expressed in terms of Digital Number (DN) values. Prototype pixels, called as training data, are chosen from known classes and every pixel in the image is classified into desired land use/land cover classes using training data. On the contrary, the unsupervised classification approach does not use prior knowledge of the information classes. The whole image is segmented into unknown classes by clustering the pixels based on their spectral similarities and labelling these classes afterwards (Richards & Jia, 2012).

An important factor that reduces the quality of the satellite image analysis is the loss of spectral information in the process of image classification caused by the insufficient representation of geographical information. The geographic information is represented as thematic maps, such as Land Use Land Cover (LULC) maps. Classification based on satellite images is used to identify these LULC classes with the support of ground data. In conventional classifiers, the study area is assumed to be composed of a number of unique, internally homogenous LULC classes that are mutually exclusive. The image pixel that corresponds to the ground is associated with single land use or land cover class in conventional hard classifiers. However, a single pixel may not be homogenous in reality and may represent more than one land cover, such as a soil pixel with sparse grassland, which can be classified as either “grassland” or “soil”. Such mixed pixels may also occur at the boundaries of different land covers and even in the presence of ground objects which are smaller than the pixel size. Also, there may exist ambiguities due to the variation in spectral responses recorded by the sensors and the corresponding ground situations. Similar spectral responses by remote sensors may correspond to dissimilar entities on the ground, while similar entities on the ground at different locations exhibit varied spectral responses. A composite spectral

response that is dissimilar to the spectral response of each of the component classes of a mixed pixel may as well be recorded. There might also be variability within land cover classes, such as vegetation in different conditions due to its age, health and water content, which can produce varied spectral responses by remote sensors. This imprecise nature or fuzziness of geographical information, arising from natural variation or through original measurements or data processing, is better represented in fuzzy-based classifiers, as opposed to conventional classifiers, through the concept of partial and multiple class memberships for pixels. In fuzzy approaches, a pixel is assigned a membership value for each class which indicates the extent to which the pixel belongs to that class. This allows one pixel to be allocated to more than one class, and the pixel is not considered as an indecomposable unit. Consequently, the uncertainty or fuzziness in geographic representation due to class mixtures, intermediate conditions, or within-class variability, is better represented in fuzzy classification algorithms. These algorithms can improve the accuracy of the classification as the loss of spectral information is reduced by a more accurate representation of geographic information (Wang, 1990; J. Zhang & Foody, 2001).

Although numerous classification techniques have been developed to maximise the accuracy of information extraction from satellite imagery, most methods utilise only spectral properties of images and ignore their spatial information (Li, Zang, Zhang, Li, & Wu, 2014). Medium-resolution satellite image sensors, such as Landsat Thematic Mapper, produce images in which pixel size is usually smaller than many landscape objects. Thus, pixels within images exhibit a high degree of spatial autocorrelation, which means the knowledge of pixels belonging to a class increases the probability of its neighbouring pixels to be in the same class (Stuckens, Coppin, & Bauer, 2000). Recently, spatial correlations between the central pixel and its neighbouring pixels have been incorporated into image classification approaches, which are termed "spatio-contextual" image classification approaches. Such approaches are efficient in distinguishing different land cover classes which produce similar spectral responses that cannot be separated only by spectral information (Li et al., 2014).

Algorithms have been developed that incorporate spatial information with per-pixel classification as well as fuzzy classifications for improving classification results. Spatial information may eliminate ambiguities due to the spectral similarity of pixels, help in recovering missing information and correcting noisy pixels during image classification (Chawla, 2010). It has been seen that, by integrating spatial-contextual information with spectral information, classifiers improve its robustness and classification accuracy (Dutta, 2009; Singha, 2013; Tso & Mather, 2009). Spatial context is generally exploited by image pre-processing to enhance the spatial properties of the image, post classification filtering using majority filters or using the probabilistic label relaxation or Markov Random Fields during classification (Richards & Jia, 2012).

1.2. Motivation and Problem Statement

Most of the remotely sensed images contain ambiguities caused by within class spectral variation that may arise due to sensor response or different conditions of the land cover. It is essential to identify and handle these noisy pixels to extract the right information about land cover through classification. Among the classifiers, Fuzzy classifiers like Fuzzy c-means (FCM) have reported to produce better classification results in terms of considering the heterogeneity of every pixel, as geographical information is more accurately represented. While FCM works well on noise-free images, it is susceptible to noise as it does not consider any information about the spatial context (Krinidis & Chatzis, 2010). Suitable use of spatial information with spectral information allows the elimination of possible ambiguities, the recovery of missing information and the correction of errors (Tso & Mather, 2009).

Pre-processing approaches to exploit spatial contextual information may lead to loss of important image details. The trade-off between smoothing and the results of segmentation cannot be controlled in this approach. Therefore, image segmentation or classification while simultaneously compensating noise insensitivities using spatial contextual information is advantageous compared to pre-processing and post-processing approaches since intermediate information can be used for correcting within-class heterogeneity (Pham & Prince, 1999). Nevertheless, rather sophisticated methods are used to integrate spatial contextual information into classification. It is possible to develop simple approaches to incorporate spatial information into classifier which may substantially improve the performance of conventional per-pixel classifiers (Gurney & Townshend, 1983).

Recently algorithms have been developed that incorporate spatial information in fuzzy classifiers by modifying the objective function of the FCM algorithm (Pham, 2001; Liew, Leung, & Lau, 2000). Pham (2001) modified the objective function of FCM to impose the spatial constraint on the behaviour of membership functions. Spatial constraint forces the membership value of a pixel to be similar to the membership values of neighbouring pixels. The advantages of such methods are that the spatial constraint function can be included into FCM, its variants, and Possibilistic c-means (PCM) algorithms in a straightforward manner with minimal changes to the resulting implementation (Pham, 2001). In few other research, an additional term that controls the effect from the neighbourhood is included in FCM objective function (Ahmed, Yamany, Mohamed, Farag, & Moriarty, 2002; S. Chen & Zhang, 2004; Krinidis & Chatzis, 2010; H. Zhang, Wang, Shi, & Hao, 2017). The neighbourhood information is incorporated by using convolution in these algorithms. Convolution here is the use of moving window or template on the image row by row and column by column to calculate the label of each pixel while incorporating its neighbour pixel information as well in the classifier. The products of the value in the image covered by the template and the corresponding template entry at a particular position are summed to get the template response or the value of the centre pixel (Richards & Jia, 1999; Schowengerdt, 2014). Such classifiers, which will be referred to as fuzzy local spatial information classifiers hereafter in this thesis, are claimed to produce enhanced classification results by providing a balance between noise removal and image detail preservation. However, these algorithms haven't been explored and optimised to their best potential. Altering the base classifiers, optimising parameters, changing the extent of neighbouring pixels used in classification may improve the classifier performance.

1.3. Research Identification

Good information processing depends on good representation methods for information and efficient processing algorithms (Wang, 1990). For precise representation of ground reality, fuzzy classifiers are used in this research. Uncertainties due to mixed pixels are handled by fuzzy classification algorithms. Possibilistic c-means (PCM) classifier (Krishnapuram & Keller, 1993, 1996), which is naturally immune to noise and outliers, is employed besides Fuzzy c-means (FCM) classifier. To incorporate spatial information into the fuzzy classifiers, simpler approach of modifying the objective function of the base classifier to include neighbourhood term is considered. Lastly, to facilitate accurate classification using representative samples of the known locations, a supervised classification approach is followed. Together with accurate representation of reality via efficient classification approaches and simpler methods to incorporate spatial information, this research is carried out to produce homogenous segmentation results with reduced edge blurring artefact.

1.4. Research Objectives

The main aim of this study was to compare the ability of fuzzy local spatial information classifiers in handling ambiguities, caused by within-class spectral variation and spectral similarity among different classes, without losing essential image details. The specific aims include:

1. To study and compare the performance of existing three FCM based fuzzy local spatial information classification algorithms for land cover classification;
2. To develop fuzzy local spatial information algorithms with PCM as the base classifier and analyse their performance in noise reduction(isolated pixels) and preservation of image details;
3. To conduct a comprehensive assessment of the PCM and FCM based local spatial information classifiers;
4. To examine the effects of some key parameters such as fuzzifier, window size and the factor that controls the impact of neighbourhood on the performance of these algorithms and optimise these values.

1.5. Research Questions

The following were the research questions that helped to meet the research objectives:

1. How fuzzy local spatial information classifiers work on remotely sensed images? Which FCM based fuzzy local spatial information classifier could be considered as a benchmark in handling ambiguities due to spectral similarity/variation and preserving image details?
2. How is local spatial information incorporated in PCM classifier? How the local spatial information PCM classifiers perform compared to the conventional PCM classifier?
3. Which fuzzy local spatial information algorithm will give better classification results? Which algorithm is better in removing noise and preserving image details?
4. What will be the optimal parameter values and window size to balance the image noise and loss of image details?

1.6. Innovation Aimed At

Innovation intended in this study:

1. To analyse the performance of local spatial information classifiers with PCM as the base classifier
2. To study the performance of the existing FCM based and the new PCM based local spatial information classification algorithms for LULC classification.

1.7. Research Approach

To develop PCM based local spatial information classifiers, the objective function of conventional PCM classifier is modified to include spatial constraints, and the parameters of the objective function are estimated by applying the minimization criteria. The satellite images of Landsat-8 and Formosat-2 are geo-registered. Supervised classification is implemented on Landsat-8 image with all the six fuzzy local spatial information classification algorithms (three FCM based and three PCM based) and the two conventional classification algorithms (FCM and PCM). Experiments are conducted to analyse the performance of local spatial information fuzzy classifiers by altering parameters and extent of neighbouring pixels (window size) used in classification. Accuracy assessment is carried out with data from Formosat-2 image as the reference.

Figure 1.1 gives the generalised methodology for research.

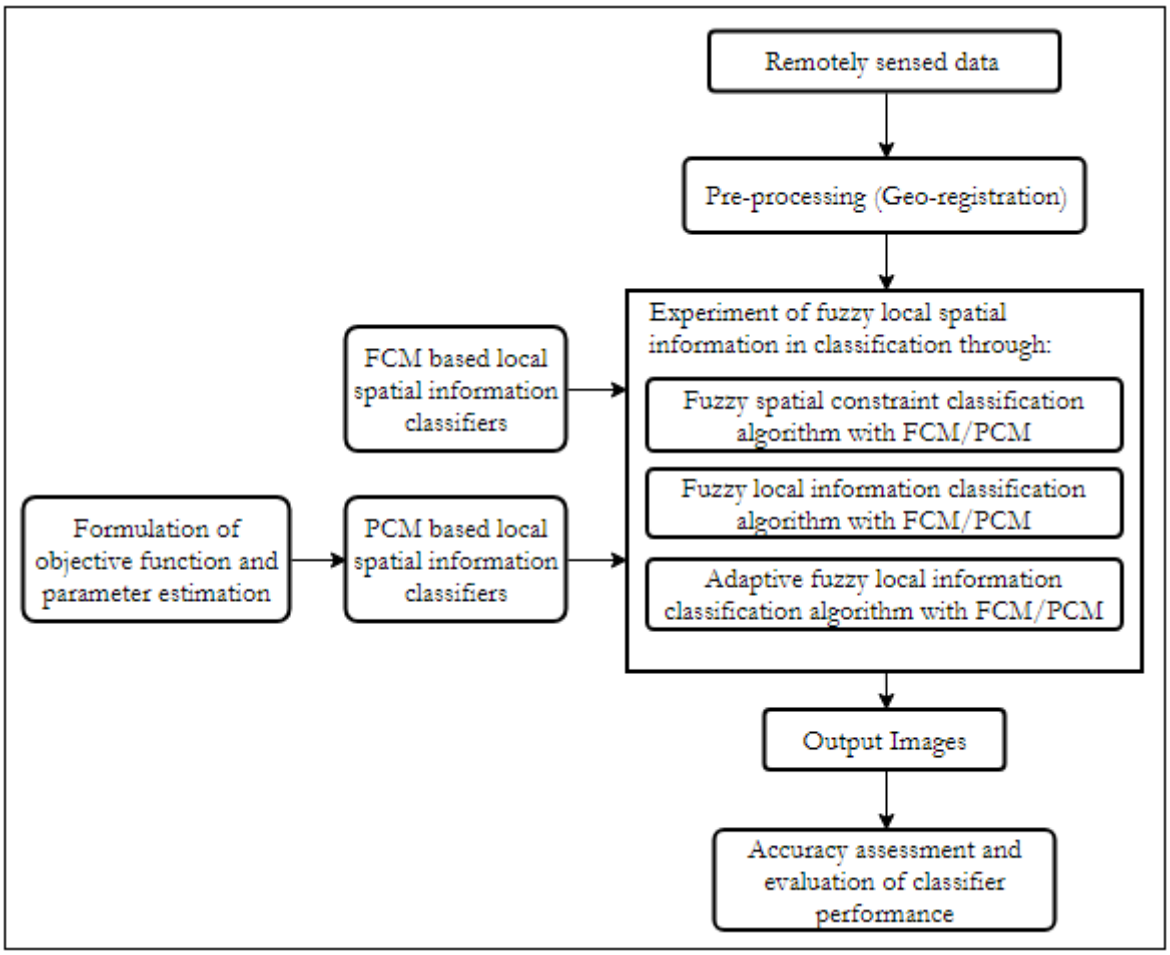


Figure 1.1: Generalised research methodology

1.8. Thesis Structure

This thesis is organised into seven chapters. ***Chapter one*** introduces the research topic, general research approach, background information, motivation and problem statement that lead to the formulation of research objectives. ***Chapter two*** discusses various works that have been done in the past concerning image classification and the use of contextual information for classification. ***Chapter three*** explains the details about different classification techniques used in this research. ***Chapter four*** gives information about the study area and the resources used for this study. ***Chapter five*** describes the methodology followed in implementing the classifiers and exploring the performance of the classifiers. ***Chapter six*** provides the results obtained by using different classifiers and the analysis on the results obtained. ***Chapter seven*** concludes the thesis with answers to the research questions and recommendations for future research.

2. BACKGROUND REVIEW

This chapter presents the necessary background review and study of related works that were required to understand and execute the research.

2.1. Remote sensing image classification

Classifying remotely sensed data to thematic maps is a challenging task because factors such as complexity of landscape in the study area, remotely sensed data selected, image classification and processing approaches used affect the classification performance. Scientists are continually coming up with advanced techniques to improve classification accuracy (Lu & Weng, 2007). It is not possible to state that a given classifier works best for all situations, as the characteristics of each image and circumstances of each study vary greatly. For that reason, the appropriate classifier should be selected by the analyst for the task at hand (Campbell & Wynne, 2011).

One of the widely used supervised image classification techniques is Maximum Likelihood classification. Maximum Likelihood Classifier (MLC) is a parametric classifier dependent on statistical theory. It assumes training data to be normally distributed and classifies pixels based on Bayesian probability formula. The probability of pixels belonging to each class is determined, and the pixel is allocated to the class for which probability is the highest (Tso & Mather, 2009).

Recently classification techniques based on artificial neural network, decision trees, and support vector machine have emphasised their superiority over traditional parametric approaches in remote sensing image classification mainly due to their non-parametric nature (Tso & Mather, 2009). They are effective in handling complex classification scenarios by incorporating advanced learning mechanisms. However, large training data is required to train artificial neural networks, decision trees etc to produce generalised outcome.

Land cover classes do not have sharp, hard boundaries and thus reality is very imprecise and heterogeneous. This heterogeneity may lead to the presence of mixed pixels. Vague pixels and mixed pixels are the pixels that show affinity with more than one information class. Conventional hard classification approaches use precise classical set theory and classify the remotely sensed data into discrete, homogeneous information classes disregarding the natural imprecise nature of the real world (Jensen, 1996). Soft classification approaches using fuzzy logic provide more information and more accurate results especially in case of coarser datasets (Lu & Weng, 2007). Fuzzy set theory, proposed by Zadeh (1965) is a framework which helps in solving knowledge representation and classification problems in ambiguous environments (Tso & Mather, 2009). Among fuzzy classification, Fuzzy *c*-means (FCM) (Bezdek, 1981; Bezdek, Ehrlich, & Full, 1984) has been widely used for land cover classifications.

2.2. Fuzzy *c*-Mean

Fuzzy *c*-means (FCM) (Bezdek, 1981; Bezdek et al., 1984) is an unsupervised iterative clustering algorithm which finds fuzzy partitions and prototypes by minimising the objective function. FCM has been implemented successfully in many applications in the field of pattern recognition and remote sensing(Atkinson, Cutler, & Lewis, 1997; Bastin, 1997; Cannon, Dave, Bezdek, & Trivedi, 1986; Fisher &

Pathirana, 1993; Foody, 1996; Ibrahim, 2004; Kumar, 2007). The similarity a point (pixel) shares with each class or cluster is represented using membership functions whose value ranges between 0 and 1. The distance of each pixel from the class centroid in the feature space is measured iteratively to get the membership value of the pixel for that class. Membership value close to zero indicates a low degree of similarity between the pixel and the class while membership value of one implies a high degree of similarity between the pixel and the class (Bezdek et al., 1984). The algorithm assumes the probabilistic constraint that the membership probabilities of data points across class should sum to one. The constraint on membership is used to avoid the trivial solution of all membership equal to zero (Dutta, 2009; Tso & Mather, 2009).

Fuzzy membership value has been found to be more accurate in representing the ground reality. Fisher & Pathirana(1993) have shown that the fuzzy membership values for each land cover yielded by FCM algorithm are similar to the actual fraction of those land covers on the ground (Bastin, 1997). Studies have been done to compare the performance of different classification algorithms in mapping land cover classes, and FCM has been found to be effective in handling mixed pixels. Atkinson, Cutler, & Lewis(1997) made a comparative analysis on the performance of artificial neural networks (ANN), mixture modelling and FCM on AVHRR (Advanced very-high-resolution radiometer) imagery for mapping sub-pixel proportion of land cover classes. The performance of ANN was found superior to that of the other two methods. However, this superiority was balanced against the requirement of accurate co-registered training data with the input imagery. FCM and mixture model doesn't require training data, and hence it was concluded that the choice of technique for mapping depends on that application in hand. Bastin (1997) compared the performance of different supervised soft classifiers to find out about the quantity and location of land use land cover in a set of coarse images. Fuzzy c-means, Linear Mixture model and probability values from maximum likelihood classification were used to obtain quantitative maps for showing the concentration of various land cover classes across an area. The FCM algorithm was found to perform best among the three in locating and handling mixed pixels.

FCM algorithm has been modified for supervised classification of remote sensing images (Foody, 2000; Ibrahim, Arora, & Ghosh, 2005; J. Zhang & Foody, 2001). Supervised classification follows three stages (J. Zhang & Foody, 2001). Training samples representing the classes under consideration are collected and descriptive statistics for each class is calculated, pixels are allocated to the classes based on the similarity measured with the derived class descriptions, and accuracy of the classification is assessed with the test samples for which reference or ground information is available. As a result, converting unsupervised FCM to supervised FCM results in a single step process where fuzzy membership values of pixels for each of the class is calculated using the class centroids determined from the training data.

Fully-fuzzy classification approaches have also been investigated by various researchers where the fuzziness in the ground data is accounted during classification process in addition to the fuzziness in the remotely sensed images (Foody, 1995; J. Zhang & Foody, 1998, 2001). This implies that the mixed pixels are accommodated during training and testing of these classifiers.

2.3. Possibilistic *c*-Means

A possibilistic approach to clustering was presented by Krishnapuram & Keller (1993) in Possibilistic c-means (PCM) algorithm. PCM is an iterative clustering algorithm similar to the FCM algorithm in which fuzzy partitions are found by minimisation of the objective function. However, the membership constraint

in FCM that forces the sum of membership value of a pixel across a class to be one is relaxed in PCM algorithm. The membership value of a pixel for a class is independent of its membership value in other classes. Because of this, the membership of pixels in a class represents the "degree of belonging or compatibility" in PCM as opposed to "degree of sharing" in FCM (Krishnapuram & Keller, 1993). PCM is said to be naturally immune to noise and outliers since the membership function corresponds to the concept of typicality and hence noise points will have a low degree of compatibility in all clusters (Krishnapuram & Keller, 1993).

Although PCM is not as popular as FCM, it has been explored for remote sensing image classification (Chawla, 2010; Foody, 2000; Ibrahim, 2004; Ibrahim et al., 2005; Kumar, 2007; Kumar, Ghosh, & Dadhwal, 2006; Sharma, 2018). There might often be areas of classes that could be excluded while training the classifiers. Foody (2000) proved that the performance of PCM is not affected by the presence of such excluded areas or untrained classes and thus it outperforms FCM in such scenarios. Since untrained classes are commonly encountered during supervised classification, PCM was suggested to be used in addition to or instead of FCM classifier to enhance the extraction of information from remote sensing imagery. Ibrahim et al. (2005) studied the performance of Maximum Likelihood Classifier (MLC), FCM classifier, and the PCM classifier in the supervised mode for land use land cover classification. PCM produced the highest accurate land cover maps when uncertainty existed in the dataset. In the study carried out by Kumar et al.(2006), it was seen that PCM gave better accuracy compared to FCM when Euclidean Norm was used.

Chapter 3 presents further details about FCM and PCM classifiers.

2.4. Spatio - contextual information for classification

Traditional per pixel classification approaches may result in salt and pepper noise in the classification maps due to ambiguities caused due to landscape complexities and spectral confusions. Incorporation of ancillary data such as spatial context might help in correcting classification confusions(Lu & Weng, 2007). Context refers to the probability of occurrence of a pixel (or a group of pixels) to a class based on the nature of pixels in the remainder of the scene (Tso & Mather, 2009). Generally, contextual classifiers are based on the use of local windows as they assume there is no relationship between the land cover types at a distance and only the pixels within the given distances affects the classification (Gurney & Townshend, 1983). The concept of moving windows are used during pre-processing, classification process or post-classification steps, where spatio-contextual information is extracted from immediate neighbourhoods of pixels by imposing a search window, computing the contextual parameter within the window and assigning that value to the central pixel (Stuckens et al., 2000). Although 3×3 window is the most applied window size, larger windows might be suitable for small pixel size and/or larger land cover entities.

In order to reduce the error in pixel-wise classification, contextual information can be exploited during pre-processing or post-processing steps. Noise is removed in the raw image in case of pre-processing approach while noise is cleared in the classified data in case of post-processing approach (Gurney & Townshend, 1983). During pre-processing, median filter or mean filter is used to exploit the spatial context. The use of filter helps in reducing noise in the image which leads to inconsistent labels in the output; or they help in improving the degree of homogeneity among brightness values in the neighbourhood which allows similar labels to be allocated to the neighbouring pixels (Richards & Jia, 2012). A straightforward way to use contextual information in the post-processing approach is the use of

majority filters (Gurney, 1981; Tso & Mather, 2009). It is applied for classification refinements by altering the label of the centre pixel by assigning to it the label of the majority of the pixels within the window. Though the accuracy of the classification is not very impressive on using this method, there is an improvement in classification results. Both pre-processing and post-processing techniques have certain disadvantages (Liew et al., 2000). In the pre-processing approach, unwanted smoothing may occur and important image details may be lost which cannot be reversed during classification. Post-processing approach works on the preliminary classifier output and, hence no information is available on how close the pixels are to the classes other than the class to which they are classified.

Two widely used methods to integrate contextual information in classification are Probabilistic Relaxation and Markov Random field (MRF) (Melgani & Serpico, 2002). These methods exploit the contextual information during the classification process as opposed to the pre-processing approach or the post-processing approach. Probabilistic Relaxation is an iterative process in which the labelling probabilities of pixels are altered depending upon the labelling probabilities of neighbouring pixels. It is assumed that the classification of image is already carried out using spectral data alone and initial labelling probabilities of pixels are available. The neighbourhood of each pixel is identified and the neighbourhood function, which is a measure of influence exerted by the neighbouring pixels on the labelling of the pixel, is used along with the initial labelling probability of the pixel for calculating new labelling probability of the pixel. Thus the spectral data (in initial labelling probabilities) and spatial context (by using neighbourhood functions) is used for classification(Richards & Jia, 2012).

MRF is a part of probability theory and has been successfully used to model spatial context in the field of pattern recognition. It is based on the assumption that there is coherence over space and the physical properties within the neighbourhood do not change drastically (Dutta, 2009). MRF models context as prior probability using smoothness assumption. The prior probability is combined with conditional probability and MAP (Maximum A Posteriori) classification is performed (Tso & Mather, 2009). MRF has been used successfully to model spatial context. However, the implementation of MRF involves challenging computational complexities which are considered difficult by many remote sensing scientists (Li et al., 2014; Moser, Serpico, & Benediktsson, 2013).

Some of the works investigated by researchers to incorporate spatial contextual information in fuzzy classification are discussed below.

Dutta (2009) used MRF to model contextual information and incorporated it with FCM algorithm resulting in contextual FCM classification algorithm. To estimate conditional energy function FCM was used, and to model contextual information as prior energy, MRF was used. The contextual FCM algorithm was found to produce improved classification accuracy compared to the conventional FCM classification approach.

Chawla (2010) developed a contextual fuzzy classification algorithm using MRF with PCM as the base classifier. Standard regularisation and Discontinuity Adaptive smoothness regularizers were employed separately to model prior energy in this research. Contextual PCM classifiers with Standard regularization and Discontinuity Adaptive (DA) smoothness prior were found to outperform the conventional PCM classifier in terms of accuracy. However, contextual PCM with DA prior was found to preserve edges.

Pham & Prince (1999) came up with an iterative Adaptive FCM (AFCM) algorithm which contains a multiplier field term that allows centroids of each class to vary across the image. First order and second order regularisation terms are used to regularise the multiplier field term to be slowly varying and smooth. The results of AFCM algorithms were found to be promising in general but doubtful in the presence of extreme noise.

Liew et al. (2000) modified the FCM objective function with a new adaptive dissimilarity index that takes into account the neighbouring pixels' influence on the central pixel within the 3x3 window around the pixel. The normalized distance which is computed for each pixel in FCM was replaced with the weighted sum of distances of pixels in the neighbourhood.

Pham (2001) introduced a spatial penalty term in the objective function of FCM. Unlike the method followed by Liew et al. (2000), this algorithm penalises the FCM objective function to constrain the behaviour of membership functions, and thus the centroid computations remain same as that in conventional FCM. The spatial penalty term discourages unlikely configurations in the membership functions and forces membership values to be similar within the neighbourhood of each pixel.

Laha, Pal, & Das (2006) proposed a two-stage scheme for land use land cover classification with the inclusion of contextual information. In the first stage, a non-contextual fuzzy rule-based classifier is developed with small training sample, and possibilistic label vector is generated as output for each pixel. In the second stage, classification decision is made for each pixel by aggregating the fuzzy decision rules in the 3x3 spatial neighbourhood of the pixel. Four different aggregation methods, three based on evidence theory and one based on the fuzzy k-NN rule, that uses information from the spatial neighbourhood are investigated for contextual decision making.

Ahmed et al. (2002) proposed an approach in which spatial neighbourhood term was introduced into the conventional FCM algorithm which forced the labelling of the neighbouring pixels to influence the labelling of each pixel. The algorithm was later called FCM-S (Fuzzy c-means spatial constraints) algorithm (S. Chen & Zhang, 2004). The computation of neighbourhood term in this approach was found to be time-consuming, and this was overcome in FCM_S1 and FCM_S2 algorithms (S. Chen & Zhang, 2004). FCM_S1 and FCM_S2 use mean filter and median filter respectively to simplify the computation of neighbourhood term. However, the output performance of FCM-S, FCM-S1 and FCM-S2 algorithms depended on an empirically selected parameter which controlled the robustness to noise and effectiveness of preserving the image details.

Krinidis & Chatzis (2010) proposed FLICM (Fuzzy local information c-means) algorithm in which a new fuzzy factor, independent of any empirically selected parameter, was introduced in FCM objective function. The fuzzy factor automatically determines the spatial and spectral relationship and hence is independent of the type of noise present in the image.

Adaptive fuzzy local information c-means algorithm (ADFLICM), similar to FLICM algorithm was proposed by H. Zhang et al. (2017). A fuzzy local similarity measure was introduced in this approach which uses spatial attraction model to describe the relationship between pixels within a local window. The similarity measure in the algorithm ensures each pixel to be influenced by its neighbouring pixels as well as its own features during classification. The ADFLICM is claimed to be advantageous over the other

mentioned algorithms as it is useful for retaining edges while removing noise because of the similarity measure introduced.

FCM-S (Ahmed et al., 2002), FLICM (Krinidis & Chatzis, 2010) and ADFLICM (H. Zhang et al., 2017) are explored in detail for satellite image classification in this research. Further details of these algorithms are explained in chapter 3.

2.5. Accuracy Assessment

Accuracy is the level of agreement between the labels assigned by the classifier and the labels accepted as true based on ground truth or reference data, known as the test data, collected by the user. Assessment of accuracy by apt methods gives the measure of classification performance. One general method for assessing accuracy in case of hard classifiers is the use of a confusion matrix or error matrix. The summary of agreement and disagreement of pixels in the classifier output and reference data is given in the form of a square matrix. This matrix is then used to derive accuracy measures such as overall accuracy, producer's accuracy, user's accuracy. Since the overall accuracy, producer's accuracy and user's accuracy do not use the information from the whole matrix but only the information from either diagonals, columns, or rows of the matrix, Cohen (1960) proposed a new index called Kappa coefficient. Kappa coefficient is statistically more accepted as it takes into consideration the chance allocation of labels (Tso & Mather, 2009).

In soft classification, the output represents the proportion of two or more classes in a pixel. For the assessment of soft classified outputs, there are no standard methods available as in the case of hard classification (Chawla, 2010). The output from the fuzzy classifiers can be hardened and conventional accuracy assessment methods like confusion matrix or kappa coefficient may be applied. However, there might be loss of information while converting soft outputs to hard outputs (Byju, 2015). Euclidean and L1 distance (Foody & Arora, 1996), entropy (Dehghan & Ghassemian, 2006), cross-entropy (Foody, 1995; J. Zhang & Foody, 2001), root mean square error (RMSE) (Foody, 2000) and correlation coefficient (Maselli, Rodolfi, & Conese, 1996) are few of the classification methods that are used for accuracy assessment when both the classified data and reference data are soft. These methods, however, are considered as indirect methods for accuracy assessment as they do not represent the actual value as in the case of traditional confusion matrix (Ibrahim, 2004). Binaghi, Brivio, Ghezzi, & Rampini (1999) put forward an approach called Fuzzy Error Matrix (FERM) for assessing the accuracy of soft classification. FERM is similar to the confusion matrix with the exception that non-negative real numbers can be the elements of the matrix. The details of FERM, RMSE and Mean Membership Difference methods are explained in chapter 4.

This chapter provided the necessary background concepts and related work that was essential in understanding and implementing the fuzzy local spatial information classifiers. It also provided an insight into the methods for accuracy assessment for soft outputs. Chapter 3 explains the mathematical concepts behind the base classifiers (FCM and PCM) and the advanced FCM based local spatial information classifiers and the PCM based local spatial information classifiers.

3. CLASSIFICATION APPROACHES

Classification approaches can be defined on the basis of training samples used (supervised and unsupervised), assumption of parameters on data (parametric and non-parametric), pixel information used (per pixel classifier, sub-pixel classifier, object-oriented classifier), number of outputs for each spatial element (hard and soft classifier) and on the basis of spatial information (spectral classifier, contextual classifier, spectral-contextual classifier) (Kamavisdar, Saluja, & Agrawal, 2013). This chapter describes the mathematical details of the conventional soft classifiers (FCM and PCM) and the six spectral-contextual classifiers implemented and explored in this research.

3.1. Fuzzy *c*-Means classifier

Fuzzy *c*-means (Bezdek et al., 1984) is originally an unsupervised clustering algorithm that creates fuzzy partitions of given data (remote sensing images) using fuzzy memberships. The fuzzy membership value denotes the similarity a data point (pixel) shares with the cluster/class. Each pixel in the image will have a membership in each cluster, and the membership value close to one denotes a high degree of similarity of the pixel to the cluster while membership value close to zero denotes a low degree of similarity of the pixel to the cluster.

The optimal fuzzy clusters are obtained by minimising the FCM objective function given by the equation (3.1) subject to the constraints defined in (3.2) (Ibrahim, 2004; Tso & Mather, 2009).

$$J_{FCM}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C} \mu_{ij}^{m} D\,(x_j, v_i) \tag{3.1}$$

$$\mu_{ij} \in [0,1] \text{ for all i and j} \tag{3.2a}$$

$$\sum_{j=1}^{N} \mu_{ij} > 0 \text{ for all i} \tag{3.2b}$$

$$\sum_{i=1}^{C} \mu_{ij} = 1 \text{ for all j} \tag{3.2c}$$

Where,

V is the matrix of cluster centres with its elements denoted by v_i;
v_i is the mean vector for cluster i;
U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;
μ_{ij} is the membership value of j^{th} pixel for cluster i;
N is the total number of pixels;
C is the number of clusters;
m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;
$D(d_{ij}^2)$ is the squared distance between the j^{th} pixel value x_j and cluster mean v_i;

The distance metrics (D or d_{ij}^2) used here is Euclidean. This can be replaced by Mahalanobis distance for the calculation of which fuzzy covariance matrix is required (Tso & Mather, 2009).

The fuzzy weight (*m)* in equation (3.1) controls the degree of fuzziness in the output. As *m* tends to infinity the outputs produced become completely fuzzy, and as *m* tends to unity the outputs produced become increasingly hard or crisp. Seeing that the choice of optimal value of *m* in various researches differs (Cannon et al., 1986; C. F. Chen & Lee, 2001; Pal & Bezdek, 1995), an experimental strategy is adopted in this study for optimising the value of *m* for high performance oriented segmentation.

Equation (3.1) has been found to be optimal for U and V in equations (3.3) and (3.4) respectively.

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{\frac{1}{m-1}}} \tag{3.3}$$

$$v_i = \frac{\sum_{j=1}^{N} \mu_{ij}^m \, x_j}{\sum_{j=1}^{N} \mu_{ij}^m} \tag{3.4}$$

The FCM classifier assumes that all the information classes have been specified and each pixel can be described entirely by its membership values in the defined classes. From (3.3), it can be seen that the membership value of a pixel in a class is dependent on the distance of the pixel with the class prototype as well as the distance of the pixel to all the remaining class prototypes. Thus, the membership values produced by FCM are analogous to the relative proportions of available classes (Foody, 2000;Tso & Mather, 2009).

FCM is essentially unsupervised clustering algorithm. It is executed iteratively by updating the equations in (3.3) and (3.4) until their values between successive iterations converge. However, FCM, PCM and its variants may also be adapted for supervised classification (Foody, 2000). In such case, the cluster centroids (V) are estimated from the input training samples in place of equation (3.4), and the membership values are calculated in a single step with equation (3.3).

3.2. Possibilistic *c*-Means classifier

Krishnapuram & Keller (1993) observed that due to the probabilistic constraint in (3.2c) of the FCM algorithm, the performance of FCM and its derivatives in noisy environments is compromised. Possibilistic *c*-means (PCM) algorithm, which works on the possibilistic theory, was proposed in which probabilistic constraint (3.2c) of the FCM is relaxed. The possibility of a data point (pixel) to belong to a class/cluster or the typicality of the data point in the cluster is represented by the membership values obtained with PCM. The membership value of a data point in a class is independent of the membership value of the data point in other clusters, and thus it is absolute unlike the relative membership value in the FCM algorithm.

The objective function of the FCM is modified to formulate the PCM objective function as in (3.5) subject to constraints in (3.6).

$$J_{PCM}(U, V) = \sum_{j=1}^{N} \sum_{i=1}^{C} \mu_{ij}^m D\left(x_j, v_i\right) + \sum_{i=1}^{C} \eta_i \sum_{j=1}^{N} (1 - \mu_{ij})^m \tag{3.5}$$

$$\mu_{ij} \in [0,1] \quad \text{for all i and j} \tag{3.6a}$$

$$\sum_{j=1}^{N} \mu_{ij} > 0 \text{ for all i} \quad (3.6b)$$

$$\max_{i} \mu_{ij} > 0 \text{ for all j} \quad (3.6c)$$

Where,

V is the matrix of cluster centres with its elements denoted by v_i;

v_i is the mean vector for cluster i;

U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;

μ_{ij} is the membership value of j^{th} pixel for cluster i;

N is the total number of pixels;

C is the number of clusters;

m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;

$D(d_{ij}^2)$ is the squared distance between the j^{th} pixel value x_j and cluster mean v_i. Euclidean distance is considered in this study;

η_i is a regularising term (a suitable positive number) added to avoid the trivial solution of all the membership values equal to zero. The first term in the objective function forces the distances from the feature vector to the cluster prototype to be low as possible and the second term forces the membership value μ_{ij} to be as large as possible. That is, the pixels with a high degree of typicality to a cluster (possibility of a pixel to belong to that cluster/class) will have high membership value in the cluster while pixels with a low degree of typicality will have low membership in all the clusters.

Membership function that optimises the objective function (3.5) is given in (3.7).

$$\mu_{ij} = \frac{1}{1 + \left(\frac{d_{ij}^2}{\eta_i}\right)^{\frac{1}{m-1}}} \quad (3.7)$$

η_i is the "bandwidth" or "resolution" or "scale" parameter (Krishnapuram & Keller, 1996) and controls the shape and size of the cluster/class. It is the distance at which the membership value of a pixel in the class becomes 0.5, and hence it should be selected depending on the possibility distribution for each cluster (distribution of pixels in a cluster). If all the clusters are identical, η_i can be set to the same value (Ibrahim, 2004). The definition of η_i in (3.8), which is proportional to the average fuzzy intra-cluster distance for cluster i, has been found to work well. Generally, the value of K in (3.8) is chosen to be 1 (Krishnapuram & Keller, 1993).

$$\eta_i = K \frac{\sum_{j=1}^{N} \mu_{ij}^m d_{ij}^2}{\sum_{j=1}^{N} \mu_{ij}^m} \quad (3.8)$$

FCM membership values obtained with (3.3) are used in this research as initial membership values for the calculation of η_i in (3.8).

The interpretation of *m* in PCM is different from that in FCM. Increasing *m* values in FCM signifies increased sharing of the pixel among all the available clusters, while increasing *m* values in PCM signifies increased possibility of all the pixels to belong to a given cluster.

The PCM algorithm, similar to the FCM algorithm, iteratively updates the membership values and cluster centres until the values converge. The formula for updating the cluster centre for PCM remains same as that of FCM which is defined in (3.4).

As the land use/land cover classes of interest considered for this study are known, the clustering algorithms (FCM, PCM, FCM based local spatial information classifiers and the PCM based local spatial information classifiers) implemented in this research follow supervised classification approach. Hence the formulas to calculate cluster centres are not included in the following sections as the cluster centres are computed from the input training data.

3.3. FCM based local spatial information classifiers

Spatial contextual information has been incorporated into conventional FCM classification algorithms by various approaches, a few of which are discussed in Section 2.4 of the thesis. This section describes the details of three different FCM based spatio-contextual classification algorithms FCM-S(Ahmed et al., 2002), FLICM (Krinidis & Chatzis, 2010), ADFLICM(H. Zhang et al., 2017) algorithms that are implemented and explored in this study. Spatial information has been incorporated in these algorithms by relatively simpler approach of modifying the FCM objective function to include the spatial constraint term in a straightforward manner, which was the motivation behind them to be incorporated with PCM classifier as well in this research. Section 3.4 describes the details of the three corresponding PCM based spatio-contextual classification algorithms developed.

3.3.1. Fuzzy *c*-Means Spatial constraint (FCM-S)

Ahmed et al. (2002) proposed a modified Fuzzy *c*-means algorithm in which labelling of a pixel is influenced by the labels in the immediate neighbourhood for compensating intensity inhomogeneities in the images. The objective function of the standard FCM function is modified to include a spatial constraint term. The neighbourhood effect in the algorithm acts a regularizer and forces piecewise-homogeneous labelling.

The objective function of FCM_S is defined in (3.9).

$$J_{FCMS}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C} \mu_{ij}^{m} D\left(x_j, v_i\right) + \frac{a}{N_R}\sum_{j=1}^{N}\sum_{i=1}^{C} \mu_{ij}^{m} \sum_{r\in N_j} D(x_r, v_i) \quad (3.9)$$

Where,

V is the matrix of cluster centres with its elements denoted by v_i;

v_i is the mean vector for cluster i;

U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;

μ_{ij} is the membership value of j^{th} pixel for cluster i;

N is the total number of pixels;

C is the number of clusters;

m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;

N_j is the set of neighbour pixels falling into the window around the pixel j;

N_R is the cardinality of N_j;

a is the parameter that controls the effect from the neighbourhood;

$D(x_j, v_i)$ (d_{ij}^2) is the squared distance between the j^{th} pixel value x_j and cluster mean v_i ;
$D(x_r, v_i)$ (d_{ir}^2) is the squared distance between the pixel value x_r and the cluster mean v_i where x_r represents the r^{th} neighbour in the neighbouring window of x_j ;

The calculation of membership values is performed using (3.10).

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{d_{ij}^2 + \frac{a}{N_R} \sum_{r \in N_i} d_{ir}^2}{d_{kj}^2 + \frac{a}{N_R} \sum_{r \in N_i} d_{kr}^2} \right)^{\frac{1}{m-1}}} \tag{3.10}$$

The neighbourhood term is calculated for each pixel and is subsequently used in estimating the pixel's membership value for each class.

3.3.2. **Fuzzy Local Information *c*-Means (FLICM)**

Krinidis & Chatzis (2010) proposed an algorithm which uses a fuzzy local grey level and spatial similarity measure for noise insensitiveness and image detail preservation. Unlike the FCM-S algorithm, FLICM algorithm is independent on empirically adjusted parameters (parameter *a* in FCM-S) to control the effect from the neighbourhood. FLICM uses fuzzy factor which automatically determines the relationship between the spatial and grey level information in the image. The spatial distance between the centre pixel and the neighbouring pixels in the local window is used to control the influence of the neighbouring pixels. The fuzzy factor of j^{th} pixel for i^{th} class is given by (3.11).

$$G_{ij} = \sum_{r \in N_j} \frac{1}{ed_{jr} + 1} (1 - \mu_{ir})^m D(x_r, v_i) \tag{3.11}$$

Where,
$D(x_r, v_i)$ (d_{ir}^2) is the squared distance between the pixel value x_r and the cluster mean v_i where x_r represents the r^{th} neighbour of x_j ;
ed_{jr} is the spatial Euclidean distance between the centre pixel j and the neighbouring pixels r;
μ_{ir} is the degree of membership of the r^{th} neighbour pixel (of centre pixel j) in cluster i;
m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;
N_j is the set of neighbour pixels falling into the window around the pixel j;

The FLICM algorithm is initialized using the FCM algorithm, which means that the FCM output membership values are used for the calculation of G_{ij}.

The objective function of FLICM algorithm is given in (3.12)

$$J_{FLICM}(U, V) = \sum_{j=1}^{N} \sum_{i=1}^{C} \mu_{ij}^m D(x_j, v_i) + G_{ij} \tag{3.12}$$

Where,
V is the matrix of cluster centres with its elements denoted by v_i;
v_i is the mean vector for cluster i;
U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;

μ_{ij} is the membership value of j^{th} pixel for cluster i;
N is the total number of pixels;
C is the number of clusters;
$D(x_j, v_i)$ (d_{ij}^2) is the squared distance between the j^{th} pixel value x_j and cluster mean v_i;

The member membership value of FLICM is calculated by (3.13)

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{d_{ij}^2 + G_{ij}}{d_{kj}^2 + G_{kj}} \right)^{\frac{1}{m-1}}} \tag{3.13}$$

The fuzzy factor, which adaptively determines the spatial information from a pixel's neighbourhood defined within the local window, is used to calculate the membership value of FLICM algorithm. The resulting membership values thus incorporate grey level information as well as the adaptively changing spatial information with respect to every pixel in the image.

3.3.3. **Adaptive Fuzzy Local Information *c*-Means (ADFLICM)**

ADFLICM algorithm, proposed by H. Zhang et al.(2017), uses fuzzy similarity measure based on spatial attraction model. The weighing factor for each of the neighbourhood is adaptive to the image content without any empirically set parameters. The new fuzzy local similarity measure provides a balance between "insensitiveness to noise and reduction of edge blurring artefact simultaneously" (H. Zhang et al., 2017) as it influences the classification of a pixel to be affected by its neighbouring pixels and its own features simultaneously.

Spatial attraction models are used to describe the spatial correlation between the pixels in the image. In ADFLICM, FCM objective function is modified to include local spatial and grey level information. The Spatial Attraction (SA) between two pixels *j* and *r* with respect to the class *i* is described in (3.14).

$$SA_{jr} = \sum_{r \in N_j} \frac{\mu_{ij} \times \mu_{ir}}{ed_{jr}^2} \tag{3.14}$$

Where,
N_j is the set of neighbour pixels falling into the window around the pixel j;
r denotes the neighbour pixels of j in N_j;
μ_{ij} is the degree of membership of the j^{th} pixel in cluster i;
μ_{ir} is the degree of membership of the r^{th} neighbour pixel (of centre pixel j) in cluster i;
ed_{jr} is the spatial distance between the centre pixel j and the neighbouring pixels r. In this research Euclidean distance is selected;

The new local similarity measure defined in (3.15) utilises the spatial attraction model in (3.14). Initial membership values, obtained from the output of the FCM algorithm, are used for the calculation of SA_{jr}.

$$S_{jr} = \begin{cases} SA_{jr} j \neq r \\ 0 \;\; j = r \end{cases} \tag{3.15}$$

The objective function for ADFLICM is given in (3.16).

$$J_{ADFLICM}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C} \mu_{ij}^{m}\left[D(x_j, v_i) + \frac{1}{N_R}\sum_{r\in N_j}(1 - S_{jr})D(x_r, v_i)\right] \quad (3.16)$$

Where,
V is the matrix of cluster centres with its elements denoted by v_i;
v_i is the mean vector for cluster i;
U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;
μ_{ij} is the membership value of j^{th} pixel for cluster i;
N is the total number of pixels;
C is the number of clusters;
m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;
N_j is the set of neighbour pixels falling into the window around the pixel j;
N_R is the cardinality of N_j;
$D(x_j, v_i)$ (d_{ij}^2) is the squared distance between the j^{th} pixel value x_j and cluster mean v_i ;
$D(x_r, v_i)$ (d_{ir}^2) is the squared distance between the pixel value x_r and the cluster mean v_i where x_r represents the r^{th} neighbour of x_j ;

Membership value is calculated using (3.17)

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{d_{ij}^2 + \frac{1}{N_R}\sum_{r\in N_i}(1-S_{jr})d_{ir}^2}{d_{kj}^2 + \frac{1}{N_R}\sum_{r\in N_i}(1-S_{jr})d_{kr}^2}\right)^{\frac{1}{m-1}}} \quad (3.17)$$

3.4. PCM based local spatial information classifiers

The base classifier in FCM-S, FLICM and ADFLICM is modified with PCM to develop PCM-S, PLICM and ADPLICM classification algorithms respectively. The mathematical details of these classifiers are explained in this section. The objective functions for the algorithms are formulated and are minimised in a fashion similar to the standard FCM/PCM classification algorithms. The parameter value (μ_{ij}) for which the objective function is optimized is then estimated.

3.4.1. Possibilistic c-Means Spatial constraint (PCM-S)

The objective function of the PCM algorithm is modified to develop PCM-S algorithm (3.18). The spatial constraint term is added to the objective function of the standard PCM algorithm.

$$J_{PCMS}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C} \mu_{ij}^{m} D(x_j, v_i) + \sum_{i=1}^{C}\eta_i\sum_{j=1}^{N}(1-\mu_{ij})^m + \frac{a}{N_R}\sum_{j=1}^{N}\sum_{i=1}^{C}\mu_{ij}^{m}\sum_{r\in N_j} D(x_r, v_i) \quad (3.18)$$

Where,
V is the matrix of cluster centres with its elements denoted by v_i;
v_i is the mean vector for cluster i;
U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;
μ_{ij} is the membership value of j^{th} pixel for cluster i;

N is the total number of pixels;
C is the number of clusters;
m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;
N_j is the set of neighbour pixels falling into the window around the pixel j;
N_R is the cardinality of N_j;
a is the parameter that controls the effect from the neighbourhood;
$D(x_j, v_i)$ (d_{ij}^2) is the squared distance between the j^{th} pixel value x_j and cluster mean v_i ;
$D(x_r, v_i)$ (d_{ir}^2) is the squared distance between the pixel value x_r and the cluster mean v_i where x_r represents the r^{th} neighbour in the neighbouring window of x_j ;
η_i value is calculated using (3.8) as in the case of PCM algorithm;

In order to obtain the membership equation, the objective function has to be minimised with respect to U. As the rows and columns of U are independent of each other, minimising the objective function with respect to U is equivalent to minimising the individual objective function in (3.19) with respect to μ_{ij} provided that the resulting membership value lies in the interval [0,1] (Krishnapuram & Keller, 1993).

$$J_{PCMS}^{ij}(\mu_{ij}, v_i) = \mu_{ij}^m d_{ij}^2 + \eta_i(1-\mu_{ij})^m + \frac{a}{N_R}\mu_{ij}^m d_{ir}^2 \tag{3.19}$$

Partial differentiating (3.19) with respect to μ_{ij} and setting it to zero yields the following equation for membership value.

$$\mu_{ij} = \frac{1}{1+\left(\frac{d_{ij}^2 + \frac{a}{N_R}\sum_{r\in N_i} d_{ir}^2}{\eta_i}\right)^{\frac{1}{m-1}}} \tag{3.20}$$

3.4.2. **Possibilistic Local Information *c*-Means (PLICM)**

PLICM algorithm is formulated as in (3.21).

$$J_{PLICM}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C}\mu_{ij}^m\left[D(x_j, v_i) + G_{ij}\right] + \sum_{i=1}^{C}\eta_i\sum_{j=1}^{N}(1-\mu_{ij})^m \tag{3.21}$$

Where,
V is the matrix of cluster centres with its elements denoted by v_i;
v_i is the mean vector for cluster i;
U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;
μ_{ij} is the membership value of j^{th} pixel for cluster i;
N is the total number of pixels;
C is the number of clusters;
m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;
$D(x_j, v_i)$ (d_{ij}^2) is the squared distance between the j^{th} pixel value x_j and cluster mean v_i ;
η_i is estimated using the formula in (3.8).

The fuzzy factor G_{ij}, is calculated using (3.11) similar to the FLICM algorithm. PCM is executed to get the initial membership values used for calculating G_{ij} and η_i.
The local minima of J_{PLICM} (U, V) with respect to U is attained for the membership in (3.22)

$$\mu_{ij} = \frac{1}{1 + \left(\frac{d_{ij}^2 + G_{ij}}{\eta_i}\right)^{\frac{1}{m-1}}} \tag{3.22}$$

3.4.3. **Adaptive Possibilistic Local Information *c*-Means (ADPLICM)**

ADPLICM classifier uses the similarity measure with spatial attraction model as defined in the case of ADFLICM algorithm. Spatial attraction (SA_{jr}) and similarity measure (S_{jr}) equations for ADPLICM are same as that of ADFLICM algorithm which corresponds to (3.14) and (3.15). The membership values to calculate SA_{jr} and η_i are initialised using the standard PCM algorithm.

The PCM objective function modified to develop ADPLICM algorithm is defined in (3.23).

$$J_{ADPLICM}(U, V) = \sum_{j=1}^{N}\sum_{i=1}^{C} \mu_{ij}^{m}\left[D(x_j, v_i) + \frac{1}{N_R}\sum_{r\in N_j}(1 - S_{jr})D(x_r, v_i)\right] + \sum_{i=1}^{C}\eta_i\sum_{j=1}^{N}(1-\mu_{ij})^m \tag{3.23}$$

Where,
V is the matrix of cluster centres with its elements denoted by v_i;
v_i is the mean vector for cluster i;
U is a C x N fuzzy partition matrix representing the membership values (μ_{ij}) of each class per pixel;
μ_{ij} is the membership value of j^{th} pixel for cluster i;
N is the total number of pixels;
C is the number of clusters;
m is the fuzzy weight, which controls the level of fuzziness and its value lie between 1 and infinity;
N_j is the set of neighbour pixels falling into the window around the pixel j;
N_R is the cardinality of N_j;
$D(x_j, v_i)$ (d_{ij}^2) is the squared distance between the j^{th} pixel value x_j and cluster mean v_i ;
$D(x_r, v_i)$ (d_{ir}^2) is the squared distance between the pixel value x_r and the cluster mean v_i where x_r represents the r^{th} neighbour in the neighbouring window of x_j ;

The final membership values are estimated using (3.24).

$$\mu_{ij} = \frac{1}{1 + \left(\frac{d_{ij}^2 + \frac{1}{N_R}\sum_{r\in N_i}(1-S_{jr})d_{ir}^2}{\eta_i}\right)^{\frac{1}{m-1}}} \tag{3.24}$$

The mathematical concepts of the classifiers implemented as part of the study has been explained in this chapter. Chapter 4 gives the details about the datasets used for experimenting the classification approaches mentioned in this chapter and also the resources used for the implementation of these classifiers. Further, chapter 5 describes the overall methodology that was followed for implementing and analysing the performance of the classifiers developed.

4. STUDY AREA AND RESOURCES USED

This chapter gives details about the study area, datasets and other resources used in the research.

4.1. Study area and datasets used

The study area chosen for the research is located in the Haridwar district of Uttarakhand State in India. The area extends from latitudes 29°48'48"N to 29°53'14"N and longitudes 78°9'56"E to 78°14'43"E. The area is diverse in terms of LULC classes present.

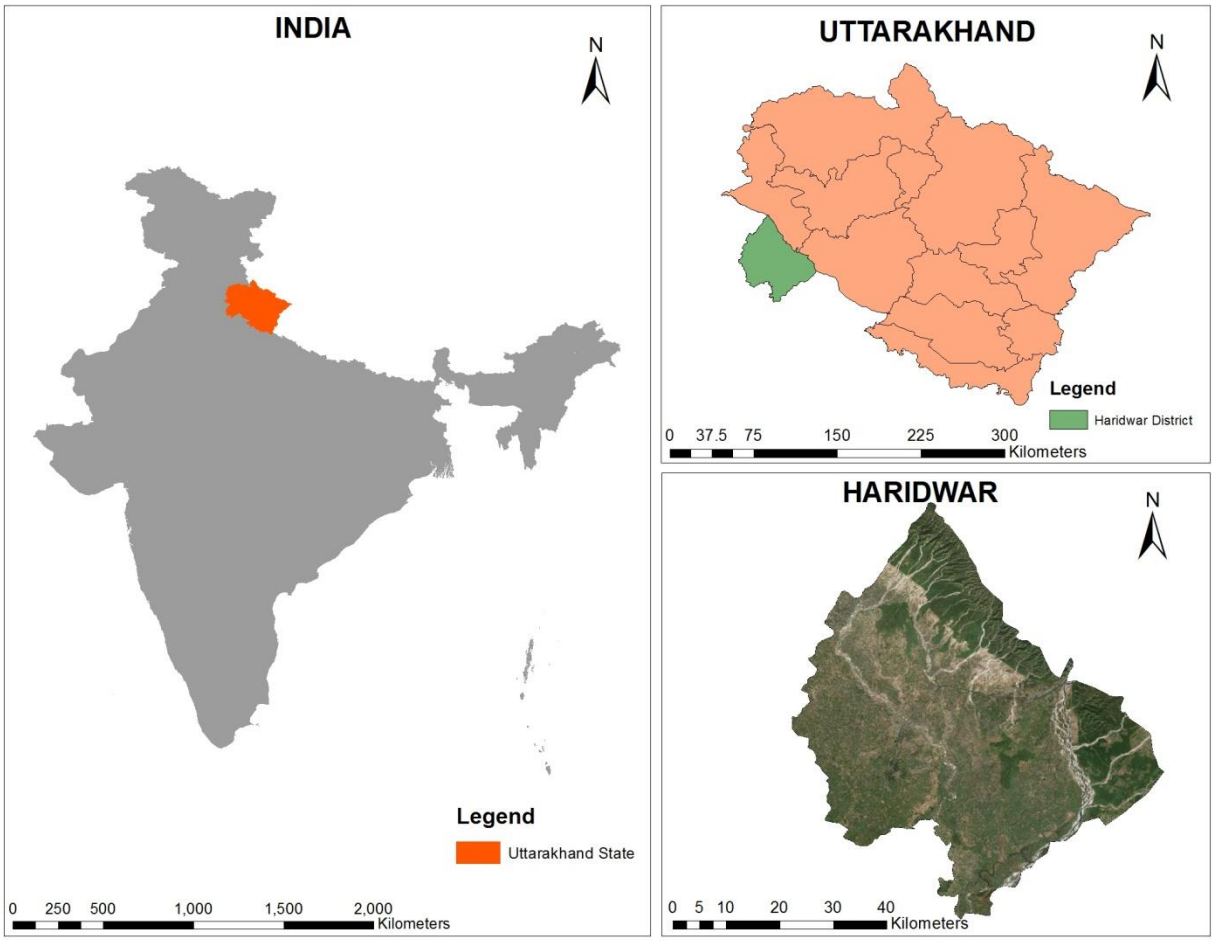


Figure 4.1: Study Area

The primary motive behind using the chosen study area was that the field data with identified LULC classes and the datasets captured around the same time for classification and accuracy assessment was available. The LULC classes present in the study area were Dense Forest, Eucalyptus plantation, Grassland, Riverine sand, Water and Wheat plantation.

Remote sensing data from Landsat-8 sensor at 30m spatial resolution and Formosat-2 sensor at 8m spatial resolution have been used in this study. The classification algorithms have been executed on Landsat-8

data, and Formosat-2 image has been used to create reference map for accuracy assessment. All the bands of the images were used for classification and reference map generation. Figure 4.2 shows the satellite images of the study area with the six LULC classes.

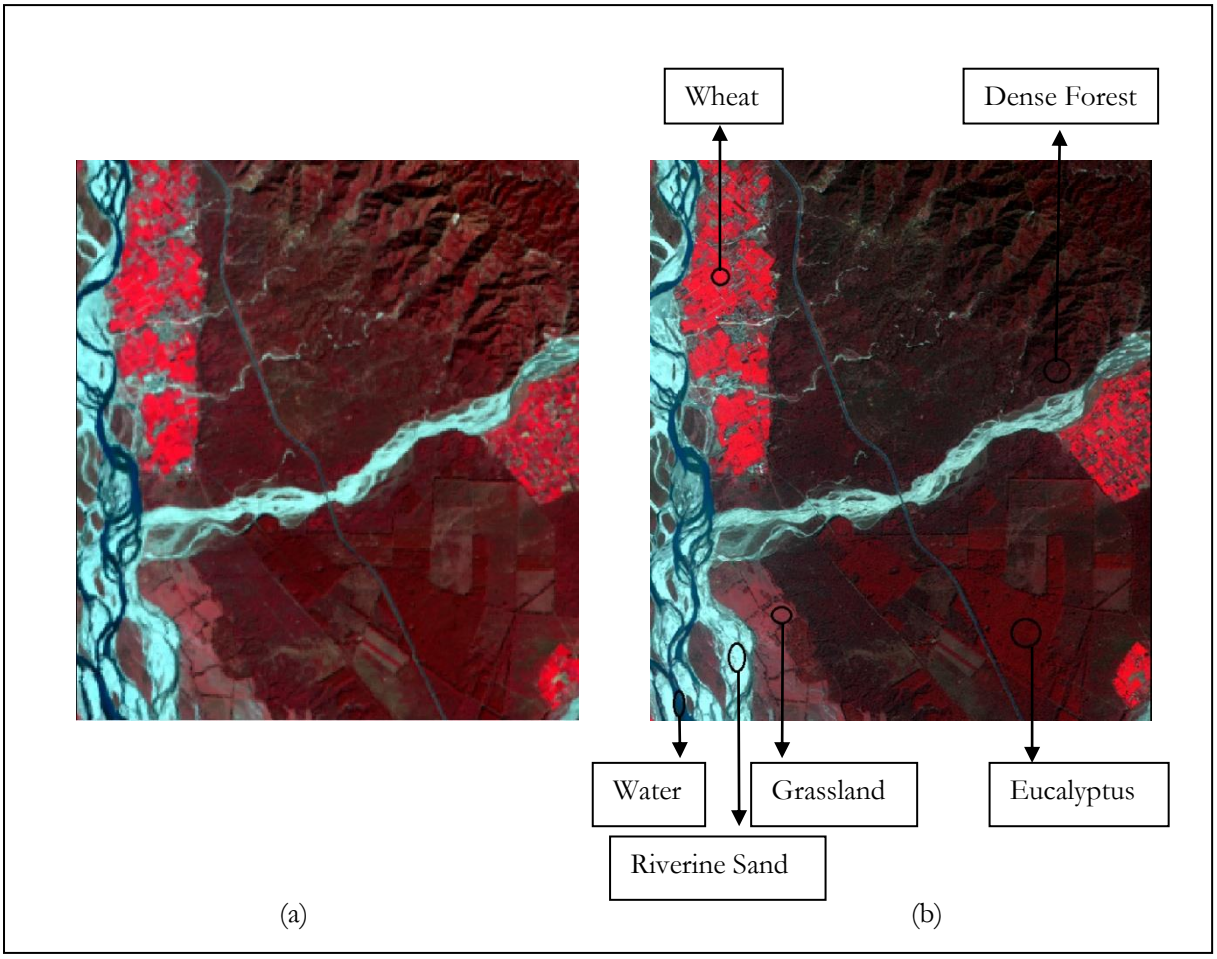


Figure 4.2: The datasets with identified LULC classes. (a) Landsat-8 image used for classification. (b) Formosat-2 image with identified LULC classes in the study area

The specifications of the datasets are summarised in Table 4.1.

Specification	Landsat-8	Formosat-2
Spatial resolution (multispectral)	30m	8m
Spectral Resolution	8 Bands	4 Bands
	Band 1- Coastal	Band 1- Blue
	Band 2- Blue	Band 2- Green
	Band 3- Green	Band 3- Red
	Band 4- Red	Band 4- NIR
	Band 5- NIR	
	Band 6- SWIR1	P is Panchromatic (2m)
	Band 7- SWIR2	
	Band 9- Cirrus	
	Band 8 is Panchromatic (15m)	
Sensor footprint	170km x 185km	24km x 24km
Revisit interval	16 Days	Daily
Date on which image was acquired	February 12, 2015	February 21, 2015

Table 4.1: Specifications of the datasets used

4.2. Software used

The classification algorithms and accuracy assessment methods were implemented in Python 3.7. The libraries used for the development, execution and accuracy assessment of the algorithms were Rasterio, Numpy, Matplotlib, Math. The image registration and collection of test data were done using ERDAS Imagine 2014.

5. METHODOLOGY

The primary aim of this research was to develop and study the performance of FCM based and PCM based local spatial information fuzzy classification algorithms. This chapter explains the steps followed for the development, execution and assessment of the performance of these algorithms. Figure 5.1 illustrates the methodology adopted. The subsequent sections of the chapter explain the details of each of the steps in the methodology which lead to the accomplishment of the research objectives mentioned in 1.4.

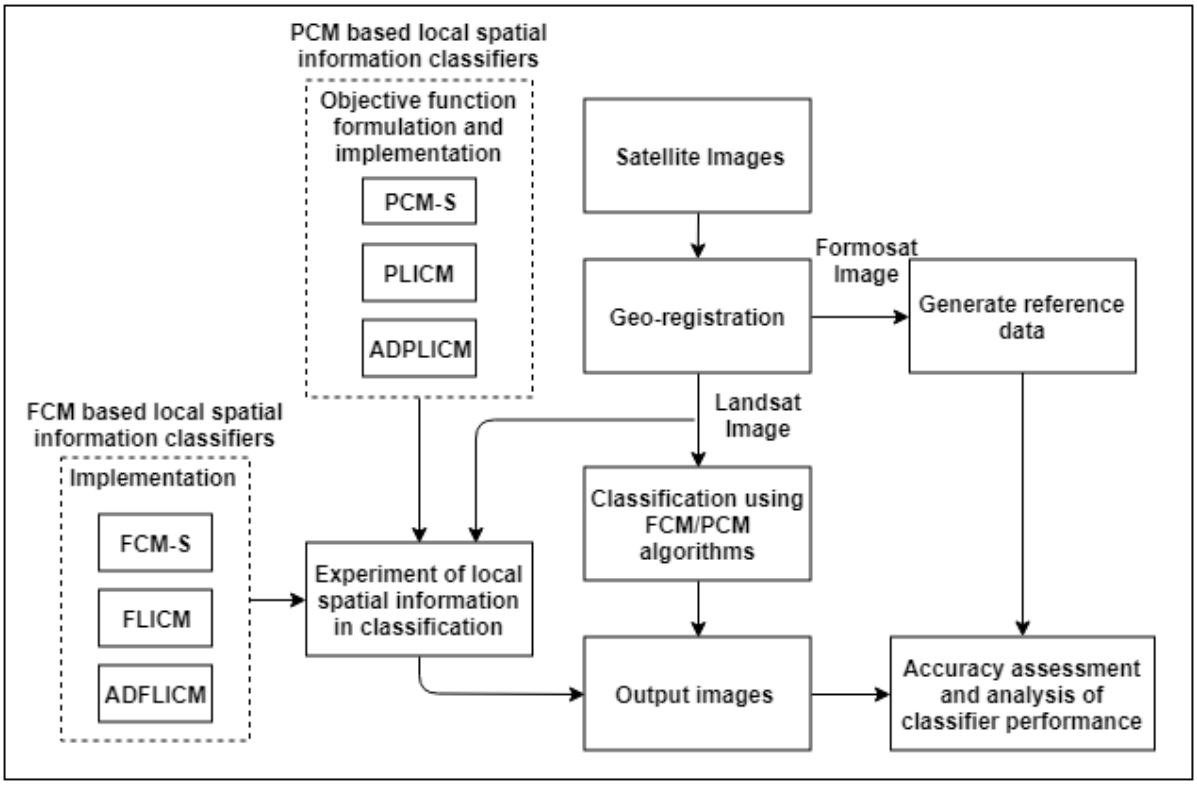


Figure 5.1: The methodology adopted for the study

5.1. Image Registration

Registering together the images captured for the same region at different times or by different sensors is necessary when pixel by pixel comparison is to be made. The images can be registered with each other by separately registering each image with the map coordinates, or by registering an image (known as the slave image) with the other chosen master image. The later approach is known as Image-to-Image registration (Richards & Jia, 2012). In this research, the accuracy of the classification algorithms on Landsat-8 image (with 30m spatial resolution) was compared with the reference image created using the Formosat-2 image. The images were needed to be aligned to a common coordinate system for monitoring subtle changes between them. The geo-referenced Formosat-2 image of 8m spatial resolution was first resampled to 10m spatial resolution using Nearest Neighbour Resampling so that the pixel sizes in the classified Landsat-8 image and the reference Formosat-2 image was in the ratio 3:1. Image-to-Image registration was then performed using the Formosat-2 image chosen as the master image.

5.2. Classification

The image has been separately classified using the FCM, PCM, FCM-S, FLICM, ADFLICM, PCM-S, PLICM and ADPLICM algorithms. Experiments have been conducted to analyse the performance of the algorithms when all the classes are considered/few classes are considered. The details of the experiments conducted and the corresponding results obtained are present in chapter 6. The accuracy of the algorithms were assessed using Mean Membership Difference, Fuzzy Error Matrix and Root Mean Square Error methods, whose details are explained in section 5.5.

5.2.1. Classification using FCM based local spatial information classifiers

The FCM-S, FLICM and ADFLICM algorithms have been implemented to analyse their performance for remote sensing image classification. The algorithms were developed for supervised classification and executed on the input dataset. The basic steps for the three FCM based local spatial information fuzzy classifiers (FCM-S, FLICM and ADFLICM), whose mathematical details are explained in 3.3, are given below. The flow chart for FCM based local spatial information classification algorithms is shown in figure 5.2.

Step 1: Get the input image and training data.
Step 2: Initialise the values for fuzzy factor (m), window size, *a* (for FCM-S algorithm).
Step 3: Calculate the class mean from the training data.
Step 4: Calculate the distance of each pixel to the class means (using Euclidean Norm).
Step 5: Calculate the neighbourhood effect.
Step 6: Calculate the membership values of pixels for each class.

5.2.2. Classification using PCM based local spatial information classifiers

The PCM-S, PLICM and ADPLICM algorithms have been developed by modifying the standard PCM algorithm as mentioned in section 3.4. Once the objective functions are formulated, membership functions that optimise the objective functions are estimated. The general algorithm for PCM based local spatial information classifiers involves the following steps:

Step 1: Get the input image and training data.
Step 2: Initialise the values for fuzzy factor (m), window size, *a* (for PCM-S algorithm).
Step 3: Calculate the class mean from the training data.
Step 4: Calculate the distance of each pixel to the class means (using Euclidean Norm).
Step 5: Calculate the neighbourhood effect.
Step 6: Estimate the η_i using class membership values initialised with the FCM algorithm .
Step 7: Calculate the final membership values of pixels for each class.

Figure 5.3 demonstrates the flow chart for the three PCM based local spatial information fuzzy classifiers (PCM-S, PLICM and ADFLICM) developed in this research.

The python code for the implementation of one of the algorithm implemented (ADPLICM) is presented in Appendix A of the thesis for reference.

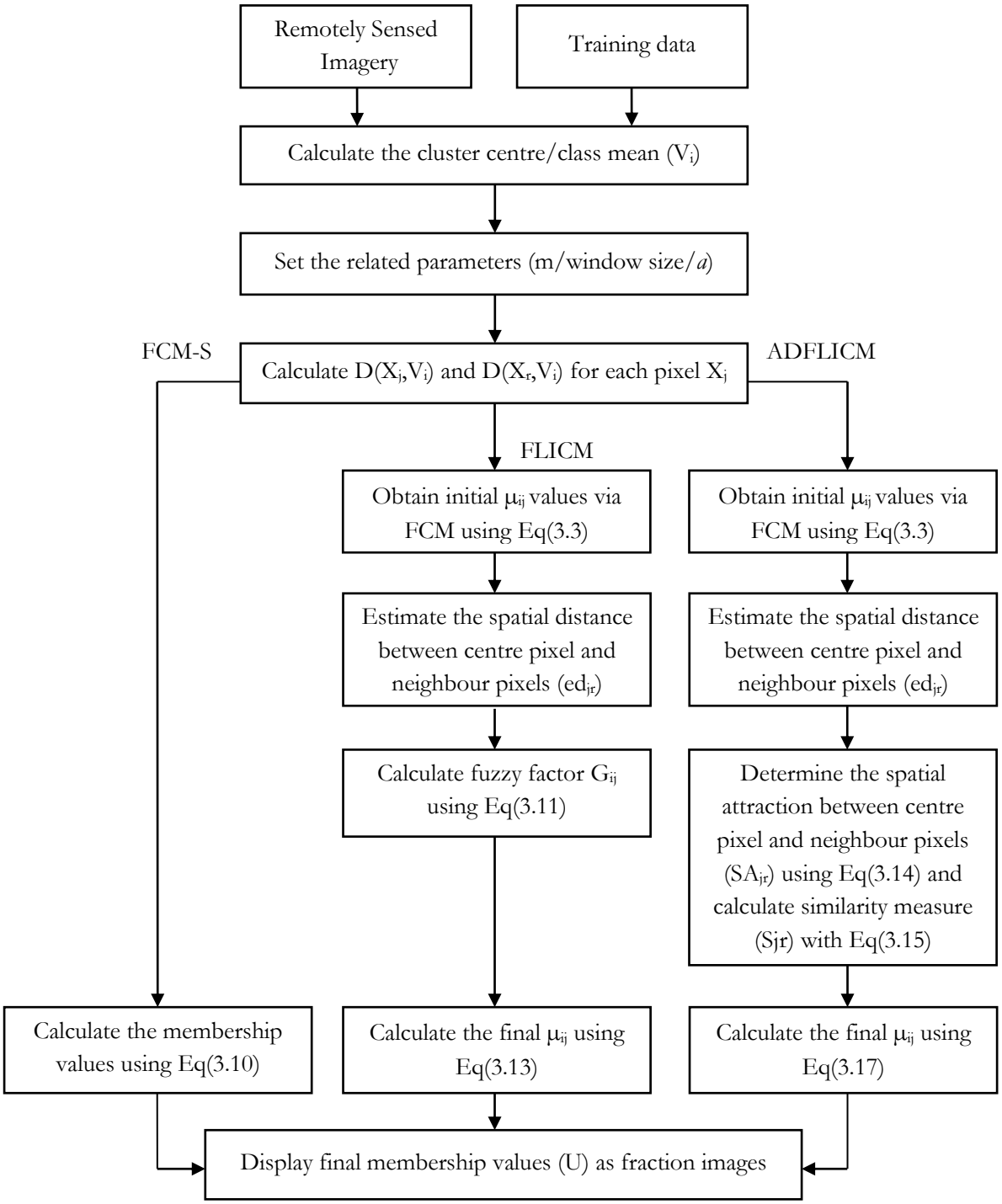


Figure 5.2: FCM based local spatial information classification algorithms

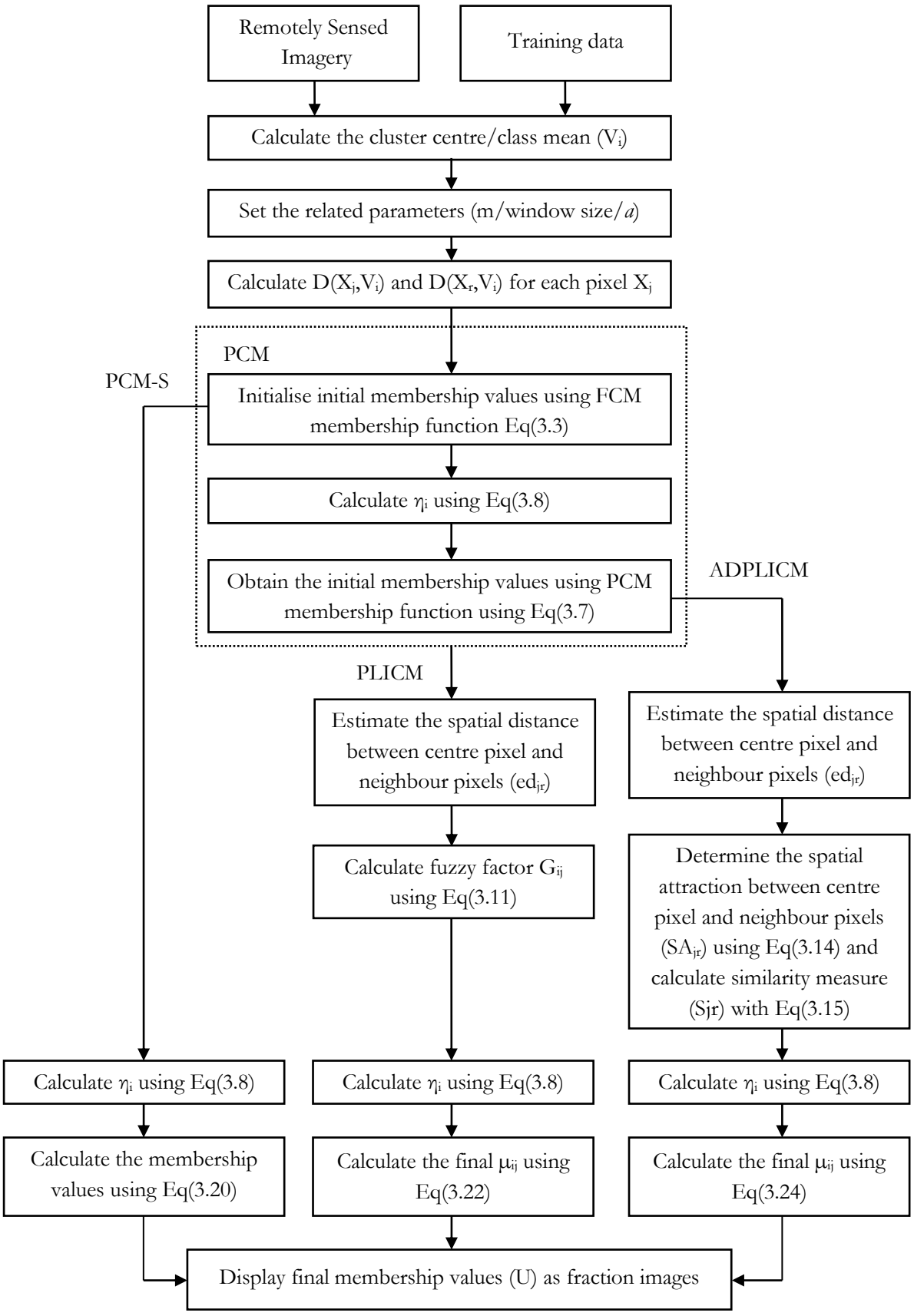


Figure 5.3: PCM based local spatial information classification algorithms

5.3. Reference data generation

Often, finer resolution images are considered as reference for coarser resolution input images. Reference data is assumed to be more correct compared to the image being used for classification and hence are usually referred to as "ground truth" data. However, the reference data may not represent the "true" value (Congalton, 2009). Reference values generated from finer resolution images are basically assumed close to ground reality compared to the coarser resolution images because of their higher spatial resolution which makes the features in the image more distinguishable.

In this research, for comparing the results of the soft classified outputs from Landsat-8 images, soft classified reference data was required. A suitable soft classification algorithm was needed to be applied on the finer resolution Formosat-2 image for the creation of soft reference data. While the Landsat-8 image was classified using spatio-contextual algorithms, classifiers that use spatial information in addition to spectral information was preferred for the creation of reference data as well. Such classifiers were not commercially available and hence reference images were created using FCM-S, FLICM, and ADFLICM algorithms on the Formosat-2 image. When all the information classes are defined, FCM is reported to produce most accurate class compositions (Foody, 2000), and hence only FCM based local spatial information classification algorithms were used for the creation of reference images. The performances of the developed algorithms on Landsat-8 image were compared by using each of the reference image generated.

5.4. Optimization of parameters

The parameter *m, window size* and *a* (in case of FCM-S and PCM-S) for all the local spatial information fuzzy classification algorithms have been optimised for refining their classification performance. An experimental approach has been followed for optimisation in which the accuracy of the classification algorithms was measured for varied values of the selected parameter keeping the remaining parameters constant. The optimal value of the parameter under consideration was chosen for which the highest classification accuracy was obtained. The details of the methods used for assessing the accuracy of the classification algorithms are mentioned in section 5.5.

5.5. Accuracy Assessment

Accuracy assessment is an essential stage in the image classification procedure as it provides information about the quality of the classification process as well as it acts as a means for comparing the performance of various classification techniques. Mean Membership Difference method, Root Mean Square Error (RMSE) and Fuzzy Error Matrix (FERM) method have been used in this study for parameter optimisation and comparison of the performance of the classification algorithms implemented.

5.5.1. Mean Membership Difference

Mean Membership Difference method is an independent measure of accuracy assessment as it is defined using the output of the classified image alone, while the other techniques such as FERM, Root Mean Square Error (RMSE), Correlation Coefficient etc. depend on the classified output and the reference output. The membership values of the pixels belonging to a certain class will be high in the fraction image generated for that class while the membership values of the pixels not belonging to the class will be low in the fraction image. Using this concept, Mean Membership Difference method is designed as a quantitative measure of accuracy. The steps for Mean Membership method is given below (Byju, 2015):

1. Identify pixel coordinates of a set of pure pixels from homogenous regions in the input image for all the defined classes.
2. To calculate Mean Membership Difference of class 1 in Figure 5.4, consider the fraction image for class 1.
3. Calculate the mean pixel values (membership values) of the set of pixels identified in step 1 for each class in the selected fraction image. Let the mean pixel values for class 1 to class 6 be denoted by M1, M2, M3, M4, M5, M6 respectively.
4. Calculate the membership value difference between the class under consideration (class 1) and the remaining classes in the fraction image (i.e. D12=M1-M2, D13=M1-M3, D14=M1-M4, D15=M1-M5, D16=M1-M6).
5. The mean of the values calculated in step 4 gives the Mean Membership Difference of the class under consideration (i.e. for class 1 it is (D12+D13+D14+D15+D16)/5).
6. Estimate the Mean Membership Difference values for all the remaining classes in a similar way.

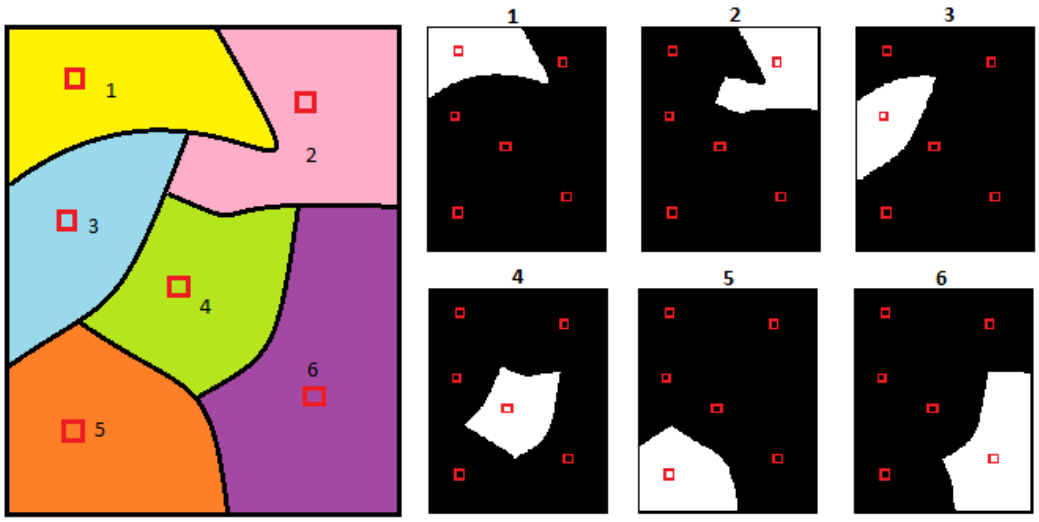


Figure 5.4: Sample input image along with the output fraction images generated for the six classes

Since pure pixel values from homogeneous regions in the input image are selected as test pixels, the mean membership value of the class under consideration will be close to 1, and the mean membership values for the remaining classes will be close to 0 in the fraction image. This will result in a mean membership difference value tending to 1. Thus, the results with higher Mean Membership Difference value (tending to 1) are considered to have better accuracy.

5.5.2. Fuzzy Error Matrix

Error matrix is a square matrix which indicates the agreement and disagreement between the reference data (pixels) and classified data (pixels). The columns of the matrix represent the reference data while the rows represent the classified data. The Fuzzy Error Matrix or FERM (Binaghi et al., 1999) has a similar layout as that of conventional error matrix with the exception that it can have non-negative real numbers instead of non-negative integer values. The elements of the FERM represent the class proportions in the reference image and classified image (Ibrahim, 2004).

If $\acute{R}_n$ represents the fuzzy set of reference data for class *n* and $\acute{C}_m$ represents the fuzzy set of classified data for class *m* with membership values of test sample *x* denoted by $\mu_{\acute{R}n}(x)$ and $\mu_{\acute{C}m}(x)$ respectively, fuzzy error matrix M is constructed as given in equation (5.1). The layout for FERM is shown in Table 5.1.

$$M_{(m,n)} = |\acute{C}_m \cap \acute{R}_n| = \sum_{x \in X} \min\left(\mu_{\acute{C}m}(x), \mu_{\acute{R}n}(x)\right) \tag{5.1}$$

Where, *x* is a set of sample data.

Class Data	**Reference Data**				**Total grades**
	Class 1	**Class2**	**.....**	**Class c**	
Class 1	$M_{(1,1)}$	$M_{(1,2)}$	.....	$M_{(1,c)}$	C_1
Class 2	$M_{(2,1)}$	$M_{(2,2)}$	.....	$M_{(2,c)}$	C_2
.....	.....	.....	.....	.....	.....
Class c	$M_{(c,1)}$	$M_{(c,2)}$	.....	$M_{(c,c)}$	C_c
Total grades	R_1	R_2	.....	R_c	

Table 5.1: Fuzzy error matrix

In the Fuzzy Error Matrix illustrated in Table 5.1, $M_{(m,n)}$ denote the member of FERM in the m^{th} class of the soft classified output and n^{th} class of the soft reference data. R_n and C_m represent the total grade of membership assigned to the n^{th} class for the reference data and m^{th} class for the classified data.

The descriptive indices such as Overall Accuracy (OA), Producer's Accuracy (PA) and User's Accuracy (UA) are estimated similar to the traditional confusion matrix from the Fuzzy Error Matrix. OA is the measure of the total match between the reference data and classified data. It is calculated by dividing the sum of the major diagonal with the total grade of membership in the reference data. PA is a measure of omission error, which defines how much data is excluded from the class to which it belongs. UA is a measure of commission error, which defines how much data is wrongly included in the class to which it does not belong. The formulas to calculate OA, PA for class *j* and UA for class *j* are given in (5.2), (5.3) and (5.4) respectively.

$$OA = \frac{\sum_{j=1}^{c} M_{(j,j)}}{\sum_{j=1}^{c} R_j} \tag{5.2}$$

$$PA_j = \frac{M_{(j,j)}}{R_j} \tag{5.3}$$

$$UA_j = \frac{M_{(j,j)}}{C_j} \tag{5.4}$$

The Average Producer's Accuracy (AA_p) and Average User's Accuracy (AA_u) is then calculated using (5.5) and (5.6) correspondingly.

$$AA_p = \frac{\sum_{j=1}^{c} PA_j}{c} \tag{5.5}$$

$$AA_u = \frac{\sum_{j=1}^{c} UA_j}{c} \tag{5.6}$$

5.5.3. **Root Mean Square Error (RMSE)**

Root Mean Square Error is a method used to find the error between the expected value and the output value. RMSE helps to estimate the accuracy of the classifiers by measuring the difference between the membership values in the classified image and the reference image. RMSE value cannot be negative and a value towards 0 denotes less deviation of classified output with the reference data. The global RMSE is calculated by (5.7) (Dutta, 2009).

$$\mathrm{RMSE(Global)} = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{C}\left(\mu C_{ij} - \mu R_{ij}\right)^2}{N \times C}} \tag{5.7}$$

where μC_{ij} and μR_{ij} are the membership values of pixel j in class i for the classified and reference image respectively , N is the dimension (number of rows and columns) , C is the number of classes . The formula to calculate the RMSE error per class is given in (5.8).

$$RMSE(Per\ Class) = \sqrt{\frac{\sum_{i=1}^{N}\left(\mu C_{ij} - \mu R_{ij}\right)^2}{N}} \tag{5.8}$$

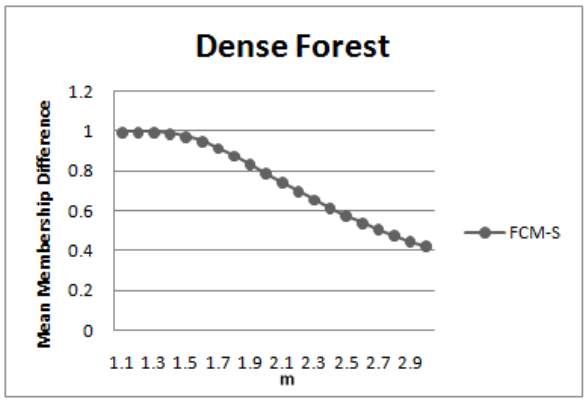
where μC_{ij} and $\mu R_{ij}$are the membership values of pixel j in class i for the classified and reference image respectively, N is the dimension(number of rows and columns).

6. RESULTS AND ANALYSIS

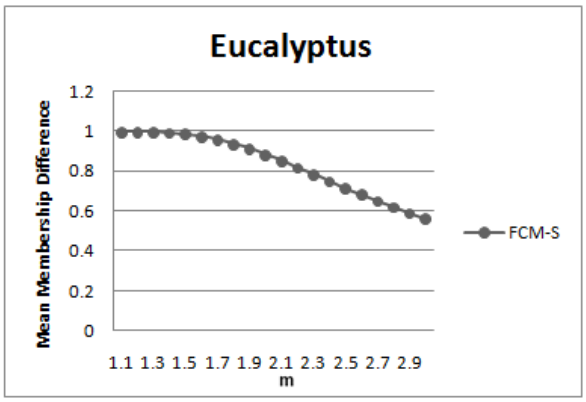
This chapter presents the results of the experiments conducted and the analysis made on the results for meeting the research objectives. Several tests were carried out to study the performance of all the algorithms described in section 5.2 for land use/land cover classification on the chosen dataset. The efficiencies of the algorithms were measured by accuracy assessment on the classified output images with the reference data generated. Initially, the parameters of the algorithms were optimised to exploit the potential of each of the classifiers. Section 6.1 explains the approach used for parameter optimisation. Once the parameters were optimised, three different scenarios were executed to examine the performance of the local spatial information classification algorithms. The scenarios executed were the (1) supervised classification with all the available classes in the study area, (2) supervised classification with few of the available classes (or the supervised classification with untrained classes), and (3) supervised classification with a single class. The results and analysis of the scenarios one, two and three executed with the algorithms are presented in sections 6.2, 6.3 and 6.4 respectively. Section 6.5 describes an additional scenario experimented with the local spatial information algorithms in removing noisy pixels in the input image. Section 6.6 concludes the chapter with the discussions about the general behaviour of the algorithms based on the results analysed for each scenario tested.

6.1. Parameter optimization

The optimisation of parameters is essential in bringing out the best performance of each of the classifiers. The parameters *m,* window size and *a* (for FCM-S and PCM-S algorithms) were optimised by repeat testing and fixing the values for which maximum accuracies were obtained for the classification algorithms. Optimal ranges for the parameters were obtained using the mean membership difference method. Further, FERM, RMSE and visual interpretation of the classified outputs with the reference map generated were used to get the optimal values of each of the parameters. Figure 6.1 shows the result of mean membership difference method for FCM-S algorithm in optimising *m* value when all the classes were selected for supervised classification. Figure 6.2 shows the results of FERM for the same.



(a)



(b)

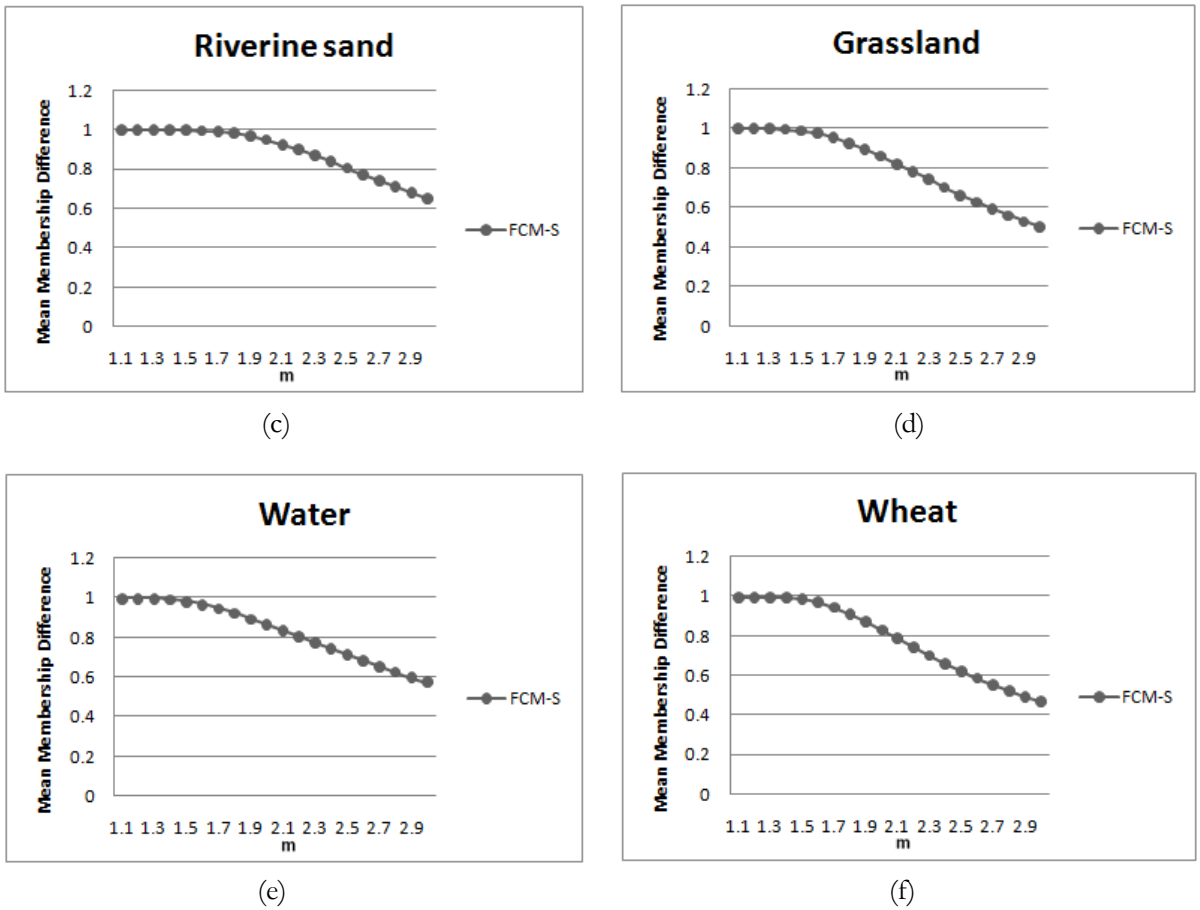


Figure 6.1: Variation in mean membership difference values with respect to *m* for FCM-S classifier with all the classes ((a)-(f))

From the graphs in Figure 6.1, the optimal range for *m* value considering all the classes for FCM-S can be seen to be between *m* = 1.1 to *m* = 1.6. The output fraction images for every class was carefully examined for varied values of *m*. Based on the results of FERM in Figure 6.2, the optimal value of *m* in this case was chosen to be 1.5. The average UA, average PA and the OA can be seen to be higher for the classification output when *m* = 1.5.

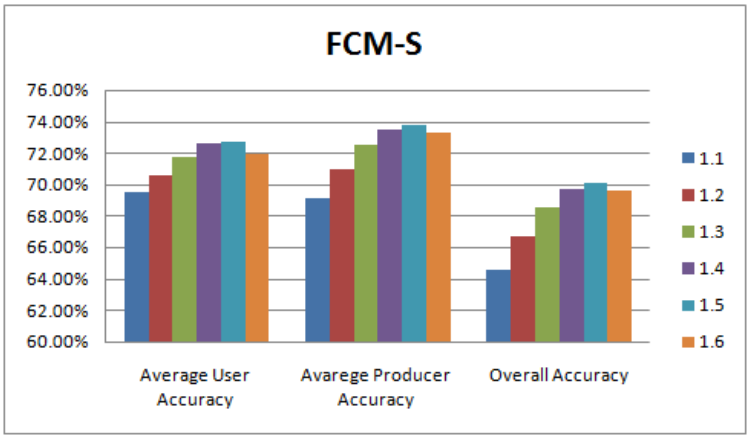


Figure 6.2: Variation in accuracy indices of FERM with respect to *m* for FCM-S classifiers with all the classes

It has been observed that the optimal value of *m* for classifiers varied when the number of classes selected for supervised classification was varied. Hence different *m* values were used for the same classifier in each of the three scenarios executed.

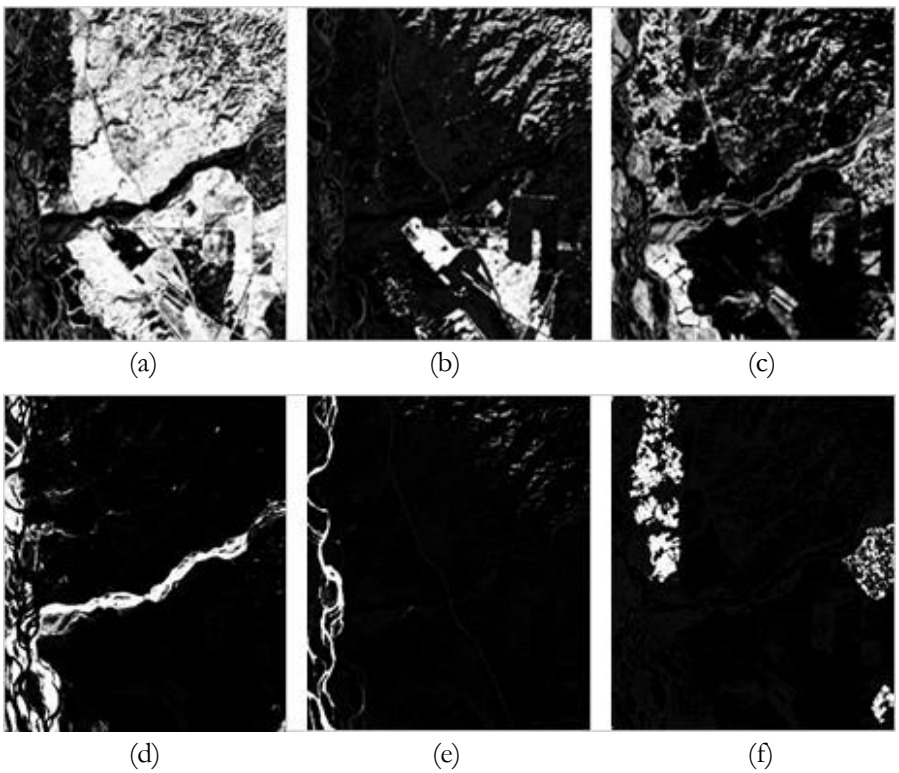
Apart from *m* value, the window size for all the local spatial information classifiers (FCM-S, FLICM, ADFLICM, PCM-S, PLICM, and ADPLICM) and the *a* value for FCM-S and PCM-S were chosen experimentally using mean membership difference method, FERM/RMSE and visual examination. The parameter *a* was chosen by trial and error based on the nature of the classes to be classified. The parameter values for classifiers, optimised for supervised classification with all the classes, are summarised in Table 6.1.

	FCM	**PCM**	**FCM-S**	**PCM-S**	**FLICM**	**PLICM**	**ADFLICM**	**ADPLICM**
m	1.7	1.8	1.5	2.2	1.7	2.2	1.5	1.8
a	-	-	2	0.5	-	-	-	-
window size	-	-	3	3	3	3	3	3

Table 6.1: The parameter values optimised for the conventional classifiers and the local spatial information classifiers for supervised classification with all the classes

6.2. Supervised classification with all the classes

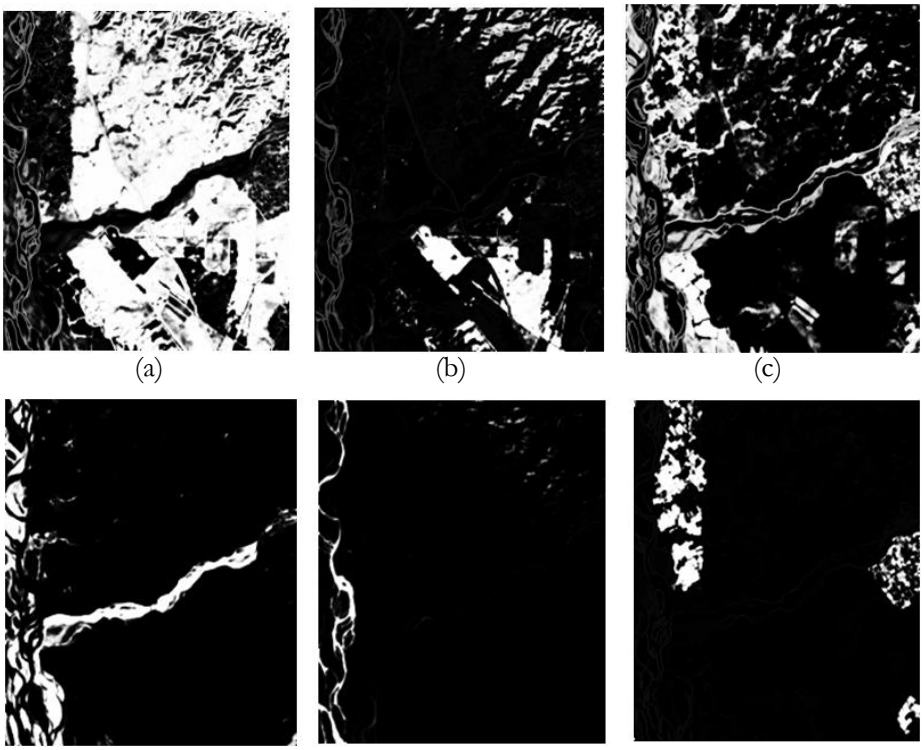
Supervised classifications with the six LULC classes (Dense Forest, Eucalyptus plantation, Grasslands, Riverine sand, Water, and Wheat) were performed using FCM, FCM-S, FLICM, ADFLICM, PCM, PCM-S, PLICM, and ADPLICM classification algorithms. The fraction images generated for each of the classifiers are presented in Figure 6.3 to Figure 6.10.



(a) (b) (c)

(d) (e) (f)

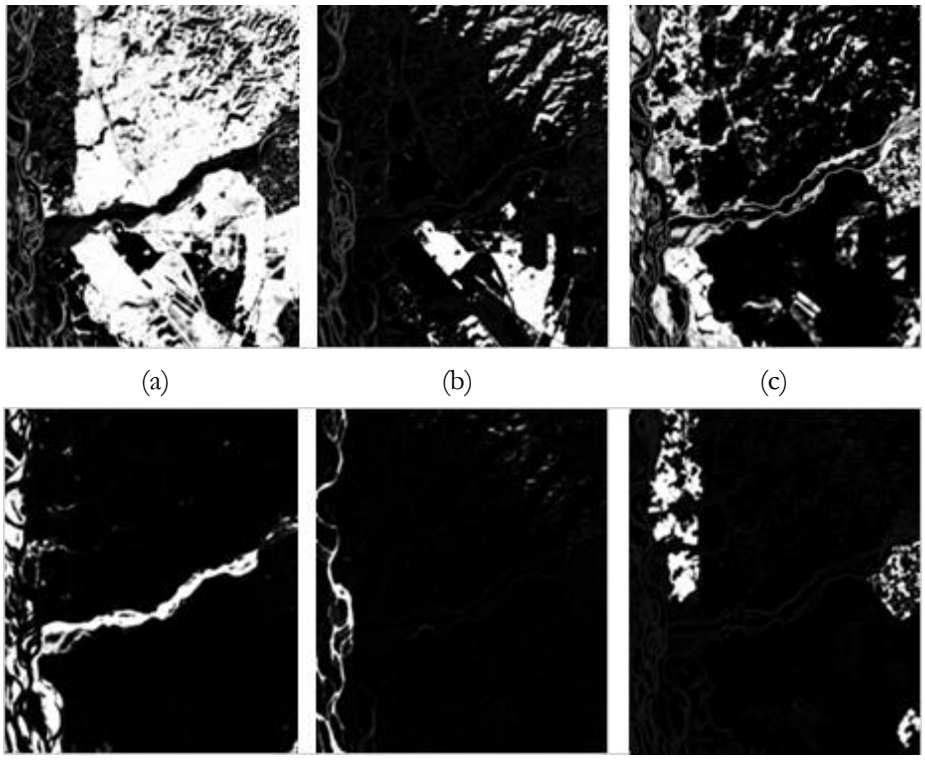
Figure 6.3: Output fraction images of FCM classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat



(a) (b) (c)

(d) (e) (f)

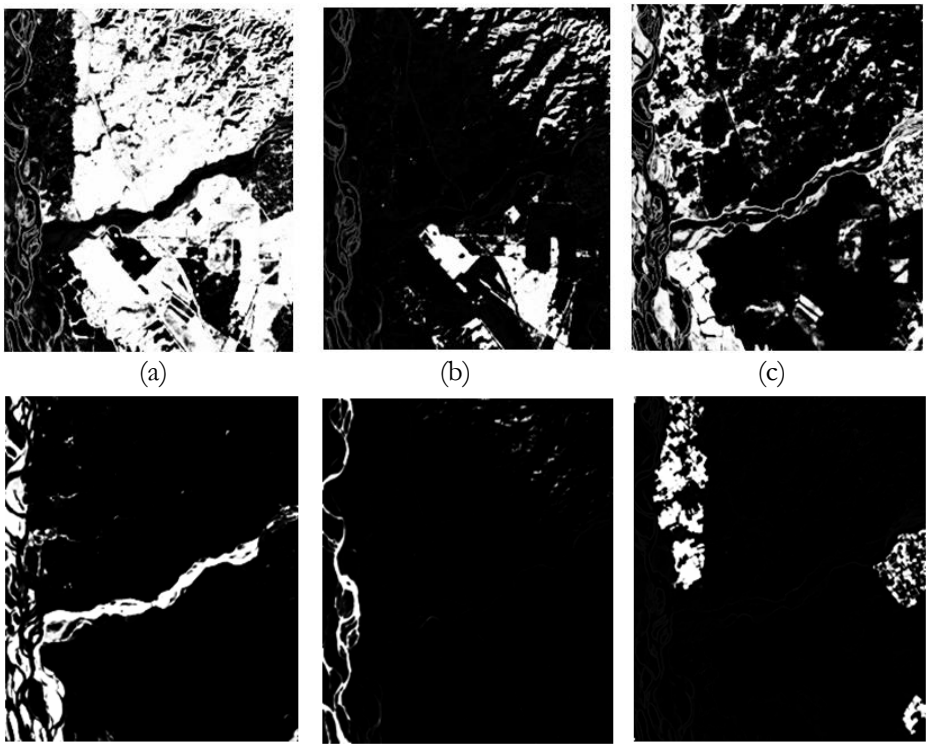
Figure 6.4: Output fraction images of FCM-S classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat



(a) (b) (c)

(d) (e) (f)

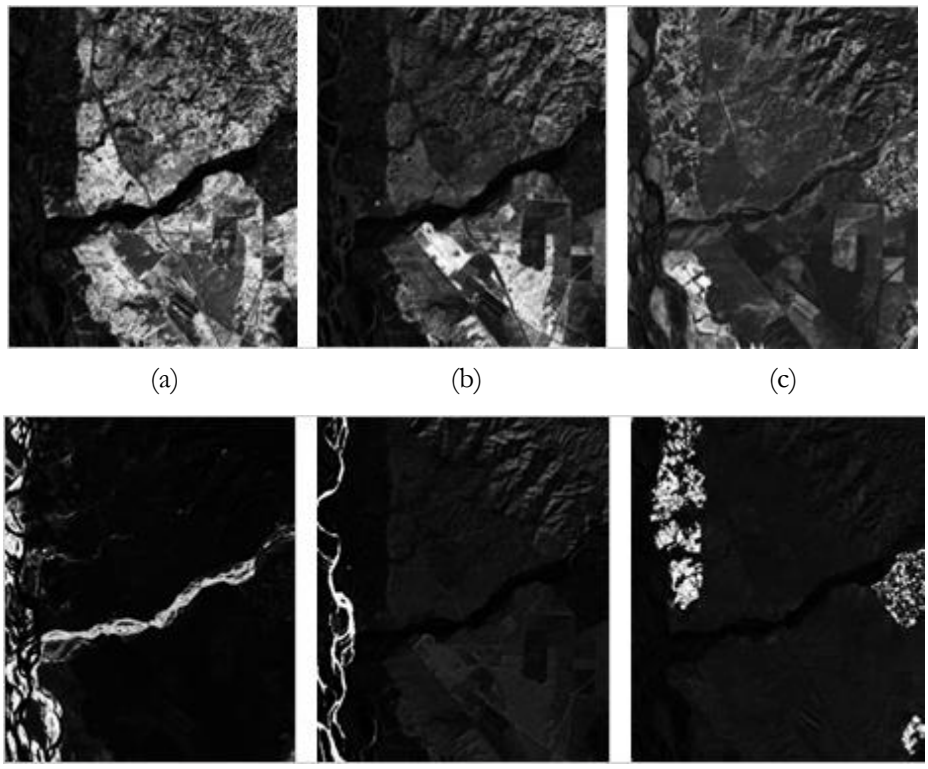
Figure 6.5: Output fraction images of FLICM classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat



(a) (b) (c)

(d) (e) (f)

Figure 6.6: Output fraction images of ADFLICM classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat



(a) (b) (c)

(d) (e) (f)

Figure 6.7: Output fraction images of PCM classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat

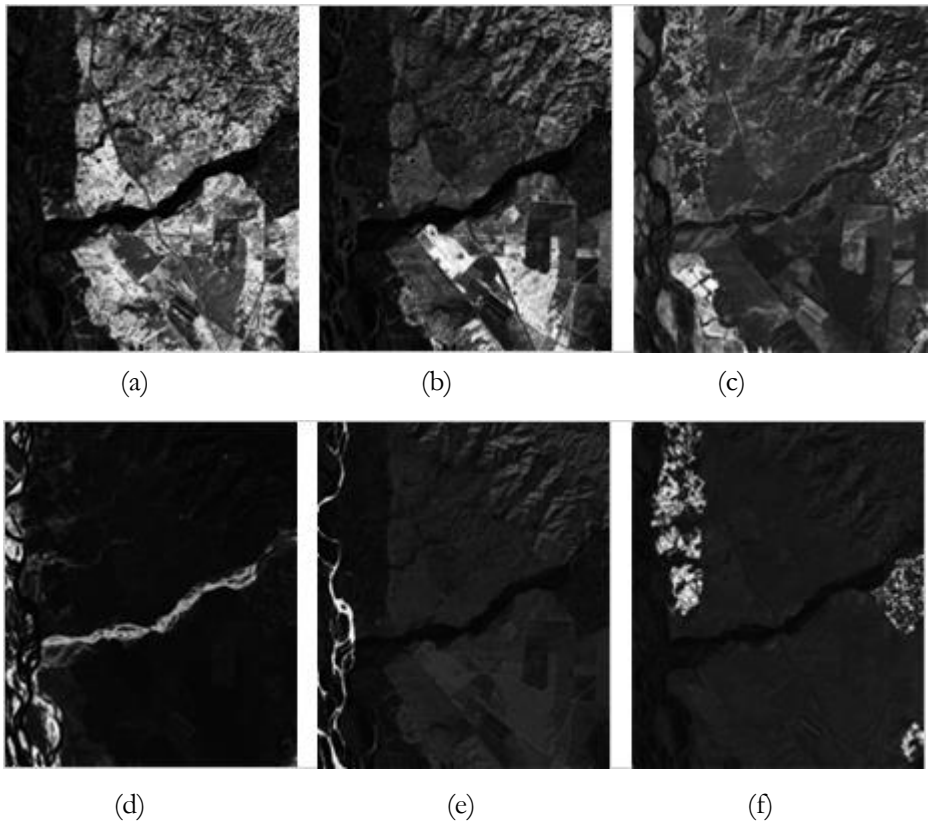


Figure 6.8: Output fraction images of PCM-S classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat

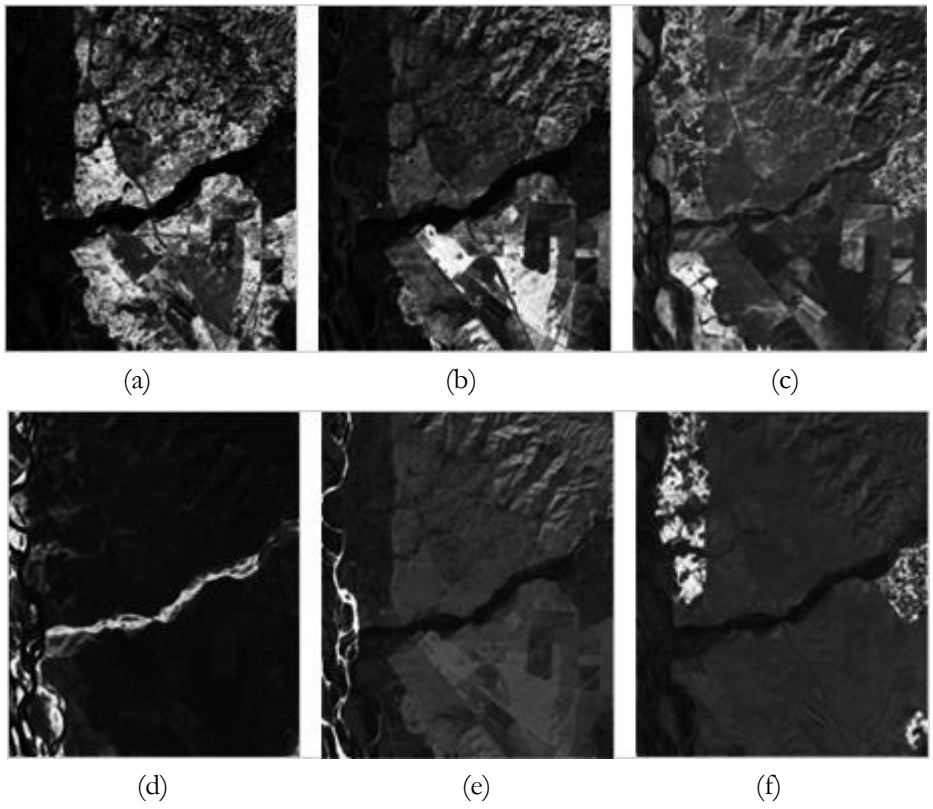


Figure 6.9: Output fraction images of PLICM classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat

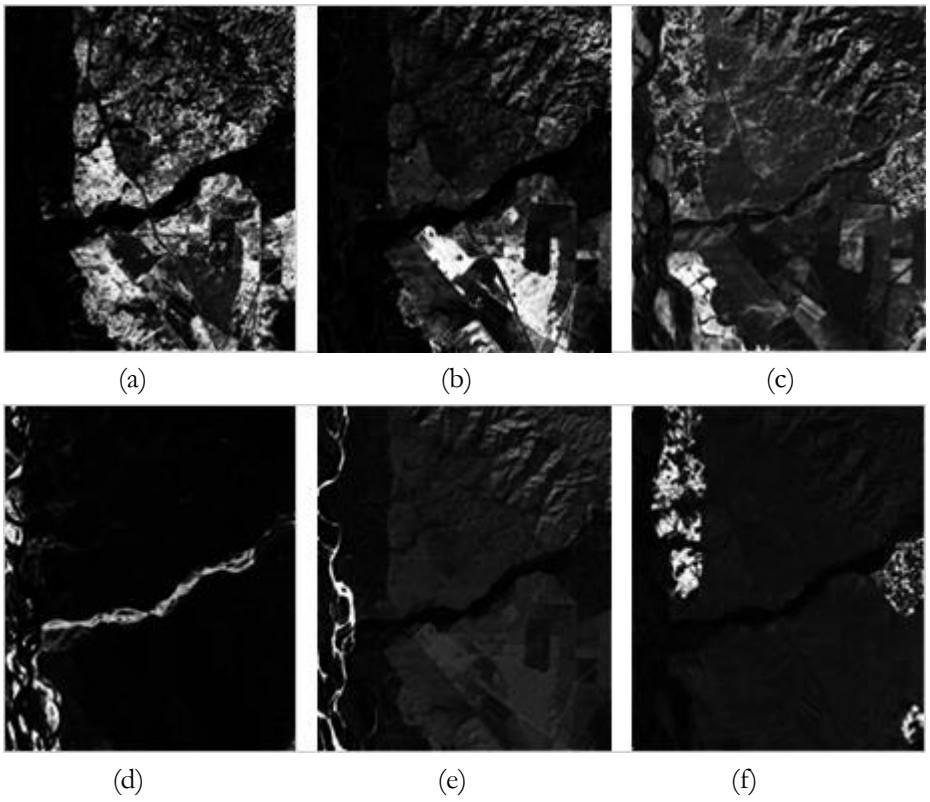


Figure 6.10: Output fraction images of ADPLICM classifier for the classes (a) Dense Forest (b) Eucalyptus (c) Grassland (d) Riverine sand (e) Water (f) Wheat

The classified output images were quantitatively analysed using FERM separately with each of the reference image created through FCM-S, FLICM and ADFLICM algorithms onFormosat-2 image. The approach used for the creation of the reference image is described in section 5.3. The results of the accuracy assessments with reference images generated from FCM-S, FLICM and ADFLICM algorithms are summarised in Table 6.2, Table 6.3 and Table 6.4 respectively. Test samples were carefully collected to include isolated pixels (vague pixels in homogenous surrounding), pixels from the homogenous regions of classes, and pixels from edges and boundaries. Overall accuracies were calculated for each of the algorithms (1) while taking all the random test samples for accuracy assessment, (2) while taking only the homogenous pixels for testing, (3) while taking isolated pixels as test samples (4) while taking the edges and boundary pixels. The overall accuracies for the four different cases discussed above are denoted by OA, OA1, OA2 and OA3 correspondingly in Table 6.2 to Table 6.4.

	FCM	FCM-S	FLICM	ADFLICM	PCM	PCM-S	PLICM	ADPLICM
OA	67.56	70.07	69.83	69.66	64.43	52.15	52.61	44.90
OA1	83.13	84.25	85.08	83.63	76.82	63.63	67.32	62.01
OA2	41.99	70.69	74.22	63.14	34.84	36.25	35.35	22.86
OA3	70.36	58.47	55.24	63.64	73.86	51.21	49.53	41.26

Table 6.2: The overall accuracies (in percentage) of FERM obtained for the conventional fuzzy classification algorithms and fuzzy local spatial information classification algorithms when reference image was created with FCM-S algorithm

	FCM	FCM-S	FLICM	ADFLICM	PCM	PCM-S	PLICM	ADPLICM
OA	65.11	67.56	67.56	67.42	61.41	48.51	49.04	42.54
OA1	82.25	84.21	85.55	84.26	75.50	61.94	66.07	61.22
OA2	35.61	65.79	69.52	58.45	28.77	29.78	28.57	18.11
OA3	68.46	54.16	50.67	59.19	70.81	46.78	45.23	38.39

Table 6.3: The overall accuracies (in percentage) of FERM obtained for the conventional fuzzy classification algorithms and fuzzy local spatial information classification algorithms when reference image was created with FLICM algorithm

	FCM	FCM-S	FLICM	ADFLICM	PCM	PCM-S	PLICM	ADPLICM
OA	67.22	69.28	68.82	69.04	63.74	50.61	51.40	44.31
OA1	83.74	85.08	85.31	84.36	76.93	63.07	66.98	61.79
OA2	39.27	68.28	72.00	60.73	32.43	33.64	32.73	21.45
OA3	70.65	56.75	53.19	62.53	73.47	49.36	48.42	40.63

Table 6.4: The overall accuracies (in percentage) of FERM obtained for the conventional fuzzy classification algorithms and fuzzy local spatial information classification algorithms when reference image was created with ADFLICM algorithm

The overall accuracies (OA, OA1, OA2 and OA3) obtained for the conventional fuzzy classifiers and fuzzy local spatial information classifiers with each reference image were noticed to be analogous. Figure 6.11 illustrates the overall accuracy values obtained for the eight classifiers. As there was less variance in the corresponding entries in tables 6.2, 6.3 and 6.4, the respective values in the tables were averaged to construct the graph in Figure 6.11.

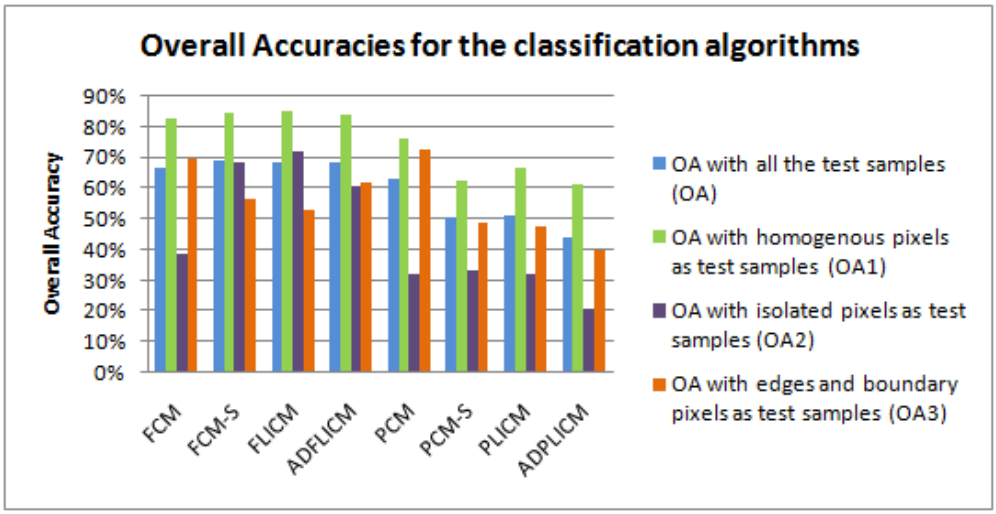


Figure 6.11: Overall Accuracies (in percentage) of classification algorithms when tested with different sample sets

The FCM based algorithms were seen to perform better than the PCM based algorithms in terms of overall accuracies, which meant that the FCM based classifiers produced class proportions that were in close agreement with the class proportions in the reference images. The local spatial information classification algorithms FCM-S, FLICM and ADFLICM showed greater overall accuracies while all the random test sample pixels were used for accuracy assessment. Although the FCM-S and FLICM were effective in removing isolated pixels (Figure 6.11 and Figure 6.12) and their classification performances were slightly higher than the performance of conventional FCM classifiers in homogenous regions, the algorithms produced smoother outputs which resulted in the loss of image details such as edges and boundaries (Figure 6.13). ADPLICM algorithm showed acceptable performance among the FCM based local spatial information classifiers in terms of handling isolated pixels while retaining the image details at the same time (Figure 6.13).

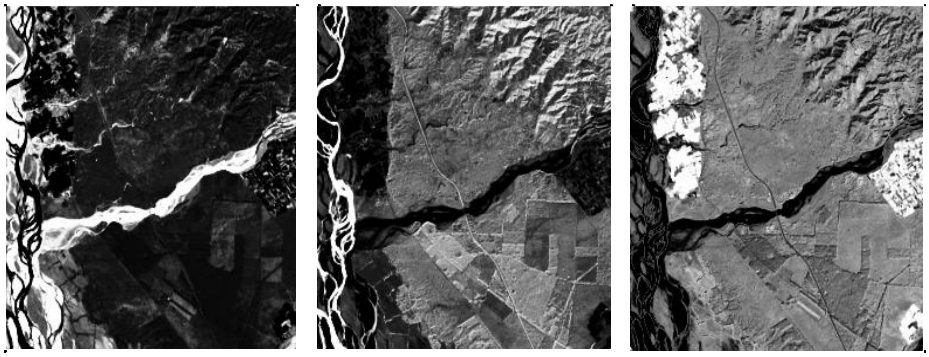
Figure 6.12: Example of fuzzy local spatial information classifier in removing isolated pixels. (a) Reference fraction image for eucalyptus class (b) FCM classifier output fraction image with isolated pixels (c) FLICM classifier output fraction image where isolated pixels are handled

Figure 6.13: Example of classifier in image detail preservation. (a) Reference fraction image for eucalyptus class (b) FLICM classifier output fraction image with vague edges(c) ADFLICM classifier output fraction image in which edges are preserved

From Figure 6.11, it was observed that PCM classifier showed comparable accuracies to FCM classifier and FCM based local spatial information classifiers while the accuracies of PCM based local spatial information classification algorithms were lower than the other algorithms.

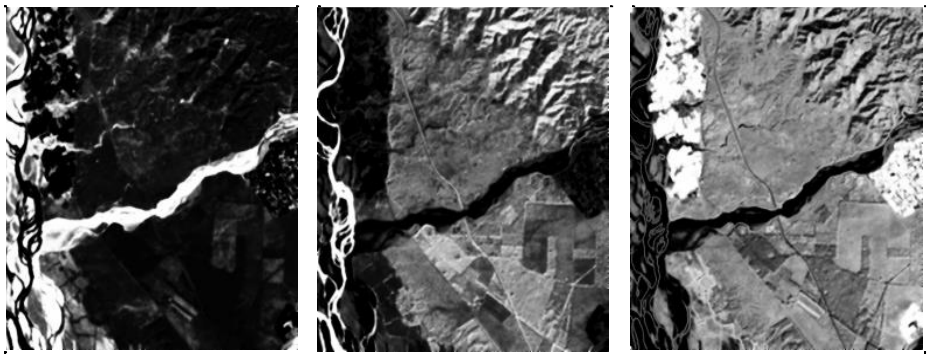
6.3. Supervised classification with few classes

To understand the performance of the classifiers in the presence of untrained classes, few classes were needed to be excluded from training the classifiers. The classes chosen for studying the performance of the FCM based and PCM based classifiers were 'Riverine sand', 'Water' and 'Wheat' classes. The training data for the remaining three classes in the study area were omitted while executing different supervised classifications on the dataset. The output fraction images generated for each of the eight classifiers (FCM, FCM-S, PCM, PCM-S, PLICM, ADPLICM) are shown in Figure 6.14 to Figure 6.22. The value chosen for *m* were 1.5, 1.2, 1.5, and 1.4 respectively for PCM, PCM-S, PLICM and ADPLICM algorithms. The parameter values presented in Table 4.1 were used for the remaining algorithms.



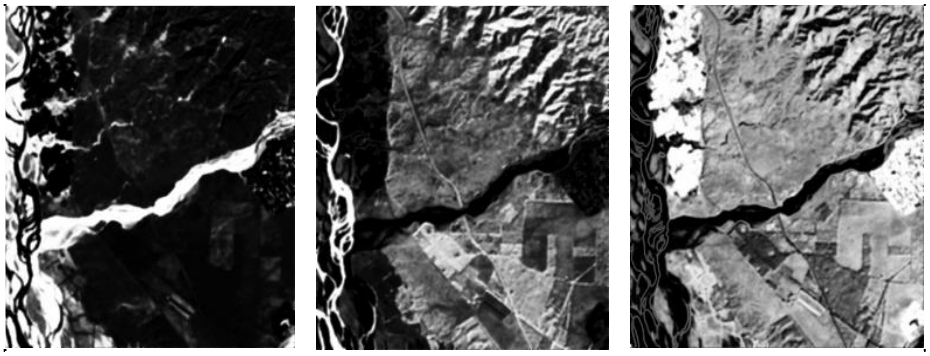
(a) (b) (c)

Figure 6.14: Output fraction images of FCM classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



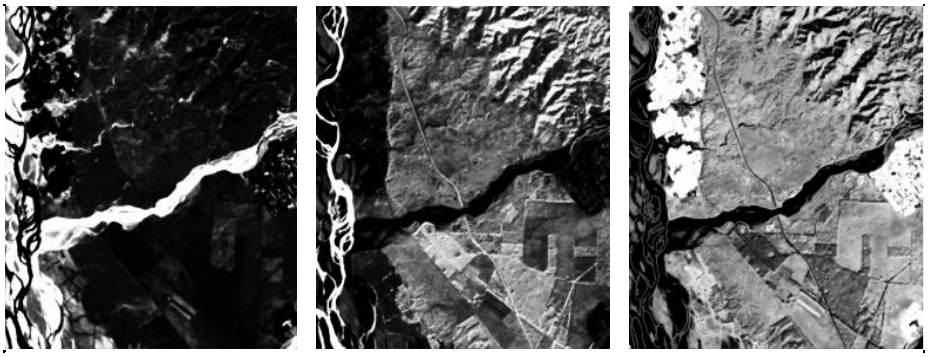
(a) (b) (c)

Figure 6.15: Output fraction images of FCM-S classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



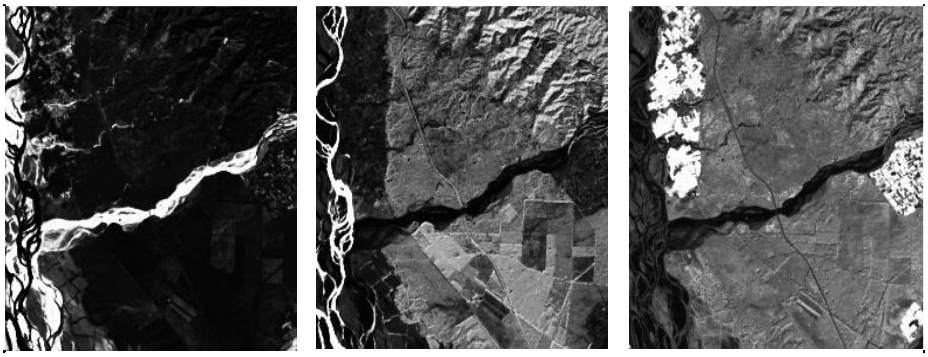
(a) (b) (c)

Figure 6.16: Output fraction images of FLICM classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



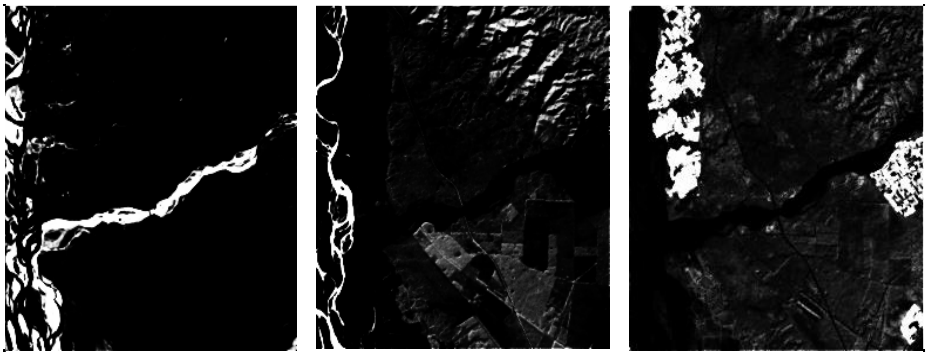
(a) (b) (c)

Figure 6.17: Output fraction images of ADFLICM classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



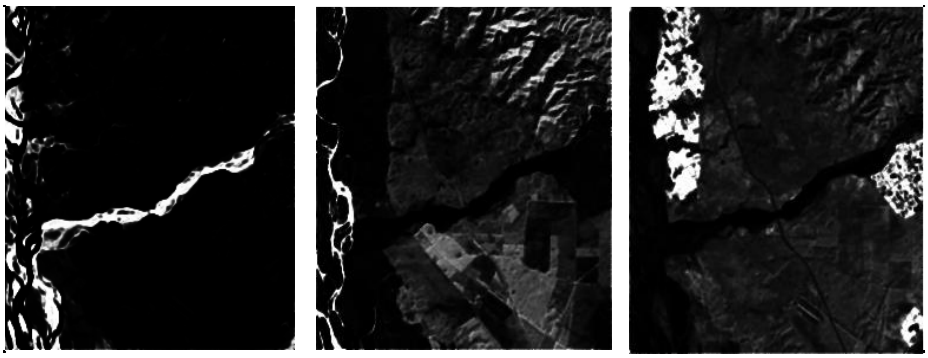
(a) (b) (c)

Figure 6.18: Output fraction images of PCM classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



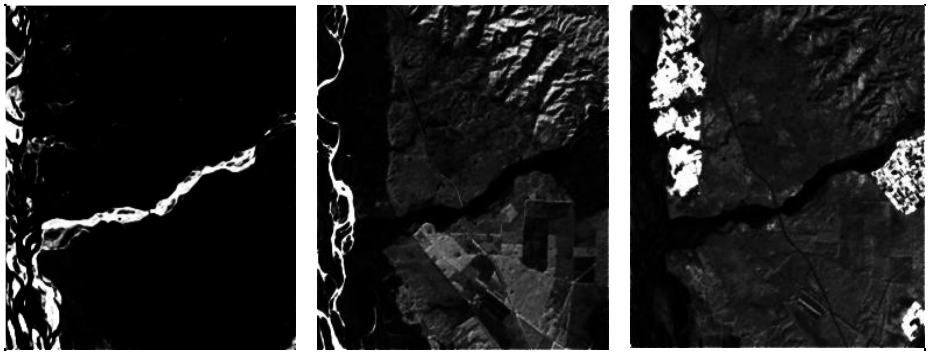
(a) (b) (c)

Figure 6.19: Output fraction images of PCM-S classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



(a) (b) (c)

Figure 6.20: Output fraction images of PLICM classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes



(a) (b) (c)

Figure 6.21: Output fraction images of ADPLICM classifier for the classes (a) Riverine sand (b) Water (c) Wheat in the presence of untrained classes

The global and class wise RMSE errors were estimated for the classifiers for iterations with each reference map. All the test samples were used for calculating the RMSE errors. The average class wise RMSE and global RMSE values from multiple iterations are summarised for each classifier in Table 6.5, and the corresponding graph is presented in Figure 6.22.

	FCM	FCM-S	FLICM	ADFLICM	PCM	PCM-S	PLICM	ADPLICM
Riverine sand	0.262	0.227	0.220	0.249	0.203	0.050	0.010	0.016
Water	0.356	0.336	0.338	0.358	0.358	0.216	0.276	0.243
Wheat	0.412	0.459	0.486	0.472	0.382	0.293	0.205	0.240
Global RMSE	0.349	0.354	0.365	0.371	0.324	0.212	0.199	0.197

Table 6.5: RMSE values estimated on the outputs of different classifiers when supervised classification was applied in the presence of untrained classes

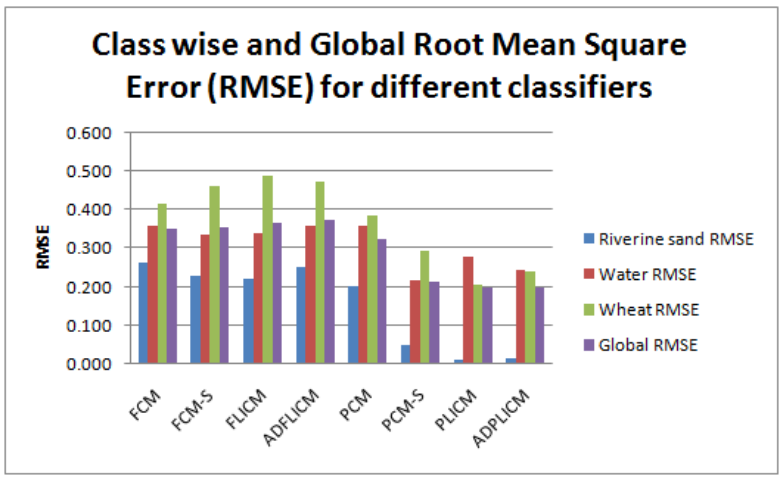


Figure 6.22: Result of Accuracy assessment of the classifiers based on RMSE values

The PCM based local spatial information algorithms exhibited lower RMSE values which meant that there was less disparity between the output fraction images generated from the classifiers and the corresponding reference fraction images. The overall RMSE values obtained for the three local spatial information classifiers were almost comparable though PLICM classifier was seen to be slightly advantageous over the

other two algorithms in terms of Global RMSE value. While the PLICM algorithm could be considered to work best in this scenario, the class wise RMSE for water class and visual interpretation suggested that there was significant loss in membership values for water pixels while classifying the image with PLICM algorithm. To compare the performance of PCM based local spatial information classifiers in image detail retention, RMSE value for test pixels from borders and edges were taken for validation. The Global RMSE values obtained for each of the PCM bases local spatial information classifier outputs are depicted in Figure 6.23.

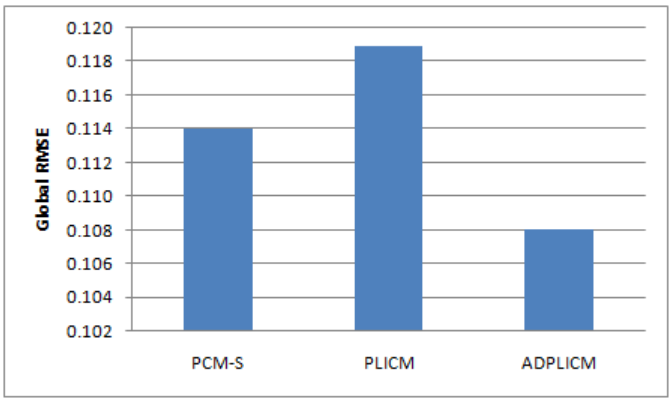


Figure 6.23: Global RMSE values estimated on the outputs of PCM based local spatial information classifiers to compare the performance of the algorithms in image detail retention

The graph in Figure 6.23, affirmed that the image details were lost when classifying the input dataset with the PLICM classifier. ADPLICM algorithm was seen to perform better among the PCM based local information classifiers in retaining the image details.

6.4. Supervised classification with single class

The conventional FCM algorithm and the FCM based local spatial information classification algorithms cannot be used to extract single class from satellite images because of the membership constraint defined in equation (3.3c). This constraint forces the membership values of the pixels in the single output fraction image generated to be 1. For the extraction of single LULC class from the input dataset, conventional PCM and the PCM based local spatial information classification algorithms were used (PCM-S, PLICM, and ADPLICM). The training data for a single class was given as input to the classifiers and output fraction images for the conventional PCM and the three local convolution information classification algorithms were generated. For varied values of *m,* mean membership difference value was calculated to estimate the performance of the classification algorithms. Figure 6.24 shows the graph of mean membership value with varied *m* values for the classifiers.

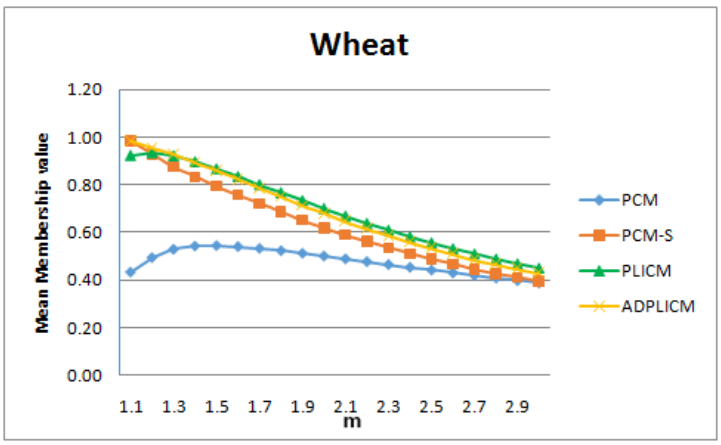
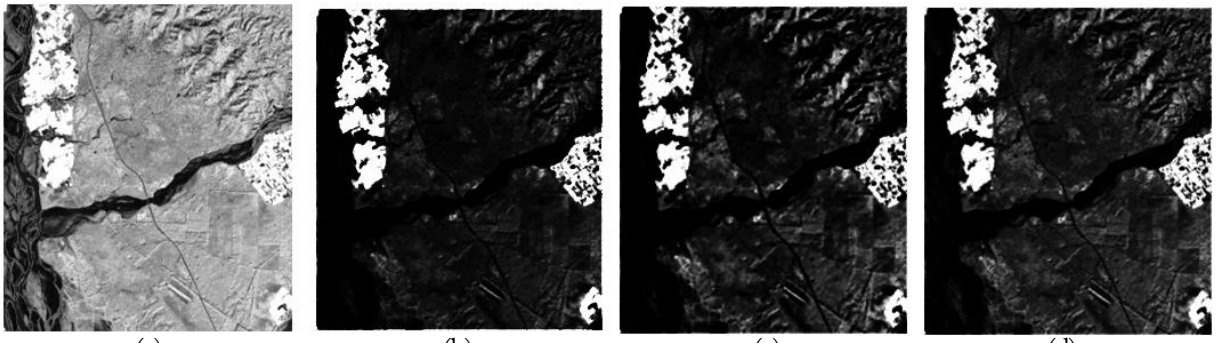


Figure 6.24: Mean membership difference values of conventional PCM and the PCM based local spatial information algorithms for varying *m* values

Figure 6.24 illustrates the superiority of the PCM based local spatial information classification algorithms against the PCM algorithm. For each value of *m*, the mean membership difference values of the three local spatial information classification algorithms (PCM-S, PLICM and ADPLICM) were higher. The output fraction images generated for each of the PCM based classification algorithms are presented in Figure 6.25, and the corresponding RMSE values obtained are given in Table 6.6.



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

Figure 6.25: The output fraction images for Wheat class from (a) PCM classifier, (b) PCM-S classifier, (c) PLICM classifier, and (d) ADPLICM classifier

	PCM	PCM-S	PLICM	ADPLICM
Global RMSE	0.515	0.379	0.270	0.279

Table 6.6: Global Root Mean Square Error for PCM based classifiers when single class ('Wheat') is extracted

Based on the RMSE value, PLICM/ADPLICM was observed to be the better performing classifier in extracting 'Wheat' class.

6.5. Local spatial information algorithm in handling noisy image

The local spatial information algorithms were experimented with a noisy image to evaluate their noise tolerance ability. The Formosat-2 image was corrupted by noisy pixels which had very different DN value from the surrounding pixels. The noisy pixels were generated by changing the DN values of random pixels in the image to '255' in all the bands and hence they are displayed as 'white' pixels in the noisy input image (Figure 6.26). It has been observed from the experiment conducted in section 6.2 that the fuzzy local

spatial information algorithms were capable of removing the isolated pixels and FCM-S and FLICM were observed to perform well in removing isolated pixels. In this experimentation, the FCM-S and its PCM counterpart PCM-S classifiers were used to observe the effect of local spatial information classification algorithms in removing noise. These algorithms were chosen because in FCM-S and PCM-S algorithms the effect from the neighbouring pixels for classification could be controlled, unlike the other local spatial information classifiers, by changing the parameter value of *a*. This is useful in controlling the classifier's tolerance to noise based on the noise intensity.

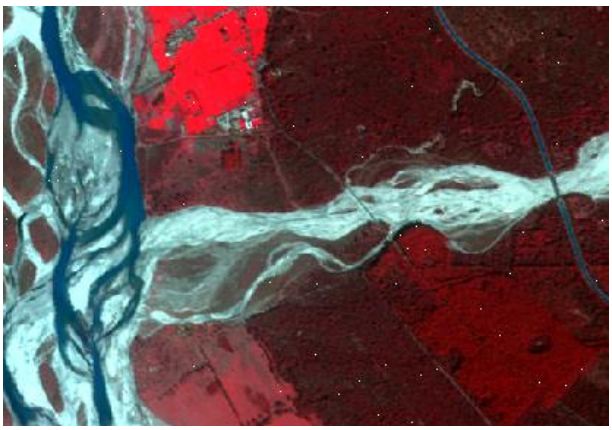
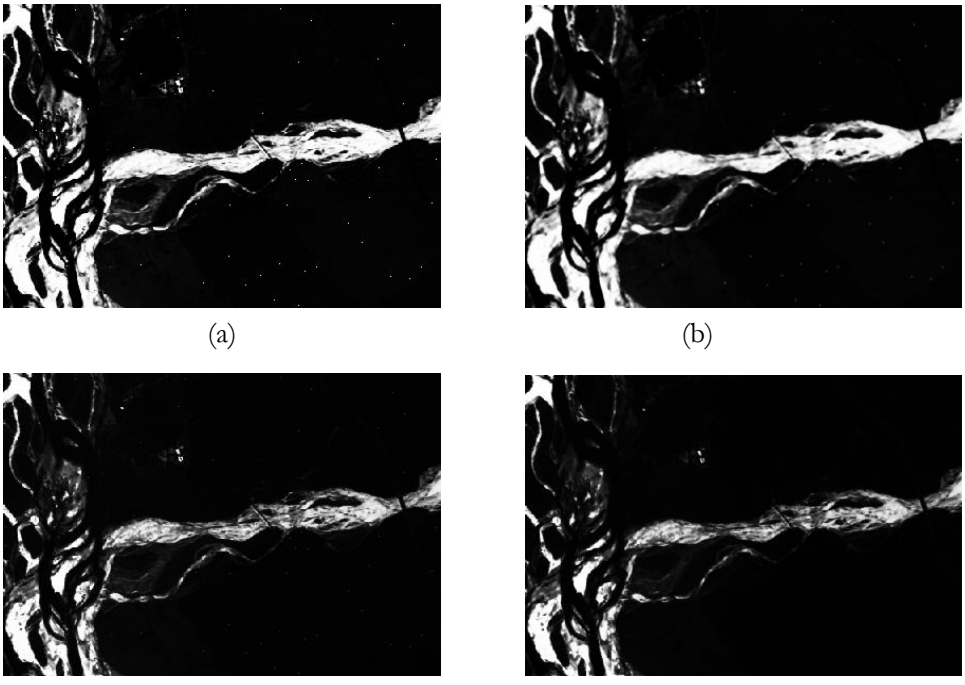


Figure 6.26: Formosat-2 image corrupted with noise

The input noisy image was classified with FCM, FCM-S, PCM and PCM-S classifiers with the training data of three classes 'Riverine sand', 'Water' and 'Wheat'. The fraction images of 'Riverine sand' class were examined for all the classifiers to understand their performance in the presence of noise. The outputs generated are furnished in Figure 6.27. The parameter *a* was set to 2 for FCM-S and 0.2 for PCM-S. The *m* value of 1.6 was used for all classifiers.



(a) (b)

(c) (d)

Figure 6.27: The fraction image generated for Riverine sand class when the noisy image is classified with (a) FCM classifier (b) FCM-S classifier (c) PCM classifier (d) PCM-S classifier

From the outputs, it can be seen that the noisy pixels were evident in the fraction image of class 'Riverine sand' in FCM classifier. This was because the noisy pixels were spectrally similar to the class 'Riverine sand' relative to the other classes and hence the FCM classifier had assigned a higher membership value to these pixels in 'Riverine sand' class. FCM-S classifier was found to reduce the noisy pixels in the output compared to the FCM classifier. The effect of noisy pixels was not much evident in the output image obtained with PCM based classifiers. This is because the PCM classifiers are naturally immune to noise and outliers as suggested by Krishnapuram & Keller (1993) in their work. The output of PCM-S was observed to have almost negligible noise.

6.6. Discussions

The choice of reference data and the accuracy assessment methods hugely influence the analysis and conclusions drawn from a study. There are no standard accuracy assessment methods available for quantitatively measuring the results of fuzzy classification outputs. This section discusses a few general observations on the results and the accuracy matrices obtained during the optimisation of parameters and the execution of the scenarios described in sections 6.2, 6.3, 6.4 and 6.5.

It was not possible to make a general comparison of the performance FCM based local spatial information classifiers with that of PCM based local spatial information classifiers since the fundamental behaviours of their corresponding base classifiers FCM and PCM were different. The fraction outputs in the FCM classifiers illustrated how well every image pixel was accommodated among the existing classes, while the fraction images of the classes in the PCM classifier focused on allocating every image pixel to that class. Due to this reason, the FCM classifier performed well when all the classes were defined, and the performance degraded when few classes were excluded from training. The performance of the PCM classifier was higher in the presence of untrained classes as the membership value in PCM classifier is calculated using absolute measure and not using relative measure. The FCM-based spatial information classifiers and PCM-based spatial information classifiers worked well in situations where their respective base classifiers performed better. Hence rather than a comparison of FCM-based local spatial information classifiers with PCM-based local spatial information classifiers, the assessments were done on the local spatial information classifiers with respect to the corresponding base classifier.

The examination on the optimisation of parameters brought to notice that the optimal values of *m* varied for the same classifier when the number of classes was altered. The meaning of *m* is different in FCM based classifiers and PCM based classifiers. Increasing the value of *m* in FCM increased the level of fuzziness in the output. This meant that on increasing *m* value, the sharing of pixels among all the available classes increased in the FCM classifier. But increasing *m* in case of PCM up to a certain value of *m,* which was selected as the optimal value for this research, the degree of a pixel to belong to the specific class increased. This was in accordance with the theory of the parameter *m* for PCM specified by Krishnapuram & Keller (1996). The other parameters *a* and *window size* for local spatial information classification algorithms had to be selected based on the classes to be extracted. Bigger values for *window size* meant more number of neighbouring pixels in determining the membership values of the centre pixels. Larger window size caused smoothened output and produced loss of information especially in the classes with small spatial extent ('Water', 'Wheat'). The parameter *a* in FCM-S and PCM-S algorithm controlled the weight of the information from the neighbouring pixels used for estimating the membership value of the central pixel. The value of *a* and *window size* were optimised in this research by considering all the classes.

From the execution of first scenario (section 6.2), it was observed that all the FCM-based local spatial information classifiers produced more or less similar overall accuracies and they were superior to the overall accuracy of the FCM classifier. Nevertheless, in the classes with lesser spatial extent like the 'Water' class, near the edges and boundaries there was loss of image details in the output fraction images compared to the outputs from the FCM classifier. This was due to smoothing at narrow parts of the class. In classes with larger spatial extent, such as 'Dense Forest' and 'Eucalyptus class' the local spatial information algorithms showed improved performance as the membership of the pixels were estimated by using the information from the neighbouring pixels thereby reducing the occurrence of isolated pixels in the output fraction images. Among the FCM-based classifiers, ADFLICM was found to retain image details.

Through the experiments in the second, third and fourth scenarios (section 6.3, section 6.4 and section 6.4), the performance of PCM based local spatial information algorithms were observed. The results from each of these scenarios demonstrated that the spatial information classifier outperformed the PCM classifier for classifications in the presence of untrained classes and classification for the extraction of a single class. The PCM classifier is naturally immune to noise and outliers and hence efficient in classification even for classification with few classes. Incorporating spatial information seemed to have further improved the performance of PCM classifier by forcing the classifier to assign a higher membership value to the pixels similar to a class and lower memberships to the outliers.

7. CONCLUSION AND RECOMMENDATIONS

This chapter presents the conclusion of the research, answers to research questions and recommendations for future work. Section 7.1 provides the conclusions drawn from studies carried out during the research. Section 7.2 gives answers to the research questions. Section 7.3 discusses few recommendations for further study.

7.1. Conclusion

The use of ancillary data such as spatial contextual information, texture information etc. to improve the classification performance is a much-researched area. The primary aim of this study was to investigate the effect of incorporating spatial contextual information in the FCM and PCM classifiers. The performance of the three FCM based local spatial information classification algorithms FCM-S, FLICM and ADFLICM algorithms and the three PCM based local spatial information classification algorithms, PCM-S, PLICM and ADPLICM classifiers developed in this research were examined. The performances of the local spatial information algorithms were compared with the respective base classifiers FCM and PCM. Multiple experiments were conducted to analyse the performances of the classifiers.

The local spatial information algorithms were observed to have few advantages over the conventional fuzzy classification algorithms. The local spatial information classifiers showed better accuracies in large homogenous classes as they were effective in handling the isolated pixels that occurred due to spectral confusions. Among the FCM-based local spatial information classifiers, FCM-S and FLICM algorithms showed higher overall accuracies. But the minute details on the images for the narrow and small classes such as in 'Water' and 'Wheat' class were lost due to over-smoothing. The ADFLICM algorithm was effective in handling isolated pixels as well as in preserving image details. The nature of the local spatial information classifiers was found to be similar in the case of FCM based algorithms and PCM based algorithms. Like in the case of FCM based classifiers, ADPLICM algorithm showed better performance in retaining the image details compared to the PCM-S and PLICM algorithms. On the whole, the PCM based local spatial information classification algorithms (PCM-S, PLICM and ADPLICM) outperformed the conventional PCM classifier in terms of RMSE values.

7.2. Answers to research questions

Q: How fuzzy local spatial information classifiers work on remotely sensed images? Which FCM based fuzzy local spatial information classifier works better in handling ambiguities due to spectral similarity/variation and preserving image details?

A: Fuzzy local spatial information classifiers were found to be effective in land use/land cover classification, and they produced output fraction images comparable to or having higher accuracies than the conventional classifiers. The choice of the classifier, however, should be based on the nature of the classes to be classified since local spatial information classifiers could possibly cause loss of information due to smoothing in classes with narrower or smaller spatial extent. The isolated pixels in the land use/land cover classes in homogenous areas were seen to be handled effectively by the local spatial

information classifiers. The ADFLICM classifier was found to work better in simultaneously handling the ambiguities caused to spectral similarity/variation and preserving image details.

Q: How is local spatial information incorporated in PCM classifier? How the local spatial information PCM classifiers perform compared to the conventional PCM classifier?

A: Local spatial information was incorporated in the PCM classifier by modifying the objective function of the conventional PCM classification algorithm to include a term that controls the effect from the neighbouring pixels. The details on developing PCM based spatial information classification algorithms are presented in Section 3.4 of the thesis. The local spatial information classifiers PCM-S, PLICM and ADPLICM were found to produce better classification outputs compared to the conventional PCM classifier.

Q: Which fuzzy local spatial information algorithm will give better classification results? Which algorithm is better in removing noise and preserving image details?

A: While it is not possible to comment which classifier gave better classification results all the time, it was observed that among FCM-based local spatial information classifier, ADFLICM algorithm was better in preserving image details. But the ability of the ADFLICM algorithm in noise removal was not as effective as that of FCM-S/FLICM algorithm but better than that of FCM classifier. The PCM based local spatial information algorithms PCM-S, PLICM and ADPLICM showed similar performance, yet ADPLICM algorithm was observed to be effective in preserving image details compared to the other two algorithms.

Q: What will be the optimal parameter values and widow size to balance the image noise and loss of image details?

A: The optimal parameter values estimated for supervised classification with all the classes are presented in Table 6.1.

7.3. Recommendations

- The performance of the PCM classifier and PCM-based local spatial information classifiers were found to be dependent on their initialization. The membership values for estimation of η_i was obtained thorough FCM classifier in this research. Other methods for determining accurate η_i estimate for each class with the available training data can be explored.
- The performance of the conventional PCM classifier was seen to improve after incorporating spatial information. The weight from the neighbouring pixels on the centre pixel were either calculated based on fixed parameter *a* in PCM-S and adaptively calculated in PLICM and ADPLICM with measures that use spatial Euclidean distance between the centre pixel and neighbouring pixels. Better methods such as the use of correlation between the pixels could be employed to control the effect from neighbouring pixels on the centre pixel.
- Methods for creation of soft reference data and accuracy assessment of the soft classified outputs are still open research areas

LIST OF REFERENCES

Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A., & Moriarty, T. (2002). A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, *21*(3), 193–199. https://doi.org/10.1109/42.996338

Atkinson, P. M., Cutler, M. E. J., & Lewis, H. (1997). Mapping sub-pixel proportional land cover with AVHRR imagery. *International Journal of Remote Sensing*, *18*(4), 917–935. https://doi.org/10.1080/014311697218836

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. https://doi.org/10.1080/014311697216847

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers.

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences*, *10*(2), 191–203. https://doi.org/https://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*(9), 935–945. https://doi.org/10.1016/S0167-8655(99)00061-6

Byju, A. P. (2015). *Non-Linear Separation of classes using a Kernel based Fuzzy c-Means (KFCM) Approach*. University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from https://library.itc.utwente.nl/papers_2015/msc/gfm/pbyju.pdf

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to Remote Sensing, Fifth Edition*. New York, United States: Guilford Publications. Retrieved from http://ebookcentral.proquest.com/lib/itc/detail.action?docID=843851

Cannon, R. L., Dave, J. V, Bezdek, J. C., & Trivedi, M. 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. https://doi.org/10.1109/TGRS.1986.289598

Chawla, S. (2010). *Possibilistic c - means - spatial contextual information based sub - pixel classification approach for multi - spectral data*. University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from http://www.itc.nl/library/papers_2010/msc/gfm/chawla.pdf

Chen, C. F., & Lee, J. (2001). The validity measurement of fuzzy c-means classifier for remotely sensed images. In *Proceedings of the 22nd Asian Conference on Remote Sensing* (pp. 208–210). Singapore.

Chen, S., & Zhang, D. (2004). Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*. https://doi.org/10.1109/TSMCB.2004.831165

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Congalton, R. (2009). *Assessing the Accuracy of Remotely Sensed Data*. Boca Raton: CRC Press. https://doi.org/10.1201/9781420055139

Dehghan, H., & Ghassemian, H. (2006). Measurement of uncertainty by the entropy: application to the classification of MSS data. *International Journal of Remote Sensing*, *27*(18), 4005–4014. https://doi.org/10.1080/01431160600647225

Dutta, A. (2009). *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). Retrieved from http://www.itc.nl/library/papers_2009/msc/gfm/dutta.pdf

Fisher, P. F., & Pathirana, S. (1993). The ordering of multitemporal fuzzy land cover information derived from landsat MSS data. *Geocarto International*, *8*(3), 5–14. https://doi.org/10.1080/10106049309354415

Foody, G. M. (1995). Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *50*(5), 2–12. https://doi.org/https://doi.org/10.1016/0924-2716(95)90116-V

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. https://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*. https://doi.org/10.1016/S0098-3004(99)00125-9

Foody, G. M., & Arora, M. K. (1996). Incorporating mixed pixels in the training, allocation and testing stages of supervised classifications. *Pattern Recognition Letters*, *17*(13), 1389–1398. https://doi.org/10.1007/978-3-319-49421-0_2

Gurney, C. M. (1981). The use of contextual information to improve land cover classification of digital remotely sensed data. *International Journal of Remote Sensing*, *2*(4), 379–388. https://doi.org/10.1080/01431168108948372

Gurney, C. M., & Townshend, J. (1983). The use of contextual information in the classification of remotely sensed data. *Photogrammetric Engineering & Remote Sensing*, *49*.

Ibrahim, M. A. (2004). *Evaluation of soft classifiers for remote sensing data*. IIT Roorkhe. Retrieved from http://shodhbhagirathi.iitr.ac.in:8081/jspui/image/pdf/web/viewer.html?file=/jspui/bitstream/123456789/1556/1/EVALUTION OF SOFT CLASSIFIERS FOR REMOTE SENSING DATA.pdf

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. https://doi.org/10.1080/01431160500057806

Jensen, J. R. (1996). *Introductory Digital Image Processing: A Remote Sensing Perspective (2nd Edition)*. Upper Saddle River, N.J. Prentice Hall,N.J.

Kamavisdar, P., Saluja, S., & Agrawal, S. (2013). A survey on image classification approaches and techniques. *International Journal of Advanced Research in Computer and Communication Engineering*, *2*(1), 1005–1009. https://doi.org/10.23883/IJRTER.2017.3033.XTS7Z

Krinidis, S., & Chatzis, V. (2010). A robust fuzzy local information C-Means clustering algorithm. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, *19*(5), 1328–1337. https://doi.org/10.1109/TIP.2010.2040763

Krishnapuram, R., & Keller, J. M. (1993). A Possibilistic Approach to Clustering. *IEEE Transactions on Fuzzy Systems*. https://doi.org/10.1109/91.227387

Krishnapuram, R., & Keller, J. M. (1996). The possibilistic C-means algorithm: Insights and recommendations. *IEEE Transactions on Fuzzy Systems*, *4*(3), 385–393. https://doi.org/10.1109/91.531779

Kumar, A. (2007). *Investgation in sub-pixel classification approaches for land use and land cover mapping*. IIT Roorkee. Retrieved from http://shodhbhagirathi.iitr.ac.in:8081/jspui/image/pdf/web/viewer.html?file=/jspui/bitstream/123456789/1601/1/INVESTIGATION IN SUB-PIXEL CLASSIFICATION APPROACHES FOR LAND USE AND LAND COVER MAPPING.pdf

Kumar, A., Ghosh, S. K., & Dadhwal, V. K. (2006). A comparison of the performance of fuzzy algorithm versus statical algorithm based sub-pixel classifier for remote sensing data. In *ISPRS Commission VII Mid-term synopsium "Remote Sensing: From pixels to processes"* (pp. 1–5). Enschede, Netherlands. Retrieved from http://www.isprs.org/proceedings/XXXVI/part7/PDF/153.pdf

Laha, A., Pal, N. R., & Das, J. (2006). Land Cover Classification Using Fuzzy Rules Through Evidence Theory. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(6), 1633–1641.

Li, M., Zang, S., Zhang, B., Li, S., & Wu, C. (2014). A review of remote sensing image classification techniques: The role of Spatio-contextual information. *European Journal of Remote Sensing*, *47*(1), 389–411. https://doi.org/10.5721/EuJRS20144723

Liew, A. W. C., Leung, S. H., & Lau, W. H. (2000). Fuzzy image clustering incorporating spatial continuity. *IEE Proceedings - Vision, Image, and Signal Processing*, *147*(2), 185. https://doi.org/10.1049/ip-vis:20000218

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2004). *Remote sensing and image interpretation* (5th ed). New York: Wiley & Sons.

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870. https://doi.org/10.1080/01431160600746456

Maselli, F., Rodolfi, A., & Conese, C. (1996). Fuzzy classification of spatially degraded thematic mapper data for the estimation of sub-pixel components. *International Journal of Remote Sensing*, *17*(3), 537–551. https://doi.org/10.1080/01431169608949026

Melgani, F., & Serpico, S. B. (2002). A statistical approach to the fusion of spectral and spatio-temporal

contextual information for the classification of remote-sensing images. *Pattern Recognition Letters*, *23*(9), 1053–1061. https://doi.org/10.1016/S0167-8655(02)00052-1

Moser, G., Serpico, S. B., & Benediktsson, J. A. (2013). Land-cover mapping by markov modeling of spatial-contextual information in very-high-resolution remote sensing images. In *Proceedings of the IEEE* (Vol. 101, pp. 631–651). https://doi.org/10.1109/JPROC.2012.2211551

Pal, N. R., & Bezdek, J. C. (1995). On Cluster Validity for the Fuzzy c-Means Model. *IEEE Transactions on Fuzzy Systems*. https://doi.org/10.1109/91.413225

Pham, D. L. (2001). Spatial models for fuzzy clustering. *Computer Vision and Image Understanding*, *84*(2), 285–297. https://doi.org/10.1006/cviu.2001.0951

Pham, D. L., & Prince, J. L. (1999). An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities. *Pattern Recognition Letters*, *20*(1), 57–68. https://doi.org/10.1016/S0167-8655(98)00121-4

Richards, J. A., & Jia, X. (1999). Geometric Enhancement Using Image Domain Techniques. In *Remote Sensing Digital Image Analysis: An Introduction* (pp. 113–132). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-03978-6_5

Richards, J. A., & Jia, X. (2012). *Remote Sensing Digital Image Analysis: An Introduction*. Berlin Heidelberg: Springer-Verlag.

Sabins, F. F. (1997). *Remote sensing: principles and interpretation* (3rd ed). New York: W.H. Freeman and Co.

Schowengerdt, R. A. (2014). Spatial Transforms. In *Revitalizing Industrial Growth in Pakistan: Trade, Infrastructure, and Environmental Performance* (pp. 53–65). https://doi.org/doi:10.1596/978-1-4648-0028-3_ch3

Sharma, S. (2018). *Role of hybrid spectral similarity measures for semi-supervised fuzzy classifier*. University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from http://www.itc.nl/library/papers_2018/msc/gfm/sharma.pdf

Singha, M. (2013). *Study the effect of discontinuity adaptive MRF models in fuzzy based classifier*. University of Twente Faculty of Geo-Information and Earth Observation (ITC). Retrieved from http://www.itc.nl/library/papers_2013/msc/gfm/singha.pdf

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. https://doi.org/10.1016/S0034-4257(99)00083-8

Tso, B., & Mather, P. M. (2009). *Classification methods for remotely sensed data*. London and New York: Taylor & Francis. https://doi.org/10.4324/9780203303566

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

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

Zhang, H., Wang, Q., Shi, W., & Hao, M. (2017). A Novel Adaptive Fuzzy Local Information C -Means Clustering Algorithm for Remotely Sensed Imagery Classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(9), 5057–5068. https://doi.org/10.1109/TGRS.2017.2702061

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. https://doi.org/10.1080/014311698214479

Zhang, J., & Foody, G. M. (2001). Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches. *International Journal of Remote Sensing*, *22*(4), 615–628. https://doi.org/10.1080/01431160050505883

APPENDIX A

IMPLEMENTATION OF THE ALGORITHMS IN PYTHON: ADPLICM algorithm

.

The python code for ADPLICM algorithm developed is presented here for reference.

```
import numpy
import math

def ADPLICM_out(image, fuzzifier, training_data, window):
    m = fuzzifier
    v = training_data
    cl = v.shape[0] # No of classes for classifcation
    M=image.width
    N=image.height
    nb=image.count
    limit = window//2

    img_band=numpy.zeros((nb,N,M))
    for i in range(nb):
        img_band[i,]=image.read(i+1)

    EU_D2=numpy.zeros((cl,N,M))
    NB_D2=numpy.zeros((cl,N,M))

# Euclidean distance between the pixels and cluster centers

    for i in range(cl):
        Difference=numpy.zeros((N,M))
        for j in range(nb):
            Difference=Difference+((img_band[j,]-v[i,j])**2)
        EU_D2[i,]=Difference

#Initializing the membership value using PCM

    #initial membership values

    Membership_val=numpy.zeros((cl,N,M))

    for i in range(cl):
        Denominator =numpy.zeros((N,M))
        for j in range (cl):
            Denominator = Denominator + ((EU_D2[i,]/EU_D2[j,])**(1/(m-1)))
        Membership_val[i,]=1/Denominator

    #Calculating parameter value (eta) for PCM
```

```
    H=numpy.zeros(cl)
    temp1=numpy.zeros((N,M))
    temp2=numpy.zeros((N,M))

    for i in range(cl):
        temp2 = Membership_val[i,]**m
        temp1 = temp2*EU_D2[i,]
        H[i] = temp1.sum()/temp2.sum()

    #Calculating new membership value

    Membership_val_pcm=numpy.zeros((cl,N,M))
    Denom =numpy.zeros((N,M))

    for i in range(cl):
        Denom = 1 + ((EU_D2[i,]/H[i])**(1/(m-1)))
        Membership_val_pcm[i,]=1/Denom

# calculation of neighbor factor for each pixel

    for i in range(cl):
        for n in range(limit,N-limit):
            for m1 in range(limit,M-limit):
                nb_difference=0
                count=0
                for wn in range(-limit, limit+1):
                    for wm in range(-limit,limit+1):
                        if wn==0 and wm==0:
                            continue
                        else:
                            Distance_2 =(wn**2)+(wm**2)
                            Sir= (Membership_val_pcm[i,n,m1]*
Membership_val_pcm[i,n+wn,m1+wm])/Distance_2
                            nb_difference=nb_difference+((1-Sir)*EU_D2[i,n+wn,m1+wm])
                            count=count+1
                NB_D2[i,n,m1]= (1/count)*nb_difference


#Calculating parameter value (eta) for ADPLICM

    H1=numpy.zeros(cl)
    temp11=numpy.zeros((N,M))
    temp21=numpy.zeros((N,M))

    for i in range(cl):
        temp21 = Membership_val_pcm[i,]**m
        temp11 = temp21*(EU_D2[i,])
```

```
        H1[i] = temp11.sum()/temp21.sum()

# Calculation of membership values

    Membership_val_final=numpy.zeros((cl,N,M))
    Denominator1 =numpy.zeros((N,M))

    for i in range(cl):
        Denominator1 = 1 + (((EU_D2[i,]+NB_D2[i,])/H1[i])**(1/(m-1)))
        Membership_val_final[i,]=1/Denominator1

    return (Membership_val_final)
```